Understanding the Strength of General-Purpose Cutting Planes

Marco Molinaro

Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Gérard Cornuéjols (Chair)
Sanjeeb Dash
Santanu Dey
François Margot
R. Ravi

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Algorithms, Combinatorics and Optimization.

Contents

Acknowledgements

First and foremost, I must thank my advisor, Gérard Cornuéjols, for his unending support, advice, patience, flexibility and mentorship. I would also like to thank my committee members Sanjeeb Dash, Santanu Dey, François Margot and R. Ravi for their thoughtful comments and help throughout the dissertation process.

I have been very fortunate to be a part of the ACO group at Carnegie Mellon. I would like to thank its faculty and students for all that they have thought me. I must specially thank Anupam Gupta for all the inspiring conversations. I am also indebted to Eduardo Laber, Ruy Milidiú and Marcus Poggi at PUC-Rio for their continuing guidance since my undergraduate years.

I must thank all my collaborators, that have made the process of doing research much more fun. I would also like to thank the research groups at IBM T. J. Watson and IBM Almaden, specially Oktay Günlük and David Woodruff, for my great summer internship experiences. I am also very grateful for the support I have received during my graduate studies from NSF, ONR and IBM.

Finally, I would not have been able to reach this stage in life without the constant support of all my friends, my parents Elizabeth and Rogério Molinaro, my sister Bruna Molinaro, and my wife Ishani Aggarwal.

Abstract

Cutting planes for a mixed-integer program are linear inequalities which are satisfied by all feasible solutions of the latter. These are fundamental objects in mixed-integer programming that are critical for solving large-scale problems in practice. One of the main challenge in employing them is that there are limitless possibilities for generating cutting planes; the selection of the strongest ones is crucial for their effective use. In this thesis, we provide a principled study of the strength of general-purpose cutting planes, giving a better understanding of the relationship between the different families of cuts available and analyzing the properties and limitations of our current methods for deriving cuts.

We start by analyzing the strength of disjunctive cuts that generalize the ubiquitous split cuts. We first provide a complete picture of the containment relationship of the split closure, second split closure, cross closure, crooked cross closure and t-branch split closure. In particular, we show that rank-2 split cuts and crooked cross cuts are neither implied by cross cuts, which points out the limitations of the latter; these results answer questions left open in [56, 65]. Moreover, given the prominent role of relaxations and their computational advantages, we explore how strong are cross cuts obtained from basic and 2-row relaxations. Unfortunately we show that not all cross cuts can be obtained as cuts based on these relaxation, answering a question left open in [56]. One positive message from this result, though, is that cross cuts do not suffer from the limitations of these relaxations.

Our second contribution is the introduction of a probabilistic model for comparing the strength of families of cuts for the continuous relaxation. We employ this model to compare the important split and triangle cuts, obtaining results that provide improved information about their behavior. More precisely, while previous works indicated that triangle cuts should be much stronger than split cuts, we provide the first theoretical support for the effect that is observed in practice: for most instances,

these cuts have the same strength.

In our third contribution, we study the multi-dimensional infinite relaxation introduced by Gomory and Johnson in the late 60's, which has been an important tool for analyzing and obtaining insights on cutting planes. The celebrated Gomory-Johnson's 2-Slope Theorem gives a sufficient condition for a cut to be facet defining from the 1-row infinite relaxation. We provide an extension of this result for the k-row case, for arbitrary k, which we call the $(k+1)$-*Slope Theorem*. Despite increasing interest in understanding the multi-row case, no such extension was known prior to our work. This result, together with the relevance of 2-slope functions for the 1-dimensional case, indicates that $(k+1)$-slope functions might lead to strong cuts in practice.

In our fourth contribution, we consider cuts that generalize Gomory fractional cuts but take into account upper bounds imposed on the variables. More specifically, we revisit the *lopsided cuts* obtained recently by Balas and Qualizza via a disjunctive procedure. We give a geometric interpretation of these cuts, viewing them as cuts for the infinite relaxation that are strengthened by a geometric lifting procedure. Using this perspective, we are able to generalize these cuts to obtain a family of cuts which has on one end the GMI cut, and on the other end the lopsided cuts. We show that all these cuts are "new", namely they are all facets of the infinite relaxation with upper bounded basic variable. We conclude by presenting preliminary experimental results, which unfortunately shows that these cuts decrease in importance as they move away from the GMI inequality, complementing the experimental results from Balas and Qualizza.

In our final contribution, we further explore properties and characterizations of split cuts, focusing on a general model of mixed-integer corner relaxation. The backbone of this work is a description of the split cuts for this relaxation from the perspective of cut-generating functions; this essentially establishes the equivalence of split cuts and (a generalization of) the k-cuts [50]. As our previous result, this characterization is obtained using the geometric lifting idea, illustrating its flexibility as a tool for analyzing cuts. As a consequence, we show that every split cut for a corner relaxation is the restriction of a split cut for the mixed-integer infinite relaxation, which further indicates the universality of the latter. As another consequence, we construct a pure-integer set with arbitrarily weak split closure, giving a pure-integer counterpart of the mixed-integer construction from [27].

Part I

Introduction and Known Results

CHAPTER 1

Integer Programming and Cutting Planes

Mixed-integer programming is a modeling framework that can be naturally used to describe problems involving discrete decisions. A *mixed-integer program* (MIP) comprises a linear function to be maximized (or minimized) subject to a collection of linear constraints over the decision variables, with additional constraints enforcing that some variables need to be integral. More precisely, given rational[1] vectors $c \in \mathbb{Q}^n$ and $d \in \mathbb{Q}^m$, a matrix $A \in \mathbb{Q}^{n \times r}$ and a vector $b \in \mathbb{Q}^r$, the goal is to find a setting (x^*, y^*) of the decision variables (x, y) that solve the following problem:

$$\begin{aligned} \max\ & cx + dy \\ \text{s.t. }\ & Ax + By \leq b \\ & x \in \mathbb{Z}^n. \end{aligned} \tag{1.1}$$

The set

$$P \triangleq \{(x, y) \in \mathbb{Z}^n \times \mathbb{R}^m : Ax + By \leq b\} \tag{1.2}$$

is called the set of *feasible solutions* of (1.1).

Despite their apparent simplicity, MIPs can be used to model a multitude of problems and processes of very different nature [108]. A prototypical problem that

[1] MIPs with irrational data can also be defined. However, some of the results concerning the structure of MIPs and cutting planes depend on this rationality assumption, so to avoid further technicalities we will only deal with rational MIPs

can be modeled by a MIP is the *Traveling Salesman Problem*, which can be described as follows: a salesperson wants to visit N different cities to sell its products; the goal is to find the best order for visiting the cities so as to minimize the total distance traveled. However, many other situations, even those including non-linear objectives and constraints, may be modeled using this framework [114].

1.1 Solving MIPs

Given the modeling power of these programs, it is not surprising that solving a general MIP is computationally hard [97]. Nonetheless, the usefulness of MIPs has driven intensive research aiming at either solving special classes of MIPs (e.g. [100, 103, 74]) or deriving heuristics that perform well in instances that appear in practice. Throughout this thesis, we focus on the latter objective. Although many strategies for solving general MIPs have been studied [96], by far the most popular one involves a combination of two ideas: *branch-and-bound* and *cutting planes*. This strategy is typically dubbed *branch-and-cut* and is implemented in most general purpose MIP solvers like CPLEX, EXPRESS or GUROBI .

Both these ideas rely on the following observation: linear programs, that is, MIPs without integrality constraints, can be much more easily solved [37]. It then becomes interesting to consider the *linear programming (LP) relaxation* of the MIP (1.1) where integrality constraints are dropped:

$$\begin{aligned} &\max\ cx + dy \\ &\ \text{s.t.}\ \ Ax + By \leq b, \end{aligned} \tag{1.3}$$

and its feasible set

$$P^{LP} \triangleq \{(x, y) \in \mathbb{R}^{n+m} : Ax + By \leq b\}.$$

Although solving the above linear relaxation gives some information about the optimal solution of the original MIP (and in particular provides an upper bound on the optimal value of the latter), the obtained solution (x^{LP}, y^{LP}) might not be feasible for (1.1) due to the non-integrality of x^{LP}; the necessity of excluding this point from consideration motivates the following.

In order to move towards a feasible solution to the original MIP, a typical *branch-and-bound* strategy [108] splits the space into regions $R_1, R_2, \ldots, R_k$ whose union contains all the feasible solutions P, but excludes the fractional solution (x^{LP}, y^{LP}).

An LP relaxation over each region R_i is computed, and those that have too small an optimal value are discarded. The procedure then continues on each of the remaining regions, further subdividing them based on the optimal solution of the linear relaxations.

Of central importance for this thesis is the concept of *cutting planes* [108]. The idea is to strengthen the linear relaxation (1.3) by adding new linear inequalities $fx + gy \leq h$ which do not "cut off" any of the feasible solutions for (1.1), that is, every solution $(x, y) \in P$ satisfies $fx + gy \leq h$. The inequality $fx + gy \leq h$ is called a *valid cut*, *cutting plane* or *valid inequality* for (1.1).

Although these valid inequalities cannot eliminate all the spurious points in $P^{LP} \setminus P$, we have the following fundamental property: if we augment the linear relaxation P^{LP} by adding *all* valid inequalities and solve the obtained program, this gives an optimal solution for the original problem (1.1). Therefore, it is conceivable to solve an MIP by solving its linear relaxation and progressively augmenting it with more and more valid inequalities; this is the *(pure) cutting plane* method. The pure cutting plane method has some theoretical appeal, but even for problems of medium size this method is currently taken to be impractical. However, another way of using valid inequalities is to incorporate them in a branch-and-bound procedure, where the linear relaxations of the different regions are strengthened through the addition of cuts; this yields the *branch-and-cut* procedure.

We remark, though, that there are infinitely many valid inequalities for (1.1), which touches upon the central theme of this thesis: *which valid inequalities should we use?* This choice heavily depends on our ability to recognize how *strong* a valid inequality is. Our main goal in this thesis is precisely to obtain a better understanding of the strength of valid inequalities.

1.2 The Role of Cutting Planes in Solving MIPs

The combination of branch-and-bound with cutting planes tailored for specific problems has been largely successfully in solving special classes of MIPs. A typical example of the efficiency of this method is illustrated by the Traveling Salesman Problem, where instances with almost 100,000 locations are solved to optimality; these techniques have enough power to find the optimal tour over all 24,978 cities, towns, and villages in Sweden [46].

The combination of branch-and-bound with *general purpose* cutting planes is also currently successful in solving some large scale MIPs. The state of affairs, however, was not always like this. Until the early 90s "the research community was unanimous: In order to solve integer programs of meaningful sizes, one had to exploit the structure of the underlying combinatorial problem" [48]. In particular, general purpose cutting planes were seen as mostly of theoretical interest. However, in the mid 90s a breakthrough came with a series of papers by Balas, Ceria and Cornuéjols (and also Natraj), where they obtained striking computational results using these cuts within the branch-and-bound framework [14, 15, 16]. These papers show the efficacy of *GMI* cuts (introduced back in the 60s [84]) and the newly discovered *lift-and-project* cuts, *when properly employed*. Both of these cuts are part of the family of cuts called *split cuts*, which will play a central role in this thesis. This development has since revived the importance of understanding the functioning of general purpose cutting planes. See [102] for an excellent survey of the evolution of MIP solvers.

CHAPTER 2

Perspectives on Cutting Planes

In this section we start a more formal development of cutting planes. We assume familiarity with basic polyhedral theory and integer programming; see [108, 113] for background material.

Here we can classify the study and development of cutting planes into roughly four different perspective: *disjunctive cuts*, *intersection cuts*, *cuts for corner and related relaxations* and *cuts from other relaxations*. One interesting feature is that there is a large interplay between these perspectives and most cuts can be analyzed and obtained using either of them; however, one perspective may present a much clearer picture of certain properties of the cuts under study.

2.1 Disjunctive Cuts

In order to make the discussion more precise, we will also need the following formal definition of a *mixed-integer set*.

Definition 1. *A* mixed-integer set *is a pair of the form* $Q = (K, \mathbb{Z}^m \times \mathbb{R}^n)$*, where* K *is a convex set in* $\mathbb{R}^{m+n}$*. In the special case* $n = 0$*, we call this pair an* integer set*. We will also often make use of the identification* $Q \equiv K \cap \mathbb{Z}^m \times \mathbb{R}^n$*.*

Typically such mixed-integer set Q will be used to represent the feasible solutions of a MIP, and K is set to be its linear relaxation. From now on, we formally redefine the set of feasible solutions given in (1.2) to be the appropriate mixed-integer set

$P \triangleq (P^{LP}, \mathbb{Z}^n \times \mathbb{R}^m)$. This slightly cumbersome definition is required because some of the constructions depend on the specific LP relaxation used, and not only on the set of feasible solutions.

In a highly influential paper, Balas introduced the concept of *disjunctive programming* [12] (see also [13]). This is a modeling framework whose main idea is to handle more explicitly constraints of the form "either or", present in many optimization problems. Indeed, these constraints are implicitly present in the integrality constraints of MIP (1.1): in any feasible solution $(x, y) \in P$, we have, say, *either* $x_1 \leq 0$ *or* $x_1 \geq 1$.

2.1.1 Split Cuts: the Simplest Disjunctive Cuts

Although we do not attempt to discuss disjunctive programming in full generality here, we illustrate how the simple connection with MIPs outlined above can be employed to give the most useful cuts in practice (as mentioned in Chapter 1): *split cuts*.

Consider a mixed-integer set $P = (P^{LP}, \mathbb{Z}^n \times \mathbb{R}^m)$ (identified with the feasible set (1.2)). Given an integral vector $\pi \in \mathbb{Z}^n$ and an integer $\gamma \in \mathbb{Z}$, let

$$D(\pi, \gamma) \triangleq \{(x, y) \in \mathbb{R}^{n+m} : \pi x \leq \gamma\} \cup \{(x, y) \in \mathbb{R}^{n+m} : \pi x \geq \gamma + 1\} \tag{2.1}$$

denote the set of points (x, y) that either satisfy $\pi x \leq \gamma$ or $\pi x \geq \gamma + 1$. Such set is called a *split disjunction* for the mixed integer set P. Geometrically, we are considering the parallel hyperplanes $\pi x = \gamma$ and $\pi x = \gamma + 1$ and, informally, focusing on points on one side of one hyperplane, and on the opposite side of the other hyperplane.

Again because of the integrality constraints, every feasible solution $(x, y) \in P$ belongs to $D(\pi, \gamma)$ (note that the example in the beginning of this section corresponds to taking $\pi = (1, 0, \ldots, 0)$ and $\gamma = 0$). Since these solutions also belong to the linear relaxation P^{LP}, we have the inclusion $P \subseteq P^{LP} \cap D(\pi, \gamma)$. Importantly, the set $P^{LP} \cap D(\pi, \gamma)$ is typically strictly smaller than P^{LP}, and in this case the disjunction $D(\pi, \gamma)$ effectively excludes spurious points from $P^{LP} \setminus P$; recall that this is exactly what we want from cutting planes.

We call a linear inequality $fx + gy \leq h$ a *split cut* for P with respect to the disjunction $D(\pi, \gamma)$ if it is valid for $P^{LP} \cap D(\pi, \gamma)$ [47]. An important observation that will be used in later sections is that adding together all split cuts with respect to $D(\pi, \gamma)$ gives the convex hull $\text{conv}(P^{LP} \cap D(\gamma, \pi))$, namely

$$\text{conv}(P^{LP} \cap D(\pi, \gamma)) = P^{LP} \cap \bigcap \{(x, y) \in \mathbb{R}^{n+m} : fx + gy \leq h\},$$

where the last intersection is taken over all split cuts $fx + gy \leq h$ with respect to $D(\pi, \gamma)$.

Split cuts in this form were introduced by Cook, Kannan and Schrijver [47] based on the disjunctive programming work of Balas. Despite its simplicity, the split cut family encompasses many other well-known families of cuts [104]. In particular, as mentioned in Section 1.2, the GMI and lift-and-project cuts responsible for the major improvement of MIP solvers in the 90s are all split cuts. Another notable example is that split cuts are equivalent to the *mixed-integer rounding cuts* [107], which were introduced independently in [108][1].

Because of its power and simplicity, split cuts have received much attention from researchers, and a large body of work concerning split cuts is already available. In particular, the practical success of split cuts is partially due to the fact that there are very efficient ways of generating a “small” collection of good split cuts. Notice again that there are infinitely many split cuts, even with respect to a single split disjunction. The first significant advance in the generation of split cuts was obtained by Balas and Peregaard [20], which established a connection between lift-and-project cuts and pivots over the simplex tableaux to provide an efficient cut generator[2]. Exciting recent advances have been obtained by Dash and Goycoolea [59], Bonami [39] and Fischetti and Salvagnin [80] to obtain even better results. These procedures are able to harness almost the full power of split cuts in a computationally efficient way (see also [23]).

There has also been much work in measuring and understanding the strength of split cuts and comparing it with other families of cuts, both empirically and theoretically. Since this is the central topic of this thesis, we devote Chapter 3 to explore these results in more detail.

2.1.2 More General Disjunctive Cuts

Given that we already have a refined understanding of split cuts and can harness their power, the next step is generalize this construction to obtain even stronger

[1]MIR cuts can be seen as a derivation of split cuts using the *simple set relaxation* explored in Section 2.4.2.

[2]This connection is related to the fact that essentially every split cut can be obtained from basic relaxations [3].

cuts. The development of split cuts outlined above can indeed be extended to give more general *disjunctive cuts*. Instead of working with the disjunctions of the form $D(\pi, \gamma)$, one can actually consider a general disjunction $D = \bigcup_{i=1}^{k} D_i$ (where each D_i is, say, a polyhedron) satisfying that no feasible solution is excluded by it: $P \subseteq D$.

Arguably the simplest such disjunction, after the split one, is the *cross disjunction*: given $\pi^1, \pi^2 \in \mathbb{Z}^n$ and $\gamma_1, \gamma_2 \in \mathbb{Z}$, define

$$D(\pi^1, \pi^2, \gamma_1, \gamma_2) \triangleq D(\pi^1, \gamma_1) \cap D(\pi^2, \gamma_2).$$

The fact that $P \subseteq D(\pi^1, \pi^2, \gamma_1, \gamma_2)$ follows again either directly from the presence of integrality constraints in (1.1) or from the fact that $P \subseteq D(\pi^i, \gamma_i)$ for $i = 1, 2$. A linear inequality is a *cross cut* for P with respect to the disjunction $D(\pi^1, \pi^2, \gamma_1, \gamma_2)$ if it is valid for $P^{LP} \cap D(\pi^1, \pi^2, \gamma_1, \gamma_2)$ [56].

Importantly, cross cuts provide a tighter relaxation of P: notice for instance that $P^{LP} \cap D(\pi^1, \pi^2, \gamma_1, \gamma_2) \subseteq P^{LP} \cap D(\pi^i, \gamma_i)$ for $i = 1, 2$. Therefore, cross cuts are hopeful candidates to provide even better performance than split cuts, while its simple structure means that it is possible to obtain a good understanding of some of its important properties.

Another important family of disjunctive cuts is the *crooked cross cuts* family [56]. Given $\pi^1, \pi^2 \in \mathbb{Z}^n$ and $\gamma_1, \gamma_2 \in \mathbb{Z}$ define the sets

$$\begin{aligned}
D_1^c(\pi^1, \pi^2, \gamma_1, \gamma_2) &= \{(x, y) \in \mathbb{R}^{n+m} : \pi^1 x \leq \gamma_1, (\pi^2 - \pi^1)x \leq \gamma_2 - \gamma_1\},\\
D_2^c(\pi^1, \pi^2, \gamma_1, \gamma_2) &= \{(x, y) \in \mathbb{R}^{n+m} : \pi^1 x \leq \gamma_1, (\pi^2 - \pi^1)x \geq \gamma_2 - \gamma_1 + 1\},\\
D_3^c(\pi^1, \pi^2, \gamma_1, \gamma_2) &= \{(x, y) \in \mathbb{R}^{n+m} : \pi^1 x \geq \gamma_1 + 1, \pi^2 x \leq \gamma_2\},\\
D_4^c(\pi^1, \pi^2, \gamma_1, \gamma_2) &= \{(x, y) \in \mathbb{R}^{n+m} : \pi^1 x \geq \gamma_1 + 1, \pi^2 x \geq \gamma_2 + 1\}.
\end{aligned}$$

We call the set

$$D^c(\pi^1, \pi^2, \gamma_1, \gamma_2) \triangleq \bigcup_{i=1}^{4} D_i^c(\pi^1, \pi^2, \gamma_1, \gamma_2)$$

a *crooked cross disjunction* for P. A linear inequality is a *crooked cross cut* for P if it is valid for $P^{LP} \cap D$ for some crooked cross disjunction D.

Cross and crooked cross cuts have obtained recent attention, starting with a paper by Dash, Dey and Günlük [56]. This direction was further explored theoretically and computationally in [57, 65], showing that these cuts are indeed promising and are able to improve over split cuts in practical instances.

The final family of disjunctive cuts that we mention here is the direct generalization of split and cross cuts, called *t-branch cuts* [101]. Consider an integer t together with $\pi^i \in \mathbb{Z}^m$ and $\gamma_i \in \mathbb{Z}$ for $i = 1, \ldots, t$. The set $D(\pi^1, \ldots, \pi^t, \gamma_1, \ldots, \gamma_t)$ given by

$$D(\pi^1, \ldots, \pi^t, \gamma_1, \ldots, \gamma_t) \triangleq \bigcap_{i=1}^{t} D(\pi^i, \gamma_i)$$

is called a *t-branch split disjunction* for P. The fact that $\mathbb{Z}^n \times \mathbb{R}^m \subseteq D(\pi^i, \gamma_i)$ implies that $P \subseteq D(\pi^1, \ldots, \pi^t, \gamma_1, \ldots, \gamma_t)$. A linear inequality is a *t-branch split cut* for P with respect to a t-branch split disjunction D if it is valid for $P^{LP} \cap D$.

In Section 3.2.1 we discuss known results about the strength of all these disjunctive cuts.

2.2 Intersection and Generalized Intersection Cuts

One can see disjunctive cuts from a slightly different perspective. Consider again, for instance, the split disjunction $D(\pi, \gamma)$ given by equation (2.1). Now notice that we can see this disjunction as the removal of what is between two parallel hyperplanes: $D(\pi, \gamma) = \mathbb{R}^{n+m} \setminus S(\pi, \gamma)$, where the *split set* $S(\pi, \gamma)$ is defined as $\{(x, y) \in \mathbb{R}^{n+m} : \gamma < \pi x < \gamma + 1\}$. Therefore, the effect of split cuts for P with respect to $D(\pi, \gamma)$ is

$$\text{conv}(P^{LP} \cap D(\pi, \gamma)) = \text{conv}(P^{LP} \setminus S(\pi, \gamma)),$$

namely removing the set $S(\pi, \gamma)$ from P^{LP} and then convexifying. Notice that the set $S(\pi, \gamma)$ is convex. We can see cross cuts in a similar way: given a cross disjunction $D(\pi^1, \pi^2, \gamma_1, \gamma_2)$, its effect on P^{LP} is $\text{conv}(P^{LP} \setminus (S(\pi^1, \gamma_1) \cup S(\pi^2, \gamma_2)))$, where now the removed set $S(\pi^1, \gamma_1) \cup S(\pi^2, \gamma_2)$ is typically non-convex. One difficulty when trying to generate cuts this way (or using disjunctions directly) is that it is hard to know which cuts are valid for, say, $P^{LP} \setminus (S(\pi^1, \gamma_1) \cup S(\pi^2, \gamma_2))$. Intersection cuts deal with this difficulty by using two observations: i) it uses a *simplicial* conic relaxation $\mathcal{B}^{LP}$ instead of P^{LP}; ii) removes only *convex* sets from this relaxation.

2.2.1 Intersection Cuts

We start by describing the conic relaxation $\mathcal{B}^{LP}$, the so called *basic relaxation*. The idea is simple: we drop linear inequalities from the definition of P^{LP} and only keep a *linearly independent* set of them. More precisely, a *basic relaxation* of (1.2) is given by a set of the form

$$\mathcal{B} = \{(x, y) \in \mathbb{Z}^n \times \mathbb{R}^m : A'x + B'y \le b'\}, \tag{2.2}$$

where the rows of the matrix $[A', B']$ are linearly independent [3]. We also use $\mathcal{B}^{LP}$ to denote the linear relaxation of $\mathcal{B}$:

$$\mathcal{B}^{LP} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : A'x + B'y \leq b'\}.$$

Clearly $\mathcal{B}$ is a relaxation of P and $\mathcal{B}^{LP}$ is a relaxation of $\mathcal{B}^{LP}$, namely $P \subseteq \mathcal{B}$ and $P^{LP} \subseteq \mathcal{B}^{LP}$.

It is not difficult to verify that $\mathcal{B}^{LP} \subseteq \mathbb{R}^{n+m}$ is a full dimensional displaced simplicial cone, namely there a point v (the apex) and a set of linearly independent vectors $r^1, r^2, \ldots, r^{n+m}$ (the generators) such that $\mathcal{B}^{LP} = v + \text{cone}(r^1, r^2, \ldots, r^{n+m})$ ($\text{cone}(V)$ denotes the conic hull of a set V of vectors). It is instructive to think of the apex v as the optimal solution of the linear relaxation P^{LP}, which we want to cut off.

Now we describe the sets S that are allowed to be removed from the relaxation $\mathcal{B}^{LP}$ in order to obtain intersection cuts. We say that a set $S \subseteq \mathbb{R}^{n+m}$ is a *convex lattice-free set* with respect to $\mathbb{Z}^n \times \mathbb{R}^m$ if is satisfies two properties: 1) S is convex; 2) S does not contain any point of $\mathbb{Z}^n \times \mathbb{R}^m$ in its interior, namely $\text{int}(S) \cap Z^n \times \mathbb{R}^m = \emptyset$. We will also focus on convex lattice-free sets which contain the apex v in their interior. Again the motivation is that $\mathcal{B}^{LP} \setminus \text{int}(S)$ contains all the points of the original feasible set P but exclude the "current solution" v. Notice, in particular, that the set $S(\pi^1, \gamma_1)$ is a lattice-free set (and so is its topological closure). Any cut valid for $\mathcal{B}^{LP} \setminus \text{int}(S)$, for some basic relaxation $\mathcal{B}^{LP}$ and convex lattice-free set S containing the apex of $\mathcal{B}^{LP}$, is called an *intersection cut* [12].

The main advantage of working with the displaced simplicial cone $\mathcal{B}^{LP}$ instead of P^{LP} is that for any convex lattice-free set S with $v \in \text{int}(S)$, the remainder $\mathcal{B}^{LP} \setminus \text{int}(S)$ gives you exactly one cut; more precisely, there is exactly one inequality $fx + gy \leq h$ such that

$$\text{conv}(\mathcal{B}^{LP} \setminus \text{int}(S)) = \mathcal{B}^{LP} \cap \{(x, y) \in \mathbb{R}^{n+m} : fx + gy \leq h\}.$$

Importantly, it is usually easy to obtain an expression for this cut.

To see the geometry driving both of these observations, let us focus on the important case where all extreme rays of $\mathcal{B}^{LP}$ intersect the boundary ∂S, or more precisely, for all $i \in \{1, 2, \ldots, n+m\}$ there is $\lambda_i > 0$ such that $v + \lambda_i r^i \in \partial S$; let p^i be such point $v + \lambda_i r^i$. Then there is exactly one inequality $fx + gy \leq h$ which simultaneously is unsatisfied by v and which all points $p^1, p^2, \ldots, p^{n+m}$ satisfy at equality (notice that the fact that $\mathcal{B}^{LP}$ is simplicial implies that the points p^i's are affinely independent) and $\text{conv}(\mathcal{B}^{LP} \setminus \text{int}(S))$ is exactly $\mathcal{B}^{LP} \cap \{(x, y) \in \mathbb{R}^{n+m} : fx + gy \leq h\}$. Moreover, given the knowledge of the generators r^i's and S, one can easily find the intersection points p^i's (say, by solving linear programs, or exploring the specific structure of S) and then compute the cut $fx + gy \leq h$ by "fitting" it to the these points (say, by solving a system of linear equations). Actually, if we rewrite $\mathcal{B}^{LP}$ in a corner relaxation form, introduced in the next section, the cut can be computed in an even more efficient manner.

Intersection cuts were introduced by Balas [11] and in conjunction with the corner relaxation and related relaxations, has been the focus of much of the studies in the theory and practice of cutting planes. We defer the discussion of some of the results regarding intersection cuts to Section 2.3.2, where we formally define the corner relaxation.

2.2.2 Generalized Intersection Cuts

Although working with the conic relaxation $\mathcal{B}^{LP}$ offers the advantage of simplifying the actual cut generation procedure (once the relaxation and the set S are chosen), some power might be lost in the process of relaxing P^{LP} to $\mathcal{B}^{LP}$. Indeed, it has been shown recently that for some problems the cuts obtained from the construction $\mathcal{B}^{LP} \setminus \text{int}(S)$ (even running over all $\mathcal{B}^{LP}$ and S) are very weak [52]. In order to obtain reasonable cuts for such problems using this intersection cuts machinery, it then becomes necessary to iterate the process: start with the linear relaxation P^{LP}, obtain relaxations of the form $\mathcal{B}^{LP}$ and add some intersection cuts back to P^{LP} to obtain a stronger linear relaxation P_2^{LP} of P; then repeat the process starting with P_2^{LP} to obtain the stronger linear relaxation P_3^{LP}, and so on. Notice that the choice of cuts used in, say, the first iteration of this process affects the further relaxations P_i^{LP}'s, and it is hard to predict the behavior of the whole procedure. Moreover, this iterative nature often leads to numerical instabilities.

In order to obtain stronger cuts directly and reduce (or potentially avoid altogether) the necessity of iterating the cut generation procedure, Balas and Margot recently proposed the so called *generalized intersection cuts* [18]. The idea behind generalized intersection cuts is the following: given a convex lattice-free set S and any relaxation $\tilde{P}^{LP}$ of P^{LP}, we can still look at (the extreme points/rays of) the intersection $\tilde{P}^{LP} \cap \partial S$ (in the above discussion where $\tilde{P}^{LP} = \mathcal{B}^{LP}$, the extreme points of this intersection are exactly the p^i's) and construct cuts based on these intersection points (for instance, by selecting $n + m$ affinely independent such points and finding the unique linear equation containing them).

The main difficulty of a practical cut generation via such procedure is the choice of the appropriate relaxation $\tilde{P}^{LP}$, dealing with the potentially large number of intersection points that need to be maintained, and how to combine these points to actually generate the cuts. In [18], the authors propose to start with $\tilde{P}^{LP} = \mathcal{B}^{LP}$ as before, but then progressively adding back some of the original inequalities defining P^{LP} to obtain tighter relaxations while controlling the number of intersection points generated. Balas, Margot and Nadarajah [19] present further discussion on how to make this procedure computationally practical, and perform experiments with these cuts.

2.3 Corner and Related Relaxations

One of the difficulties of generating strong cuts for a general MIP like (1.1) is exactly its generality: there is very little structure to rely on. Because of that, one of the most successful strategies to generate cuts for (1.1) is to actually generate cuts for a *relaxation* of it, which was already hinted to in the previous section. The relaxations considered have a significantly simpler structure and hence it is easier to understand what are the strong cuts for them, namely those that we should generate in practice; also, their special structure typically allow for a more efficient cut generation.

Another important feature of deriving cuts from relaxations of a MIP is that it is more explicit what information is being used to derive such cuts. For instance, if a relaxation is relaxing the integrality of some of the variables and a cut is generated, then we know this cut did not take into account such integrality information. Some relaxations we consider indeed relax the integrality of some variables (e.g., the continuous relaxation in Section 2.3.4), while others relax the non-negativity of some variables (e.g., the corner relaxation in Section 2.3.2) and others are able to somehow ignore the specific structure of particular MIPs (e.g., the infinite relaxation in Section 2.3.3). In order to derive strengthened cuts, it is invaluable to known what additional information needs to be incorporated.

2.3.1 Generalized Corner Relaxation

We start by defining the most general family of MIPs that includes all other corner-type relaxations. Given a point $f \in \mathbb{R}^n \backslash \mathbb{Z}^n$ and sets $R, Q \subseteq \mathbb{R}^n$, define the *generalized* *corner relaxation* $\mathcal{C}(f, R, Q)$ as the set of solutions (x, s, y) to the system

$$\begin{aligned} x &= f + \sum_{r \in R} r \cdot s(r) + \sum_{q \in Q} q \cdot y(q) \\ x &\in \mathbb{Z}^n \\ y(q) &\in \mathbb{Z} \quad \forall q \in Q \\ s &\in \mathbb{R}_+^{\{R\}}, y \in \mathbb{R}_+^{\{Q\}}, \end{aligned} \tag{$\mathcal{C}(f, R, Q)$}$$

where $\mathbb{R}_+^{\{S\}}$ is the set of non-negative functions $h : S \to \mathbb{R}_+$ with finite support. We also define

$$\mathcal{C}_{LP}(f, R, Q) = \left\{ (x, s, y) : x = f + \sum_{r \in R} r \cdot s(r) + \sum_{q \in Q} q \cdot y(q), s \geq 0, y \geq 0 \right\} \tag{2.3}$$

to be the linear relaxation of $\mathcal{C}(f, R, Q)$.

A *valid cut* or *valid function* for a generalized corner relaxation $\mathcal{C}(f, R, Q)$ is a pair of non-negative functions $(\psi, \pi) \in \mathbb{R}_+^R \times \mathbb{R}_+^Q$ (where $\mathbb{R}_+^S$ denotes the set of functions from S to $\mathbb{R}_+$) such that for every $(x, s, y) \in \mathcal{C}(f, R, Q)$ we have

$$\sum_{r \in R} \psi(r) s(r) + \sum_{q \in Q} \pi(q) y(q) \geq 1. \tag{2.4}$$

We will associate the function pair (ψ, π) with the cut defined by the above inequality. With slight abuse of notation, we may use (ψ, π) to denote the set of points satisfying (2.4) (e.g., $\mathcal{C}_{LP}(f, R, Q) \cap (\psi, \pi)$).

We next explore the different special cases of generalized corner relaxations that have been considered in the literature, also mentioning the motivation behind these constructions.

2.3.2 Corner Relaxation

A *corner relaxation* is obtained from the above definition by using finite set R, Q, namely it is a mixed-integer set $\mathcal{C}(f, R, Q)$ for finite R, Q. The corner relaxation was introduced by Gomory [85, 86] and has received a lot of attention (see [44] for an excellent introduction to the corner relaxation and related notions). Importantly, although some power is lost when taking a relaxation (this will be discussed in more detail in Section 3.2), many of the prominent families of cuts are actually valid for the corner relaxation. For example, Andersen, Cornuéjols and Li [3] showed that every split cut for a general MIP (2.5), which includes the GMI and lift-and-project cuts mentioned in Chapter 1, can be obtained as a split cut from one of its corner relaxations. One can think of a corner relaxation as a basic relaxation embedded in a higher-dimensional space.

Motivation. We now motivate the construction of the (generalized) corner relaxation. For that, it will be convenient to encode the feasible set (1.2) in a different way. First, we put the MIP (1.2) in *standard form*[3]

$$\begin{aligned} Dw &= d \\ w &\geq 0 \\ w_i &\in \mathbb{Z}, \quad \forall i \in I \end{aligned} \tag{2.5}$$

by performing the following operations: 1) transforming each inequality $A_i x + B_i y \leq b_i$ into the pair of constraints $A_i x + B_i y + s_i = b_i$ and $s_i \geq 0$; 2) replacing any unrestricted variable, say x_i, by the term $x_i^+ + x_i^-$ and imposing the non-negativity $x_i^+ \geq 0$, $x_i^- \geq 0$; 3) renaming all the variables into w, with the integral variables $\{w_i\}_{i \in I}$ corresponding to the variables y. Moreover, we assume that the matrix D has full row rank: if not, then either the linear system $Dw = d$ has no solutions (which we can recognize efficiently, making the instance easy to handle), or it contains equations that are linear combinations of others, which are redundant and can be removed from the formulation.

Then given a square submatrix B of D of full rank (called a *basis*), we can rewrite (2.5) in *tableau form* with respect to B (i.e., pre-multiplying the system by B^{-1}) to obtain the equivalent system

$$\begin{aligned} w_B &= \bar{d} - \bar{N} w_N \\ w &\geq 0 \\ w_i &\in \mathbb{Z}, \quad \forall i \in I, \end{aligned}$$

where w_B are known as the *basic variables* and w_N are called the *non-basic variables.*

At this point we still have a system equivalent to (2.5). In order to finally obtain the corner relaxation of (2.5) with respect to the basis B we simply remove the

[3] See, for instance, [37] for more details on the transformations and assumptions made in this part, as well as the notation used.

non-negativity constraints for the basic variables w_B:

$$\begin{aligned} w_B &= \bar{d} - \bar{N} w_N \\ w_N &\geq 0 \\ w_i &\in \mathbb{Z}, \quad \forall i \in I. \end{aligned}$$

Further notice that if a given basic variable w_i is not required to be integral (i.e., $i \notin I$), then the equality constraint with w_i on the left-hand side is not really constraining the solution set of the system. Therefore, we can drop the equality constraints corresponding to all the basic variables which are not required to be integral. This brings us to a corner relaxation, namely system of the form $\mathcal{C}(f, R, Q)$ with finite R, Q.[4]

Cuts. It is known that valid functions, and more precisely the cuts given by (2.4) associated with them, capture every valid cut for a corner relaxation $\mathcal{C}(f, R, Q)$ that cuts off the infeasible point $(f, 0, 0)$ [44].

The first observation that strengthens the connection between the corner relaxation and intersection cuts is the following. Notice that $\mathcal{C}_{LP}(f, R, Q) \subseteq \mathbb{R}^{n+|R|+|Q|}$ is actually a displaced simplicial cone with apex $(f, 0, 0)$ and generators $\{(r, \mathbf{1}_r, 0)\}$ and $(q, 0, \mathbf{1}_q)$, where $\mathbf{1}_v : \mathbb{R}^n \to \{0, 1\}$ is the indicator function of the vector v [44]. Therefore, we can obtain intersection cuts using $\mathcal{C}_{LP}(f, R, Q)$ as relaxation.[5] (To simplify the notation, we use $\tilde{r} = (r, \mathbf{1}_r, 0)$ and $\tilde{q} = (q, 0, \mathbf{1}_q)$.)

For that, we are going to define the following fundamental objects in convex analysis that give a functional description of a convex set [111].

Definition 2 (Gauge function)**.** *Given a full dimensional convex set $K \subseteq \mathbb{R}^n$ containing the origin in its interior, the function $\gamma_K : \mathbb{R}^n \to \mathbb{R}_+$ defined by*

$$\gamma_K(r) = \inf \left\{ t > 0 : \frac{r}{t} \in K \right\}$$

is called the gauge *function of K.*

Let $S \subseteq \mathbb{R}^{n+|R|+|Q|}$ be a convex lattice-free set; in this context, by lattice-free we mean that $\operatorname{int}(S) \cap \mathbb{Z}^n \times \mathbb{R}^{|R|} \times \mathbb{Z}^{|Q|} = \emptyset$. Moreover, assume that S contains the point

[4] It is instructive to think of the point $(f, 0, 0) \in \mathbb{R}^n \times \mathbb{R}^{|R|} \times \mathbb{R}^{|Q|}$ as being the optimal solution for the linear relaxation of (2.5).

[5] Notice that $\mathcal{C}_{LP}(f, R, Q)$ is not full dimensional as the basic relaxation C^{LP} used in the previous section. However, we will still be able to obtain a cut by looking at intersection points of $\mathcal{C}_{LP}(f, R, Q) \cap \partial S$, for a convex lattice-free set S.

$(f, 0, 0)$ in its interior. Let ψ_S be the gauge function of the translated set $S - (f, 0, 0)$ (which now contains the origin in its interior), namely $\psi_S = \rho_{S-(f,0,0)}$. The inequality

$$\sum_{r \in R} \psi_S(\tilde{r}) s(r) + \sum_{q \in Q} \psi_S(\tilde{q}) y(q) \geq 1$$

is called the *intersection cut* for $\mathcal{C}_{LP}(f, R, Q)$ associated with S.

Theorem 1 ([11]). *Let $S \subseteq \mathbb{R}^{n+|R|+|Q|}$ be a convex lattice-free set containing the point $(f, 0, 0)$ in its interior. Then the cut (ψ_S, ψ_S) is valid for $\mathcal{C}(f, R, Q)$. More precisely,*

$$\mathcal{C}_{LP}(f, R, Q) \backslash \operatorname{int}(S) \subseteq \mathcal{C}_{LP}(f, R, Q) \cap \left\{ (x, s, y) : \sum_{r \in R} \psi_S(\tilde{r}) s(r) + \sum_{q \in Q} \psi_S(\tilde{q}) y(q) \geq 1 \right\}.$$

To get a feeling for the connection of this cut with intersection cuts as defined earlier, suppose that for every $r \in R$ we have $\psi_S(\tilde{r}) > 0$ and for every $q \in Q$ we have $\psi_S(\tilde{q}) > 0$, and notice that in this case the intersection between the line $\{(f, 0, 0) + \lambda \tilde{r}, \lambda \geq 0\}$ and ∂S is exactly $(f, 0, 0) + \frac{1}{\psi_S(\tilde{r})} \tilde{r}$ (and similarly for $\tilde{q}$). Then it is easy to verify that the hyperplane (in $\mathbb{R}^{n+|R|+|Q|}$) given by $\sum_{r \in R} \psi_S(\tilde{r}) s(r) + \sum_{q \in Q} \psi_S(\tilde{q}) y(q) = 1$ contains all these intersection points. Also notice that the apex $(f, 0, 0)$ is indeed cut off by this cut.

We also have a converse that says that every valid inequality for $\mathcal{C}(f, R, Q)$ is actually an intersection cut.

Theorem 2 ([43]). *Every valid cut for $\mathcal{C}(f, R, Q)$ is dominated by an intersection cut.*

The connection offered by the previous proposition means that understanding strong cuts for the relaxation $\mathcal{C}(f, R, Q)$ amounts to understanding convex lattice-free sets. The prominent role played by such sets has attracted a lot of research, see [44, 31, 106, 82, 10, 95] for some related work. We remark that actually objects related to lattice-free sets have played a major role in the geometry of numbers, including the fundamental Minkowski's Convex Body Theorem and Khintchine Flatness Theorem [91].

Although convex lattice-free sets in $\mathbb{R}^2$ are well understood, the structure of these sets in higher-dimensional spaces is still poorly understood. From a practical perspective, it is hard to know which of these sets should be used to generate cuts for a given MIP. Working with low-dimensional lattice-free sets will be a major motivation behind the *continuous relaxation* introduced in Section 2.3.4.

As a side remark, we can tie back intersection cuts to disjunctive cuts as follows. Consider a convex lattice-free set S with $(f, 0, 0)$ in its interior, and consider a linear description of this set in the form $S = \{(x, s, y) : a^j(x, s, y) \leq b_j, j \in J\}$. Then the set $\text{conv}(\mathcal{C}_{LP}(f, R, Q)\backslash\text{int}(S))$ is given exactly by applying the disjunction $\bigcup_{j \in J}\{(x, s, y) : a^j(z, s, y) \geq b_j\}$ to $\mathcal{C}_{LP}(f, R, Q)$ and then convexifying:

$$\text{conv}(\mathcal{C}_{LP}(f, R, Q) \setminus \text{int}(S)) = \text{conv}\left(\mathcal{C}_{LP}(f, R, Q) \cap \bigcup_{j \in J}\{(x, s, y) : a^j(z, s, y) \geq b_j\}\right).$$

Therefore, the intersection cut $\sum_{r \in R} \psi_S(\tilde{r})s(r) + \sum_{q \in Q} \psi_S(\tilde{q})y(q) \geq 1$ associated with S is also a disjunctive cut.

2.3.3 Infinite Relaxation

One difficulty of working with a corner relaxation is that the structure of the cuts seem to heavily depend on the specific sets R and Q used to define it. In order to reduce this dependence, define the *infinite relaxation* as the set $\mathcal{C}(f, \emptyset, \mathbb{R}^n)$. This relaxation was introduced by Gomory and Johnson [87, 88]. One can also define the *mixed-integer infinite relaxation* as the set $\mathcal{C}(f, \mathbb{R}^n, \mathbb{R}^n)$, but here we focus on the (pure-integer) infinite relaxation.

As pointed out in [44], the non-negativity assumption in the definition of a valid function is *not* without loss of generality in the context of the infinite relaxations. However, although there are functions π taking negative values that satisfy (2.4) for all feasible y, any such function must be non-negative over all *rational* vectors. Since data in mixed-integer linear programs are usually rational, it is natural to assume non-negativity in the definition of valid functions, which brings some technical benefits.

After introducing this relaxation, Gomory and Johnson [87, 88] started analyzing the structure of valid inequalities. In particular, they gave a simple characterization of all *minimal* valid inequalities, which can be seen as the only necessary inequalities to describe $\mathcal{C}_{LP}(f, \emptyset, Q)$ (see Section 3.1 for definitions). In addition, they provided what was arguably the deepest result in the theory of the infinite relaxation, the *2-Slope Theorem*, that gives a sufficient condition for valid inequalities to be a "facet" of $\mathcal{C}(f, \emptyset, Q)$ for the case $n = 1$. One of the results presented in this thesis is the extension of this theorem for general n (Chapter 6), which had been open since this seminal Gomory-Johnson paper in the 60s.

Notice that the infinite relaxation introduces a shift from the more geometric perspective to a functional analytic perspective. However, the relaxation presented next will clarify that there is actually a strict connection between these two perspectives.

2.3.4 Continuous Relaxation

As mentioned previously, one difficulty in obtaining strong cuts for the corner relaxation is that lattice-free sets in high dimensions are not well understood. The *continuous relaxation* presented next tries to overcome this difficulty. The *continuous relaxation* is defined as a set of the form $\mathcal{C}(f, R, \emptyset)$, for a finite set $R \subseteq \mathbb{R}^n$.

Given that this relaxation is really a special case of the corner relaxation where only continuous non-basic variables as present, we can still derive intersection cuts for it using convex lattice-free sets $K \subseteq \mathbb{R}^{n+|R|}$ with $(f, 0) \in \text{int}(K)$; we remark that in this context lattice-free means that $\text{int}(K) \cap (\mathbb{Z}^n \times \mathbb{R}^{|R|}) = \emptyset$. The main observation is that all the "action" is really happening only on the first n dimensions of the space, relative to the integral variables x.

More precisely, the first observation is the following: Consider a convex set K' containing K, but still such that K' is lattice-free (notice that $(f, 0) \in \text{int}(K')$). Then the valid cut $\sum_{r \in R} \psi_{K'}((r, \mathbf{1}_r))s(r) \geq 1$ generated by K' *dominates* $\sum_{r \in R} \psi_K((r, \mathbf{1}_r))s(r) \geq 1$: every point $(x, s) \in \mathbb{R}^{n+|R|}$ satisfying the latter also satisfies the former. Therefore, we can focus our attention to (inclusion-wise) *maximal* lattice-free sets containing $(f, 0)$.

So assume now that K is a maximal convex lattice-free set in $\mathbb{R}^{n+|R|}$. It is easy to see that K must be a *cylinder*, more precisely, there is $S \subseteq \mathbb{R}^n$ such that $K = S \times \mathbb{R}^{|R|}$: simply set S to be the projection of K into the first n coordinates. Because of this property, the functions ψ_K and ψ_S have the following relationship: for every $(r, r') \in \mathbb{R}^{n+|R|}$, $\psi_S(r) = \psi_K((r, r'))$. Therefore, the cut given by K is really determined by S:

$$\sum_{r \in R} \psi_K((r, \mathbf{1}_r))s(r) \geq 1 \;\equiv\; \sum_{r \in R} \psi_S(r)s(r) \geq 1.$$

The cut $\sum_{r \in R} \psi_S(r)s(r) \geq 1$ is called the *intersection cut* or *lattice-free cut* from the set S. Notice that S is lattice-free, namely $\text{int}(S) \cap \mathbb{Z}^n = \emptyset$, and that $f \in \text{int}(S)$. Theorem 2 directly gives the following connection.

Proposition 1. *For finite* $R \subseteq \mathbb{R}^n$*,* $\text{conv}(\mathcal{C}(f, R, \emptyset)) = \mathcal{C}_{LP}(f, R, \emptyset) \cap \bigcap_S \{(x, s) \in \mathbb{R}^{n+|R|} : \sum_{r \in R} \psi_S(r)s(r) \geq 1\}$*, where the intersection is taken over all convex lattice-free sets* $S \subseteq \mathbb{R}^n$ *with* $f \in \text{int}(S)$*.*

This proposition shows that in the continuous relaxation, we can focus on lattice-free sets in $\mathbb{R}^n$ instead of the potentially much larger space $\mathbb{R}^{n+|R|}$. However, the dimension n itself is typically quite large. The next idea is that, bearing some loss, we can focus on the case where n is small, say 1 or 2, by simply dropping most of the equations defining $\mathcal{C}(f, R, \emptyset)$.

2-Row continuous relaxations were studied recently by Andersen et. al [4], and this has spurred intensive research on the topic (see [44]). Despite the very simplified structure of such relaxation, one compelling feature is the following: a famous bad example for split cuts introduced by Cook, Kannan and Schrijver [47] can be handled using a single cut obtained from a 2-row continuous relaxation [4] (see Section 3.2 for more details).

Liftings. As mentioned in the beginning of Section 2.3, one way to strengthen the cuts obtained from relaxations is by incorporating information of the original MIP that was discarded in the process. In particular, given a valid inequality for the continuous relaxation, we can obtain a stronger inequality that is valid for a corner relaxation by imposing back integrality on some of the variables. This procedure is called *lifting*. Very interesting results about the structure of these liftings have been recently obtained in [72, 71, 28, 45, 29, 33]. We remark that the idea of strengthening inequalities by considering additional integrality conditions is not new, having already been explored by Balas and Jeroslow [17] with the *monoidal strengthening* procedure.

We also point out that a continuous version of the infinite relaxation was considered by Borozan and Cornuéjols [40]. They show that in their model, there is an equivalence (via the intersection cut procedure) between valid cuts and lattice-free sets in $\mathbb{R}^n$. Again lifting procedures can be used to give strong inequalities for the infinite relaxation. Therefore, these lifting procedures are able to establish a connection between the geometric perspective of intersection cuts and the functional perspective of cuts for the infinite relaxation. In Chapters 7 and 8 we explore further such connections.

2.4 Other Relaxations

In this section we present relaxations that are not directly related to the corner relaxation but are very useful for deriving cuts. For this section it will be convenient to work with MIPs in the standard form (2.5).

2.4.1 k-Row Relaxation

Consider the mixed-integer set P given by the MIP (2.5) in standard form. A *k-row relaxation* of P is obtained by combining the equality constraints into k equalities. More precisely, such relaxation is a set of the form

$$P(M) = \{w : MDw = Md, w \geq 0, w_i \in \mathbb{Z}\ \forall i \in I\},$$

for a $k \times \ell$ matrix M (where ℓ is the number of equality constraints in (2.5)).

Although such relaxations greatly simplify the structure of the problem, which also lends to efficient cut generation procedures, they are surprisingly powerful. For example, Nemhauser and Wolsey [107] showed that every split cut can be obtained as a split cut for a 1-row relaxation. More results concerning this relaxation are presented in Section 3.2.3.

2.4.2 Simple Mixed-integer Sets

Another different idea for using relaxations to derive cuts is to consider very simple mixed-integer sets for which we can obtain (usually) the description of *all* facets; a (relaxation for a) general MIP is then "embedded" in these simple sets.

The most basic example of this approach is to look at the following 2-dimensional set

$$Q^1 = \{(u, v) \in \mathbb{Z} \times \mathbb{R} : u + v \geq \beta, v \geq 0\},$$

for some $\beta \in \mathbb{R}$. The inequality $[\beta]u + v \geq [\beta]\lceil\beta\rceil$, where $[\beta] = \beta - \lfloor\beta\rfloor$ denotes the fractional part of β, is valid for Q^1. In fact, it is the only missing facet from its linear relaxation.

Proposition 2 ([114]). $\operatorname{conv}(Q^1) = \{(u, v) \in \mathbb{R}\times\mathbb{R} : u+v \geq \beta, [\beta]u+v \geq [\beta]\lceil\beta\rceil, v \geq 0\}$.

Although the set Q^1 is not directly a relaxation of (2.5), we can still "embed" the latter in the former [104, 55]. For example, let $fw \geq h$ be a valid inequality for (2.5); for example, we can take $fw = h$ to be one of the rows of the system $Dw = d$, or a linear combination of these rows. The second step is to transform the coefficient vector f into $\tilde{f}$ as follows. Let $I' \subseteq I$ be a subset of the indices of integral variables. Then for every $i \in I'$, we obtain the coefficient $\tilde{f}_i$ by rounding up f_i; for every $i \notin I'$, we obtain the coefficient $\tilde{f}_i$ by setting it equal to 0 if $f_i < 0$, and equal to f_i otherwise.

It is easy to check that the cut $\tilde{f}w \geq h$ is actually dominated by $fw \geq h$, and so the former is valid for (2.5). This gives the following relaxation of (2.5):

$$Q = \left\{ w : \begin{array}{c} \sum_{i \in I'} \tilde{f}_i w_i + \sum_{i \notin I'} \tilde{f}_i w_i \geq h \\ w_i \in \mathbb{Z} \;\; \forall i \in I' \\ w_i \geq 0 \;\; \forall i \notin I' \end{array} \right\}.$$

Notice the resemblance of this set with Q^1. More precisely, by the definition of $\tilde{f}$, for every feasible solution $w \in Q$ we have that $\sum_{i \in I'} \tilde{f}_i w_i$ is integral and that $\sum_{i \notin I'} \tilde{f}_i w_i$ is non-negative. Setting $u = \sum_{i \in I'} \tilde{f}_i w_i$ and $v = \sum_{i \notin I'} \tilde{f}_i w_i$ gives a correspondence between the points in Q and those in Q^1 (for β set to h). Under this correspondence, the inequality $[\beta]u + v \geq [\beta]\lceil\beta\rceil$ for Q^1 becomes the valid inequality

$$[h] \sum_{i \in I'} \tilde{f}_i w_i + \sum_{i \notin I'} \tilde{f}_i w_i \geq [h]\lceil h \rceil$$

for Q, which is also valid for the original program (2.5).

We remark that improved embedding procedures can be used, which can lead to stronger inequalities (see [114, 55]). For instance, one can use this improved procedure to obtain the *mixed-integer rounding (MIR)* inequality with respect to a given base inequality $fw \geq h$. Recall that it was shown that MIR inequalities are equivalent to split cuts [107].

This approach for generating cuts shifts the difficulty from understanding the facial structure of a MIP (or some of its relaxation) into finding the right simple set to be considered and then properly embedding the original MIP into this set. Thus, in some ways, this approach is complementary to the approaches that we have been discussing so far. One attractive feature is that one can generate cuts for specific optimization problems by abstracting an appropriate simple set similar to Q^1 which captures the main features of the problem.

The approach of using simple sets for deriving cuts has been very successful, in particular in obtaining generalizations of MIR inequalities. For instance, extensions of Q^1 with multiple integer variables were used to obtain the *2-step MIR* [60] and *n-step MIR* [98], an extension with multiple constraints connected through a single continuous variable was used to obtain the *mixing MIR* inequalities [92], and an extension with bounded variables was used to obtain *mingling* inequalities [7, 8].

We also highlight that this approach can be used to obtain cuts for *non-linear* mixed-integer programs by using simple non-linear mixed-integer sets, such as the *conic* [9] and *n-step conic* [105] MIR inequalities. The possibility of generating cuts for non-linear programs using simple sets had already been observed in [92], by incorporating the non-linearity in the embedding step.

CHAPTER 3

Strength of Cuts

The infinitude of possibilities for generating cuts motivates the main theme of this thesis: in order to generate effective cuts, we need to understand the strength of the different families available and the properties and limitations of our current methods for deriving cuts.

Before presenting some ways of measuring the strength of cuts and known results relative to them, we discuss the interplay between theory and practice in this context. The end goal of studying cutting planes is to eventually obtain improvement in the efficiency of solving MIPs in practice. Therefore, it is natural that the empirical performance of cuts is a very important factor when studying the strength of cuts. However, relying on experiments to determine the strength of cuts has two major drawbacks: the results are highly dependent on the problem instances used, and also on how the cuts are generated throughout the experiments.[1] Therefore, we feel that analytical results about the strength of cuts are very important, specially when they can relate back to experimental results. Moreover, the theoretical understanding of cuts is also important for providing insights on how to improve on the currently available cuts.

Aiming at bridging the theoretical and practical performance of cuts for the continuous relaxation is the focus of one of our results, presented in Chapter 5. Even though the model considered is not fully realistic, it is able to partially explain the empirical performance of some families of cuts better than previous models.

[1]Looking at the *closure* of a family of cuts avoids the second problem, although at this point it is not always feasible to take such approach.

3.1 Measures of Strength

In this section we present some of the common ways of measuring the strength of families of cuts.

Minimality, extremality and facetness. This is the most traditional way of capturing the strength of cuts. Consider a generic feasible set P of a MIP as in (1.2). Recall the standard definitions of *minimal valid* and *facet defining inequalities* for $\mathrm{conv}(P)$ (see for instance [108]). With slight overload in notation, we also use the term *facet* to refer to facet defining inequalities.

The concept of minimal valid and facet defining inequalities defines a hierarchy of inequalities: minimal valid inequalities are a subset of valid inequalities, and facet defining inequalities are a subset of the former. Importantly, one does not lose any power by focusing on the smaller of these classes:

$$\mathrm{conv}(P) = \bigcap \{(x, y) \in \mathbb{R}^{n+m} : fx + gy \leq h\},$$

where the intersection runs over all *facets* $fx + gy \leq h$ of $\mathrm{conv}(P)$. Therefore, focusing on trying to generate facets of $\mathrm{conv}(P)$ has been a standard practice. Indeed, in the 80s and 90s there was a flurry of work deriving facet defining inequalities for specific problems, where we highlight the success obtained for the Traveling Salesmen Problem [5].

The notions of minimal valid and facet defining inequalities can also be extended to the an infinite relaxation $\mathcal{C}(f, \emptyset, \mathbb{R}^n)$ in a reasonably natural way [87, 88, 89]. We say that a valid function π is *minimal* if there is no other valid function π' such that $\pi' \leq \pi$ with $\pi'(r) < \pi(r)$ for some $r \in \mathbb{R}^k$; this is essentially the same as in the finite-dimensional case. Now, to make the definition of a facet in this context more transparent, given a valid function π we define $E(\pi)$ to be the set of all feasible solutions $(x, y) \in \mathcal{C}(f, \emptyset, \mathbb{R}^n)$ that satisfy π at equality, namely $\sum_{r \in \mathbb{R}^n} \pi(r) y(r) = 1$. Then π is *facet defining* if there is no other valid function π' such that $E(\pi') \supsetneq E(\pi)$, that is, which has strictly more contact points with $\mathcal{C}(f, \emptyset, \mathbb{R}^n)$. Finally, we say that the valid function π is *extreme* is there are no valid functions π^1 and π^2 such that $\pi = (\pi^1 + \pi^2)/2$.

It is known that a facet defining function is extreme, which in turn is a minimal function [88, 89]. Therefore, these notions define a 3-level hierarchy of valid functions for the infinite relaxation. Unfortunately, unlike in the finite-dimensional case, it is not clear if it suffices to consider facet defining functions (or even extreme functions).

Nonetheless, being a facet (or extreme) is taken to be an indication of strength of a valid function [89].

Another distinction is that in the finite-dimensional case, the corresponding definitions of facet defining inequality and extreme inequality coincide. As far as we know, it is an open question whether they also coincide in the context of the infinite-relaxation. Another distinction is that in the finite-dimensional case, the corresponding definitions of facet defining inequality and extreme inequality coincide. As far as we know, it is an open question whether they also coincide in the context of the infinite-relaxation.

Closure inclusion. The notion of minimal/extreme/facet defining inequalities is interesting for studying the strength of a particular cut, but not as useful for comparing the strength among families of cuts. One reason for that is that not all cuts within a family (e.g. split cuts) will be, say, facet defining for a given MIP. One way of performing such comparison is by considering the *closure* of the families of interest.

More precisely, consider a mixed-integer set $P = (K, \mathbb{Z}^m \times \mathbb{R}^n)$. Given a family $\mathcal{F}$ of valid cuts for P, the $\mathcal{F}$*-closure* of P, denoted by $\mathcal{F}(P)$, is the mixed-integer set $(K', \mathbb{Z}^m \times \mathbb{R}^n)$ obtained by adding to K all the cuts in $\mathcal{F}$:

$$K' = K \cap \bigcap \{(x, y) \in \mathbb{R}^{n+m} : fx + gy \leq h\},$$

where the last intersection is taken over all cuts $fx + gy \leq h$ in $\mathcal{F}$.

By validity of the cuts, we always have that $\text{conv}(P) \subseteq \mathcal{F}(P)$ and the smaller the closure is, the better an approximation to $\text{conv}(P)$ it gives. Therefore, given two families $\mathcal{F}$ and $\mathcal{F}'$ of cuts (for instance, cross cuts and split cuts), the closure inclusion $\mathcal{F}(P) \subseteq \mathcal{F}'(P)$ indicates that the family $\mathcal{F}$ is stronger than $\mathcal{F}'$. Unfortunately, it can happen that neither $\mathcal{F}(P) \subseteq \mathcal{F}'(P)$ nor $\mathcal{F}'(P) \subseteq \mathcal{F}(P)$, in which case more refined methods are needed to provide further information.

We note that this inclusion comparison can be performed between any two relaxations of P (say, taking the intersection of all corner relaxations versus $\mathcal{F}(P)$, for some family $\mathcal{F}$).

Integrality gap. This type of measure is by far the most widely used to report the performance of cuts in experimental results. Given a MIP (1.1) and a specific relaxation, the *integrality gap*[2] is given by

$$\frac{\text{[value when optimizing over the relaxation]} - \text{[optimal value of MIP]}}{\text{optimal value of MIP}}.$$

Another related measure is the *percentage of integrality gap closed*, which compares the capacity that a given relaxation has of closing the integrality gap when compared to the actual integer hull:

$$\frac{\text{[value when optimizing over the relaxation]} - \text{[value when optimizing over the LP relaxation]}}{\text{[optimal value of MIP]} - \text{[value when optimizing over the LP relaxation]}}.$$

Blow-up measure. One can develop a more refined measure of strength than the closure inclusion defined above, at least for problems of *blocking type*. A closed, convex set $X \subseteq \mathbb{R}^n_+$ is said to be of *blocking type* if for every $x \in X$, whenever $y \geq x$ then we have $y \in X$ [113]. Although this is a restrictive property, we remark that, for instance, the projection of the continuous relaxation onto the space of the s variables is of blocking type (see [51]). For a set X of blocking type and for a scalar $\alpha \in \mathbb{R}_+$, define $\alpha X = \{\frac{x}{\alpha} : x \in X\}$. Notice that αX contains X for all $\alpha \geq 1$.

Consider a non-empty set $X \subseteq \mathbb{R}^n_+$ of blocking type and let $A, B \subseteq \mathbb{R}^n_+$ be closed convex relaxations of X; notice that A and B are also of blocking type. In order to compare A and B, define $\text{bu}(A, B)$ as the amount that A has to be "blown up" in order to contain B, namely

$$\text{bu}(A, B) = \inf\{\alpha : \alpha A \supseteq B\},$$

see [83].[3] Informally, the larger $\text{bu}(A, B)$ is, the larger B is compared to A. In particular, if A contains B then we have $\text{bu}(A, B) \leq 1$, otherwise we have $\text{bu}(A, B) > 1$. Therefore, this measure refines the closure inclusion discussed above.

This is also somewhat related to the integrality gap measure. To see this, define the *multiplicative gap* between A and B with respect to the cost vector c as:

$$\text{gap}(A, B, c) = \frac{\inf\{cs : s \in A\}}{\inf\{cs : s \in B\}}, \tag{3.1}$$

[2] Although we use this measure with a general relaxation and not only with the linear relaxation, we still use the term *integrality* gap.

[3] We define $\text{bu}(A, B) = \infty$ if there is no finite α such that $\alpha A \supseteq B$.

where we define $\text{gap}(A, B, c) = \infty$ if $\inf\{cs : s \in B\} = 0$. It was shown in [27] that, whenever A is polyhedral[4], then the quantity $\text{bu}(A, B)$ equals the worst gap over all non-negative cost vectors c:

$$\text{bu}(A, B) = \sup_{c \in \mathbb{R}^m_+} \{\text{gap}(A, B, c)\}. \tag{3.2}$$

Therefore, if there is any direction/cost where B is far away from A (so B is a loose relaxation in this direction), the value of this measure becomes large. This sensitivity to individual directions is one of the drawbacks of this measure of strength.

Rank. Yet another way of comparing the strength of two families of cuts is by looking at their relative *rank*. Consider a mixed-integer set P and a family $\mathcal{F}$ of valid cuts for it. The *ith $\mathcal{F}$-closure* of P, denoted by $\mathcal{F}^i(P)$, is obtained by iterating the closure procedure i times. More formally, it is defined by the recursion $\mathcal{F}^i(P) = \mathcal{F}(\mathcal{F}^{i-1}(P))$ with $\mathcal{F}^1(P) = \mathcal{F}(P)$. An inequality valid for the ith closure $\mathcal{F}^i(P)$ is called a *rank-i $\mathcal{F}$*-inequality.

Then given two families $\mathcal{F}$ and $\mathcal{F}'$ of cuts, we can look at the rank of $\mathcal{F}'$ relative to $\mathcal{F}$, namely the minimum i such that $\mathcal{F}^i(P) \subseteq \mathcal{F}'(P)$. If such value is large, then it shows that it takes several rounds of cuts from the family $\mathcal{F}$ to obtain the same effect given by one round of cuts from $\mathcal{F}'$, which indicates that $\mathcal{F}$ is significantly weaker than $\mathcal{F}'$.

We remark that this notation is also related to the length of *cutting planes proofs* (see for instance [110]), a topic that we will not explore here.

Other measures. A few other measures, which we will not cover in detail here, have also been used in the literature. For example, we can cite the *merit index* and *intersection index* measures introduced by Gomory and Johnson in [89], as well as the *shooting experiments*, proposed originally by Kuhn, used later by Gomory and developed further in [90].

3.2 Known Strength Results

In this section we survey some of the results known about the strength of the different types of cuts and relaxations presented in Chapter 2.

[4]In Chapter 5 we show that this polyhedrality assumption is not needed.

3.2.1 Disjunctive Cuts

Split cuts (or equivalently GMI, MIR, or strengthened lift-and-project cuts) is the most important family of cuts used in current MIP solvers [102]. The strength of these cuts was further substantiated by the controlled experimental works of Balas and Saxena [23] (see also [64, 77, 78]). There, the authors consider a procedure for generating split cuts that allow them to approximately optimize over the split closure. They employ this procedure over instances of the standard MIPLIB 3.0 problem set and show that, for the purpose of optimization, the split closure gives a very good approximation of the integer hull: the percentage of integrality gap closed is, on average, at least 82%. We remark that this procedure was not meant to be used in practice, and in fact it is not efficient enough for this task. However, recent works [59, 39, 80] have been able to obtain comparable approximations to the split closure in a practical manner.

Despite their strong empirical performance, Cook, Kannan and Schrijver [47] showed that in some cases split cuts can be weak. More precisely, they constructed a MIP with a valid cut that cannot be obtained by the ith split closure, for *any* i (see also [34, 68, 101] for generalizations). Also, Basu, Cornuéjols and Margot [27] constructed an example (based on the continuous relaxation from Section 2.3.4) where the split closure is arbitrarily bad when compared to the integer hull according to the blow-up measure. These results show a big disconnect between the practical and the theoretical performance of split cuts.

Interestingly, Dash, Dey and Günlük [56] pointed out that the bad examples of Cook, Kannan and Schrijver and of Basu, Cornuéjols and Margot can be handled by using a single cross cut. This implies that, for these example, the cross closure is strictly stronger (in terms of inclusion) than the iterated application of the split closure. Li and Richard [101] constructed examples for $t > 2$ where the t-branch split closure is strictly stronger than the iterated application of the cross (or 2-branch split) closure. Subsequently, it was constructed in [63] examples for all $t > k > 0$ where the t-branch split closure is strictly stronger than the iterated application of the k-branch split closure.

In addition, it is known that the 3-branch split closure is always contained in the crooked cross closure, which, in turn, is always contained in the cross closure [57, 56]. However, these two dominance relationships were not known to be strict prior to our work. In [57] the authors show that there are crooked cross cuts that cannot be obtained by a *single* cross cut; however, this result does not rule out the possibility that the cross closure (which contains potentially infinitely many cuts) is *always* *equal* to the crooked cross closure. The question whether the crooked cross closure can be strictly stronger than the cross closure was then left as an open question in [57].

Despite the similarities between cross cuts and rank-2 split cuts (intuitively both use two split disjunctions), Dash, Günlük and Vielma remarked (in Section 4.1 of [65]) that although the Cook-Kannan-Schrijver example gives a cross cut that cannot be obtained via rank-2 split cuts, it was not known if in fact the cross closure is always stronger than the second split closure. Another related question that remained open is whether the 3-branch split closure can be strictly contained in the crooked cross closure.

Addressing these open questions is the focus of our work presented in Chapter 4, where we provide the complete picture of the closure-inclusion relationship of these generalizations of split cuts (see Figure 4.1).

3.2.2 Corner and Related Relaxations

Strength of corner itself. The study of cuts valid for the corner and related relaxation dates back from the 60s and has extensive associated literature. Instead of surveying cuts for the corner relaxation, here we focus on a slightly different idea: how well do corner relaxations approximate a MIP. Notice that if corner relaxations provide a poor approximation, then cuts generated from them will also be weak.

One result that indicates that not much power is lost when considering corner relaxations was obtained by Andersen, Cornuéjols and Li [3], and roughly states that every split cut for a general MIP can be obtained as a split cut from one of its corner relaxations. However, Fischetti and Monaci [79] recently conducted experiments which indicate that unfortunately the corner relaxation obtained *from the optimal LP* solution has often a large integrality gap. The authors also consider the so called *strict corner relaxation* and show that it seems to have a significantly smaller integrality gap.[5] Complementing these experimental results, Cornuéjols, Michini and Nannicini [52] show that for the stable set problem, even if *all* corner relaxations are intersected, the relaxation obtained still has a large integrality gap.

In Chapter 4 we show that the result of Andersen, Cornuéjols and Li does not

[5]The authors in the paper use different names for the relaxations. What we are calling *strict corner relaxation* they call *corner relaxation*, and what we are calling *corner relaxation* they call *group relaxation*

hold for cross cuts, namely we cannot obtain all of them from basic relaxations.[6] This gives more indication about the potential weakness of these relaxations. On the up side, this disconnect also puts cross cuts forth as an option to overcome the weakness of the corner relaxations.

We pointed out earlier that it is possible to relax the MIP (2.5) by dropping some of the equations to obtain a corner relaxation with only k rows, typically for a small value of k. Indeed, the famous GMI cuts are obtained from the unassuming 1-row corner relaxation. Based on this surprising fact, some authors have tried to understand the strength of 1-row corner relaxations and if we are already harnessing all that they have to offer [62, 81]. Their experimental results indicate that many times the GMI cuts are actually the only useful cuts that can be obtained from these relaxations. The experiments conducted in [50], where the authors consider a variant of GMI cuts that is also obtained from a 1-row corner relaxation, also indicate that when the MIP has many constraints these cuts seem to have their effectiveness reduced.

Continuous relaxation. The results mentioned above indicate that it is not enough to work with 1-row corner-type relaxations. Hoping to remedy this situation, there has been a recent push for understanding how cuts for the multi-row version of the corner relaxation behave [69].[7] The continuous relaxation proved to be a fertile model for exploring these multi-row cuts [4].

Although earlier papers had already started studying the minimal valid inequalities for the continuous relaxation, the paper of Basu et. al [27] was one of the first to explore in more detail the strength of the cuts obtained from different lattice-free sets. To be more precise, they consider 2-row continuous relaxations (i.e. with $n = 2$). As mentioned in Section 2.3.4, all cuts for this relaxation can be obtained from 2-dimensional lattice-free sets. Actually Louveaux, Weismantel and Wolsey [4] showed that all such cuts can be obtained by lattice-free sets of 3 types: split sets, triangles and quadrilaterals. We remark that split cuts for the continuous relaxation are the same as intersection cuts obtained from split sets (see for instance Theorem 19 in Chapter 8).

Consider a continuous corner relaxation $\mathcal{C} = \mathcal{C}(f, R, \emptyset)$. Let SC, TC and QC denote respectively its split, triangle and quadrilateral closures, which are by defini-

[6] Actually this result holds for any family of cuts which includes cross cuts, such as crooked cross cuts and t-branch split cuts with $t \geq 2$

[7] Another possibility is to consider 1-row relaxations that are stronger than the corner relaxation, see for instance [6, 7].

tion obtained by adding to $\mathcal{C}_{LP}(f, R, \emptyset)$ all cuts given by the corresponding type of lattice-free sets. Finally, to make things formal let $\overline{\mathcal{C}}$ denote the projection of $\text{conv}(\mathcal{C})$ onto the s-space; let $\overline{SC}$ be the projection of SC onto the s-space, and define $\overline{TC}$ and $\overline{QC}$ similarly.[8]

Basu et. al [27] first show that one always has $\overline{TC} \subseteq \overline{SC}$ and $\overline{QC} \subseteq \overline{SC}$, so triangle and quadrilateral cuts are (with respect to closure inclusion) stronger than split cuts. They also show that $\text{bu}(\overline{\mathcal{C}}, \overline{TC}) \leq 2$ and $\text{bu}(\overline{\mathcal{C}}, \overline{QC}) \leq 2$, indicating that triangle and quadrilateral cuts provide a fair approximation to $\text{conv}(\mathcal{C})$. However, they also show that, for a specially constructed instance of the continuous relaxation, the blow-up measure $\text{bu}(\overline{TC}, \overline{SC})$ can be made to be arbitrarily large. In particular, this shows that $\text{bu}(\overline{\mathcal{C}}, \overline{SC})$ can be made arbitrarily large, and hence split cuts provide a poor approximation to the integer hull.

These results indicate that triangle and quadrilateral cuts can be particularly useful.[9] Indeed, the bad example for the split closure given by Cook, Kannan and Schrijver can be handled using triangle cuts [4]. These observation have then motivated experiments with these cuts (and generalizations), for example the results by Espinoza [75], Balas and Qualizza [21], Basu et. al [26] and Dey et. al [67]. Unfortunately, the results obtained from these experiments show that in practice only a mild improvement over split cuts is typically obtained. This again shows a gap between the theoretical and empirical results.

Developing new ways of measuring the strength of cuts that try to help explaining this situation is the focus of our work presented in Chapter 5.

Infinite relaxation. There has been plenty of research trying to understand properties of strong valid inequalities for the infinite relaxation. In the same paper where they introduced this relaxation, Gomory and Johnson [87, 88] completely characterized minimal valid inequalities in terms of subadditive functions satisfying additional properties. They also discussed methods for obtaining extreme inequalities.

Another particularly interesting result obtained by Gomory and Johnson in these papers is the celebrated 2-Slope Theorem. This theorem states that every minimal valid function which has 2 slopes is extreme for the 1-row infinite relaxation ($n = 1$). One interesting feature is that the valid inequality for the infinite relaxation corre-

[8] Notice that this projection is injective for these sets and that no information is really lost. See the beginning of Chapter 5 for further justification.

[9] Also, among these three classes, triangle and quadrilateral cuts are the only ones that are not necessarily valid for a 1-dimensional corner relaxation, see Section 3.2.3 below.

sponding to the classic GMI inequality satisfies the assumptions of this theorem [44], which further substantiates its strength. The functions corresponding to other well known inequalities such as the 2-step MIR are also 2-slope. Moreover, shooting experiments performed in [90, 61, 76] seem to indicate that cuts from 2-slope functions are indeed important for the corner and related relaxations. All these observations substantiate that the 2-Slope Theorem is indeed capturing a meaningful property of strong valid inequalities.

Moving away from the 1-row version of this relaxation, Dey and Richard [69] revisited the 2-row infinite relaxation ($n = 2$). There, they provide tools for proving the minimality of a valid inequality. The authors also construct families of extreme inequalities by extending the techniques used by Gomory and Johnson. Cornuéjols and Molinaro [53] also studied this relaxation and extended the 2-Slope Theorem for this setting, giving the *3-Slope Theorem*. Extending this theorem, and more generally the understanding of extreme functions, to the case of general n is the focus of our work presented in Chapter 6, where we prove the $(n + 1)$-*Slope Theorem* for the n-row infinite relaxation. Based on the discussion above, this result indicates that the $(k + 1)$-slope functions might be a promising class to generate strong cuts that use information from multi-row corner-type relaxations.

3.2.3 k-Row relaxation

As mentioned previously, although such relaxations greatly simplify the structure of the problem, they are surprisingly powerful even for small k. For example, Nemhauser and Wolsey [107] showed that every split cut can be obtained as a split cut for a 1-row relaxation.

In a similar vein, Dash, Dey and Günlük [56] showed that every cross cut (resp. crooked cross cut) can be obtained as a cross cut (resp. crooked cross cut) for a 3-row relaxation. However, they left as an open question whether these cuts can be obtained using 2-row relaxations instead. Again, the ability of considering as small number of rows as possible is interesting from the perspective of efficiently generating good cuts. Moreover, if crooked cross cuts from 2-row relaxations yielded all crooked cross cuts, then crooked cross cuts would be equivalent to cuts from all 2-row continuous relaxations of a MIP [56].

We address this open question in our result presented in Chapter 4.

Part II

New Results

CHAPTER 4

The Relative Strength of Generalizations of Split Cuts

In previous chapters we contended that split cuts are among the most important cutting planes for general MIPs, and therefore it is interesting to obtain a good understanding of its generalizations. In this chapter we compare cuts obtained from different generalizations of split cuts; in particular, we analyze the strength of the closure of all of the disjunctive cuts introduced in Section 2.1, namely split, cross, crooked cross and general t-branch split cuts, as well as cuts obtained from multi-row and basic relaxations. We refer the reader to Section 3.2.1 for further motivation and survey of related work.

On the technical side, in all the results presented here we need to show that adding possibly infinitely many cuts of a specific family is not strong enough. For that, we develop tools that allow us to exhibit the weakness of one cut at a time and then patch things together to obtain the weakness of the whole collection of cuts.

We remark that the Cook, Kannan and Schirjver example [47] is similar in spirit to some of the constructions that we use. However, in their proof, the authors make use of the fact that the split closure is polyhedral, which essentially allows them to consider the effect of only finitely many valid inequalities. In contrast, it is not known whether the cross/crooked closure or the intersection of the relaxations that we consider are polyhedral. This forces us to develop tools that directly tackle the interaction of potentially infinitely many cuts.

Organization of the chapter. First we recall some formal definitions needed throught the chapter and state our main results, even if still not completely formally. Section 4.3 presents a simple but main technical tool, dubbed the "Height Lemma", for handling infinitely many cuts. In Sections 4.4, 4.5, and 4.6 we compare the closures of multi-branch split cuts and crooked cross cuts. In the last two sections we compare the strength of cross cuts with cuts obtained from multi-row and basic relaxations.

Acknowledgments. This chapter is joint work with Sanjeeb Dash and Oktay Günlük.

4.1 Statement of Results

We now state in more detail the main results obtained in this chapter. For that, recall from Section 3.1 that the given a family $\mathcal{F}$ of valid cuts for a mixed-integer set P, the *$\mathcal{F}$-closure* of P, denoted by $\mathcal{F}(P)$, is given by the intersection of all cuts in $\mathcal{F}$:

$$\mathcal{F}(P) \triangleq \bigcap \{(x, y) \in \mathbb{R}^{n+m} : fx + gy \leq h\},$$

where the intersection is taken over all cuts $fx + gy \leq h$ in $\mathcal{F}$. Also recall that the *ith $\mathcal{F}$-closure* of P, denoted by $\mathcal{F}^i(P)$, is obtained by iterating the closure procedure i times, that is, $\mathcal{F}^i(P) = \mathcal{F}(\mathcal{F}^{i-1}(P))$ and $\mathcal{F}^1(P) = \mathcal{F}(P)$.

We say that a family of cuts *dominates* another if for every mixed-integer set, the closure of the first family of cuts for the set is contained in the closure of the second family of cuts for the same set. We say that the dominance is *strict* if there are examples where the elementary closure of the first family is strictly contained in the elemetary closure of the second family.

Multi-branch split cuts. Recall that it is known that 3-branch split cuts dominate crooked cross cuts which, in turn, dominate cross cuts [57, 56]. However, these two dominance relationships were not known to be strict prior to our work. In [57] the authors show that there are crooked cross cuts that cannot be obtained by a *single* cross cut; however, this result does not rule out the possibility that the cross closure (which contains potentially infinitely many cuts) is *always equal* to the crooked cross closure.

In our first set of results in this chapter, we establish that 3-branch split cuts strictly dominate crooked cross cuts which, in turn, strictly dominate 2-branch split cuts.

Theorem 1. *There is a mixed-integer set such that its crooked cross closure is strictly contained in its cross closure.*

Theorem 2. *There is a mixed-integer set such that its* 3*-branch split closure is strictly contained in its crooked cross closure.*

Similarly, Dash et al. remark (Section 4.1 of [65]) that although there are cross cuts (Example 2 of [47]) that cannot be obtained via rank-2 split cuts, it is not known if in fact the cross closure strictly dominates the second split closure. This question is relevant in their computational procedure for generating cross cuts. We also answer this question and show that cross cuts and rank-2 split cuts are not comparable.

Theorem 3. *For every finite integer $t > 0$, there is a mixed-integer set whose second split closure is strictly contained in its t-branch split closure.*

In Figure 4.1 we summarize the dominance relationships between these closures, with a plain arrow from A to B if the closure A dominates the closure B, and a crossed arrow from A to B if the closure A does not dominate the closure B (in the sense that for some mixed-integer set, the first closure is not contained in the second closure). When both types of arrows are present between a pair of closures, then one closure strictly dominates the other. Dashed arrows indicate results known prior to this work and solid arrows indicate results obtained in this work. In the figure, we denote the closure of t-branch split cuts with tBC for $t = 1, 2, 3$ and we use 4^+BC for all $t > 3$. We denote the crooked cross cut closure by CCC and use SC^2 to denote the second split closure. Note that the displayed arrows can be used to infer the relationship between any pair of the closures considered.

Cuts from relaxations. In the next set of results, we analyze the strength of cuts arising from the k-row relaxation defined in Section 2.4.1 and the basic relaxation defined in Section 2.2.

Dash, Dey and Günlük [56] showed that every cross cut (resp. crooked cross cut) can be obtained as a cross cut (resp. crooked cross cut) from a 3-row relaxation. However, they left as an open question whether these cuts can also be obtained from 2-row relaxations. They also note that if crooked cross cuts can be obtained as crooked cross cuts from 2-row relaxations, then crooked cross cuts would be equivalent to cuts

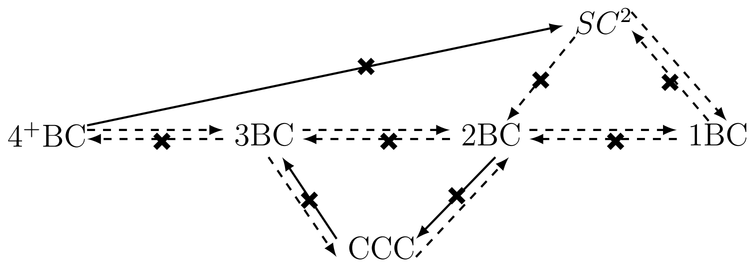

Figure 4.1: Comparing multi-branch split cuts with crooked cross cuts and rank 2 split cuts

from all 2-row continuous group relaxations of the set. In this chapter we answer this question.

Theorem 4. *There is a mixed-integer set such that its cross cut closure cannot be obtained by all cuts from its* 2*-row relaxations.*

Since the crooked cross closure is contained in the cross closure, the above theorem directly implies the following.

Corollary 1. *There is a mixed-integer set such that its crooked cross cut closure cannot be obtained by all cuts from its* 2*-row relaxations.*

Finally, we also show that unlike split cuts, t-branch split cuts in general cannot always be obtained from basic relaxations.

Theorem 5. *There is a mixed-integer set such that its cross cut closure cannot be obtained by all cuts from its basic relaxations.*

In Figure 4.2 we show some of the dominance relationships between these closures. We denote the closure of cuts from k-row relaxations by kR for $k = 1, 2, 3$ and we use CCC to denote crooked cross cuts, CC to denote cross cuts (2-branch split cuts) and SC to denote split cuts (1-branch split cuts). We use BR to denote cuts from basic relaxations. The fact that $2R$ does not dominate $3R$ follows from the fact that $2R$ does not dominate CC. The fact that $1R$ does not dominate $2R$ can be proved using the example of Cook, Kannan and Schrijver [47] where the integer hull has infinite split rank. It is shown in [4] that the integer hull in this example can be obtained from a 2-row relaxation. On the other hand, it is possible (and nontrivial) to show that all cuts from 1-row relaxations of this example are split cuts, and thus cannot yield the integer hull. We believe that CCC does not dominate $1R$ but we cannot prove this.

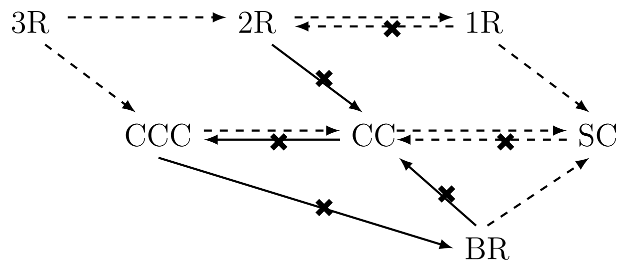

Figure 4.2: Comparing cuts from multi-row and basic relaxations with multi-branch split cuts

4.2 Preliminaries

In this chapter we will mostly work with mixed-integer sets of the following form: given rational matrices A, G, b with dimensions $r \times m$, $r \times n$ and $r \times 1$, respectively, the polyhedron P^{LP} is given by

$$P^{LP} = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n \ : \ Ax + Gy = b, \ y \geq 0\}, \tag{4.1}$$

and the associated mixed-integer set is given by

$$P = (P^{LP}, \mathbb{Z}^m \times \mathbb{R}^n). \tag{4.2}$$

Recall we will often use the identification $P \equiv P^{LP} \cap (\mathbb{Z}^m \times \mathbb{R}^n)$.

4.2.1 Split Cuts and More General Disjunctive Cuts

We recall some definitions from Sections 2.1 and 2.2. A *split disjunction* for the mixed-integer set P is a set of the form

$$D(\pi, \gamma) = \big\{(x, y) \in \mathbb{R}^{m+n} \ : \ \pi x \leq \gamma\big\} \cup \big\{(x, y) \in \mathbb{R}^{m+n} \ : \ \pi x \geq \gamma + 1\big\}$$

for some $\pi \in \mathbb{Z}^m, \gamma \in \mathbb{Z}$. Define the *split set* associated with the disjunction $D(\pi, \gamma)$ as

$$S(\pi, \gamma) = \{(x, y) \in \mathbb{R}^{m+n} \ : \ \gamma < \pi x < \gamma + 1\} = \mathbb{R}^{m+n} \setminus D(\pi, \gamma).$$

Split cuts. A linear inequality is a *split cut* for P with respect to the disjunction $D(\pi,\gamma)$ if it is valid for $P^{LP} \cap D(\pi,\gamma)$ [47]. For an integer $k \geq 1$, we use $SC^k(P)$ to denote the kth split closure of P. Notice that by definition the split closure of P is given by

$$SC(P) = \bigcap_{(\pi,\gamma)\in\mathbb{Z}^{m+1}} \operatorname{conv}(P^{LP} \cap D(\pi,\gamma)) = \bigcap_{(\pi,\gamma)\in\mathbb{Z}^{m+1}} \operatorname{conv}(P^{LP} \setminus S(\pi,\gamma)).$$

t-branch split cuts. Consider an integer t together with $\pi^i \in \mathbb{Z}^m$ and $\gamma_i \in \mathbb{Z}$ for $i = 1,\ldots,t$. The set $D(\pi^1,\ldots,\pi^t,\gamma_1,\ldots,\gamma_t)$ given by

$$D(\pi^1,\ldots,\pi^t,\gamma_1,\ldots,\gamma_t) = \bigcap_{i=1}^{t} D(\pi^i,\gamma_i) = \mathbb{R}^{m+n} \setminus \bigcup_{i=1}^{t} S(\pi^i,\gamma_i) \tag{4.3}$$

is called a *t-branch split disjunction* for P [101]. A linear inequality is a *t-branch split cut* for P with respect to a t-branch split disjunction D if it is valid for $P^{LP} \cap D$. We use $tBC(P)$ to denote the t-branch split closure of P, which is then given by

$$tBC(P) = \bigcap_{(\pi^1,\gamma_1),\ldots,(\pi^t,\gamma_t)\in\mathbb{Z}^{m+1}} \operatorname{conv}(P^{LP} \cap D(\pi^1,\ldots,\pi^t,\gamma_1,\ldots,\gamma_t)).$$

Notice that in the case $t = 1$ we have $1BC(P) = SC(P)$.

In [56], 2-branch split disjunctions are called *cross disjunctions*, and 2-branch split cuts are called *cross cuts*. In this case, we have the equivalent definition of the *cross closure* as

$$CC(P) = \bigcap_{(\pi^1,\gamma_1),(\pi^2,\gamma_2)\in\mathbb{Z}^{m+1}} \operatorname{conv}(P^{LP} \setminus (S(\pi^1,\gamma_1) \cup S(\pi^2,\gamma_2))).$$

Crooked cross cuts. Given $\pi^1, \pi^2 \in \mathbb{Z}^m$ and $\gamma_1, \gamma_2 \in \mathbb{Z}$ we define the sets

$$\begin{aligned}
D_1^c(\pi^1,\pi^2,\gamma_1,\gamma_2) &= \{(x,y) \in \mathbb{R}^{m+n} : \pi^1 x \leq \gamma_1, (\pi^2 - \pi^1)x \leq \gamma_2 - \gamma_1\},\\
D_2^c(\pi^1,\pi^2,\gamma_1,\gamma_2) &= \{(x,y) \in \mathbb{R}^{m+n} : \pi^1 x \leq \gamma_1, (\pi^2 - \pi^1)x \geq \gamma_2 - \gamma_1 + 1\},\\
D_3^c(\pi^1,\pi^2,\gamma_1,\gamma_2) &= \{(x,y) \in \mathbb{R}^{m+n} : \pi^1 x \geq \gamma_1 + 1, \pi^2 x \leq \gamma_2\},\\
D_4^c(\pi^1,\pi^2,\gamma_1,\gamma_2) &= \{(x,y) \in \mathbb{R}^{m+n} : \pi^1 x \geq \gamma_1 + 1, \pi^2 x \geq \gamma_2 + 1\}.
\end{aligned}$$

We call the set $D^c(\pi^1,\pi^2,\gamma_1,\gamma_2) = \bigcup_{i=1}^4 D_i^c(\pi^1,\pi^2,\gamma_1,\gamma_2)$ a *crooked cross disjunction* for P. A linear inequality is a *crooked cross cut* for P if it is valid for $P^{LP} \cap D^c$ for some crooked cross disjunction D^c. We use $CCC(P)$ to denote the crooked cross closure of P, which is again given by

$$CCC(P) = \bigcap_{(\pi^1,\gamma_1),(\pi^2,\gamma_2)\in\mathbb{Z}^{m+1}} \mathrm{conv}(P^{LP} \cap D^c(\pi^1, \pi^2, \gamma_1, \gamma_2)).$$

4.2.2 Relaxations of Mixed-integer Sets

k-row relaxation. Consider a polyhedral set P as in (4.2). Recall that a *k-row relaxation* of P is obtained by combining the r equality constraints defining the set into $k \leq r$ equalities. More precisely, it is the mixed-integer set $P(M) = (P^{LP}(M), \mathbb{Z}^m \times \mathbb{R}^n)$ where $P^{LP}(M) = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n : MAx + MGy = Mb, y \geq 0\}$ for a $k \times r$ matrix M. Any inequality valid for $P(M)$ is called a *cut from a k-row relaxation.*

Basic relaxation. For this relaxation we need to consider a polyhedron defined in inequality form. Let $P^{LP} = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n : Ax + Gy \leq b\}$ where A, G and b have $r \geq m + n$ rows, and consider the associated mixed-integer set $P = (P^{LP}, \mathbb{Z}^m \times \mathbb{R}^n)$. For a subset $J \subseteq \{1, \ldots, r\}$ of row indices, we use A_J to denote the submatrix of A consisting of the rows of A corresponding to the indices in J. We define G_J and b_J similarly. Then a *basic relaxation* of P is obtained by keeping in the linear relaxation only linearly independent constraints, namely it is a mixed-integer set of the form $P_{[J]} = \{(x, y) \in \mathbb{Z}^m \times \mathbb{R}^n : A_J x + G_J y \leq b_J\}$ for some $J \subseteq \{1, 2, \ldots, r\}$ such that the matrix $[A_J \; G_J]$ has full-row rank. Any inequality valid for $P_{[J]}$ is called *a cut from a basic relaxation.*

4.2.3 Notation

We use $\|\cdot\|$ to denote the ℓ_2 norm. Given a point $x \in \mathbb{R}^n$ and a positive real $r > 0$, we use $B(x, r) = \{y \in \mathbb{R}^n : \|x - y\| < r\}$ to denote the ball centered at x with radius r. For a set $S \subseteq \mathbb{R}^n$ we use $\mathrm{aff}(S)$ to denote the affine hull of S. Given a set of vectors $V \subseteq \mathbb{R}^n$ we use $\mathrm{span}(V)$ to denote the subspace spanned by V. Given a matrix $M \in \mathbb{R}^{n\times m}$, we use $\mathrm{rowspan}(M)$ to denote the subspace spanned by the rows of M.

4.3 Height Lemma

In preparation for the proof of our results we present the main technical tool used, called *Height Lemma* (this generalizes a similar result in [63]). Intuitively this lemma states the following: consider a collection (of arbitrary cardinality) of full dimensional pyramids, all sharing the same base. If we have a uniform lower bound on the height of the pyramids, plus the property that their apexes are not arbitrarily far from each other, then the intersection of all these pyramids contains a point outside of the common base. The motivation is that these pyramids will later represent what is 'left over' of P when we employ a subset of the cuts of interest, so this result allow us to talk about the left over of P when we add all these cuts together. In the formal statement below, the points $s^1, s^2, \ldots, s^n$ form the base of the pyramids and the points in Q are the apexes.

Lemma 1 (Height Lemma)**.** *Let $(a, b) \in \mathbb{R}^n \times \mathbb{R}$ with $a \neq 0$ and let $s^1, s^2, \ldots, s^n$ be affinely independent points in the hyperplane $\{x \in \mathbb{R}^n : ax = b\}$. Take $b' > b$ and $U \geq 0$ and define $Q = \{x \in \mathbb{R}^n : ax \geq b', \|x\| \leq U\}$. If $Q \neq \emptyset$, then there exists a point x in $\bigcap_{q \in Q} \mathrm{conv}(s^1, s^2, \ldots, s^n, q)$ satisfying the strict inequality $ax > b$.*

Proof. Let $H = \{x \in \mathbb{R}^n : ax = b\}$ and $S = \mathrm{conv}(s^1, s^2, \ldots, s^n)$. We say that a point x that satisfies $ax > b$ has positive *height*; so our goal is to find a point in $\bigcap_{q \in Q} \mathrm{conv}(s^1, s^2, \ldots, s^n, q)$ with positive height. To simplify the notation, we assume without loss of generality that $\|a\| = 1$.

Clearly S is an $(n-1)$-dimensional simplex contained in H and, by comparing dimensions, the affine hull of S equals H. Consider a point x^* in the relative interior of S, and let $r > 0$ be such that the ball $B(x^*, r) \cap H$ is contained in S. Let U' be an upper bound on the norm of the points in S (this exists as S is bounded). We show that the point $x^* + (b' - b)\frac{r}{U+U'}a$ belongs to $\bigcap_{q \in Q} \mathrm{conv}(s^1, s^2, \ldots, s^n, q)$, which gives the desired result.

Consider $q \in Q$ and let q^* denote its orthogonal projection into H, namely $q^* = q - b''a$ for $b'' = b' - b$. The idea is to show that x^* can be written as a convex combination $\alpha q^* + (1 - \alpha)y^*$ for some point y^* in S (see Figure 4.3). Then replacing q^* by q in this expression, we get by convexity that $\alpha q + (1 - \alpha)y^* = x^* + \alpha b'' a$ belongs to $\mathrm{conv}(s^1, s^2, \ldots, s^n, q)$ and has positive height. Importantly, our construction will guarantee that we can bound α from below independently of the choice of q.

To make this construction, consider the line $\{q^* + \lambda(x^* - q^*) : \lambda \in \mathbb{R}\}$ passing though the points q^* and x^*, and notice that it lies in the hyperplane H. This

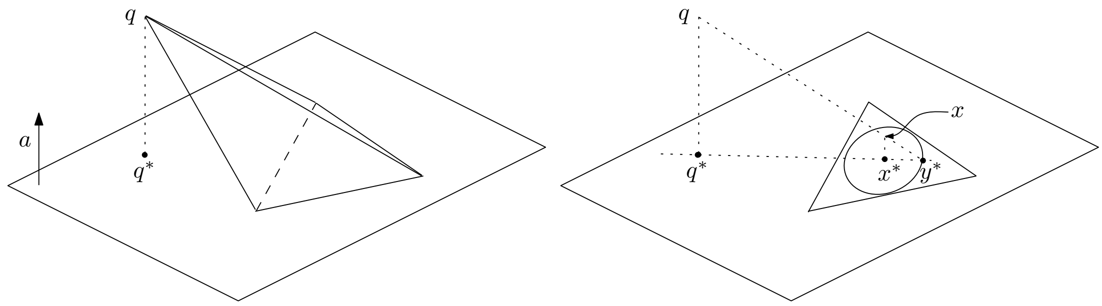

Figure 4.3: The left picture shows $\text{conv}(s_1, s_2, \ldots, s_n, q)$ and the hyperplane $H = \{x \in \mathbb{R}^n : ax = b\}$. The right picture shows the construction of y^* and the point x which belongs to $\text{conv}(s_1, s_2, \ldots, s_n, q)$ and has positive height, namely it satisfies $ax > b$.

line intersects the boundary of the closed ball $\bar{B}(x^*, r) \cap H$ in two points, so let y^* denote such point which is farthest from q^* (notice that this point belongs to S); specifically, we have $y^* = q^* + \lambda^*(x^* - q^*)$ for $\lambda^* = 1 + \frac{r}{\|x^* - q^*\|}$, and notice that $\|y^* - q^*\| = \lambda^* \|x^* - q^*\| = r + \|x^* - q^*\|$. Rearranging, we can write explicitly x^* as a convex combination of q^* and y^*: $x^* = \alpha q^* + (1 - \alpha) y^*$ for $\alpha = \frac{r}{\|y^* - q^*\|} \in [0, 1]$. As mentioned previously, we get that the point $\alpha q + (1-\alpha) y^* = x^* + \frac{r}{\|y^* - q^*\|} b'' a$ belongs to $\text{conv}(s^1, s^2, \ldots, s^n, q)$. Using the triangle inequality, we get that

$$\frac{r}{\|y^* - q^*\|} b'' \geq \frac{r}{\|y^*\| + \|q^*\|} b'' \geq \frac{r}{U + U'} b''.$$

Using convexity we conclude that the point $x^* + b'' \frac{r}{U+U'} a$ belongs to $\text{conv}(s^1, s^2, \ldots, s^n, q)$. Since the point is independent of q, it belongs to $\bigcap_{q \in Q} \text{conv}(s^1, s^2, \ldots, s^n, q)$ and the result follows.

□

Also note that in the proof we do not use the property that the norms of the points in Q are bounded, we only use the fact that their projection on H has bounded norm. It is therefore possible to generalize this result slightly to unbounded Q that has bounded projection on H.

By employing an affine transformation, this lemma also carries over to affine subspaces of $\mathbb{R}^n$.

Corollary 2. *Let $A \in \mathbb{R}^n$ be an affine subspace of dimension k. Fix $(a, b) \in \mathbb{R}^n \times \mathbb{R}$ such that $a \neq 0$ and let $s^1, s^2, \ldots, s^k \in \mathbb{R}^n$ be affinely independent points in $A \cap \{x \in$*

$\mathbb{R}^n : ax = b\}$. Take $b' > b$ and $U \geq 0$ and define $Q = \{x \in A : ax \geq b', \|x\| \leq U\}$. If $Q \neq \emptyset$, then there exists a point x in $\bigcap_{q \in Q} \text{conv}(s^1, s^2, \ldots, s^n, q)$ satisfying the strict inequality $ax > b$.

For a vector $v \in \mathbb{R}^n$, a $n \times n$ matrix M and a set $S \subseteq \mathbb{R}^n$, let $S - v = \{s - v : s \in S\}$ and $MS = \{Ms : s \in S\}$. To see that the corollary follows from Lemma 1, let M be an $n \times n$ matrix with determinant one such that $M(A - s^1) = \mathbb{R}^k \times \{0\}^{n-k}$. Applying this affine transformation and subsequently removing the last k coordinates, the corollary reduces to the previous lemma applied to objects in $\mathbb{R}^k$ (points in Q are mapped to points in $\mathbb{R}^k \times \{0\}^{n-k}$ with bounded norm).

4.4 Crooked cross Closure Versus Cross Closure

In this section we prove Theorem 1 by constructing a polyhedral set P whose cross closure $CC(P)$ is strictly contained in its crooked cross closure $CCC(P)$. One important component of the construction is a triangle that cannot be covered by a cross set.

Theorem 6 ([58]). *There exists a rational triangle $T^* \subseteq \mathbb{R}^2$ satisfying the following: (i) T^* does not contain integer points in its interior; (ii) T^* contains the points $(0, 0)$, $(1, 0)$, $(0, 1)$ in its boundary; (iii) there is $\delta > 0$ such that for any pair of split sets S_1, S_2 for $\mathbb{Z}^2$, the set $T^* \setminus (S_1 \cup S_2)$ has area at least δ.*

Let T^* be such a triangle and let x^* be a point in the interior of T^*, say its centroid (which has rational coordinates). In this section we work with the polyhedron P^{LP} defined as

$$P^{LP} = \left\{(x, y) \in \mathbb{R}^2 \times \mathbb{R} : (x, y) \in \text{conv}\,(T^* \times \{0\}) \cup (x^* \times \{1\})\right\},$$

and the mixed-integer set $P = (P^{LP}, \mathbb{Z}^2 \times \mathbb{R})$. We also define $T_\epsilon \triangleq P^{LP} \cap \{x \in \mathbb{R}^3 : x_3 = \epsilon\}$ for $\epsilon \geq 0$ and define T^*_ϵ to be the projection of T_ϵ onto the first 2 coordinates. We next obtain the following result.

Lemma 2. *The inequality $x_3 \leq 0$ is valid for $CCC(P)$.*

Proof. Notice that $T^*_0 = T^*$ and $T^*_1 = x^*$, and as the latter belongs to the interior of T^*, we conclude that T_ϵ is contained in the interior of T^* for all $\epsilon > 0$. As T^* does not contain any integer points in its interior, $T^*_\epsilon \cap \mathbb{Z}^2 = \emptyset$ for all $\epsilon > 0$ and therefore $\text{conv}(P) \subseteq T^* \times \{0\}$. Consequently, the inequality $x_3 \leq 0$ is valid for $\text{conv}(P)$.

To conclude the proof, we recall the fact that the convex hull of any polyhedral mixed-integer set in $\mathbb{Z}^2 \times \mathbb{R}$ is given by crooked cross cuts [57]. In particular, $\text{conv}(P) = CCC(P)$ and consequently $x_3 \leq 0$ is valid for $CCC(P)$, concluding the proof. $\square$

We next show that the inequality $x_3 \leq 0$ is not valid for $CC(P)$; since $CC(P)$ always contains $CCC(P)$ – which equals $\text{conv}(P)$ in this section – such a result would imply that $CC(P)$ strictly contains $CCC(P)$, as desired. We start by showing that a single cross disjunction cannot imply the cut $x_3 \geq 0$.

Lemma 3. *There exists $\epsilon^* > 0$ such that for any pair of split sets S_1, S_2 for P, the set $T_{\epsilon^*} \setminus (S_1 \cup S_2)$ is non-empty.*

Proof. Notice that $\text{area}(T^*_0) > \text{area}(T^*_1) = 0$ and $\text{area}(T^*_\epsilon)$ is continuous as a function of ϵ. Let $\delta > 0$ be given by Theorem 6 and take $\epsilon^* > 0$ such that $\text{area}(T^*_{\epsilon^*}) \geq \text{area}(T^*_0) - \delta/2$; the existence of ϵ^* is guaranteed by the Intermediate Value Theorem.

Let S^*_1 denote the projection of the split set S_1 onto the first 2 (integer) coordinates, and notice that $S_1 = S^*_1 \times \mathbb{R}$ and that S^*_1 is a split set for $(T^*, \mathbb{Z}^2)$. Define S^*_2 similarly. It then follows that $T_{\epsilon^*} \setminus (S_1 \cup S_2)$ is non-empty if and only if $T^*_{\epsilon^*} \setminus (S^*_1 \cup S^*_2)$ is non-empty; we prove the latter. Theorem 6 guarantees that the set $T^* \setminus (S^*_1 \cup S^*_2)$ has area at least δ, and so $T^*_{\epsilon^*} \setminus (S^*_1 \cup S^*_1)$ has area at least $\delta/2$. Therefore $T_{\epsilon^*} \setminus (S_1 \cup S_2)$ is non-empty. $\square$

Together with the previous lemma, the Height Lemma directly implies that the cut $x_3 \leq 0$ is not valid for the cross closure of P; the proof is exactly the same as in Lemma 8 and is omitted.

Lemma 4. *The inequality $x_3 \leq 0$ is not valid for $CC(P)$.*

Employing Lemmas 2 and 4 we obtain Theorem 1:

Theorem 1 (restated). *$CCC(P) \subsetneq CC(P)$.*

4.5 Crooked Cross Cuts Versus 3-branch Split Cuts

In this section we prove Theorem 2 by constructing an integer set $P = (P^{LP}, \mathbb{Z}^3)$ such that $3BC(P) = \text{conv}(P) = \emptyset$ but $CCC(P) \neq \emptyset$. We define the polyhedron P^{LP}

to be the intersection of a specific octahedron with the unit cube, i.e.,

$$P^{LP} = \left\{ x \in [0,1]^3 : \sum_{i \in I} x_i - \sum_{i \notin I} x_i \leq |I| - \frac{1}{2}, \forall I \subseteq \{1,2,3\} \right\}.$$

Notice that P is the empty set.

We first claim that $3BC(P) = \emptyset$. To see this, consider the 3-branch split disjunction $D = D(e^1, e^2, e^3, 0, 0, 0)$, where e^i is the ith unit vector in $\mathbb{R}^3$. Notice that x belongs to D iff $x_i \notin (0,1)$ for all $i = 1,2,3$, and therefore $x \in P^{LP} \cap D$ if and only if x is a 0-1 vector. Therefore $P^{LP} \cap D = \emptyset$. Since $3BC(P) \subseteq P^{LP} \cap D$, the claim follows.

Now we need to show that $CCC(P) \neq \emptyset$; in particular, we show that $(1/2, 1/2, 1/2)$ belongs to $CCC(P)$. For that, we need the following characterization of the crooked cross closure.

Theorem 7. *([57, Theorem 3.1]) For any mixed-integer set* $\tilde{P} = (\tilde{P}^{LP}, \mathbb{Z}^n \times \mathbb{R}^l)$,

$$CCC(\tilde{P}) = \bigcap_{\pi^1, \pi^2 \in \mathbb{Z}^n} \operatorname{conv}\left(\tilde{P}^{LP} \cap \{(x,y) : \pi^1 x \in \mathbb{Z}, \pi^2 x \in \mathbb{Z}\}\right).$$

Lemma 5. *The point* $(1/2, 1/2, 1/2)$ *belongs to* $CCC(P)$.

Proof. Consider an arbitrary pair of vectors $\pi^1, \pi^2 \in \mathbb{Z}^3$ and define $P^{LP}_{\pi^1,\pi^2} = \operatorname{conv}(P^{LP} \cap \{x \in \mathbb{R}^3 : \pi^1 x \in \mathbb{Z}, \pi^2 x \in \mathbb{Z}\})$. Given Theorem 7, it suffices to show that

$$(1/2, 1/2, 1/2) \in P^{LP}_{\pi^1,\pi^2}. \tag{4.4}$$

For that, let $v \in \mathbb{R}^3$ be a non-zero vector orthogonal to π^1 and π^2. We will prove (4.4) for the case $v_1 \neq 0$; the proof for the cases $v_2 \neq 0$ or $v_3 \neq 0$ is similar. The idea in the analysis is that the set $\{x \in \mathbb{R}^3 : \pi^1 x \in \mathbb{Z}, \pi^2 x \in \mathbb{Z}\}$ contains all lines in the direction of v that pass through an integer point. We are interested in the lines that cross the intersection of P^{LP} with the plane $x_1 = 1/2$; therefore, it suffices to project $\mathbb{Z}^3$ onto this plane along v and analyze the obtained set of points Λ, and show that $\operatorname{conv}(P^{LP} \cap \Lambda)$ contains $(1/2, 1/2, 1/2)$.

Define the integer points $w^1 = (0, -\lfloor \frac{v_2}{2} \rfloor, -\lfloor \frac{v_3}{2} \rfloor)$ and $w^2 = (1, 1 + \lfloor \frac{v_2}{2} \rfloor, 1 + \lfloor \frac{v_3}{2} \rfloor)$; clearly $w^j \pi^i \in \mathbb{Z}$ for $i, j \in \{1,2\}$. Now consider the points $u^1 = w^1 + v/2$ and $u^2 = w^2 - v/2$, which lie in the plane $x_1 = 1/2$. We can use the fact that v is orthogonal to π^1, π^2 to deduce that $u^j \pi^i \in \mathbb{Z}$ for $i, j \in \{1,2\}$. Also, notice that u^j_2 and u^j_3 belong to the interval $[0,1]$ for $j \in \{1,2\}$. Now any point in $[0,1]^3$ with one component equal to 1/2 is contained in P^{LP}, and therefore so are u^1, u^2. Therefore, these points belong to $P^{LP}_{\pi^1,\pi^2}$. By convexity of $P^{LP}_{\pi^1,\pi^2}$, the point $(u^1 + u^2)/2 = (1/2, 1/2, 1/2)$ also belongs to it, which concludes the proof of the lemma. □

The fact that $3BC(P) = \emptyset \neq CCC(P)$ then concludes the proof of Theorem 2:

Theorem 2 (restated). *$3BC(P) \subsetneq CCC(P)$.*

4.6 t-branch Split Closure Versus Second Split Closure

In this section we prove Theorem 3, which states that there is an integer set P whose t-branch split closure is not contained in its second split closure. More specifically, we will work with the integer set $P = (P^{LP}, \mathbb{Z}^{n+1})$ where P^{LP} is based on a distorted simplex and is defined as

$$P^{LP} = \left\{ x \in \mathbb{R}^{n+1} \ : \ \sum_{i=1}^{n} x_i + 2x_{n+1} \leq n + 2 - \epsilon, \quad x_i \geq \epsilon, \ i = 1, \dots, n \right\},$$

where $\epsilon > 0$ is a small scalar defined in the proof of Lemma 7 below (see Figure 4.4).

We claim that the cut $x_{n+1} \leq 0$ is valid for $SC^2(P)$. First notice that Chvátal-Gomory cuts [108] for P can be obtained by rounding the right-hand sides of the constraints above. Since every Chvátal-Gomory cut is also a split cut, we observe that

$$\begin{aligned} SC(P) &\subseteq \left\{ x \in \mathbb{R}^{n+1} \ : \ \sum_{i=1}^{n} x_i + 2x_{n+1} \leq n + 1, \quad x_i \geq 1, \ i = 1, \dots, n \right\} \\ &\subseteq \left\{ x \in \mathbb{R}^{n+1} \ : \ x_{n+1} \leq \frac{1}{2} \right\}. \end{aligned}$$

Again by using Chvátal-Gomory cuts, we get that

$$SC^2(P) \subseteq \{ x \in \mathbb{R}^{n+1} \ : \ x_{n+1} \leq 0 \},$$

which proves the claim. Since $P \subseteq SC^2(P)$, we get the following.

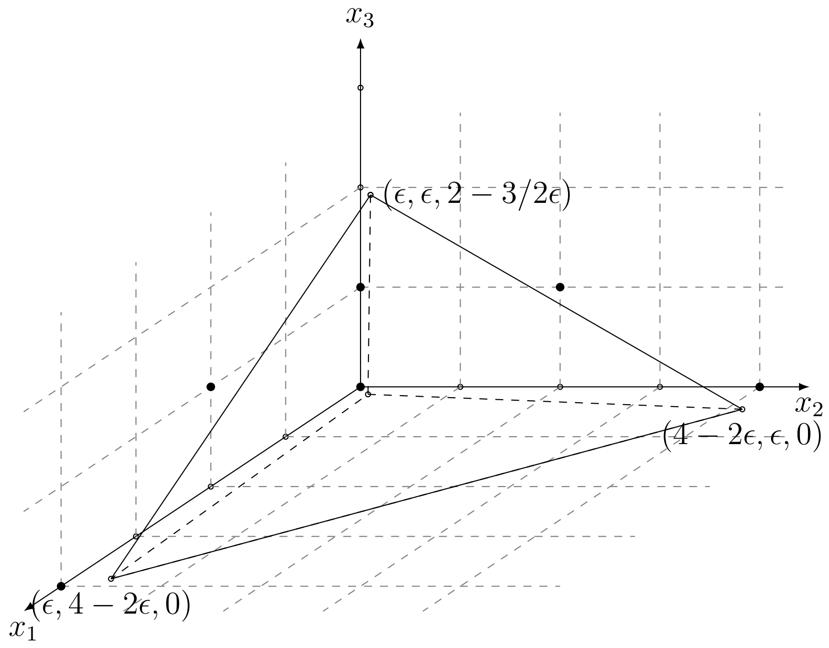

Figure 4.4: The set P when $n = 2$

Lemma 6. *The inequality $x_{n+1} \leq 0$ is valid for $SC^2(P)$. Furthermore, it is facet defining as it contains the following $n + 1$ affinely independent points in P:*

$$s_1 = (2, 1, \dots, 1, 0), s_2 = (1, 2, \dots, 1, 0), \dots, s_n = (1, 1, \dots, 2, 0), s_{n+1} = (1, 1, \dots, 1, 0). \tag{4.5}$$

We next argue that the inequality $x_{n+1} \leq 0$ is not valid for the t-branch split closure of P when $t < n$. First we show that a single t-branch split cut cannot imply the cut $x_{n+1} \leq 0$. The main tool used is the fact that simplices cannot be covered by a small collection of split sets. More precisely, define the simplex

$$\Delta_n = \left\{ x \in \mathbb{R}^n \ : \ \sum_{i=1}^{n} x_i \leq n, \quad x_i \geq 0, \ i = 1, \dots, n \right\}.$$

Theorem 8 ([63]). *For every integer $n > 0$, there exists a constant $\delta > 0$ such that the volume of the n-dimensional simplex Δ_n not covered by any collection of $n - 1$ split sets is at least δ.*

Lemma 7. *Let $S_1, \ldots, S_t$ be a collection of split sets for P with $t < n$ and let $S = \bigcup_{i=1}^t S_i$ be their union. Then the set $P^{LP} \setminus S$ contains a point x such that $x_{n+1} = 1$.*

Proof. Consider the slice of P^{LP} with $x_{n+1} = 1$, namely

$$\begin{aligned} T &\triangleq P^{LP} \cap \{x \in \mathbb{R}^{n+1} : x_{n+1} = 1\} \\ &= \left\{ x \in \mathbb{R}^{n+1} : \sum_{i=1}^n x_i \leq n - \epsilon, \quad x_{n+1} = 1, \quad x_i \geq \epsilon, \ i = 1, \ldots, n \right\}. \end{aligned}$$

We will show that if $t < n$ then $T \setminus S \neq \emptyset$, which proves the lemma. Let $\pi^i \in \mathbb{Z}^{n+1}$ and $\gamma_i \in \mathbb{Z}$ be such that $S_i = \{x \in \mathbb{R}^{n+1} : \gamma_i < \pi^i x < \gamma_i + 1\}$. Notice that

$$T \cap S_i = T \cap \left\{ x \in \mathbb{R}^{n+1} : \gamma_i - \pi^i_{n+1} < \sum_{i=1}^n \pi^i x_i < \gamma_i - \pi^i_{n+1} + 1 \right\}$$

and therefore, $T \cap S_i = T \cap (S_i^* \times \mathbb{R})$, where S_i^* is the split set $S(\pi^i, \gamma_i - \pi^i_{n+1})$ contained in $\mathbb{R}^n$. Let $S^* = \bigcup_{i=1}^n S_i^*$ and observe that $T \cap S = T \cap (S^* \times \mathbb{R})$. Let T^* denote the projection of T onto the first n coordinates, namely

$$T^* = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i \leq n - \epsilon, \quad x_i \geq \epsilon \ i = 1, \ldots, n \right\},$$

and notice that $T \setminus S \neq \emptyset$ if and only if $T^* \setminus S^* \neq \emptyset$, so it suffices to prove the latter.

Now notice that T^* is a perturbation of the simplex Δ_n, depending on ϵ. Choosing $\epsilon > 0$ small enough, we get the volume of $T^* \setminus S^*$ arbitrarily close to the volume of $\Delta_n \setminus S^*$, which is strictly positive by Theorem 8. This implies that $T^* \setminus S^*$ is non-empty, which concludes the proof of the lemma. □

Applying the height lemma, we can make a statement about the simultaneous effect of every possible collection of $t < n$ split sets $S_1, S_2, \ldots, S_t$ on P.

Lemma 8. *For $t < n$, the inequality $x_{n+1} \leq 0$ is not valid for the t-branch split closure of P.*

Proof. Let $\mathcal{S}_t$ denote the family of t-branch split sets for $I = \mathbb{Z}^{n+1}$, namely sets of the form $\bigcup_{i=1}^t S_i$ where each S_i is a split set for I. To prove the lemma, we show that $\bigcap_{S \in \mathcal{S}_t} \operatorname{conv}(P^{LP} \setminus S)$ contains a point x with $x_{n+1} > 0$.

For each $S \in \mathcal{S}_t$, let x^S be the point given by Lemma 7. As $x^S \in P^{LP}$ and $x^S_{n+1} = 1$, we have $||x^S|| \leq n+1$ and so we can apply the Height Lemma with parameters $a = (0, 0, \ldots, 0, 1)$, $b = 0$, $b' = 1$, $U = n+1$, and $s_1, s_2, \ldots, s_{n+1}$ defined in (4.5) to get that $\bigcap_{S \in \mathcal{S}_t} \text{conv}(s_1, s_2, \ldots, s_{n+1}, x^S)$ contains a point x with $x_{n+1} > 0$. Notice that for each $S \in \mathcal{S}_t$ we have $\text{conv}(s_1, s_2, \ldots, s_{n+1}, x^S) \subseteq \text{conv}(P^{LP} \setminus S)$ (since the integer points $s_1, s_2, \ldots, s_{n+1}$ belong to $P \subseteq \text{conv}(P^{LP} \setminus S)$), which implies that $\bigcap_{S \in \mathcal{S}_t} \text{conv}(P^{LP} \setminus S)$ contains a point x with $x_{n+1} > 0$. This concludes the proof. □

Using Lemmas 6 and 8 we now prove Theorem 3.

Theorem 3 (restated). *For any positive integer $t < n$, $SC^2(P)$ is strictly contained in $tBC(P)$.*

Proof. To simplify the notation define

$$A = \left\{ x \in \mathbb{R}^{n+1} \; : \; \sum_{i=1}^{n} x_i + 2x_{n+1} \leq n+1, \quad x_{n+1} \leq 0, \quad x_i \geq 1, \; i = 1, \ldots, n \right\}.$$

We have already showed that

$$SC^2(P) \subseteq A. \tag{4.6}$$

We claim that A is an integral polyhedron. To see that, first note that it is defined by $n+2$ inequalities in $\mathbb{R}^{n+1}$, and therefore it has at most $n+2$ extreme points that are obtained by intersecting all but one of the defining hyperplanes. It can be checked that the only fractional point that can be obtained by intersecting $n+1$ of these hyperplanes is obtained by excluding the inequality $x_{n+1} \leq 0$. The corresponding point, however violates $x_{n+1} \leq 0$ and therefore is not an extreme point of the polyhedron, thus proving the claim.

This integrality, together with the fact that $P \subseteq A \subseteq P^{LP}$, implies that $A = \text{conv}(P)$. Moreover, since $SC^2(P) \supseteq P$, this also gives that the containment in (4.6) actually holds as equality, and hence $\text{conv}(P) = SC^2(P)$. By Lemma 8 we then conclude that $SC^2(P)$ is strictly contained in $tBC(P)$. □

Further, as P^{LP} is defined by $n+1$ linearly independent linear inequalities in $n+1$ variables, P is a basic relaxation of itself, and therefore $\text{conv}(P)$ can be obtained by cuts from basic relaxations. This yields the following corollary.

Corollary 3. *For any positive integer $t < n$, the set of points satisfying all cuts from basic relaxations of P is strictly contained in $tBC(P)$.*

4.7 Cross Cuts from Basic Relaxations

In this section we prove Theorem 5 by constructing a mixed-integer set P with the property that the intersection of all cuts from its basic relaxations does not dominate its cross closure. We will work with the polyhedron (see Figure 4.5)

$$P^{LP} = \Big\{(x,w) \in \mathbb{R}^2 \times \mathbb{R} \ : \ -x_1 - x_2 + w \le 0, \quad x_1 + x_2 + w \le 2, \quad -x_1 + x_2 + w \le 1, \quad x_1 - x_2 + w \le 1 \Big\}, \tag{4.7}$$

and define $P = (P^{LP}, \mathbb{Z}^2 \times \mathbb{R})$. For $j = 1,2,3,4$, let P_j^{LP} denote the relaxation of P^{LP} obtained by dropping the jth constraint in (4.7); also let $P_j = (P_j^{LP}, \mathbb{Z}^2 \times \mathbb{R})$.

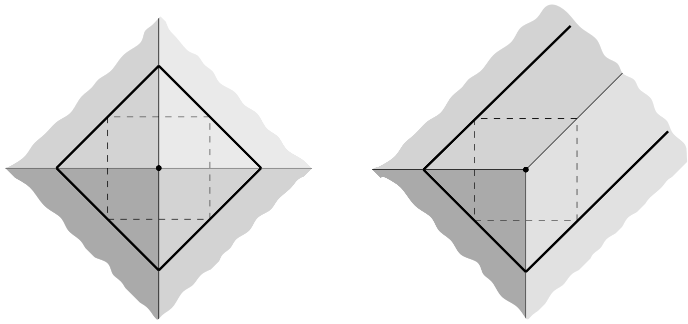

Figure 4.5: The left picture shows P along the (x_1,x_2)-plane, with the unit square $[0,1]^2$ in dashed lines and the intersection of P with the plane $t = 0$ in bold. The right picture shows the basic relaxation P_2, which gives rise to the set $P_2^I = P_2 \cap I$.

As $P^{LP} \subseteq \mathbb{R}^3$ is defined by four linearly independent constraints, the sets P_j for $j = 1,2,3,4$ give all the basic relaxations of P. Thus we want to show that $\bigcap_{j=1}^4 \text{conv}(P_j) \not\subseteq CC(P)$. For that, we show that $w \le 0$ is a cross cut for P but it is not valid for $\bigcap_{j=1}^4 \text{conv}(P_j)$.

Lemma 9. *The inequality $w \le 0$ is a valid cross cut for P.*

Proof. We will show that $w \le 0$ is a cross cut for P derived from the cross disjunction $D(e^1, e^2, 0, 0) = \mathbb{R}^3 \setminus (S_1 \cup S_2)$ where e^i is the ith unit vector in $\mathbb{R}^3$ and S_1 is the split set $\{(x,w) \in \mathbb{R}^2 \times \mathbb{R} : 0 < x_1 < 1\}$ and S_2 is the split set $\{(x,w) \in \mathbb{R}^3 : 0 < x_2 < 1\}$.

This statement would be false only if there exists some point (x, w) belonging to both P^{LP} and $D(e^1, e^2, 0, 0)$ with $w > 0$. But if (x, w) belongs to P^{LP} and $w > 0$, the inequalities in (4.7) immediately imply that $0 < x_1 + x_2 < 2$ and $-1 < x_1 - x_2 < 1$. Therefore (see Figure 4.6)

$$\left\{(x, w) \in P^{LP} : w > 0\right\} \subseteq S_1 \cup S_2 = \mathbb{R}^3 \setminus D(e^1, e^2, 0, 0),$$

and hence (x, w) does not belong to $D(e^1, e^2, 0, 0)$. The result then follows. □

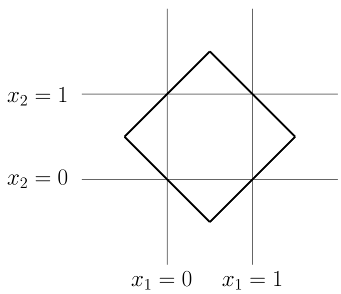

Figure 4.6: The set $\{x \in \mathbb{R}^2 : 0 < x_1 + x_2 < 2, -1 < x_1 - x_2 < 1\}$ is the interior of the depicted quadrilateral.

Next we show that this cut cannot be obtained from basic relaxations.

Lemma 10. *The inequality $w \leq 0$ is not valid for $\bigcap_{j=1}^{4} \text{conv}(P_j)$.*

Proof. Observe that the points

$$p_1 = (1, 1, 0),\ p_2 = (0, 0, 0),\ p_3 = (1, 0, 0),\ p_4 = (0, 1, 0)$$

all belong to P and therefore to $\text{conv}(P_j)$ for $j = 1, \ldots, 4$. Also, the points

$$q_1 = (0, 0, 1),\ q_2 = (1, 1, 1),\ q_3 = (0, 1, 1),\ q_4 = (1, 0, 1)$$

belong to, respectively, P_1, P_2, P_3 and P_4 (q_j violates only the jth constraint defining P^{LP}). But $(p_j + q_j)/2 = (1/2, 1/2, 1/2)$ for $j = 1, \ldots, 4$, and therefore the point $(1/2, 1/2, 1/2)$ belongs to $\bigcap_{j=1}^{4} \text{conv}(P_j)$ but violates $w \leq 0$. □

Theorem 5 follows from the previous two lemmas:

Theorem 5 (restated). *Let $\{P_j\}_{j \in J}$ denote the set of basic relaxations of P. Then*

$$P^{LP} \cap \left(\bigcap_{j \in J} \text{conv}(P_j) \right) \not\subseteq CC(P).$$

4.8 Cross Cuts that Cannot be Obtained from 2-row Relaxations

In this section we prove Theorem 4, namely we exhibit a mixed-integer set such that the intersection of all cuts from its 2-row relaxations does not dominate its cross closure. The polyhedron we work with in this section is

$$P^{LP} = \Big\{(x,y) \in \mathbb{R}^2 \times \mathbb{R}^4 \ : x_1 = \frac{1}{2} + \frac{1}{2}y_1 - \frac{1}{2}y_4, \quad x_2 = \frac{1}{2} + \frac{1}{2}y_1 - \frac{1}{2}y_3, \quad -y_1 - y_2 + y_3 + y_4 = 0, \quad y \geq 0 \Big\}, \tag{4.8}$$

and the associated mixed-integer set $P = (P^{LP}, \mathbb{Z}^2 \times \mathbb{R}^4)$.

Observation 1. *The set P contains the points $p^k = (x^k, y^k)$ for $k = 1, \ldots, 4$ given by*

$$\begin{aligned} x^1 &= (0,0), y^1 = (0,2,1,1) \\ x^2 &= (1,1), y^2 = (2,0,1,1) \\ x^3 &= (0,1), y^3 = (1,1,0,2) \\ x^4 &= (1,0), y^4 = (1,1,2,0). \end{aligned} \tag{4.9}$$

Moreover, the points p^1, p^2, and p^3 are affinely independent.

For convenience, we define

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

so that $P^{LP} = \{(x,y) \in \mathbb{R}^2 \times \mathbb{R}^4 : Dx - Ay = b, y \geq 0\}$. We use A_i to denote the ith row of A. Note that $\text{rank}(A) = 3$, and so $\dim(P) \leq 3$. On the other hand, P contains the affinely independent points p^1, p^2, p^3 and $(1/2, 1/2, 0, 0, 0, 0)$, and so $\dim(P) = 3$.

P^{LP} can be obtained from the polyhedron in (4.7) by: (i) introducing slack variable y_i to convert the ith ($i = 1, \ldots, 4$) inequality to an equation, e.g., $-x_1 - x_2 + w + y_1 = 0$; (ii) Replacing w in the second to the fourth equations by $x_1 +$

$x_2 - y_1$ (obtained from the first equation) and, (iii) subtracting the third and fourth equations from the second equation, and then dividing the third and fourth equations by 2. It follows from the above operations that there is a one-to-one correspondence between the solutions of (4.7) and (4.8). For any solution (x_1, x_2, w) of (4.7), one gets a solution $(x_1, x_2, y_1, \ldots, y_4)$ of (4.8) by keeping x_1, x_2 unchanged and letting $y_1, \ldots, y_4$ stand for the slacks of the inequalities in (4.8). Conversely, for any solution $(x_1, x_2, y_1, \ldots, y_4)$ of (4.8), $(x_1, x_2, x_1 + x_2 - y_1)$ or $(x_1, x_2, 1 - (y_1 + \cdots + y_4)/4)$ is a solution of (4.7). The latter claim follows from the fact that adding up the four constraints in (4.7) (after introducing the slack variables) yields $4w+y_1+y_2+y_3+y_4 = 4$.

Any 2-row relaxation of P is of the form

$$P(M) = (P^{LP}(M), \mathbb{Z}^2 \times \mathbb{R}^4), \quad P^{LP}(M) = \left\{(x,y) \in \mathbb{R}^2 \times \mathbb{R}^4 : MDx - MAy = Mb, y \geq 0\right\}$$

for a 2×3 matrix M. To prove Theorem 4, we will show that

$$P^{LP} \cap \Big(\bigcap_{M \in \mathbb{R}^{2\times 3}} \operatorname{conv}(P(M)) \Big) \not\subseteq CC(P).$$

Before starting, we observe that it is sufficient to consider matrices M that have full row rank.

Lemma 11. *For any $M \in \mathbb{R}^{2\times 3}$, there is a rank 2 matrix $M' \in \mathbb{R}^{2\times 3}$ such that* $\operatorname{conv}(P(M')) \subseteq \operatorname{conv}(P(M))$.

Proof. Clearly there exists a rank 2 matrix $M' \in \mathbb{R}^{2\times 3}$ such that $\operatorname{rowspan}(M') \supseteq \operatorname{rowspan}(M)$. It is easy to verify that such M' satisfies $P^{LP}(M') \subseteq P^{LP}(M)$, and hence $\operatorname{conv}(P(M')) \subseteq \operatorname{conv}(P(M))$. □

We start by showing that the inequality $cy \geq 4$, where $c = (1, 1, 1, 1)$, is a cross cut for P. Notice that the inequality $cy \geq 4$ translates to the inequality $w \leq 0$ for the polyhedron (4.7).

Lemma 12. *The inequality $cy \geq 4$ is a cross cut for P.*

Proof. We will show that $cy \geq 4$ is a cross cut for P derived from the cross disjunction $D(e^1, e^2, 0, 0) = \mathbb{R}^6 \setminus (S_1 \cup S_2)$, where e^i is the ith unit vector in $\mathbb{R}^6$, $S_1 = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^4 : 0 < x_1 < 1\}$ and $S_2 = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^4 : 0 < x_2 < 1\}$.

This statement would be false only if there exists some point $(\bar{x}, \bar{y})$ belonging to both P^{LP} and $D(e^1, e^2, 0, 0)$ with $c\bar{y} < 4$. But if $(\bar{x}, \bar{y})$ belongs to P^{LP} and $c\bar{y} < 4$, then $(\bar{x}, \bar{w})$ with $\bar{w} = 1 - cy/4$ is a solution of (4.7) with $\bar{w} > 0$. As in the proof of Lemma 9, we then infer that $0 < \bar{x}_1 + \bar{x}_2 < 2$ and $-1 < \bar{x}_1 - \bar{x}_2 < 1$, hence $(\bar{x}, \bar{y}) \in S_1 \cup S_2 = \mathbb{R}^6 \setminus D(e^1, e^2, 0, 0)$ and thus $(\bar{x}, \bar{y})$ does not belong to $D(e^1, e^2, 0, 0)$. The result then follows. □

We will next show that there exists a point $(\bar{x}, \bar{y}) \in P^{LP} \cap \big(\bigcap_{M \in \mathbb{R}^{2\times 3}} \text{conv}(P(M))\big)$ such that $c\bar{y} < 4$, and hence the cross cut $cy \geq 4$ is not valid for this set. To this end, we will show that for any $M \in \mathbb{R}^{2\times 3}$ we can construct a point $(x(M), y(M)) \in P^{LP} \cap \text{conv}(P(M))$ such that $cy(M) \leq 3$. We will then apply the Height Lemma using these points and a common base formed by points p^1, p^2 and p^3 presented in Observation 1. The following lemma, whose proof is deferred to Section 4.8.1, shows the existence of the points mentioned above.

Lemma 13. *Consider a matrix $M \in \mathbb{R}^{2\times 3}$ of rank 2. Then, there is a point (x, y) with the following properties: (i) $(x, y) \in P^{PL} \cap \text{conv}(P(M))$; (ii) $cy \leq 3$; (iii) $\|(x, y)\| \leq 6$.*

Using Lemma 13 we next prove Theorem 4:

Theorem 4 (restated). *The crooked cross cut closure of P cannot be obtained by all cuts from its 2-row relaxations. More precisely,*

$$P^{LP} \cap \Big(\bigcap_{M \in \mathbb{R}^{2\times 3}} \text{conv}\left(P(M)\right)\Big) \not\subseteq CC(P).$$

Proof. Consider a matrix $M \in \mathbb{R}^{2\times 3}$. Using Lemma 13 (and Lemma 11 if necessary), find a point $(x(M), y(M))$ in $P^{LP} \cap \text{conv}(P(M))$ such that $cy(M) \leq 3$ and $\|(x(M), y(M))\| \leq 6$. Also, for $i = 1, 2, 3$, the affinely independent points p^i in Observation 1 belong to $P^{LP} \cap \text{conv}(P(M))$ and satisfy $cy^i = 4$. Then applying Corollary 2 (with $A = \text{aff}(P^{LP})$, $a = (0, 0, -c)$, $b = -4$ and $b' = -3$), we conclude that the set

$$Q \;=\; \bigcap_{M \in \mathbb{R}^{2\times 3}} \text{conv}\left(p^1, p^2, p^3, (x(M), y(M))\right)$$

contains a point (x^*, y^*) satisfying $cy^* < 4$. Note that it is possible to apply Corollary 2 because the dimension of $\text{aff}(P)$ is 3.

Since $P^{LP} \cap (\bigcap_{M \in \mathbb{R}^{2\times 3}} \text{conv}(P(M)))$ contains Q, it also contains (x^*, y^*). This shows that the cut $cy \geq 4$ is not valid for this set; together with Lemma 12, this concludes the proof of Theorem 4. □

4.8.1 Proof of Lemma 13

Let $M \in \mathbb{R}^{2\times 3}$ be a rank 2 matrix. Since $\text{rank}(A) = 3$, this implies that $\text{rank}(MA) = 2$. We will construct the points $(x(M), y(M))$ satisfying the properties of the lemma in three steps. In the first step, we will construct points in $\text{aff}(P^{LP})$, which violate $cy \geq 4$, but do not belong to $P(M)$; informally, these points almost belong to $P^{LP} \cap P(M)$, except that they do not satisfy the required non-negativity condition. In the second step, we create two directions d^1 and d^2 in order to 'correct' the points constructed in the first step. In the final step, we use these directions to correct the points created in the first step, obtaining the desired point $(x(M), y(M))$ in $P^{LP} \cap P(M)$ but still violating $cy \geq 4$.

Step 1. Consider the points $(x^i, y^i) \in P$ for $i = 1, \ldots, 4$ from Observation 1, and recall that they all satisfy $cy^i = 4$. Since they belong to P, we have $Dx^i - Ay^i = b$ for $i = 1, \ldots, 4$. Moreover, since $Ac = 0$, we have $Dx^i - A(y^i - c/2) = b$ for all i, which then implies $MDx^i - MA(y^i - c/2) = Mb$ for all i. In other words, the points $(x^i, \bar{y}^i) = (x^i, y^i - c/2)$ $(i = 1, \ldots, 4)$ satisfy the equations defining both P^{LP} and $P(M)$ but violate one non-negativity inequality each, as

$$\begin{aligned}
\bar{y}^1 &= (0, 2, 1, 1) - c/2 = (-1, 3, 1, 1)/2 \\
\bar{y}^2 &= (2, 0, 1, 1) - c/2 = (3, -1, 1, 1)/2 \\
\bar{y}^3 &= (1, 1, 0, 2) - c/2 = (1, 1, -1, 3)/2 \\
\bar{y}^4 &= (1, 1, 2, 0) - c/2 = (1, 1, 3, -1)/2.
\end{aligned} \tag{4.10}$$

Note that each point above has exactly one negative coefficient which equals $-1/2$, and the remaining coefficients are strictly positive and at least $1/2$. These four points also violate the inequality $cy \geq 4$, as $c \cdot c = 4$ and therefore, $(x^i, y^i - c/2)$ satisfies $c(y^i - c/2) = 2$.

Step 2. We now define the 'correcting' directions $d^1, d^2 \in \mathbb{R}^4$. To do so, recall that $\text{rowspan}(A)$ has dimension 3 and by assumption $\text{rowspan}(MA)$ is a 2-dimensional subspace of $\text{rowspan}(A)$. If $A_3 \notin \text{rowspan}(MA)$, let $i^* = 3$, and if $A_3 \in \text{rowspan}(MA)$, let $i^* \in \{1, 2\}$ be the index such that A_{i^*} does not belong to $\text{rowspan}(MA)$. Notice that the rows of MA together with A_{i^*} span exactly $\text{rowspan}(A)$.

Now define $d^1, d^2 \in \mathbb{R}^4$ to be solutions of the following two systems of four equations each (the coefficient η is specified later):

$$\begin{bmatrix} MA \\ A_{i^*} \\ c \end{bmatrix} d^1 = \begin{bmatrix} \mathbf{0} \\ 1 \\ \eta \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} MA \\ A_{i^*} \\ c \end{bmatrix} d^2 = \begin{bmatrix} \mathbf{0} \\ -1 \\ \eta \end{bmatrix}. \tag{4.11}$$

As the rows of MA together with A_{i^*} span exactly rowspan(A) and the vector c is orthogonal to the rows of A and hence to rowspan(A), the matrix in the left-hand side of equations in (4.11) (which is the same) is invertible. Therefore, these systems have unique nonzero solutions.

We will show that for some η, there are scalars $\lambda^1, \lambda^2 > 0$ such that $\lambda^1 d^1$ and $\lambda^2 d^2$ are nonzero vectors satisfying the following properties:

1. $MA(\lambda^1 d^1) = MA(\lambda^2 d^2) = 0$.
2. There exists an $\alpha \in (0, 1)$ such that $A(\alpha\lambda^1 d^1 + (1 - \alpha)\lambda^2 d^2) = 0$.
3. $\max_i \lambda^1 d^1_i = 1/2 = \max_i |\lambda^1 d^1_i|$ and $\max_i \lambda^2 d^2_i = 1/2 = \max_i |\lambda^2 d^2_i|$.
4. $c\lambda^1 d^1 \leq 1$ and $c\lambda^2 d^2 \leq 1$.

The motivation for these properties is the following: (i) the first and second properties will ensure that the 'corrected' vectors $(x^i, \bar{y}^i + (\alpha d^1 + (1 - \alpha)d^2))$ still satisfy all the constraints of P and $P^I(M)$, except for the non-negativity conditions; (ii) we will use the third property to argue that there is an index i such that the corresponding corrected vector satisfies the non-negativity conditions, and hence belongs to $P^{LP} \cap P(M)$; (iii) the fourth property will ensure that the corrected vector does not satisfy the inequality $cy \geq 4$.

Note that Properties 1 and 2 hold for all λ independent of the choice of η. Property 1 follows directly from the first two equations in both systems in (4.11). For Property 2, since $A_{i^*} d^1 = 1$ and $A_{i^*} d^2 = -1$, we have that $A_{i^*}(d^1 + d^2) = 0$. Therefore, $d^1 + d^2$ is orthogonal to the rows of MA and to A_{i^*}, and hence to the rows of A (as rows of MA and A_{i^*} span rowspan(A)).

In order to obtain Properties 3 and 4 we need to rescale the vectors d^1 and d^2, but notice that this operation preserves Properties 1 and 2. We consider two cases depending on whether A_3 belongs to rowspan(MA) or not, and set the coefficient η appropriately.

Case 1: $A_3 \in \text{rowspan}(MA)$**.** Set $\eta = 0$. In this case, the last constraint in both systems in (4.11) (which are identical) guarantee that $\lambda^1 d^1$ and $\lambda^2 d^2$ satisfy Property 4 for all λ^1, λ^2.

We now consider Property 3 for a rescaling of d^1; the proof for d^2 is identical. Since A_3 belongs to rowspan(MA), the first two constraints in the first system in (4.11) guarantee that $A_3 d^1 = 0$, and therefore $d^1_1 + d^1_2 = d^1_3 + d^1_4$. The last constraint implies that $d^1_1 + d^1_2 + d^1_3 + d^1_4 = 0$. In addition, $d^1 \neq 0$ as $A_{i^*} d^1 \neq 0$. Therefore, $d^1_1 + d^1_2 = d^1_3 + d^1_4 = 0$ and hence $\max_i d^1_i = \max_i |d^1_i|$, so we can multiply d^1 by an appropriate positive scalar λ^1 so that $\max_i \lambda^1 d^1_i = 1/2$. The vector $\lambda^1 d^1$ then satisfies Properties 1,2,3, and 4.

Case 2: $A_3 \notin \text{rowspan}(MA)$**.** Set $\eta = 1$. In this case $i^* = 3$, namely both systems in (4.11) contain a constraint of the form $A_3 d = \pm 1$ (instead of the implied constraint $A_3 d = 0$ in the previous case). Adding the third and fourth constraints in the first system in (4.11), we get $d^1_3 + d^1_4 = 1$. Subtracting the third constraint from the fourth constraint, we get $d^1_1 + d^1_2 = 0$. Therefore $\max_i d^1_i = \max_i |d^1_i| \geq 1/2$. We can then rescale d^1 by $\lambda^1 \in (0, 1]$ so that $\lambda^1 d^1$ satisfies Property 3. Further, $\lambda^1 d^1$ satisfies Property 4, since $cd^1 \leq 1$. Therefore, $\lambda^1 d^1$ satisfies Properties 1,2,3 and 4.

As for d^2, adding and subtracting constraints as in the case of d^1, we see that $d^2_3 + d^2_4 = 0$ and $d^2_1 + d^2_2 = 1$. Once again we can scale d^2 so that it satisfies all properties.

Step 3. Consider the vectors $\lambda^1 d^1$ and $\lambda^2 d^2$ from the previous step. Let $i = \text{argmax}_k\, d^1_k$ and $j = \text{argmax}_k\, d^2_k$. As $\lambda^1 d^1$ is nonzero, and because of Property 3, we have $\lambda^1 d^1_i = 1/2$ and $\bar{y}^i + \lambda^1 d^1 \geq 0$. Property 1 implies that $MDx - MA(\bar{y}^i + \lambda^1 d^1) = MDx - MA\bar{y}^i = Mb$, and therefore $(x^i, \bar{y}^i + \lambda^1 d^1)$ belongs to $P(M)$ (but not to P^{LP}, since we can still have $Dx^i - A(\bar{y}^i + \lambda^1 d^1) \neq b$). Also, Property 4 implies that $c(\bar{y}^i + \lambda^1 d^1) \leq 3$, and hence the point does not satisfy the inequality $cy \geq 4$. Similarly, Properties 1 and 3 imply that $(x^j, \bar{y}^j + \lambda^2 d^2) \in P(M)$, and $c(\bar{y}^j + \lambda^2 d^2) \leq 3$.

Finally, by Property 2 there is an $\alpha \in (0, 1)$ such that the point

$$(x(M), y(M)) \triangleq \alpha(x^i, \bar{y}^i + \lambda^1 d^1) + (1 - \alpha)(x^j, \bar{y}^j + \lambda^2 d^2)$$

satisfies $Dx(M) - Ay(M) = Dx^i - A\bar{y}^i = b$. Therefore, this point $(x(M), y(M))$ belongs to $P^{LP} \cap \text{conv}(P(M))$. In addition, we clearly have $cy(M) \leq 3$, and it is easy to verify that $\|(x(M), y(M))\| \leq 6$. This concludes the proof of Lemma 13.

CHAPTER 5

Probabilistic Analysis of the Strength of the Split and Triangle Closures

In this chapter we study the strength of the split and triangle cuts for the *continuous relaxation* or *relaxed corner polyhedra* (RCPs) introduced in Section 2.3.4. Recall from Section 3.2.2 that Basu et. al [27] showed examples of RCPs whose split closure can be arbitrarily worse than their triangle closure, under the blow-up measure given by (3.2). However, also recall that despite experiments carried out by several authors [75, 21, 26, 67], the usefulness of triangle cuts in practice has fallen short of its theoretical strength.

In order to understand this issue, we consider two types of measures between the closures: the 'worst-cost' blow-up measure, where we look at the direction where the split closure has the largest gap, and the 'average-cost' measure which takes an average over all directions. Moreover, we consider a natural model for generating random RCPs. Our first result is that, under the worst-cost measure, a random RCP has a weak split closure with reasonable probability. This shows that the bad examples given by Basu et. al are not pathological cases. However, when we consider the average-cost measure, with high probability both split and triangle closures obtain a very good approximation of the integer hull of the RCP. The above result holds even if we replace split cuts by the simple split or Gomory cuts. This gives an indication that split/Gomory cuts are indeed as useful as triangle cuts, except in specific situations.

Two recent papers address the fundamental question of comparing the strengths of triangle and split cuts from a probabilistic point of view. He et al. [93] use the same random model for generating RCP's, but a different measure to compare the strength of cuts, comparing the random coefficients of the inequalities induced by the randomness of the rays. Their analysis does not consider the important triangles of Type 3. Although the results cannot be directly compared, their paper also indicates that split cuts perform at least as well as some classes of triangles.

Del Pia et al. [109] base their analysis on the lattice width of the underlying convex set. They show that the importance of triangle cuts generated from Type 2 triangles (the same family which was considered in [27]) decreases with decreasing lattice width, on average. They also have results for triangles of Type 3 and for quadrilaterals.

Our approach is very different from these two papers.

Organization of the chapter. We start off by recalling some definitions relative to RCPs, and then in Section 5.2 we introduce the probabilistic model that we use for generating random RCPs and the measure of strength used to compare closures, including the average-cost measure. In Section 5.3 we show that with constant probability, random RCP has its split closure significantly weaker that its triangle closure according to the worst-cost measure. In Section 5.4, we show that under the average-cost measure, with high probability even the simple-splits closure provides a good approximation of the integer hull. Finally, in Section 5.5 we discuss the extension of the latter result to a mixed-integer version of the relaxed corner polyhedra.

Acknowledgments. This chapter is joint work with Amitabh Basu and Gérard Cornuéjols.

5.1 Preliminaries

Recall that the *continuous relaxation* or *relaxed corner polyhedron* (RCP) is a MIP of the form

$$\begin{aligned} &\min cs \\ &x = f + \sum_{j=1}^{n} r^j s_j \\ &x \in \mathbb{Z}^m, \\ &s \geq 0, s \in \mathbb{R}^n. \end{aligned} \tag{RCP}$$

Notice that we are now keeping an objective function as well, and not only focusing on the set of feasible solutions. Therefore, an RCP is defined by the vectors $f, r^1, \ldots, r^n$ and the cost vector c. We call a tuple $\langle f, r^1, r^2, \ldots, r^n \rangle$ an *ensemble*. Given an ensemble $\mathcal{E}$ and cost vector c, we use $RCP(\mathcal{E}, c)$ to denote the corresponding RCP.

Intersection cuts. We recall the definition of intersection cuts for the relaxed corner polyhedron from Section 2.3.4. Consider a convex lattice-free set $X \subseteq \mathbb{R}^m$ (namely $\text{int}(X) \cap \mathbb{Z}^m = \emptyset$) that contains f in its interior. The function $\psi_X : \mathbb{R}^m \to \mathbb{R}$ is defined as the gauge function of $K - \{f\}$, namely

$$\psi_X(r) = \inf \left\{ \lambda > 0 : f + \frac{r}{\lambda} \in X \right\}. \tag{5.1}$$

The inequality

$$\sum_{j=1}^{n} \psi_X(r^j) s_j \geq 1$$

is valid for $RCP(\mathcal{E}, c)$ and it is called an *intersection cut* or *lattice-free cut*. Conversely, it was shown in [40] that all (minimal) cuts for $RCP(\mathcal{E}, c)$ can be obtained in this way.

One important family of inequalities is derived from sets X which are (the closure of) split sets, that is, X is of the form $\{x \in \mathbb{R}^m : \gamma \leq \pi x \leq \gamma + 1\}$ for some $\pi \in \mathbb{Z}^m$ and $\gamma \in \mathbb{Z}$. These are denoted as *split cuts*. Also of particular importance are *simple-split* cuts, that is, cuts from sets of the form $\{x \in \mathbb{R}^m : \gamma \leq e^i x \leq \gamma + 1]\}$, where e^i is the ith canonical vector in $\mathbb{R}^m$. Finally, in the two-dimensional case $m = 2$, a cut obtained from sets X that are a triangle.

Working on the s-space. It will be much more convenient for us to work only on the s-space, instead of the original (x, s)-space of (RCP). So define $P(\mathcal{E})$ as the projection of the feasible region of $RCP(\mathcal{E}, c)$ onto the s-space. We also use $P_L(\mathcal{E})$ to denote the projection of the linear relaxation of $RCP(\mathcal{E}, c)$ onto the s-space.

Similarly, we define the projected version of the split closure, denoted by $S(\mathcal{E})$, as the intersection of the relaxation $P_L(\mathcal{E})$ and all valid split cuts for $RCP(\mathcal{E}, c)$. Define the projected simple-split closure $G(\mathcal{E})$ and (for the case $m = 2$) the projected triangle closure $T(\mathcal{E})$ similarly.

One justification why we can restrict to these projected sets is the following. By the construction of RCP's, optimizing a function of the form cs over the feasible region of $RCP(\mathcal{E}, c)$ yields the same value as optimizing this function over the projected region $P(\mathcal{E})$; a similar observation holds for all projected sets defined above. Since the measures of strength that we use are based on the gap(.,.) construction, these observations imply that it suffices to focus on the gap between the projected regions.

Not only working on the s-space is more natural, but also gives us the following nice property: the recession cone of $P(\mathcal{E})$ is $\mathbb{R}^n_+$ (see [51]) and hence $P(\mathcal{E})$ is of blocking type. Since $P(\mathcal{E}) \subseteq S(\mathcal{E})$, this implies that $S(\mathcal{E})$ is also of blocking type, and so are the closures $G(\mathcal{E})$ and $T(\mathcal{E})$.

5.2 Random Model and Measures of Strength

Random model. Let $\mathcal{D}^m_n$ denote the distribution of ensembles $\langle f, r^1, \ldots, r^n \rangle$ where f is picked uniformly from $[0, 1]^m$ and each of $r^1, \ldots, r^n$ is picked independently and uniformly at random from the set of unit vectors in $\mathbb{R}^m$, i.e. the rays $r^1, \ldots, r^n$ are sampled uniformly from $\mathbb{S}^{n-1}$. We make a note here that the rays in an RCP can be assumed to be unit vectors, by suitably scaling the cost coefficients. Hence, in our model, we assume the rays are sampled from the set of unit vectors. When the dimension is 2, we write $\mathcal{D}_n$ for the distribution, omitting the superscript.

Measures of strength. Recall some definitions from Section 3.1. A closed, convex set $X \subseteq \mathbb{R}^n_+$ is said to be of blocking type if $y \geq x \in X$ implies $y \in X$. Given two sets A and B of blocking type, the *gap* between these sets with respect to the cost vector c is defined as

$$\text{gap}(A, B, c) = \frac{\inf\{cs : s \in A\}}{\inf\{cs : s \in B\}}, \tag{5.2}$$

whenever both numerator and denominator are finite. Notice that this value is greater than 1 if A is contained in B. We define the gap to be $+\infty$ if A is empty or $\inf\{cs : s \in B\} = 0$.

Based on this idea, we can define the *worst-cost* measure between the two sets A and B as the worst possible gap over all non-negative cost vectors:

$$\text{wc}(A, B) = \sup_{c \in \mathbb{R}^m_+} \{\text{gap}(A, B, c)\} \tag{5.3}$$

$$= \sup_{c \in [0,1]^m} \{\text{gap}(A, B, c)\}, \tag{5.4}$$

where the second equation follows from the fact that the ratios are preserved under positive scaling of the cost vectors. Note that for convex sets of blocking type, only non-negative cost vectors have bounded optimum, motivating why we restrict ourselves to this case.

For any convex set X of blocking type, define $\alpha X = \{\frac{x}{\alpha} : x \in X\}$. Recall that the blow-up measure between A and B is

$$\text{bu}(A, B) = \inf\{\alpha : \alpha A \supseteq B\}.$$

It was shown in [27] that, whenever A is polyhedral, then have the equivalence $\text{wc}(A, B) = \text{bu}(A, B)$. In Section 5.6 we prove a generalization of this result that does not rely on the polyhedrality of any of these sets.

Now we define another (more robust) measure of strength which tries to capture the average strength with respect to different costs. Consider a distribution $\mathcal{C}$ over vectors in $\mathbb{R}^n_+$. Then, the *average-cost* measure between A and B is defined by

$$\text{avg}(A, B, \mathcal{C}) = \mathbb{E}_{c \sim \mathcal{C}}\left[\text{gap}(A, B, c)\right]. \tag{5.5}$$

5.3 Worst-cost Measure in $\mathbb{R}^2$

The main result of this section is that, for a significant fraction of the RCP's in the plane, $S(\mathcal{E})$ is significantly worse than $T(\mathcal{E})$ based on the worst-cost measure.

Theorem 9. *For any $\alpha \geq 1$ and $\beta \in [0, 1]$, a random ensemble $\mathcal{E} \sim \mathcal{D}_n$ satisfies*

$$\Pr\left(\text{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \alpha\right) \geq \left[1 - 2\left(1 - g(\frac{\beta}{4\alpha})\right)^n\right]\left[\frac{1-\beta}{\alpha} - \frac{1-\beta^2}{4\alpha^2}\right],$$

n	α	β	Pr
100	1.5	0.37	25.7%
100	2	0.43	16.7%
500	2	0.16	33.6 %
500	3	0.22	21.3%
1000	2	0.01	37.7%
1000	3	0.14	25.0 %
1000	4	0.17	18.2 %
$+\infty$	2	0	43.75 %
$+\infty$	4	0	30.56 %

Table 5.1: Values of the bound of Theorem 9 for different values of n and approximation factor α. The value of β in every entry was chosen empirically and attempts to optimize the bound.

where

$$g(x) = \left(\frac{x}{0.75 - (2 - \sqrt{2})x} - \frac{x}{1 - (2 - \sqrt{2})x} \right).$$

Notice that this bound increases as n grows. In the limit $n \to \infty$, and using the optimal choice $\beta \to 0$, the bound becomes $1/\alpha - 1/(4\alpha^2)$. To obtain an idea about the probabilities in the above theorem, Table 5.1 presents the bound obtained for different values of n and α.

The way to prove this result is to consider a particular (deterministic) ensemble $\langle f, r^1, r^2 \rangle$ which is 'bad' for the split closure and show that it appears with significant probability in a random ensemble. We employ the following monotonicity property to transfer the 'badness' to the whole RCP. The proof appears at the end of Section 5.6.

Lemma 14. *Consider an ensemble* $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$ *and let* $\mathcal{E}' = \langle f, r^{i_1}, r^{i_2}, \ldots, r^{i_k} \rangle$ *be a sub-ensemble of it. Then* $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \mathrm{wc}(T(\mathcal{E}'), S(\mathcal{E}'))$.

5.3.1 A Bad Ensemble for the Split Closure

First, we introduce the following notation: Given a point f and a ray r, we say that $f + r$ *crosses* a region $R \subseteq \mathbb{R}^n$ if there is $\lambda \geq 0$ such that $f + \lambda r \in R$.

In this part we will focus on ensembles $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f \in (0,1)^2$, and $f + r^1$ and $f + r^2$ cross the open segment connecting $(0,0)$ to $(0,1)$. The high-level idea is the following. Suppose that r^1 an r^2 have x_1-value equal to -1 and consider a lattice-free triangle T containing the points $f + r^1$ and $f + r^2$, and also containing f in its interior. This triangle gives an inequality which is at least as strong as $s_1 + s_2 \geq 1$, hence we have a lower bound of 1 for minimizing $s_1 + s_2$ over the triangle closure $T(\mathcal{E})$. However, further assume that the angle between rays r^1 and r^2 is large. Then we can see that any split that contains f in its interior will have a very large coefficient for either s_1 or s_2. More specifically, suppose that there is a large M such that, for every inequality $\psi(r^1)s_1 + \psi(r^2)s_2 \geq 1$ coming from a split, we have $\max\{\psi(r^1), \psi(r^2)\} \geq M$. Then the point $(s_1, s_2) = (1/2M, 1/2M)$ satisfies every such inequality and hence is feasible for the split closure $S(\mathcal{E})$; this gives an upper bound of $2/M$ for minimizing $s_1 + s_2$ over the split closure. Then using the choice of $c = [1, 1]$ in the maximization in (5.3) gives $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq M/2$.

The following lemma, whose proof is presented in Section 5.8, formalizes the observation that if r^1 and r^2 are spread out then the split closure is weak.

Lemma 15. *Consider an ensemble $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f = (f_1, f_2) \in (0, 1)^2$, $r^1 = c_1(-1, t_1)$ and $r^2 = c_2(-1, t_2)$ with $c_1, c_2 \geq 0$ and $t_1 \geq t_2$. Moreover, assume that both $f + r^1$ and $f + r^2$ cross the left facet of the unit square. Then*

$$\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E})\} \leq \frac{2}{t_1 - t_2}.$$

Now we are ready to establish the main lemma of this section, which exhibits bad ensembles for the split closure.

Lemma 16. *Consider an ensemble $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f = (f_1, f_2) \in (0, 1)^2$. Suppose that $f + r^1$ crosses the open segment connecting $(0, 1 - \epsilon)$ and $(0, 1)$ and that $f + r^2$ crosses the open segment connecting $(0, 0)$ and $(0, \epsilon)$, for some $0 \leq \epsilon < 1/2$. Then $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq (1 - 2\epsilon)/2f_1$.*

Proof. Let $v^1 = (-1, t_1)$, $v^2 = (-1, t_2)$ and let $c_1, c_2 \geq 0$ be such that $r^1 = c_1 v^1$ and $r^2 = c_2 v^2$. By the assumptions on the rays, we have $t_1 \geq t_2$.

Consider the rays $\underline{v}^1 = (-1, \underline{t}_1)$ and $\underline{v}^2 = (-1, \underline{t}_2)$ such that $f + \underline{v}^1$ crosses the point $(0, 1 - \epsilon)$ and $f + \underline{v}^2$ crosses the point $(0, \epsilon)$ (see Figure 5.1).

Notice that $t_1 \geq \underline{t}_1 \geq \underline{t}_2 \geq t_2$, implying that $t_1 - t_2 \geq \underline{t}_1 - \underline{t}_2$. Moreover, using similarity of the triangles with vertices f, $(0, 1 - \epsilon)$, $(0, \epsilon)$ and f, $f + \underline{v}^1$, $f + \underline{v}^2$, we obtain that $\underline{t}_1 - \underline{t}_2 = \frac{(1-2\epsilon)(1+f_1)}{f_1}$. Therefore, $t_1 - t_2 \geq (1 - 2\epsilon)/f_1$.

Employing Lemma 15 over $\mathcal{E}$ gives $\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E})\} \leq 2f_1/(1 - 2\epsilon)$. In contrast, $\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in T(\mathcal{E})\} \geq 1$, because of the inequality

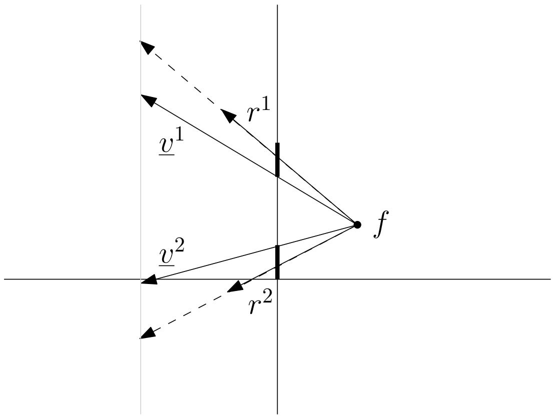

Figure 5.1: This figure illustrates the construction used in the proof of Lemma 16. The vectors $r^1, r^2, \underline{v}^1$ and $\underline{v}^2$ are represented by solid arrows, while the vectors v^1 and v^2 are represented by dashed arrows.

$c_1 s_1 + c_2 s_2 \geq 1$ derived from the lattice-free triangle with vertices $f + v^1$, $f + v^2$ and $f - (\gamma, 0)$ for some small $\gamma > 0$. Notice that such γ exists because $f + v^1$ and $f + v^2$ do not cross the points $(0, 1)$ and $(0, 0)$ respectively. Using the cost vector $c = [c_1, c_2]$, we obtain the desired bound $\text{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq (1 - 2\epsilon)/2f_1$. □

5.3.2 Probability of Bad Ensembles

Using the ensemble constructed in the previous section and the monotonicity property from Lemma 14, we now analyze the probability that a random ensemble $\mathcal{E} \sim \mathcal{D}_n$ is bad for the split closure. Let Δ denote the triangle in $\mathbb{R}^2$ with vertices $(0, 0)$, $(0, 1)$, $(1/2, 1/2)$.

Lemma 17. *Let $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$ be a random ensemble from $\mathcal{D}_n$, where $f = (f_1, f_2)$. Then for all $\bar{f} = (\bar{f}_1, \bar{f}_2) \in \Delta$ and all $\epsilon \in (0, 1/2)$, we have*

$$\Pr\left(\text{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \frac{1 - 2\epsilon}{\bar{f}_1} \Big| f = \bar{f}\right) \geq 1 - 2\left(1 - g(\bar{f}_1)\right)^n,$$

where

$$g(x) = \left(\frac{x}{1 - \epsilon - (2 - \sqrt{2})x} - \frac{x}{1 - (2 - \sqrt{2})x}\right).$$

Proof. Let us call *portals* the open segment connecting $(0, 1-\epsilon)$ and $(0,1)$ and the open segment connecting $(0,\epsilon)$ and $(0,0)$. Due to Lemmas 14 and 16 it suffices to bound from below the probability that a random ensemble has rays r^i and r^j such that $f + r^i$ crosses one portal and $f + r^j$ crosses the other portal.

Consider a ray r^i; the probability that $f+r^i$ crosses the open segment connecting $(0,1-\epsilon)$ and $(0,1)$ equals to $\theta/2\pi$, where θ is the angle between the vectors $(0,1-\epsilon)-\bar{f}$ and $(0,1)-\bar{f}$. We prove the following lower bound for θ in the appendix.

Claim 1. $\theta \geq g(\bar{f}_1)$.

Therefore, the probability that $\bar{f}+r^i$ crosses the open segment connecting $(0,1-\epsilon)$ and $(0,1)$ is at least $g(\bar{f}_1)$. By symmetry, we can also prove that the probability that $\bar{f}+r^i$ crosses the open segment connecting $(0,\epsilon)$ and $(0,0)$ is also at least $g(\bar{f}_1)$; this bounds also holds for this case because it is independent of $\bar{f}_2$.

Let B_1 denote the event that no ray of $\mathcal{E}$ crosses the open segment connecting $(0,1-\epsilon)$ and $(0,1)$ and let B_2 denote the even that no ray of $\mathcal{E}$ crosses the open segment connecting $(0,\epsilon)$ and $(0,0)$. Using our previous bound we obtain that $\Pr(B_1) \leq (1-g(\bar{f}_1))^n$, and the same lower bound holds for $\Pr(B_2)$. Notice that the probability that $\mathcal{E}$ has rays r^i and r^j such that $f+r^i$ and $f+r^j$ cross distinct portals is $1-\Pr(B_1 \vee B_2)$; from union bound we get that this probability is at least $1-2(1-g(\bar{f}_1))^n$. This concludes the proof of the lemma. □

5.3.3 Proof of Theorem 9

In order to conclude the proof of Theorem 9 we need to remove the conditioning in the previous lemma. To make progress towards this goal, for $t \in [0,1/2]$ let $\Delta_t = \Delta \cap \{(x_1,x_2) : x_1 \leq t\}$. It is easy to see that the area of Δ_t equals $(1-t)t$. Now it is useful to focus on the set $\Delta_t \setminus \Delta_{\beta t}$, for some $\beta \in [0,1]$, since we can bound the probability that a uniform point lies in it and Lemma 17 is still meaningful. Using the independence properties of the distribution $\mathcal{D}_n$ we get that for every $\beta \in [0,1]$ and $\epsilon \in (0,1/2)$ a random ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle \sim \mathcal{D}_n$ satisfies:

$$
\begin{aligned}
&\Pr\left(\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \frac{1-2\epsilon}{2t} \middle| f \in \Delta\right) \\
&\geq \Pr\left(\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \frac{1-2\epsilon}{2t} \middle| f \in \Delta_t \setminus \Delta_{\beta t}\right) \Pr\left(f \in \Delta_t \setminus \Delta_{\beta t} \middle| f \in \Delta\right) \\
&\geq \left[1 - 2\left(1-g\left(\beta t\right)\right)^n\right] \cdot 4 \cdot \left[(1-t)t - (1-\beta t)\beta t\right],
\end{aligned}
$$

where the first inequality follows from the fact that $\Delta_t \setminus \Delta_{\beta t} \subseteq \Delta$ and the second inequality follows from the fact that $\beta t \leq f_1 \leq t$ and that the function $g(x)$ is increasing in x.

Finally, notice that this bound holds for all four 90-degree rotations of Δ around the point $(1/2, 1/2)$; this is because of the symmetries of $\mathcal{D}_n$. Thus, by law of total probability we can remove the last conditioning. Using $\epsilon = 1/4$ and $\alpha = 1/4t$ we then obtain Theorem 9. We remark that we fixed the value of ϵ in order to simplify the expression in the theorem and that the value 1/4 was chosen experimentally in order to obtain good bounds specially for reasonably small values of n.

Since $T(\mathcal{E})$ is a relaxation of $P(\mathcal{E})$, as a corollary of the theorem we obtain a bound on the probability that the split closure is bad for random RCP's.

Corollary 4. *For any $\alpha \geq 1$ and $\beta \in [0,1]$, a random ensemble $\mathcal{E} \sim \mathcal{D}_n$ satisfies*

$$\Pr\left(\mathrm{wc}(P(\mathcal{E}), S(\mathcal{E})) \geq \alpha\right) \geq \left[1 - 2\left(1 - g(\frac{\beta}{4\alpha})\right)^n\right]\left[\frac{1-\beta}{\alpha} - \frac{1-\beta^2}{4\alpha^2}\right],$$

5.4 Average-cost Measure

For $\epsilon > 0$ we define the product distribution $\mathcal{P}_\epsilon$ over $[\epsilon, 1]^n$ where a vector is obtained by taking each of its n coefficients independently uniformly in $[\epsilon, 1]$. In this section we show that $avg(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon)$ is small for most ensembles $\mathcal{E}$ in $\mathcal{D}_n^m$.

Theorem 10. *Fix reals $\epsilon > 0$ and $\alpha > 1$ and an integer $m > 0$. Then for large enough n,*

$$\Pr_{\mathcal{E} \sim \mathcal{D}_n^m}\left(\mathrm{avg}(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon) \leq \alpha\right) \geq 1 - \frac{1}{n}.$$

We remark that the property that the cost vector is bounded away from zero in every coordinate is crucial in our analysis. This is needed because the ratio in (5.2) can become ill-defined in the presence of rays of zero cost.

The high level idea for proving the theorem is the following. Consider an ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$. Define f^1 as the integral point closest to f in l_2 norm. It is not difficult to see that for every $c \in \mathcal{P}_\epsilon$, $\min\{cs : s \in P(\mathcal{E})\}$ is lower bounded by $\epsilon|f^1 - f|$, and this is achieved when the ensemble has the ray $(f^1 - f)/|f^1 - f|$ with cost ϵ. We prove that this lower bound also holds for minimizing over $G(\mathcal{E})$ instead of $P(\mathcal{E})$. In addition, we show that for most ensembles $\mathcal{E}$, there are enough rays similar to $f^1 - f$

that have small cost. This allows us to upper bound $\min\{cs : s \in P(\mathcal{E})\}$ by roughly $\epsilon|f^1 - f|$ for most of the ensembles, which gives the desired result.

We start by proving the upper bound. For that, we need to study a specific subset of the ensembles in $\mathcal{D}^m_n$. We remark that the bounds presented are not optimized and were simplified in order to allow a clearer presentation.

5.4.1 (β, k)-Good Ensembles

Consider an ensemble $\mathcal{E} = \langle f, r^1, \dots, r^n \rangle$. At a high level, we consider special regions in $\mathbb{R}^m$ 'around' $f - f^1$, whose size depends on a parameter $\beta > 0$; then an ensemble is (β, k)-good if it has at least k rays in each of these regions.

To make this precise, let S^{m-1} denote the $(m-1)$-dimensional unit sphere in $\mathbb{R}^m$. Define $t = f^1 - f$ and let ρ be a rotation of $\mathbb{R}^m$ which maps $t/|t|$ into e^m. Let $\bar{C}(\beta)$ be the cap of the hypersphere S^{m-1} consisting of all unit vectors with dot product at least β with e^m. We also define H_i^+ as the halfspace given by $\{x \in \mathbb{R}^m : x_i \geq 0\}$ and $H_i^- = \{x \in \mathbb{R}^m : x_i \leq 0\}$. We use the halfspaces H_i^+ and H_i^- to partition $\bar{C}(\beta)$ into 2^{m-1} parts. That is, for $I \subseteq [m-1]$, let $\bar{C}_I(\beta) = \bar{C}(\beta) \cap (\bigcap_{i \in I} H_i^+) \cap (\bigcap_{i \in [m-1] \setminus I} H_i^-)$. Finally, let $C(\beta) = \rho^{-1}\bar{C}(\beta)$ and $C_I(\beta) = \rho^{-1}\bar{C}_I(\beta)$, that is, the sets obtained by applying the inverse rotation ρ^{-1}.

Using these structures, we say that $\mathcal{E}$ is (β, k)-good if for every $I \subseteq [m-1]$ there are at least k rays r^i in $C_I(\beta)$. The main property of such ensembles is that they allow us to use the following lemma.

Lemma 18. *Let R be a subset of the rays of $\mathcal{E}$ such that $R \cap C_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$. Then there is a solution $s \in P(\mathcal{E})$ supported in R such that $\sum_{i=1}^n s_i \leq \frac{|t|}{\beta}$.*

Proof. Without loss of generality assume that $R \cap C(\beta) = \{r^1, r^2, \dots, r^{n'}\}$. First we show that $t \in cone(R \cap C(\beta))$. This follows from Farkas' Lemma and the hypothesis $R \cap C_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$; the proof is deferred to the appendix.

Claim 2. *$t \in cone(R \cap C(\beta))$.*

So consider $s_1, s_2, \dots, s_{n'} \geq 0$ with $\sum_{i=1}^{n'} s_i r^i = t$. We claim that $\sum_{i=1}^{n'} s_i \leq |t|/\beta$. To see this, first notice that by definition of $C(\beta)$ we have $r(t/|t|) \geq \beta$ for all $r \in C(\beta)$. Then multiplying the equation $\sum_{i=1}^{n'} s_i r^i = t$ by t gives $\sum_{i=1}^{n'} s_i \beta |t| \leq \sum_{i=1}^{n'} s_i r^i t = tt = |t|^2$ and the claim follows.

Since $f + t = f^1$ is integral we obtain that s is a feasible solution for $P(\mathcal{E})$. This concludes the proof of the lemma. □

Using this lemma we can prove an upper bound on optimizing a cost vector in $\mathcal{P}_\epsilon$ over $P(\mathcal{E})$.

Lemma 19. *Fix $\beta, \epsilon > 0$ and an integer $k \geq 0$. Consider a (β, k)-good ensemble $\mathcal{E}$ and let $z(c) = \min\{cs : s \in P(\mathcal{E})\}$. Then*

$$\mathbb{E}_{c \sim \mathcal{P}_\epsilon}\left[z(c)\right] \leq |t| \left(p\frac{\epsilon}{\beta^2} + (1-p)\frac{1}{\beta} \right),$$

where

$$p = 1 - 2^{m-1}\left(\frac{1 - \epsilon/\beta}{1-\epsilon}\right)^k.$$

Proof. Consider a vector c which satisfies the following property: (*) for each $I \subseteq [m-1]$ there is a ray in $C_I(\beta)$ which has cost w.r.t c at most ϵ/β. Then employing Lemma 18 we obtain that $z(c) \leq |t|\epsilon/\beta^2$. Similarly, for a general vector $c \in [\epsilon, 1]^m$ we have the bound $z(c) \leq |t|/\beta$.

Now consider a vector $c \sim \mathcal{P}_\epsilon$. For a fixed I, the probability that every ray in $\mathcal{E} \cap C_I(\beta)$ has cost greater than ϵ/β is at most $((1-\epsilon/\beta)/(1-\epsilon))^k$. By union bound, c satisfies property (*) with probability at least

$$1 - 2^{m-1}\left(\frac{1 - \epsilon/\beta}{1-\epsilon}\right)^k.$$

The lemma then follows by employing the bounds on $z(c)$. □

5.4.2 Probability of Obtaining a (β, k)-Good Ensemble

In this section we estimate the probability that a random ensemble in $\mathcal{D}_n^m$ is (β, k)-good. Let

$$\bar{k} = n\frac{area(\bar{C}_\emptyset(\beta))}{area(S^{m-1})} - \sqrt{\frac{n(\ln n + m - 1)}{2}}. \tag{5.6}$$

Using some Chernoff bound arguments, we can show the following lemma. The proof is deferred to the appendix.

Lemma 20. *Consider a random ensemble $\mathcal{E} \sim \mathcal{D}_n^m$ and let $\bar{k}$ be defined as in* (5.6)*. If $\bar{k} \geq 0$, then*

$$\Pr\left(\mathcal{E} \text{ is } (\beta, \bar{k})\text{-good}\right) \geq 1 - \frac{1}{n}.$$

5.4.3 Lower Bound for Simple Splits

In this section we show that $\epsilon|t|$ is also a lower bound for optimizing any vector in $[\epsilon, 1]^n$ over $G(\mathcal{E})$.

Lemma 21. *Fix $\epsilon > 0$ and consider an ensemble $\mathcal{E}$ in $\mathcal{D}_n^m$ and a vector $c \in [\epsilon, 1]^n$. For t defined as before, we have*

$$\min\{cs : s \in G(\mathcal{E})\} \geq \epsilon|t|.$$

Proof. To prove this lemma, let $S_i \equiv \sum_{j=1}^n \psi^i(r^j)s^j \geq 1$ be the inequality for $P(\mathcal{E})$ obtained from the simple split $\{x : 0 \leq x_i \leq 1\}$. Clearly S_i is valid for $G(\mathcal{E})$. Using the definition of Minkowski's functional, it is not difficult to see that

$$\psi^i(r^j) = \frac{r_i^j}{[r_i^j \geq 0] - f_i},$$

where $[r_i^j \geq 0]$ is the function that is equal to 1 if $r_i^j \geq 0$ and equal to 0 otherwise.

Now consider the inequality $\sum_{j=1}^n \psi(r^j)s_j \geq 1$ where

$$\psi(r^j) = \frac{\sum_{i=1}^m (f_i^1 - f_i)^2 \psi^i(r^j)}{\sum_{i=1}^m (f_i^1 - f_i)^2}.$$

This inequality is a non-negative combination of the inequalities S_i and therefore is valid for $G(\mathcal{E})$. We claim that for any $c \in [\epsilon, 1]^m$, $\min\{cs : \sum_{j=1}^n \psi(r^j)s_j \geq 1\} \geq \epsilon|t|$, which will give the desired lower bound on optimizing c over $G(\mathcal{E})$.

To prove the claim recall that $\sum_{i=1}^m (f_i^1 - f_i)^2 = |t|^2$ and notice that

$$\psi(r^j) = \frac{1}{|t|^2}\sum_{i=1}^m (f_i^1 - f_i)^2 \psi^i(r^j) = \frac{1}{|t|^2}\sum_{i=1}^m \frac{(f_i^1 - f_i)^2 r_i^j}{[r_i^j \geq 0] - f_i}.$$

Employing the Cauchy-Schwarz inequality and using the fact that $|r^j| = 1$, we get

$$\psi(r^j) \leq \frac{1}{|t|^2}|r^j|\sqrt{\sum_{i=1}^m \left(\frac{(f_i^1 - f_i)^2}{[r_i^j \geq 0] - f_i}\right)^2} \leq \frac{1}{|t|^2}\sqrt{\sum_{i=1}^m \frac{(f_i^1 - f_i)^4}{([r_i^j \geq 0] - f_i)^2}}.$$

However, since f^1 is the integral point closest to f, for all i it holds that $(f_i^1 - f_i)^2 \leq ([r_i^j \geq 0] - f_i)^2$. Employing this observation on the previous displayed inequality gives $\psi(r^j) \leq 1/|t|$. Therefore, any s satisfying $\sum_{j=1}^n \psi(r^j) s_j \geq 1$ also satisfies $\sum_{j=1}^n s_j \geq |t|$. The claim then follows from the fact that every coordinate of c is lower bounded by ϵ. This concludes the proof of Lemma 21. □

5.4.4 Proof of Theorem 10

Recall that ϵ, α and m are fixed. Let β be the minimum between $\sqrt{2/\alpha}$ and a positive constant strictly less than 1; this guarantees that $\bar{C}_\emptyset(\beta) > 0$. Consider a large enough positive integer n. Let $\mathcal{E}$ be a $(\beta, \bar{k})$-good ensemble in $\mathcal{D}_n^m$, where $\bar{k}$ is defined as in (5.6). Notice that $\bar{k}$, as a function of n, has asymptotic behavior $\Omega(n)$. We assume that n is large enough so that $\bar{k} > 0$.

Now let us consider Lemma 19 with $k = \bar{k}$. The value p defined in this lemma is also function of n, now with asymptotic behavior $1 - o(1)$. Thus, if n is chosen sufficiently large we get $1 - p \leq \epsilon\beta\alpha/2$ and hence $\mathbb{E}_{c\sim\mathcal{P}_\epsilon}[z(c)] \leq |t|\epsilon\alpha$. If in addition we use the lower bound from Lemma 21, we obtain that $\text{avg}(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon) \leq \alpha$. The theorem then follows from the fact that an ensemble in $\mathcal{D}_n^m$ is $(\beta, \bar{k})$-good with probability at least $1 - 1/n$, according to Lemma 20.

5.5 Implications for the Mixed-Integer Case

In this section we consider the mixed integer model obtained from (RCP) by introducing integer 'non-basic' variables. That is we consider IP's of the form

$$\begin{aligned} &\min c^1 s + c^2 y \\ &x = f + \sum_{j=1}^n r^j s_j + \sum_{j=1}^p q^j y_j \\ &x \in \mathbb{Z}^m, \quad s \geq 0, \quad y \in \mathbb{Z}_+^p \end{aligned} \tag{MG}$$

As RCP's arise as relaxations for IP's in tableaux form, the above IP also appears in the same context and offers a possibly much tighter relaxation since it only relaxes possible non-negativity constraints of basic variables. Our goal is to understand how the results from the previous section carry over to this model.

The IP (MG) is completely defined by a cost vector c and a tuple $\mathcal{E} = \langle f, r^1, \ldots, r^n, y^1, \ldots, y^q \rangle$. With some overload in the notation we also call such tuple an ensemble. Given an ensemble $\mathcal{E}$ and a cost vector c, we use $MG(\mathcal{E}, c)$ to denote the associated mixed-integer program. As before, we work on the space of the 'non-basic' variables, hence we define $\tilde{P}(\mathcal{E})$ as the projection of the feasible region of $MG(\mathcal{E}, c)$ onto the (s, y)-space.

The random model for RCP's can be extended naturally for these mixed integer programs. That is, define the distribution $\tilde{\mathcal{D}}^m_{n,p}$ over ensembles where f is picked uniformly from $[0,1]^m$ and each of the rays $r^1, \ldots, r^n, y^1, \ldots, y^q$ is picked independently and uniformly at random from the set of unit vectors in $\mathbb{R}^m$. Similarly, we define the cost distribution $\tilde{\mathcal{P}}_\epsilon$ where a vector in $[\epsilon, 1]^{n+p}$ is obtained by selecting each coefficient independently uniformly in $[\epsilon, 1]$.

Again our goal is to study the strength of simple split cuts for these random mixed integer programs. The (unstrengthened) extension of split cuts to MG's is also direct: given a split X in $\mathbb{R}^m$, its associated split cuts is $\sum_{j=1}^n \psi_X(r^j) s_j + \sum_{j=1}^p \psi_X(q^j) y_j \geq 1$, where ψ_X is defined in (5.1). It is easy to see that such inequality is valid for $\tilde{P}(\mathcal{E})$, since it is in fact valid for the set of solutions when the integrality constraints on the 'non-basic' variables are dropped. The simple split cut of $MG(\mathcal{E})$, denoted by $\tilde{G}(\mathcal{E})$, is defined as the intersection of all simple split cuts.

With definitions at hand, we extend Theorem 10 for MG's. The proof is a fairly direct modification of the proof of Theorem 10. We give it in the appendix.

Theorem 11. *Fix reals $\epsilon > 0$ and $\alpha > 1$ and positive integers m and p. Then for n sufficiently larger than m,*

$$\Pr_{\mathcal{E} \sim \tilde{\mathcal{D}}^m_{n,p}} \left(\text{avg}(\tilde{P}(\mathcal{E}), \tilde{G}(\mathcal{E}), \tilde{\mathcal{P}}_\epsilon) \leq \alpha \right) \geq 1 - \frac{1}{n}.$$

Appendix

5.6 Characterizing the Measure of Strength

Let V be a linear vector space over the reals. Let V^* denote the (algebraic) dual space of V, i.e. the set of all linear functionals on V.

Let K be a convex cone in V. We denote the *dual cone* by $K^* = \{f \in V^* \mid f(s) \geq 0 \text{ for all } x \in K\}$. For a set $C \subseteq V$, and $\gamma \geq 0$, we denote $\gamma C = \{\gamma x \mid x \in C\}$. Let P and Q be subsets of K defined as follows, where A and B are arbitrary (possibly empty) index sets:

$$P = \{s \in K \mid p_\alpha(s) \geq 1 \quad \forall \alpha \in A\}, \;\; Q = \{s \in K \mid q_\beta(s) \geq 1 \quad \forall \beta \in B\}, \tag{5.7}$$

where $p_\alpha, q_\beta \in K^*$ for all $\alpha \in A$ and $\beta \in B$.

Define

$$\begin{aligned} d_1(P,Q) &= \begin{cases} \sup\{\gamma > 0 \mid Q \subseteq \gamma P\} & \text{if } \exists \gamma > 0 \text{ such that } Q \subseteq \gamma P \\ 0 & \text{otherwise} \end{cases} \\ d_2(P,Q) &= \inf_{\alpha \in A} \inf_{s \in Q} p_\alpha(s) \\ IG(P,Q) &= \inf_{\Phi \in K^*} \frac{\inf_{s \in Q} \Phi(s)}{\inf_{s \in P} \Phi(s)} \end{aligned}$$

with the convention that $\frac{\inf_{s \in Q} \Phi(s)}{\inf_{s \in P} \Phi(s)} = +\infty$ if $\inf_{s \in P} \Phi(s) = 0$ for some $\Phi \in K^*$.

Lemma 22. $d_2(P,Q) = d_1(P,Q)$.

Proof. ($\geq$) If $d_1(P,Q) = 0$ then the inequality holds trivially since $d_2(P,Q) \geq 0$: for every $\alpha \in A$, $p_\alpha \in K^*$ is nonnegative over $Q \subseteq K$. Consider any $\gamma > 0$ such that $Q \subseteq \gamma P$, any $s \in Q$ and any $\alpha \in A$. Since $s \in Q \subseteq \gamma P$, $\frac{s}{\gamma} \in P$. Therefore, $p_\alpha(\frac{s}{\gamma}) \geq 1$ and therefore $p_\alpha(s) \geq \gamma$. Since this relation is true for any $s \in Q$ and any $\alpha \in A$, we conclude that $\inf_{\alpha \in A} \inf_{s \in Q} p_\alpha(s) \geq \gamma$, i.e. $d_2(P,Q) \geq \gamma$. This relation holds for any $\gamma > 0$ such that $Q \subseteq \gamma P$, taking a supremum all such γ, we get $d_2(P,Q) \geq \sup\{\gamma > 0 \mid Q \subseteq \gamma P\} = d_1(P,Q)$.

($\leq$) If $d_2(P,Q) = 0$, then the inequality is trivial because $d_1(P,Q) \geq 0$ by definition. If $d_2(P,Q) = +\infty$, this means that either Q is empty or A is empty or both. If Q is empty, then $Q \subseteq \gamma P$ for all $\gamma > 0$ and so $d_1(P,Q) = +\infty$. If A is empty, $P = K$ and $\gamma P = K$ for all $\gamma > 0$ and $Q \subseteq K$, and so $d_1(P,q) = +\infty$. So we consider the case $0 < d_2(P,Q) < +\infty$. By definition, we have $d_2(P,Q) \leq p_\alpha(s)$ for all $s \in Q$ and $\alpha \in A$. Therefore, $1 \leq p(\frac{s}{d_2(P,Q)})$ for all $\alpha \in A$. This implies that $\frac{s}{d_2(P,Q)} \in P$, i.e. $s \in d_2(P,Q)P$. Since this is true for all $s \in Q$, we find that $Q \subseteq d_2(P,Q)P$. Therefore, $d_1(P,Q) = \sup\{\gamma > 0 \mid Q \subseteq \gamma P\} \geq d_2(P,Q)$. □

The above result generalizes a lemma from Basu et. al [27].

Lemma 23. $IG(P,Q) = d_1(P,Q)$.

Proof. ($\geq$) If $d_1(P,Q) = 0$ then the inequality is trivial because $IG(P,Q) \geq 0$. If $\inf_{s \in P} \Phi(s) = 0$ for all $\Phi \in K^*$, then $IG(P,Q) = +\infty$ and the inequality holds trivially. Otherwise, consider any $\Phi \in K^*$ such that $\inf_{s \in P} \Phi(s) > 0$, and $\gamma > 0$ such that $Q \subseteq \gamma P$. If Q is empty, then $\inf_{s \in Q} \Phi(s) = +\infty$ and so $\frac{\inf_{s \in Q} \Phi(s)}{\inf_{s \in P} \Phi(s)} \geq \gamma$. Otherwie, consider $s \in Q$, so $\frac{s}{\gamma} \in P$. Therefore, $\Phi(s) = \gamma \Phi(\frac{s}{\gamma}) \geq \gamma \inf_{s \in P} \Phi(s)$. Taking an infinium over all $s \in Q$ on the left hand side, we get that $\inf_{s \in Q} \Phi(s) \geq \gamma \inf_{s \in P} \Phi(s)$. So again we get that $\frac{\inf_{s \in Q} \Phi(s)}{\inf_{s \in P} \Phi(s)} \geq \gamma$. Now taking an infimum over all $\Phi \in K^*$ on the left hand side, and a supremum over all $\gamma > 0$ such that $Q \subseteq \gamma P$, we obtain $IG(P,Q) \geq d_1(P,Q)$.

($\leq$) Since $d_1(P,Q) = d_2(P,Q)$ by Lemma 22, it suffices to show that $IG(P,Q) \leq d_2(P,Q)$. If A is empty, then $d_2(P,Q) = +\infty$ and the inequality holds trivially. Otherwise, we observe that for any $\alpha \in A$, $\inf_{s \in Q} p_\alpha(s) \geq \frac{\inf_{s \in Q} p_\alpha(s)}{\inf_{s \in P} p_\alpha(s)}$ since $p_\alpha(s) \geq 1$ for all $s \in P$. Since $p_\alpha \in K^*$, we conclude that $\inf_{s \in Q} p_\alpha(s) \geq \inf_{\Phi \in K^*} \frac{\inf_{s \in Q} \Phi(s)}{\inf_{s \in P} \Phi(s)} = IG(P,Q)$. We then take an infimum over $\alpha \in A$ on the left hand side and conclude that $d_2(P,Q) \geq IG(P,Q)$. □

We define $\text{wc}(P,Q) = \frac{1}{IG(P,Q)}$. Note that this definition generalizes the definition in (5.3). Hence we conclude that

Theorem 12. $d_1(P,Q) = d_2(P,Q) = IG(P,Q) = \frac{1}{\text{wc}(P,Q)}$

Let $\mathcal{B}$ be any Hamel basis for V. We can then consider the standard "coordinatization" of any point $s \in V$ as a function $s : \mathcal{B} \to \mathbb{R}$ with finite support, i.e. $s(r) \neq 0$ for finitely many $r \in \mathcal{B}$. Moreover, any linear functional $\Phi \in V^*$ can then be represented as a function $\phi : \mathcal{B} \to \mathbb{R}$.

Let P be any subset of V. For any subset $\mathcal{K} \subseteq \mathcal{B}$, we denote $V_\mathcal{K}$ as the subspace spanned by the vectors in $\mathcal{K}$. The truncation of P to $V_\mathcal{K}$ is defined as $P_\mathcal{K} = P \cap V_\mathcal{K}$.

Consider P and Q given as in (5.7). If we think of $P_{\mathcal{K}}, Q_{\mathcal{K}}$ as subsets of the linear space $V_{\mathcal{K}}$, we can also talk about $\mathrm{wc}(P_{\mathcal{K}}, Q_{\mathcal{K}})$. More precisely, let $K_{\mathcal{K}} = K \cap V_{\mathcal{K}}$. Viewing $V_{\mathcal{K}}$ as a linear vector space in its own right and $K_{\mathcal{K}}$ as a convex cone in this linear space, we consider the algebraic dual of $V_{\mathcal{K}}$ and the dual cone $K^*_{\mathcal{K}}$. We can then define $IG(P_{\mathcal{K}}, Q_{\mathcal{K}}) = \inf_{\Phi \in K^*_{\mathcal{K}}} \frac{\inf_{s \in Q_{\mathcal{K}}} \Phi(s)}{\inf_{s \in P_{\mathcal{K}}} \Phi(s)}$ and define $\mathrm{wc}(P_{\mathcal{K}}, Q_{\mathcal{K}}) = \frac{1}{IG(P_{\mathcal{K}}, Q_{\mathcal{K}})}$. We can also represent $P_{\mathcal{K}}, Q_{\mathcal{K}}$ as subsets of the linear space $V_{\mathcal{K}}$ in the form (5.7), using linear functionals from $K^*_{\mathcal{K}}$. Let P be given as $P = \{s \in K \mid p_\alpha(s) \geq 1 \quad \forall \alpha \in A\}$. With our "coordinatization", we view each p^α as a function from $\mathcal{B}$ to $\mathcal{R}$; we consider $\bar{p}_\alpha$ as the restriction of p^α to $\mathcal{K}$. Then it is easy to check that $P_{\mathcal{K}}$ as a subset of $V_{\mathcal{K}}$ is given by $P_{\mathcal{K}} = \{s \in K_{\mathcal{K}} \mid \bar{p}_\alpha(s) \geq 1 \quad \forall \alpha \in A\}$. We do the same for $Q_{\mathcal{K}}$. This enables us to define $d_2(P_{\mathcal{K}}, Q_{\mathcal{K}})$ in the linear space $V_{\mathcal{K}}$. The definition of $d_1(P_{\mathcal{K}}, Q_{\mathcal{K}})$ needs no comment; it is a purely set theoretic definition. Of course, Theorem 12 applies to $P_{\mathcal{K}}, Q_{\mathcal{K}}$ when viewed as subsets of the linear space $V_{\mathcal{K}}$, with the appropriate definition for $d_1(P_{\mathcal{K}}, Q_{\mathcal{K}})$, $d_2(P_{\mathcal{K}}, Q_{\mathcal{K}})$, $IG(P_{\mathcal{K}}, Q_{\mathcal{K}})$ and $\mathrm{wc}(P_{\mathcal{K}}, Q_{\mathcal{K}})$ and this will be utilized in proving the next lemma.

Lemma 24. *Consider two closed convex sets sets $P \subseteq V$ and $Q \subseteq V$ of the form* (5.7). *Then for any $\mathcal{K} \subseteq \mathcal{B}$,*

$$\mathrm{wc}(P, Q) \geq \mathrm{wc}(P_{\mathcal{K}}, Q_{\mathcal{K}}).$$

Proof. If $\mathrm{wc}(P, Q) = +\infty$ there is nothing to prove. So assume $\mathrm{wc}(P, Q) = \alpha < +\infty$. Lemma 12 gives that $Q \subseteq \frac{1}{\alpha} P$. We claim that this implies that $Q_k \subseteq \frac{1}{\alpha} P_k$. To see this, consider $q \in Q_{\mathcal{K}} = Q \cap V_{\mathcal{K}}$. Since $q \in Q \subseteq \frac{1}{\alpha} P$, we get that $q \in \frac{1}{\alpha} P$. Moreover, $q \in V_{\mathcal{K}} = \frac{1}{\alpha} V_{\mathcal{K}}$. So $q \in \frac{1}{\alpha} P \cap \frac{1}{\alpha} V_{\mathcal{K}} = \frac{1}{\alpha}(P \cap V_{\mathcal{K}}) = \frac{1}{\alpha} P_{\mathcal{K}}$. Therefore, $Q_{\mathcal{K}} \subseteq \frac{1}{\alpha} P_{\mathcal{K}}$. By definition, $d_1(P_{\mathcal{K}}, Q_{\mathcal{K}}) \geq \frac{1}{\alpha}$ and therefore by Theorem 12, we have that $\mathrm{wc}(P_{\mathcal{K}}, P_{\mathcal{K}}) \leq \alpha$ and the result follows. □

We now use Lemma 24 in the special case of $V = \mathbb{R}^n$ and $K = \mathbb{R}^n_+$ to derive Lemma 14. We consider the standard basis for V, $\mathcal{B} = \{e^1, \ldots, e^n\}$.

Proof of Lemma 14. In order to simplify the notation, we assume without loss of generality that $\mathcal{E}' = \langle f, r^1, r^2, \ldots, r^k \rangle$. We let $\mathcal{K} = \{e^1, \ldots, e^k\} \subseteq \mathcal{B}$.

Consider $T(\mathcal{E})$ and recall that it is defined as $\{s \in \mathbb{R}^n : \sum_{i=1}^n \psi_T(r^i) s_i \geq 1 \ \ \forall T \in \mathcal{T}\}$, where $\mathcal{T}$ is the set of all triangles in $\mathbb{R}^2$ that contain f but no integral point in their interior. Since $\psi_T \geq 0$, it follows that $T(\mathcal{E})$ is a convex set of the form 5.7. Similarly, $S(\mathcal{E})$ is a convex set of the form 5.7. Using again its definition, we have that $T(\mathcal{E}') = \{s \in \mathbb{R}^k : \sum_{i=1}^k \psi_T(r^i) s_i \geq 1 \quad \forall T \in \mathcal{T}\}$, and therefore

$T(\mathcal{E}') = T(\mathcal{E})_{\mathcal{K}}$. The same argument can be used to show that $S(\mathcal{E}') = S(\mathcal{E})_{\mathcal{K}}$. Then employing Lemma 24 with $P = T(\mathcal{E})$ and $Q = S(\mathcal{E})$, we obtain the desired result $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \mathrm{wc}(T(\mathcal{E}'), S(\mathcal{E}'))$. □

5.7 Proof of Claim 1

Recall that θ is the angle between the vectors $(0, 1-\epsilon) - \bar{f}$ and $(0,1) - \bar{f}$. So we have

$$\theta = \arctan\left(\frac{1-\bar{f}_2}{\bar{f}_1}\right) - \arctan\left(\frac{1-\epsilon-\bar{f}_2}{\bar{f}_1}\right). \tag{5.8}$$

Recall that arctan(.) is concave in $\mathbb{R}_+$. This implies that (5.8) is minimized when $\bar{f}_2$ is minimum. Since $\bar{f} \in \Delta$, $\bar{f}_2 \geq \bar{f}_1$ and hence we have

$$\theta \geq \arctan\left(\frac{1-\bar{f}_1}{\bar{f}_1}\right) - \arctan\left(\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}\right). \tag{5.9}$$

In order to simplify the previous bound we integrate arctan and notice that its derivative can be bounded as $1/(x^2+1) \geq 1/(x+\sqrt{2}-1)^2$ for all $x \in [1, \infty)$. Thus:

$$\theta \geq \arctan\left(\frac{1-\bar{f}_1}{\bar{f}_1}\right) - \arctan\left(\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}\right) = \int_{\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}}^{\frac{1-\bar{f}_1}{\bar{f}_1}} \frac{1}{x^2+1} \geq \int_{\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}}^{\frac{1-\bar{f}_1}{\bar{f}_1}} \frac{1}{(x+\sqrt{2}-1)^2}$$
$$= -\frac{1}{x+\sqrt{2}-1}\Bigg|_{\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}}^{\frac{1-\bar{f}_1}{\bar{f}_1}} = \left(\frac{\bar{f}_1}{1-\epsilon-(2-\sqrt{2})\bar{f}_1} - \frac{\bar{f}_1}{1-(2-\sqrt{2})\bar{f}_1}\right) = g(\bar{f}_1).$$

5.8 Proof of Lemma 15

The proof follows the arguments used in Section 5.6.2 in [25]. A key step is a method for constructing a polyhedron contained in the split closure. We minimize the function $c_1 s_1 + c_2 s_2$ over this strengthening of the split closure. The resulting LP implies an upper bound on the objective value when minimizing the function over the split closure.

To obtain this polyhedron, we define some inequalities which dominate the split closure $S(\mathcal{E})$. A *pseudo-split* is the convex set between two distinct parallel lines passing through $(0,0)$ and $(0,1)$ respectively. The direction of the lines, called *direction* of the pseudo-split, is a parameter.The *pseudo-split inequality* is derived from a pseudo-split exactly in the same way as from any maximal lattice-free convex set using formula (5.1). Note that pseudo-splits are in general not lattice-free and hence do not generate valid inequalities for $RCP(\mathcal{E}, c)$. However, we can dominate any split inequality cutting f by an inequality derived from these convex sets. Indeed, consider any split S containing the fractional point f in its interior and passing through the segment joining $(0,0)$ and $(0,1)$. The pseudo-split with direction identical to the direction of S generates an inequality that dominates the split inequality derived from S, as the coefficient for any ray is smaller in the pseudo-split inequality. The condition imposed on the rays to cross the left facet of the unit square implies the following. Any split which contains f, but does not pass through the segment $(0,0),(0,1)$, is dominated by any pseudo-split passing through the segment joining $(0,0)$ and $(0,1)$. So to dominate the split closure in this case, we only need to consider the inequalities derived from the pseudo-splits.

The next lemma states that we can dominate the split closure by using only the inequalities generated by the pseudo-splits with direction parallel to the rays r^1, r^2.

Lemma 25. *Consider an ensemble $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f = (f_1, f_2) \in (0,1)^2$, $r^1 = c_1(-1, t_1)$ and $r^2 = c_2(-1, t_2)$ with $c_1, c_2 \geq 0$ and $t_1 \geq t_2$, such that both $f + r^1$ and $f + r^2$ cross the segment joining $(0,0)$ and $(0,1)$. Then any pseudo-split inequality is dominated by the convex combination of the two pseudo-splits parallel to r^1, r^2.*

Proof. Let the pseudo-split parallel to r^1 be denoted by S_1 and similarly the pseudo-split parallel to r^2 be S_2. Consider any other pseudo split S'. Consider the point $\bar{f}$ on the segment joining $(0,0)$ and $(0,1)$ be such that the segment joining f and $\bar{f}$ is parallel to the direction of S'. Let $\bar{\mathcal{E}}$ be the ensemble $\langle \bar{f}, r^1, r^2 \rangle$. We compare the inequalities generated by the convex set S' using the formula (5.1) for $P(\mathcal{E})$ and $P(\bar{\mathcal{E}})$. Let $\psi_X(r^i)$ be the coefficeint for r^i in $P(\mathcal{E})$ and $\bar{\psi}_X(r^i)$ be the coefficient for r^i in $P(\bar{\mathcal{E}})$ with respect to the convex set X.

Observation 2. *$\psi_{S'}(r^i) = \bar{\psi}_{S'}(r^i)$ for $i = 1, 2$ since the distance cut by S' on the rays r^1, r^2 does not change in the two ensembles.*

Observation 3. *$\psi_{S_1}(r^i) \geq \bar{\psi}_{s_1}(r^i)$. This is because the coefficient for r^1 remains 0 and the distance cut by S_1 on r^2 is more in ensemble $\mathcal{E}$ as compared to in ensemble $\bar{\mathcal{E}}$. By a similar argument, $\psi_{S_2}(r^i) \geq \bar{\psi}_{s_2}(r^i)$: the coefficient for r^2 remains 0 and the distance cut off on r^1 is more in $\mathcal{E}$ compared to $\bar{\mathcal{E}}$.*

We now make the following claim.

Claim 3. *There exists $0 \leq \lambda \leq 1$ such that $\bar{\psi}_{S'}(r^i) = \lambda \bar{\psi}_{S^1}(r^i) + (1-\lambda)\bar{\psi}_{S^2}(r^i)$ for $i = 1, 2$.*

Proof. We first note that $\bar{\psi}_{S_i}(r^i) = 0$ for $i = 1, 2$. This implies it suffices to show that $\frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)} + \frac{\bar{\psi}_{S'}(r^1)}{\bar{\psi}_{S^2}(r^1)} = 1$. Indeed, we can then pick $\lambda = \frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)}$. We use similarity of triangles to establish that $\frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)} + \frac{\bar{\psi}_{S'}(r^1)}{\bar{\psi}_{S^2}(r^1)} = 1$. Refer to Figure 5.2 for the following notation. In the figure, ray r^2 is extended back to intersect S_1 at D and S' at E. Note that $\frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)}$ is equal to Ff/Gf. By similarity of triangles, $Ff/Gf = Df/Ef = AC/AE$. Also, $\frac{\bar{\psi}_{S'}(r^1)}{\bar{\psi}_{S^2}(r^1)}$ is equal to $Bf/Af = CD/Af$ and by similarity of triangles, $CD/Af = CE/AE$. Since, $AC/AE + CE/AE = 1$, we have our identity. □

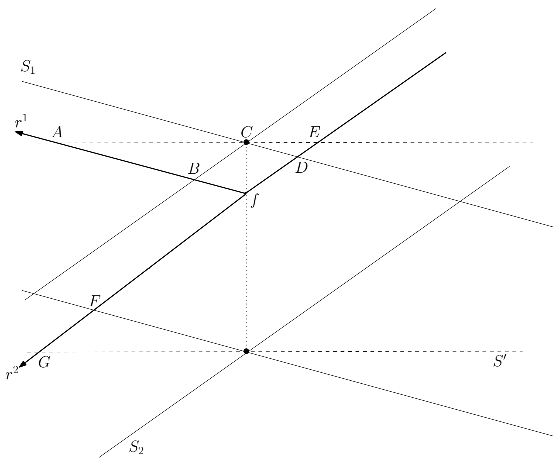

Figure 5.2: Figure for proof of Claim 3

Now combining Claim 3 and Observations 2 and 3, we obtain that the inequality $\psi_{S'}(r^1)s_1 + \psi_{S'}(r^2) \leq 1$ is dominated by the convex combination of the two inequalities $\psi_{S_1}(r^1)s_1 + \psi_{S_1}(r^2) \leq 1$ and $\psi_{S_2}(r^1)s_1 + \psi_{S_2}(r^2) \leq 1$ defined by λ from Claim 3.

□

We have thus shown that we need to consider the following LP to bound $\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E})\}$ from above.

$$\begin{array}{lll} \min & c_1 s_1 + c_2 s_2 & \\ & \psi_{S_1}(r^1)s_1 + \psi_{S_1}(r^2)s_2 & \geq 1 \\ & \psi_{S_2}(r^1)s_1 + \psi_{S_2}(r^2)s_2 & \geq 1 \\ & s \in \mathbb{R}^2_+. & \end{array} \tag{5.10}$$

We can derive the constraints corresponding to S_1, S_2. We have then to find an upper bound on the value of the following LP.

$$\begin{aligned} \min \quad & s_1 + s_2 \\ & 0 \cdot s_1 + \frac{c_2(t_1 - t_2)}{f_2 + f_1 t_1} s_2 \geq 1 \\ & \frac{c_1(t_1 - t_2)}{1 - f_2 - f_1 t_2} s_1 + 0 \cdot s_2 \geq 1 \\ & s \in \mathbb{R}^2_+ \,. \end{aligned} \tag{5.11}$$

The upper bound can be obtained by exhibiting a feasible solution :

$$s_1 = \frac{1 - f_2 - f_1 t_2}{c_1(t_1 - t_2)} \quad \text{and} \quad s_2 = \frac{f_2 + f_1 t_1}{c_2(t_1 - t_2)} \,.$$

The value of this feasible solution is

$$c_1 s_1 + c_2 s_2 = \frac{1 + f_1(t_1 - t_2)}{t_1 - t_2}.$$

Finally notice that $f + f_1 r^1 = (0, f_2 + f_1 t_1)$ and using the crossing property we get that $f_2 + f_1 t_1 \leq 1$. Similarly, $f + f_1 r^2 = (0, f_2 + f_1 t_2)$, hence $f_2 + f_1 t_2 \geq 0$. Isolating f_2 in both inequalities and chaining them we obtain $f_1(t_1 - t_2) \leq 1$. This concludes the proof of Lemma 15.

5.9 Proof of Claim 2 in Lemma 18

First we need a preliminary lemma.

Lemma 26. *Let $R' \subseteq \bar{C}(\beta)$ be such that $R' \cap \bar{C}_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$. Then $e^m \in cone(R')$.*

Proof. Consider a vector $a \in \mathbb{R}^m$ such that $ar \geq 0$ for all $r \in R'$; we claim that $a_m \geq 0$. To see this, consider the set of indices $I = \{i \in [m-1] : a_i < 0\}$. Making use of our hypothesis, there is $r' \in R' \cap \bar{C}_I(\beta)$, which then satisfies $\sum_{i \in I} a_i r'_i + \sum_{i \in [m-1] \setminus I} a_i r'_i \leq 0$. Since $ar' \geq 0$, this implies that $a_m r'_m \geq 0$. Finally, since $r' e^m \geq \beta > 0$, we obtain that $r'_m > 0$ and hence $a_m \geq 0$.

From Farkas' Lemma $e^m \in cone(R')$ iff there is no vector with $a_m r \geq 0$ for all $r \in R'$ and $a_m < 0$, so the result follows from the previous claim. ☐

In order to prove Claim 2 we can proceed as follows. Letting $R' = \rho R$, the definition of R and the fact that $\bar{C}_I(\beta) = \rho C_I(\beta)$ implies that $R' \cap \bar{C}_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$. Then Lemma 26 implies that $e^m \in cone(R')$. Since ρ^{-1} is a linear transformation, we have $t = \rho^{-1} e^m \in \rho^{-1}(cone(R')) = cone(R)$.

5.10 Proof of Theorem 11

As in the proof of Theorem 10, we need an upper bound on $z(c) = \min\{cs : s \in \tilde{P}(\mathcal{E})\}$ and a lower bound on $\min\{cs : s \in \tilde{G}(\mathcal{E})\}$.

For the upper bound, we consider solutions of (MG) with $y = 0$. We are now back to a problem of the form (RCP). We say that a tuple $\mathcal{E} = \langle f, r^1, \ldots, r^n, y^1, \ldots, y^q \rangle$ is (β, k)-good if the ensemble $\langle f, r^1, \ldots, r^n \rangle$ is (β, k)-good as defined in Section 5.4.1. We can apply Lemmas 18-20 to (β, k)-good tuples $\mathcal{E}$.

For the lower bound, we relax the integrality constraint on the y variables of (MG). We are now back to a problem of the form (RCP) with $n + p$ continuous variables. Applying Lemma 21, we get

Lemma 27. *Fix $\epsilon > 0$ and consider an ensemble $\mathcal{E}$ in $\tilde{\mathcal{D}}^m_{n,p}$ and a vector $(c^1, c^2) \in [\epsilon, 1]^{n+p}$. For $t = f^1 - f$, we have*

$$\min\{c^1 s + c^2 y : (s, y) \in \tilde{G}(\mathcal{E})\} \geq \epsilon |t|.$$

The proof of Theorem 11 now follows the proof of Theorem 10:

Let β be the minimum between $\sqrt{2/\alpha}$ and a positive constant strictly less than 1; this guarantees that $\bar{C}_\emptyset(\beta) > 0$. Consider a large enough positive integer n. Let $\mathcal{E}$ be a $(\beta, \bar{k})$-good tuple in $\tilde{\mathcal{D}}^m_{n+p}$, where $\bar{k}$ is defined as in (5.6). Notice that $\bar{k}$, as a function of n, has asymptotic behavior $\Omega(n)$. We assume that n is large enough so that $\bar{k} > 0$.

Now let us consider Lemma 19 with $k = \bar{k}$. The value p defined in this lemma is also function of n, now with asymptotic behavior $1 - o(1)$. Thus, if n is chosen sufficiently large we get $1 - p \leq \epsilon\beta\alpha/2$ and hence $\mathbb{E}_{c \sim \tilde{\mathcal{P}}_\epsilon}[z(c)] \leq |t|\epsilon\alpha$. If in addition we use the lower bound from Lemma 27, we obtain that $\operatorname{avg}(\tilde{P}(\mathcal{E}), \tilde{G}(\mathcal{E}), \tilde{\mathcal{P}}_\epsilon) \leq \alpha$. The theorem then follows from the fact that a tuple in $\tilde{\mathcal{D}}^m_{n+p}$ is $(\beta, \bar{k})$-good with probability at least $1 - 1/n$, according to Lemma 20.

5.11 Proof of Lemma 20

We assume that β, n and m are such that $\bar{k} \geq 0$.

Consider a random ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$ from $\mathcal{D}_n^m$ and let R denote the set of rays of $\mathcal{E}$. We have that

$$\Pr\left(\mathcal{E} \text{ is } (\beta, \bar{k})\text{-good }\right) = \Pr\left(\bigwedge_{I \subseteq [m-1]} |R \cap C_I(\beta)| \geq \bar{k}\right) \geq 1 - 2^{m-1} \Pr\left(|R \cap \bar{C}_\emptyset(\beta)| < \bar{k}\right), \tag{5.12}$$

where the last inequality follows from the union bound and the fact that, by symmetry, $\Pr\left(|R \cap C_I(\beta)| < \bar{k}\right)$ is the same as $\Pr\left(|R \cap \bar{C}_\emptyset(\beta)| < \bar{k}\right)$ for every $I \subseteq [m-1]$.

Due to the independence of the rays, $|R \cap \bar{C}_\emptyset(\beta)|$ behaves as a sum of n 0/1 random variables which take value 1 with probability $area(\bar{C}_\emptyset(\beta))/area(S^{m-1})$. At this point we recall the additive Chernoff bound on the tail of such distributions.

Theorem 13 (Theorem 1.1 of [73])**.** *Let $X = \sum_{i=1}^n X_i$, where X_i are random variables independently distributed in* $[0, 1]$*. Then for all* $t > 0$

$$\Pr(X < \mathbb{E}[X] - t) \leq e^{-2t^2/n}.$$

By linearity of expectation we obtain that $\mathbb{E}[|R \cap \bar{C}_\emptyset(\beta)|] = n(area(\bar{C}_\emptyset(\beta))/area(S^{m-1}))$, hence employing the previous bound with $t = \sqrt{n(\ln n + m - 1)/2}$ we obtain that

$$\Pr\left(|R \cap \bar{C}_\emptyset(\beta)| < \bar{k}\right) \leq \frac{1}{ne^{m-1}}.$$

This upper bound together with inequality (5.12) gives Lemma 20.

CHAPTER 6

$(k+1)$-Slope Theorem for the k-Dimensional Infinite Relaxation

In this chapter we consider the infinite relaxation introduced in Section 2.3.3. Recall that the latter is a further relaxation of the important *corner relaxation*, having the additional property that we obtain a *single* (infinite-dimensional) integer program that can be used to generate important cuts for essentially any finite-dimensional integer program. Also recall that there is a hierarchy of cuts for the infinite relaxation relative to their strength, with valid cuts/functions as the weakest ones, passing through minimal functions, extreme functions and reaching facets, the strongest ones in the hierarchy.

In this chapter we focus on understanding the facets of the infinite relaxation. We present sufficient conditions for a function to be a facet; these conditions generalize the 2-Slope Theorem of Gomory and Johnson [87, 88, 89] and the 3-Slope Theorem of Cornuéjols and Molinaro [53]. More precisely, we show that any minimal valid function for the k-dimensional infinite relaxation that is continuous, piecewise linear, with at most $k+1$ slopes and does not factor through a linear map with non-trivial kernel, is a facet.

Organization of the chapter. In the next section, we start by recalling some definitions regarding the infinite relaxation and stating formally our main results;

this section also presents a high-level idea of the proof strategy. Section 6.2 provides some preliminary technical results, while Section 6.3 delves into the proof of our main result. The chapter closes with Section 6.4 providing short concluding discussions.

Acknowledgments. This chapter is joint work with Amitabh Basu, Robert Hildebrand and Matthias Köppe.

6.1 Introduction

6.1.1 Basic Definitions

Recall from Section 2.3.3 the *infinite relaxation* (here we consider the projection of the infinite relaxation to the y-space, as in the previous chapter):

$$\begin{aligned} f + \sum_{r \in \mathbb{R}^k} r \cdot s_r \in \mathbb{Z}^k \\ s_r \in \mathbb{Z}_+ \ \text{ for all } r \in \mathbb{R}^k \\ s \text{ has finite support.} \end{aligned} \tag{IR}$$

Valid functions, Minimal functions, extreme functions and facets. Also recall the following concepts. A function $\pi\colon \mathbb{R}^k \to \mathbb{R}$ is *valid* for (IR) if $\pi \geq 0$ and the inequality

$$\sum_{r \in \mathbb{R}^k} \pi(r) s_r \geq 1 \tag{6.1}$$

is satisfied by every feasible solution s of (IR). A valid function π is said to be *minimal* if there is no valid function $\pi' \neq \pi$ such that $\pi'(r) \leq \pi(r)$ for all $r \in \mathbb{R}^k$.

It is intuitively clear that for every valid function there is a minimal one which dominates it. However, we could not find a proof of this statement in the literature, and so we present a proof using Zorn's Lemma in Section 6.5.

Theorem 3. *Let π be a valid function. Then there exists a minimal valid function π' such that $\pi' \leq \pi$.*

A function $\pi\colon \mathbb{R}^k \to \mathbb{R}$ is *periodic with respect to the lattice $\mathbb{Z}^k$* if $\pi(r) = \pi(r + w)$ holds for all $r \in \mathbb{R}^k$ and $w \in \mathbb{Z}^k$. We say that π satisfies the *symmetry condition* if

$\pi(r)+\pi(-f-r) = 1$ for all $r \in \mathbb{R}^k$. Finally, π is *subadditive* if $\pi(a+b) \leq \pi(a)+\pi(b)$ for all $a, b \in \mathbb{R}^k$.

Theorem 4 (Gomory and Johnson [87])**.** *Let $\pi\colon \mathbb{R}^k \to \mathbb{R}$ be a non-negative function. Then π is a minimal valid function for* (IR) *if and only if $\pi(0) = 0$, π is periodic with respect to $\mathbb{Z}^k$, subadditive and satisfies the symmetry condition.*

Although minimality reduces the number of relevant valid functions that we need to study, it still leaves too many under consideration. Inspired by the importance of facets in the finite-dimensional setting, Gomory and Johnson introduce the analogous concepts in this setting [87, 88, 89]. A valid function π is *extreme* if it cannot be written as a convex combination of two other valid functions, i.e., $\pi = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$ implies $\pi = \pi_1 = \pi_2$. It is easy to verify that extreme functions are minimal [87]. For any valid function π, let $S(\pi)$ denote the set of all feasible solutions s satisfying (IR) such that $\sum_{r\in\mathbb{R}^k} \pi(r)s_r = 1$. A valid function π is a *facet* if for every valid function π', the condition holds that $S(\pi) \subseteq S(\pi')$ implies $\pi' = \pi$. This concept was introduced by Gomory and Johnson in [89] and we prove below that if π is a facet, then it is extreme. Thus, facets can be seen as the strongest valid functions.

Lemma 28. *If π is a facet, then π is extreme.*

Proof. Suppose π is a facet and let $\pi = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$. We observe that $S(\pi) \subseteq S(\pi_1)$ and $S(\pi) \subseteq S(\pi_2)$. Let $s \in S(\pi)$. Then

$$1 = \sum_{r\in\mathbb{R}^k} \pi(r)s_r = \frac{1}{2}\sum_{r\in\mathbb{R}^k} \pi_1(r)s_r + \frac{1}{2}\sum_{r\in\mathbb{R}^k} \pi_2(r)s_r \geq \frac{1}{2} + \frac{1}{2} = 1,$$

so equality must hold throughout and in particular, $\sum_{r\in\mathbb{R}^k} \pi_i(r)s_r = 1$ for both $i = 1, 2$. Therefore $s \in S(\pi_i)$ for both $i = 1, 2$. Since π is a facet, by definition this implies $\pi = \pi_1 = \pi_2$. □

In general, constructing or even proving that a valid function is a facet or extreme can be a very difficult task. Arguably the deepest result on the infinite relaxation is a sufficient condition for facetness in the restricted setting $k = 1$, the so-called 2-Slope Theorem of Gomory and Johnson [88, 89].

Theorem 5 (Gomory–Johnson 2-Slope Theorem)**.** *Let $\pi\colon \mathbb{R} \to \mathbb{R}$ be a minimal valid function. If π is a continuous piecewise linear function with only two slopes, then π is a facet (and hence extreme).*

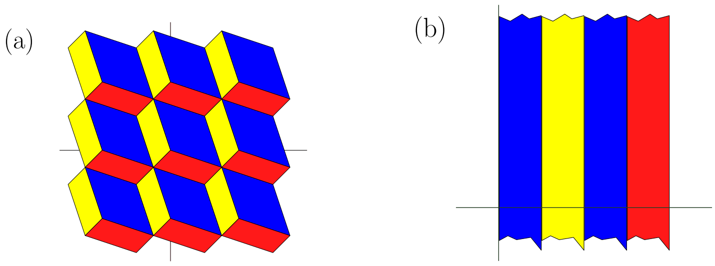

Figure 6.1: The assumptions of the Cornuéjols–Molinaro 3-slope theorem (Theorem 6). (a) A function with 3 slopes and 3 directions, where each color represents the cells where the function has a certain slope. (b) A function with 3 slopes and only 1 direction.

In addition to its theoretical appeal, this result also has practical relevance. It supplies theoretical indication about the intrinsic power of 2-slope functions, which are very effective cuts in integer programming solvers [38] (*e.g.*, GMI's). This surprising result was already known in the 1970s, and despite the increased efforts in understanding the case $k > 1$, a generalization of this result was obtained only recently for the case $k = 2$ by Cornuéjols and Molinaro [53].

Theorem 6 (3-Slope Theorem). *Let* $\pi\colon \mathbb{R}^2 \to \mathbb{R}$ *be a minimal valid function. If* π *is a continuous piecewise linear function with only 3 slopes and with 3 directions, then* π *is extreme.*

Here, the *directions* of the function refer to the direction of the edges bounding the cells of the piecewise linear function. Assuming 3 directions ensures that we do not include 3-slope functions that are constant in some direction; see Figure 6.1 for an illustration. The authors show in [53] that there exist functions with 3 slopes that are not extreme. This suggests that for $k \geq 2$, we need some additional hypothesis, over and above a $(k+1)$-slope assumption, to imply extremality.

Theorems 5 and 6 contribute a simple sufficient condition for extremality that capture *why* specific families of functions are extreme (for instance the 3-slope functions in Section 7 of [69]). Our goal in this chapter is to prove such a theorem for general k.

6.1.2 Our Results

We generalize the above results and present a sufficient condition for facetness (and therefore for extremality) of valid functions for arbitrary dimension k. For this, we generalize the notion of a function having 3 directions.

Definition 3. *A function $\theta\colon \mathbb{R}^k \to \mathbb{R}$ is* genuinely k-dimensional *if there does not exist a function $\phi\colon \mathbb{R}^{k-1} \to \mathbb{R}$ and a linear map $T\colon \mathbb{R}^k \to \mathbb{R}^{k-1}$ such that $\theta = \phi \circ T$.*

In the context of piecewise linear functions $\pi\colon \mathbb{R}^2 \to \mathbb{R}$, the above definition is related to the concept of functions with 3 directions as used in [53]. Indeed, Section 6.7 shows that the assumption of being genuinely 2-dimensional is a weaker assumption; namely, if a function has 3 directions, then it is genuinely 2-dimensional.

Theorem 7. *Let $\pi\colon \mathbb{R}^k \to \mathbb{R}$ be a minimal valid function that is piecewise linear with a locally finite cell complex and genuinely k-dimensional with at most $k+1$ slopes. Then π is a facet (and therefore extreme) and has exactly $k+1$ slopes.*

This settles an open question posed by Gomory and Johnson in [89]. We comment here that the hypothesis of being piecewise linear with a locally finite cell complex will be made precise in Section 6.2 and the definition implies that such functions are continuous. Thus, in this chapter we only consider continuous piecewise linear functions, just like the hypotheses of Theorems 5 and 6.

One direct application of Theorem 7 is to investigate the facetness of certain valid functions studied in [33]. There is a useful procedure known as the *trivial lifting procedure* which can be used to derive minimal valid functions for (IR) using the Minkowski functionals of maximal lattice-free convex sets. This procedure was studied in [33] as applied to maximal lattice-free simplices. It turns out that for a special class of such simplices, the minimal valid functions obtained will have $k+1$ slopes and Theorem 7 can be directly applied to prove that they are facets (see Section 6.8 for more details).

6.1.3 Proof Structure

The high-level structure of the proof of Theorem 7 is similar to the proof of the 2-Slope and 3-Slope Theorems presented in [89] and [53]. Let $\pi\colon \mathbb{R}^k \to \mathbb{R}$ be a valid function satisfying the assumptions of the theorem. We consider an arbitrary valid function π' such that $S(\pi) \subseteq S(\pi')$, and our goal will be to show that $\pi = \pi'$. In order to achieve this, we first prove in Section 6.3.1 that if π is affine in a region of $\mathbb{R}^k$, then π' is also affine in this region, albeit with a possibly different gradient.

The next step, Section 6.3.2, is to write a system of equations which is satisfied by the gradients of π as well as the gradients of π'. This is the most involved step in the proof. Some properties proved in [53] rely on arguments about low dimensional objects; whereas, the appropriate high dimensional generalizations require more sophisticated techniques. In particular, we make use of a topological lemma about closed coverings of the simplex, the Knaster–Kuratowski–Mazurkiewicz Lemma (Lemma 42), which in the one-dimensional case reduces to the easy fact that an interval cannot be covered by two disjoint closed sets.

The final step, Section 6.3.3, is then to prove that this system has a unique solution, which implies that the gradients of π and π' are the same. This, together with the fact that $\pi(0) = \pi'(0) = 0$, implies that $\pi = \pi'$. The proof of this last step simplifies the one presented in [53] and directly exposes the properties driving the uniqueness of the system.

6.2 Preliminaries

6.2.1 Basic Polyhedral Theory

In this section we recall some basic definitions from polyhedral theory (see [115]) as well as two simple lemmas that will be used throughout the text. The open ball of radius ϵ around a point r will be denoted by $B_\epsilon(r)$.

Definition 4. *A* polyhedral complex *in $\mathbb{R}^k$ is a collection $\mathcal{P}$ of polyhedra in $\mathbb{R}^k$ such that:*

(i) *if $P \in \mathcal{P}$, then all faces of P are in $\mathcal{P}$,*

(ii) *the intersection $P \cap Q$ of two polyhedra $P, Q \in \mathcal{P}$ is a face of both P and Q.*

Any polyhedron in a polyhedral complex $\mathcal{P}$ is called a *cell* of the complex. A cell P is *maximal* if there is no $Q \in \mathcal{P}$ containing it. The complex is *pure* if all maximal cells have the same dimension. The complex is said to be *complete* if the union of all elements in the complex is $\mathbb{R}^k$. A *subcomplex* of $\mathcal{P}$ is a subset $\mathcal{P}' \subseteq \mathcal{P}$ such that $\mathcal{P}'$ is itself a polyhedral complex.

A *polyhedral fan* is a polyhedral complex of a finite number of cones (not necessarily pointed). Given a polyhedral fan $\mathcal{F}$ and $r \in \mathbb{R}^k$, we use the notation $\mathcal{F} + r = \{ C + r \mid C \in \mathcal{F} \}$, which is a finite polyhedral complex. Given a cone $C \subseteq \mathbb{R}^k$, a *triangulation* of C is a polyhedral fan $\mathcal{P}$ such that each element of $\mathcal{P}$ is a simplicial cone and the union of all elements in $\mathcal{P}$ is C. Given a polyhedral fan $\mathcal{P}$, a *triangulation* of $\mathcal{P}$ is a polyhedral fan $\mathcal{F}$ such that for every element $P \in \mathcal{P}$, there exists a triangulation of P as a subcomplex of $\mathcal{F}$ and every element of $\mathcal{F}$ is simplicial.

Given a polyhedral complex $\mathcal{P}$ and a set $X \subseteq \mathbb{R}^k$, we use the notation $\mathcal{P} \cap X$ to denote the collection of sets $\{ P \cap X \mid P \in \mathcal{P} \text{ and } P \cap X \neq \emptyset \}$. Observe that if $\mathcal{P}$ is complete, the union of all sets in $\mathcal{P} \cap X$ is X.

With a slight abuse of notation, for any point $v \in \mathbb{R}^k$ and a polyhedral complex $\mathcal{P}$, we will use $v \in \mathcal{P}$ to denote that $v \in P$ for some element $P \in \mathcal{P}$.

Definition 5. *A polyhedral complex $\mathcal{P}$ is called* locally finite *if for every point $r \in \mathbb{R}^k$ there exists an open ball $B_\epsilon(r)$ around r, such that $\mathcal{P} \cap B_\epsilon(r)$ equals $(\mathcal{F}_r + r) \cap B_\epsilon(r)$ for some polyhedral fan $\mathcal{F}_r$ (recall that a polyhedral fan is finite by definition).*

We remind the reader that a polyhedral fan may contain cones that are not pointed. Notice that the above definition is equivalent to stating that each point in $\mathbb{R}^k$ has a neighborhood which intersects only finitely many elements of $\mathcal{P}$. In addition, using standard arguments, it is easy to see that this finite intersection extends from points to compact sets.

Proposition 3. *Let $\mathcal{P}$ be a locally finite polyhedral complex. Then for every compact set $K \subseteq \mathbb{R}^k$, only finitely many elements of $\mathcal{P}$ intersect K.*

Now we present two simple linear algebraic facts that will be useful for the next sections.

Lemma 29. *If $r^1, \ldots, r^{k+1} \in \mathbb{R}^k$ are such that* $\text{cone}(r^i)_{i=1}^{k+1} = \mathbb{R}^k$, *then every proper subset of $\{r^1, \ldots, r^{k+1}\}$ is composed of linearly independent vectors.*

Proof. It suffices to show that every k-subset of $\{r^1, \ldots, r^{k+1}\}$ is linearly independent. Without loss of generality, we will just show that $r^1, \ldots, r^k$ are linearly independent. If not, then there exists a hyperplane H containing the linear span of $r^1, \ldots, r^k$. Suppose H_+ is the half-space defined by H containing r^{k+1}. This implies that $\text{cone}(r^i)_{i=1}^{k+1} \subseteq H_+$, contradicting the fact that $\text{cone}(r^i)_{i=1}^{k+1} = \mathbb{R}^k$. $\square$

Lemma 30. *Let* $\{a^1, \dots, a^{k+1}\}$ *and* $\{b^1, \dots, b^{k+1}\}$ *be two sets of* $k+1$ *vectors in* $\mathbb{R}^k$*. Suppose* $\operatorname{cone}(a^i)_{i=1}^{k+1} = \mathbb{R}^k$*, and* $a^i \cdot b^j < 0$ *for* $i \neq j$*. Then* $\operatorname{cone}(b^j)_{j=1}^{k+1} = \mathbb{R}^k$*.*

Proof. We show that the cone $X = \operatorname{cone}(b^j)_{j=1}^{k+1} = \mathbb{R}^k$ by considering the polar cone

$$X^\circ = \{\, r \in \mathbb{R}^k \mid b^j \cdot r \leq 0, \;\; j = 1, \dots, k+1 \,\}$$

and equivalently showing that $X^\circ = \{0\}$. Consider any vector $r^0 \neq 0$. Since $\operatorname{cone}(a^i)_{i=1}^{k+1} = \mathbb{R}^k$, by Carathéodory's theorem, there exists $j \in \{1, \dots, k+1\}$ such that $-r^0 = \sum_{i \neq j} \lambda_i a^i$ with $\lambda_i \geq 0$. Since $b^j \cdot a^i < 0$ for all $i \neq j$, we have that $b^j \cdot (-r^0) < 0$, or equivalently, $b^j \cdot r^0 > 0$. Thus $r^0 \notin X^\circ$. □

6.2.2 Piecewise Linear Functions

We now give a precise definition of piecewise linear functions and related notions.

Definition 6. *Let* $\mathcal{P}$ *be a pure, complete polyhedral complex in* $\mathbb{R}^k$ *and let* $\{\mathcal{P}_i\}_{i \in I}$ *be a partition of the set of maximal cells of* $\mathcal{P}$*. Consider a function* $\theta\colon \mathbb{R}^k \to \mathbb{R}$ *such that for each* $i \in I$*, there exists a vector* $g^i \in \mathbb{R}^k$ *such that for every* $P \in \mathcal{P}_i$*, there exists a constant* δ_P *such that* $\theta(r) = g^i \cdot r + \delta_P$ *for all* $r \in P$*. Then* θ *is called a* piecewise linear function*, more specifically a* piecewise linear function with cell complex $\mathcal{P}$*, and a* piecewise linear function compatible with $\{\mathcal{P}_i\}_{i \in I}$*.*

Given a piecewise linear function θ with cell complex $\mathcal{P}$, for any maximal cell $P \in \mathcal{P}$ let $g^P \in \mathbb{R}^k$ denote the vector such that $\theta(r) = g^P \cdot r + \delta_P$ for some constant δ_P. We define the equivalence relation $\sim$ on the maximal elements of $\mathcal{P}$ according to their gradients as $P \sim P'$ if and only if $g^P = g^{P'}$. Each equivalence class defines a subcomplex $\mathcal{P}_i \subseteq \mathcal{P}$, $i \in I$ for some index set I. We say that θ has *n slopes* if $|I| = n$. The *gradient set* of a piecewise linear function is the set of vectors $\{g^i\}_{i \in I}$ corresponding to the equivalence classes $\mathcal{P}_i$, namely $g^i = g^P$ for all $P \in \mathcal{P}_i$.

6.2.3 Lipschitz Continuity

The following two lemmas assert strong continuity properties of piecewise linear, subadditive functions with a locally finite cell complex. We will use $\|\cdot\|$ to denote the Euclidean norm.

Lemma 31. *Let $\theta\colon \mathbb{R}^k \to \mathbb{R}$ be a piecewise linear function with a locally finite cell complex. Moreover, suppose $\theta(0) = 0$. Then θ is locally Lipschitz continuous at the origin, i.e., there exist $\epsilon > 0$ and $K > 0$ such that $|\theta(r)| \leq K\|r\|$ for all $r \in B_\epsilon(0)$.*

Proof. Let θ be a piecewise linear function with a locally finite cell complex $\mathcal{P}$. Since $\mathcal{P}$ is locally finite, there exists an open ball $B_\epsilon(0)$ around the origin such that $\mathcal{P} \cap B_\epsilon(0) = \mathcal{F} \cap B_\epsilon(0)$ for some complete polyhedral fan $\mathcal{F}$. This implies that there exists a *finite* subcomplex $\mathcal{P}' \subseteq \mathcal{P}$ such that $\mathcal{P} \cap B_\epsilon(0) = \mathcal{P}' \cap B_\epsilon(0)$ and every maximal element of $\mathcal{P}'$ contains the origin. Therefore, the union of all P in $\mathcal{P}'$ contains $B_\epsilon(0)$.

Since θ is piecewise linear and $\theta(0) = 0$, for every maximal element $P \in \mathcal{P}'$, there exists $g^P \in \mathbb{R}^k$ such that $\theta(r) = g^P \cdot r$ for all $r \in P$ and so $|\theta(r)| \leq \|g^P\|\|r\|$ by the Cauchy–Schwarz inequality. Let $K = \max\{\, \|g^P\| \mid P \in \mathcal{P}' \,\}$. Then $|\theta(r)| \leq K\|r\|$ for all $r \in B_\epsilon(0)$. □

The next lemma relates global continuity properties of a subadditive function with its behavior around the origin (a similar result is presented in Theorem 7.8.2 of [94]).

Lemma 32. *Let $\theta\colon \mathbb{R}^k \to \mathbb{R}$ be a subadditive function that is locally Lipschitz continuous at the origin. Then θ is (globally) Lipschitz continuous.*

Proof. We first show that there exist $K > 0$ and $\epsilon > 0$, such that given any $\tilde{r} \in \mathbb{R}^n$, $|\theta(r) - \theta(\tilde{r})| \leq K\|r - \tilde{r}\|$ for all $r \in \mathbb{R}^n$ satisfying $\|r - \tilde{r}\| < \epsilon$. Indeed, since θ is locally Lipschitz continuous at the origin, it follows that there exists $K > 0$, $\epsilon > 0$ such that, for all $r \in \mathbb{R}^n$ satisfying $\|r - \tilde{r}\| < \epsilon$, we have $|\theta(r - \tilde{r})| \leq K\|r - \tilde{r}\|$. Hence, for all $r \in \mathbb{R}^n$ satisfying $\|r - \tilde{r}\| < \epsilon$,

$$|\theta(r) - \theta(\tilde{r})| \leq \max\{\theta(\tilde{r} - r), \theta(r - \tilde{r})\} \leq K\|r - \tilde{r}\|,$$

where the first inequality follows from the subadditivity of θ.

We now show that for any $\tilde{r}, r \in \mathbb{R}^k$, $|\theta(\tilde{r}) - \theta(r)| \leq K\|\tilde{r} - r\|$. Define m to be an integer larger than $\|\tilde{r} - r\|/\epsilon$. Let $r^i = \frac{i}{m}(r - \tilde{r}) + \tilde{r}$, so $r^0 = \tilde{r}$ and $r^m = r$. Moreover, $\|r^{i+1} - r^i\| = \|\tilde{r} - r\|/m < \epsilon$. Therefore, $|\theta(r^{i+1}) - \theta(r^i)| \leq K\|r^{i+1} - r^i\|$ for all $i = 0, \dots, m - 1$. Hence,

$$|\theta(\tilde{r}) - \theta(r)| \leq \sum_{i=0}^{m-1} |\theta(r^{i+1}) - \theta(r^i)| \leq \sum_{i=0}^{m-1} K\|r^{i+1} - r^i\| = K\|\tilde{r} - r\|. \qquad \square$$

6.2.4 Properties of Genuinely k-Dimensional Functions

Now we focus on properties that we gain by imposing that a function is genuinely k-dimensional. We will need the following lemma, which is implied by Lemma 13 in [30] and is a consequence of Dirichlet's Approximation Theorem for the reals.

Lemma 33. *Let $y \in \mathbb{R}^k$ be any point and $r \in \mathbb{R}^k \setminus \{0\}$ be any direction. Then for every $\epsilon > 0$ and $\bar\lambda \geq 0$, there exists $w \in \mathbb{Z}^k$ such that $y + w$ is at distance less than ϵ from the half line $\{y + \lambda r \mid \lambda \geq \bar\lambda\}$.*

Lemma 34. *Let $\theta\colon \mathbb{R}^k \to \mathbb{R}$ be non-negative, Lipschitz continuous, subadditive and periodic with respect to the lattice $\mathbb{Z}^k$. Suppose there exist $r \in \mathbb{R}^k \setminus \{0\}$ and $\bar\lambda > 0$ such that $\theta(\lambda r) = 0$ for all $0 \leq \lambda \leq \bar\lambda$. Then θ is not genuinely k-dimensional.*

Proof. Let the Lipschitz constant for θ be K, that is, $|\theta(x) - \theta(y)| \leq K\|x - y\|$ for all $x, y \in \mathbb{R}^k$.

We will begin by showing that $\theta(\lambda r) = 0$ for all $\lambda \in \mathbb{R}$. Let $\lambda' \in \mathbb{R}$.

Suppose that $\lambda' > \bar\lambda$ and let $M \in \mathbb{Z}_+$ such that $0 \leq \lambda'/M \leq \bar\lambda$. From the hypothesis, we have that $\theta(\frac{\lambda'}{M} r) = 0$. By non-negativity and subadditivity of θ we see $0 \leq \theta(\lambda' r) \leq M\theta(\frac{\lambda'}{M} r) = 0$, and therefore, $\theta(\lambda' r) = 0$. This shows that $\theta(\lambda r) = 0$ for all $\lambda \geq 0$.

Next suppose $\lambda' < 0$. By Lemma 33, for all $\epsilon > 0$ there exists a $w \in \mathbb{Z}^k$ such that $\lambda' r + w$ is at distance less than ϵ from the half line $\{\lambda' r + \lambda r \mid \lambda \geq -\lambda'\} = \{\lambda r \mid \lambda \geq 0\}$. That is, there exists a $\tilde\lambda \geq 0$ such that $\|\lambda' r + w - \tilde\lambda r\| \leq \epsilon$. Since $\theta(\tilde\lambda r) = 0$, by periodicity and then Lipschitz continuity, we see that $0 \leq \theta(\lambda' r) = \theta(\lambda' r + w) = \theta(\lambda' r + w) - \theta(\tilde\lambda r) \leq K\epsilon$. This holds for every $\epsilon > 0$ and therefore $\theta(\lambda' r) = 0$. Thus, we have shown that $\theta(\lambda r) = 0$ for all $\lambda \in \mathbb{R}$.

Let $L = \{\lambda r \mid \lambda \in \mathbb{R}\}$. We claim that if $x - y \in L$, then $\theta(x) = \theta(y)$. Since $x - y \in L$, as shown above, $\theta(x - y) = 0$. By subadditivity, $\theta(y) + \theta(x - y) \geq \theta(x)$, which implies $\theta(y) \geq \theta(x)$. Similarly, $\theta(x) \geq \theta(y)$, and hence we have equality.

We conclude that $\theta = \phi \circ \operatorname{proj}_{L^\perp}$ for some function $\phi\colon \mathbb{R}^{k-1} \to \mathbb{R}$ and therefore θ is not genuinely k-dimensional. □

Lemma 35. *Let $\theta\colon \mathbb{R}^k \to \mathbb{R}$ be non-negative, piecewise linear with a locally finite cell complex, subadditive, periodic with respect to the lattice $\mathbb{Z}^k$ and genuinely k-dimensional with at most $k+1$ slopes and suppose that it satisfies $\theta(0) = 0$. Then θ has exactly $k+1$ slopes. Let the gradient set of θ be the vectors $g^1, \dots, g^{k+1}$. Then* *they satisfy* $\mathrm{cone}(g^i)_{i=1}^{k+1} = \mathbb{R}^k$. *Furthermore, for every* $i = 1, \ldots, k+1$, *there exists a maximal cell* $P \in \mathcal{P}_i$ *such that* $0 \in P$.

Proof. First we note that since θ is subadditive, satisfies $\theta(0) = 0$ and is piecewise linear with a locally finite cell complex, Lemmas 31 and 32 imply that θ is a Lipschitz continuous function.

We label the gradient set of θ as $g^1, \ldots, g^n$ and the corresponding subcomplexes as $\mathcal{P}_1, \ldots, \mathcal{P}_n$ where $n \leq k+1$. Without loss of generality, assume that $0 \in \mathcal{P}_i$ for $i \leq m$ and $0 \notin \mathcal{P}_i$ for $i > m$ for some $m \leq n \leq k+1$. Let $C = \{\, r \in \mathbb{R}^k \mid g^i \cdot r \leq 0, \;\; i = 1, \ldots, m \,\}$ be the polar cone of $\mathrm{cone}(g^i)_{i=1}^m$. We show that $C = \{0\}$, which implies that $\mathrm{cone}(g^i)_{i=1}^m = \mathbb{R}^k$. This would imply that $m = k+1$ and $\mathrm{cone}(g^i)_{i=1}^{k+1} = \mathbb{R}^k$. Moreover, this would imply that $0 \in \mathcal{P}_i$ for every i and so there exists a maximal cell in $\mathcal{P}_i$ containing 0.

Suppose there exists $r^0 \in C \setminus \{0\}$. Since θ has a locally finite cell complex, there exists an open ball $B_\epsilon(0)$ such that $\mathcal{P} \cap B_\epsilon(0) = \mathcal{F} \cap B_\epsilon(0)$, where $\mathcal{F}$ is a polyhedral fan where every maximal cell contains the origin. Since $0 \in \mathcal{P}_i$ for $i \leq m$ and $0 \notin \mathcal{P}_i$ for $i > m$, there exists $0 < \delta < \epsilon$ such that $B_\delta(0)$ intersects only $\mathcal{P}_1, \ldots, \mathcal{P}_m$. Let $\bar{\lambda} > 0$ such that $\lambda r^0 \in B_\delta(0)$ for all $0 \leq \lambda \leq \bar{\lambda}$. Since $r^0 \in C$, we see that $g^i \cdot r^0 \leq 0$ for all $i = 1, \ldots, m$. Since $\mathcal{F}$ is a polyhedral fan, the line segment from 0 to $\bar{\lambda} r^0$ lies completely within a cell $P' \in \mathcal{P}_i$ for some $i = 1, \ldots, m$. Thus $0 \leq \theta(\lambda r^0) = \lambda g^i \cdot r^0 \leq 0$ for all $0 \leq \lambda \leq \bar{\lambda}$. But then by Lemma 34, θ is not genuinely k-dimensional. This is a contradiction. □

6.2.5 Line Integrals

The following discussion shows that we can compute line integrals of the gradients of $(k+1)$-slope functions. We choose to restrict ourselves to functions with locally finite cell complexes. This is motivated by the necessity of excluding certain pathological cases and allows us to give a completely elementary proof. We remark that this restriction precludes handling some important functions such as the ones constructed in [32], which do not have locally finite cell complexes. More general versions of Lemma 36 below can, of course, be proved using the Lebesgue version of the fundamental theorem of calculus.

Lemma 36. *Consider a locally finite, complete polyhedral complex* $\mathcal{P}$ *in* $\mathbb{R}^k$ *and let* $\{\mathcal{P}_i\}_{i=1}^{k+1}$ *be a partition of the set of maximal cells of* $\mathcal{P}$. *Fix a point* $r \in \mathbb{R}^k$. *Then there exist* $\mu_1, \mu_2, \ldots, \mu_{k+1} \in \mathbb{R}_+$ *with* $\sum_{i=1}^{k+1} \mu_i = 1$ *such that for every function* θ *that* *is piecewise linear compatible with* $\{\mathcal{P}_i\}_{i=1}^{k+1}$ *with gradients* $g^1, \dots, g^{k+1}$ *corresponding to this partition, the following holds.*

$$\theta(r) = \theta(0) + \sum_{i=1}^{k+1} \mu_i (g^i \cdot r).$$

Proof. Let $\rho\colon [0,1] \to \mathbb{R}^k$ be the parameterization of the segment $[0, r]$ given by $\rho(\lambda) = \lambda r$. Let $\mathcal{Q}_i = \{\, \rho^{-1}(P \cap [0, r]) \mid P \in \mathcal{P}_i \,\}$. By convexity, $\mathcal{Q}_i$ is a family of intervals in $[0, 1]$ (some of these intervals could be degenerate). Moreover, since $\mathcal{P}$ is locally finite, Remark 3 guarantees that $\mathcal{Q}_i$ is a finite family. In addition, since $\mathcal{P}$ is complete, the union of the intervals in $\bigcup_{i=1}^{k+1} \mathcal{Q}_i$ equals $[0, 1]$.

Using the finiteness of the above families, let $0 = \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_n = 1$ be the end-points of the intervals in $\bigcup_{i=1}^{k+1} \mathcal{Q}_i$; i.e., each interval $[\lambda_j, \lambda_{j+1}]$ is an interval in $\mathcal{Q}_i$, for some $i = 1, \dots, k+1$. This implies that $\rho([\lambda_j, \lambda_{j+1}])$ is contained in a polyhedron in $\mathcal{P}_i$. In this case, the compatibility of θ with $\{\mathcal{P}_i\}_{i=1}^{k+1}$ gives

$$\theta(\lambda_{j+1} r) - \theta(\lambda_j r) = g^i \cdot (\lambda_{j+1} r) - g^i \cdot (\lambda_j r) = (\lambda_{j+1} - \lambda_j)(g^i \cdot r).$$

Therefore,

$$\theta(r) - \theta(0) = \sum_{j=0}^{n-1} (\theta(\lambda_{j+1} r) - \theta(\lambda_j r)) = \sum_{i=1}^{k+1} |\mathcal{Q}_i| (g^i \cdot r), \tag{6.2}$$

where $|\mathcal{Q}_i|$ is the sum of the lengths of all the intervals in $\mathcal{Q}_i$. Setting $\mu_i = |\mathcal{Q}_i|$ completes the result. □

6.3 Proof of Theorem 7

We now concentrate on a function π which satisfies the hypothesis, i.e., π is a minimal valid function that is piecewise linear with a locally finite cell complex and genuinely k-dimensional with at most $k+1$ slopes. We recapitulate properties of π that we have derived. Theorem 4 shows that $\pi(0) = 0$, π is subadditive, periodic with respect to the lattice $\mathbb{Z}^k$, and satisfies the symmetry condition. Lemma 31 and Lemma 32 show that π is Lipschitz continuous. Lemma 35 uses the assumption of π being a genuinely k-dimensional function and shows that π has exactly $k+1$ slopes. Let π be a piecewise linear function with cell complex $\mathcal{P}$ and we denote the gradient set of π by $\{\bar{g}^1, \dots, \bar{g}^{k+1}\}$ and the subcomplex corresponding to vector $\bar{g}^i$ as $\mathcal{P}_i$. Lemma 35 also shows that $0 \in \mathcal{P}_i$ for all $i = 1, \dots, k+1$.

The proof structure is guided by the so-called *Facet Theorem* proved in [87]. For the sake of completeness and because of differences in notation, we provide a statement of this theorem below and a self-contained proof in Section 6.6. For any valid function θ, let $E(\theta)$ denote the set of all pairs $(u, v) \in \mathbb{R}^k \times \mathbb{R}^k$ such that $\theta(u + v) = \theta(u) + \theta(v)$.

Theorem 8 (Facet Theorem)**.** *Let π be a minimal valid function. Suppose for every minimal valid function π', the condition holds that $E(\pi) \subseteq E(\pi')$ implies $\pi' = \pi$. Then π is a facet.*

Main Goal. We consider any minimal valid function π' such that $E(\pi) \subseteq E(\pi')$ and show that $\pi' = \pi$. By Theorem 8, this will imply that π is a facet.

Since π' is minimal, by Theorem 4, π' is non-negative, subadditive, periodic with respect to the lattice $\mathbb{Z}^k$. Moreover, $\pi'(0) = 0$ and the symmetry condition holds, i.e., $\pi'(r) + \pi'(-f - r) = 1$ for all $r \in \mathbb{R}^k$, and because of periodicity, $\pi'(w - f) = 1$ for every $w \in \mathbb{Z}^k$. Finally, the symmetry condition and non-negativity of π' implies that π' is bounded above by 1.

For the following proof, we will use θ when we wish to refer to a more general function than π'.

6.3.1 Compatibility

We show that π' is a piecewise linear function compatible with $\{\mathcal{P}_i\}_{i=1}^{k+1}$.

The idea of the proof is the following. First, using the partial additivity of π' implied by $E(\pi) \subseteq E(\pi')$, we show that π' is affine in parallelotopes around the origin. Then, we use translates of these parallelotopes to show that π' is affine in each maximal cell of $\mathcal{P}$. In fact, our arguments will imply that π' has the same gradient in every maximal cell in $\mathcal{P}_i$, which then gives the desired result.

In order to carry out the first step, we will use the following variant of a classical result in real analysis regarding Cauchy's functional equation (see [1] for a reference). The first application of this result in the context of the infinite group problem was made by Gomory and Johnson in [89]. The form in which we will use this result appears as Lemma 5.8 in [44] and was proved in [32].

Lemma 37 (Interval Lemma)**.** *Let $\theta \colon \mathbb{R} \to \mathbb{R}$ be a function bounded on every bounded interval. Given real numbers $u_1 < u_2$ and $v_1 < v_2$, let $U = [u_1, u_2]$, $V = [v_1, v_2]$, and*

$U + V = [u_1 + v_1, u_2 + v_2]$. *If* $\theta(u) + \theta(v) = \theta(u + v)$ *for every* $u \in U$ *and* $v \in V$, *then there exists* $c \in \mathbb{R}$ *such that*

$$\begin{aligned} \theta(u) &= \theta(u_1) + c(u - u_1) && \text{for every } u \in U, \\ \theta(v) &= \theta(v_1) + c(v - v_1) && \text{for every } v \in V, \\ \theta(w) &= \theta(u_1 + v_1) + c(w - u_1 - v_1) && \text{for every } w \in U + V. \end{aligned}$$

Lemma 38 (π' is linear on parallelotopes at the origin)**.** *Let* P_0 *be a cell in* $\mathcal{P}$ *containing the origin. Consider any parallelotope* $\Pi \subset P_0$ *such that:* (i) $0 \in \Pi$ *and* (ii) $\Pi + \Pi \subseteq P_0$. *Then there exists* g' *such that* $\pi'(r) = g' \cdot r$ *for all* $r \in \Pi$.

Proof. Since Π contains the origin, let $v^1, \dots, v^n$ be generating vectors of Π, namely these are linearly independent vectors such that $\Pi = \{ \sum_{i=1}^n \lambda_i v^i \mid \lambda_i \in [0,1] \text{ for } i = 1, \dots, n \}$.

We first claim that for all $r, r' \in \Pi$, we have that $\pi'(r) + \pi'(r') = \pi'(r + r')$. To see this, recall that π is affine in P_0, and hence in Π. Since $\pi(0) = 0$, π is actually linear in Π. Using the fact that $\Pi + \Pi \subseteq P_0$, we obtain that $\pi(r) + \pi(r') = \pi(r + r')$ for all $r, r' \in \Pi$; since $E(\pi) \subseteq E(\pi')$, the same holds for π', which proves the claim.

Fix any $i \in \{1, \dots, n\}$. We claim that π' is linear on the segment $[0, v^i]$. Consider the function $\phi(\lambda) = \pi'(\lambda v^i)$, which by the previous paragraph is additive over $[0,1]$, i.e., $\phi(\lambda) + \phi(\lambda') = \phi(\lambda + \lambda')$ for all $\lambda, \lambda' \in [0,1]$. Since π' (and hence ϕ) is bounded, the Interval Lemma (Lemma 37) applied to θ implies there exists a scalar α_i such that $\pi'(\lambda v^i) = \phi(\lambda) = \alpha_i \lambda + \phi(0)$ for all $\lambda \in [0,1]$. Since $\pi'(0) = 0$, we also have $\phi(0) = 0$. Therefore, $\pi'(\lambda v^i) = \alpha_i \lambda$.

As $v^1, \dots, v^n$ are linearly independent, there exists a $g' \in \mathbb{R}^k$ such that $g' \cdot v^i = \alpha_i$ for all $i = 1, \dots, n$. We claim that $\pi'(r) = g' \cdot r$ for all $r \in \Pi$. By letting $r = \sum_{i=1}^n \lambda_i v^i$, the result follows because π' is additive on Π.

$$\pi'(r) = \pi'\Big(\sum_{i=1}^n \lambda_i v^i\Big) = \sum_{i=1}^n \pi'(\lambda_i v^i) = \sum_{i=1}^n \alpha_i \lambda_i = g' \cdot \Big(\sum_{i=1}^n \lambda_i v^i\Big) = g' \cdot r. \qquad \square$$

Before we proceed, we prove a technical lemma about the continuity of π'. The motivation is to prove that π' is affine in each maximal cell of $\mathcal{P}$ by showing this for the interior of each cell, and then the full result follows by continuity.

Lemma 39. π' *is Lipschitz continuous.*

Proof. We first show that π' is locally Lipschitz continuous at 0, i.e., there exist $K > 0$, $\epsilon > 0$ such that $|\pi'(r)| \le K\|r\|$ for all $r \in B_\epsilon(0)$. Since π has a locally finite cell complex, there exists a neighborhood $B_{\epsilon_1}(0)$ of the origin satisfies $\mathcal{P} \cap B_{\epsilon_1}(0) = \mathcal{F} \cap B_{\epsilon_1}(0)$ for some complete polyhedral fan $\mathcal{F}$. We consider a triangulation $\bar{\mathcal{F}}$ of $\mathcal{F}$, i.e., $\bar{\mathcal{F}}$ contains a triangulation for each cone in $\mathcal{F}$ and every element of $\bar{\mathcal{F}}$ is simplicial. Consider any maximal simplicial cone $C \in \bar{\mathcal{F}}$ and consider $P \in \mathcal{P}$ such that $C \cap B_{\epsilon_1}(0) \subseteq P$ (note that such a P exists because $\bar{\mathcal{F}}$ is a triangulation of $\mathcal{F}$). Then there exist generators $\{v_C^1, \dots, v_C^k\}$ for C such that the parallelotope Π formed by $\{v_C^1, \dots, v_C^k\}$ is such that $\Pi + \Pi \subseteq P$. We do this construction for all maximal elements of $\bar{\mathcal{F}}$ to obtain a finite polyhedral complex of parallelotopes $\mathcal{S}$.

We now show that the union of all elements in $\mathcal{S}$ contains 0 in its interior. For every maximal element C of $\bar{\mathcal{F}}$, there exists $\epsilon_C > 0$ such that $\delta r \in \Pi_C$ for all $r \in C$ and all $0 \leq \delta \leq \epsilon_C$, where Π_C is the parallelotope in $\mathcal{S}$ corresponding to C. Observe that $\bar{\mathcal{F}}$ is complete because $\mathcal{F}$ is complete and $\bar{\mathcal{F}}$ is a triangulation of $\mathcal{F}$. Therefore, choosing $\epsilon_2 = \min\{\, \epsilon_C \mid C \in \bar{\mathcal{F}} \,\}$, the ball $B_{\epsilon_2}(0)$ is contained in the union of all the parallelotopes in $\mathcal{S}$.

From Lemma 38, for every parallelotope $\Pi \in \mathcal{S}$, there exists $g^\Pi \in \mathbb{R}^k$ such that $\pi'(r) = g^\Pi \cdot r$ for all $r \in \Pi$. Let $K = \max\{\|g^\Pi\| \mid \Pi \in \mathcal{S}\}$. By the Cauchy–Schwarz inequality, $|\pi'(r)| \leq \|g^\Pi\|\|r\| \leq K\|r\|$ for all $r \in \Pi$. Since the union of all parallelotopes in $\mathcal{S}$ contains $B_{\epsilon_2}(0)$ in its interior, $|\pi'(r)| \leq K\|r\|$ for all $r \in \mathbb{R}^n$ satisfying $\|r\| < \epsilon_2$, i.e., π' is locally Lipschitz continuous at the origin.

Since π' is a subadditive function that is locally Lipschitz continuous at the origin, Lemma 32 shows that π' is (globally) Lipschitz continuous. □

The following lemma will be the main tool for using translates of patches to prove that π' is affine in the maximal cells of $\mathcal{P}$.

Lemma 40 (Finite path of patches)**.** *Let $P \subseteq \mathbb{R}^k$ be a full-dimensional polyhedron and $\Pi \subseteq \mathbb{R}^k$ be a full-dimensional parallelotope with $0 \in \Pi$. Let x, y be points that lie in* $\operatorname{int}(P)$*. Then there exist a number $0 < \epsilon \leq 1$, an integer m, and points $x^0 = x, x^1, x^2, \dots, x^m = y \in P$ such that:*

(i) $x^j + \epsilon\Pi \subseteq P$ *for* $j = 0, \dots, m$,

(ii) $(x^j + \epsilon\Pi) \cap (x^{j+1} + \epsilon\Pi)$ *is non-empty for* $j = 0, \dots, m-1$,

where $\epsilon\Pi = \{\, \epsilon x \mid x \in \Pi \,\}$.

Proof. After a linear change of coordinates, we can assume that the parallelotope Π is the unit cube $[0, 1]^k$.

Since $x, y \in \operatorname{int}(P)$, then there exists $\delta > 0$ such that both $B_\delta(y)$ and $B_\delta(x)$ lie within P. Choose $0 < \epsilon \leq 1$ so that $\epsilon\Pi \subset B_\delta(0)$. Therefore, $x + \epsilon\Pi \subseteq P$ and $y + \epsilon\Pi \subseteq P$. Let $m > \|y^0 - x^0\|_\infty/\epsilon$ be an integer.

Let

$$x^j = x^0 + \frac{j}{m}(y^0 - x^0) \quad \text{for } j = 1, \ldots, m;$$

thus $x^m = y^0$. Since $x + \epsilon\Pi \subseteq P$ and $y + \epsilon\Pi \subseteq P$, by convexity $x^j + \epsilon\Pi \subseteq P$ for all $j = 0, \ldots, m$. In particular, $x^j \in P$ for all $j = 0, \ldots, m$. Moreover $\|x^{j+1} - x^j\|_\infty < \epsilon \leq 1$, and thus $(x^j + \epsilon\Pi) \cap (x^{j+1} + \epsilon\Pi)$ is non-empty. □

Lemma 41 (π' is affine on each maximal cell of π)**.** *Let $P_0 \in \mathcal{P}_i$ be a maximal cell containing the origin and let Π be a full-dimensional parallelotope with $0 \in \Pi \subseteq P_0$ such that $\pi'(x) = g' \cdot x$ for all $x \in \Pi$. Let P be a maximal cell in $\mathcal{P}_i$ and $\bar{x} \in \operatorname{int}(P)$. Then $\pi'(x) = g' \cdot (x - \bar{x}) + \pi'(\bar{x})$ for all $x \in P$.*

Proof. First consider $\bar{y} \in \operatorname{int}(P)$. Let ϵ and $x^0 = \bar{x}, \ldots, x^m = \bar{y} \in P$ be the data from applying Lemma 40 on x, y, P and Π. Fix any $j \in \{0, \ldots, m\}$ and consider an arbitrary $s \in \epsilon\Pi$. Since $P \in \mathcal{P}_i$, $\pi(x^j + s) - \pi(x^j) = \bar{g}^i \cdot s = \pi(s)$, where the second equality follows from $\Pi \subseteq P_0 \in \mathcal{P}_i$. Therefore, $\pi(x^j + s) = \pi(x^j) + \pi(s)$ and so the pair (x^j, s) is in $E(\pi) \subseteq E(\pi')$. Therefore, $\pi'(x^j + s) = \pi'(x^j) + \pi'(s)$, and thus $\pi'(x^j + s) = \pi'(x^j) + g' \cdot s$. Thus π', restricted to each $x^j + \epsilon\Pi$, is an affine function with gradient g', which we write as

$$\pi'(x) = g' \cdot (x - \bar{x}) + \alpha_j \quad \text{for } x \in x^j + \epsilon\Pi \tag{6.3}$$

for some real number α_j.

For all $j = 0, 1, \ldots, m$, we prove that $\alpha_j = \pi'(\bar{x})$ and therefore $\pi'(x) = g' \cdot (x - \bar{x}) + \pi'(\bar{x})$ holds for all $x \in x^j + \epsilon\Pi$. We do this by induction on j. For $j = 0$, this holds since $\bar{x} = x^0$. Now let $j + 1 > 0$ and assume $\alpha_j = \pi'(\bar{x})$. Let z be any point in the intersection $(x^j + \epsilon\Pi) \cap (x^{j+1} + \epsilon\Pi)$, which is non-empty by Lemma 38. By evaluating (6.3) for j and $j + 1$ at $x = z$, we see that in fact $\alpha_{j+1} = \pi'(\bar{x})$. Therefore, in particular, $\pi'(\bar{y}) = \pi'(x^m) = a'(\bar{y} - \bar{x}) + \pi'(\bar{x})$.

This shows that for every $x \in \operatorname{int}(P)$, $\pi'(x) = g' \cdot (x - \bar{x}) + \pi'(\bar{x})$. By Lemma 39, π' is continuous, and therefore the equation extends from $\operatorname{int}(P)$ to all of P. □

Proposition 4. *The function π' is a piecewise linear function compatible with $\{\mathcal{P}_i\}_{i=1}^{k+1}$.*

Proof. Fix $i \in \{1, \ldots, k + 1\}$. Since π satisfies the hypotheses of Lemma 35, there exists a maximal cell $P_0 \in \mathcal{P}_i$ containing the origin. Since P_0 is a full-dimensional polyhedron containing the origin, there exists a full-dimensional parallelotope Π with $0 \in \Pi$ and $\Pi + \Pi \subseteq P_0$. Let g' be the vector from Lemma 38 such that $\pi'(r) = g' \cdot r$ for $r \in \Pi$. Define $\tilde{g}^i = g'$. Now let P be any maximal cell in $\mathcal{P}_i$ and pick any $y \in \operatorname{rel\,int}(P)$. By Lemma 41,

$$\pi'(r) = \tilde{g}^i \cdot (r - y) + \pi'(y) = \tilde{g}^i \cdot r + \delta_P$$

for $r \in P$, where we set $\delta_P = \pi'(y) - \tilde{g}^i \cdot y$. Thus π' is a piecewise linear function compatible with $\{\mathcal{P}_i\}_{i=1}^{k+1}$. □

Notice that this compatibility implies that there exist vectors $\tilde{g}^1, \tilde{g}^2, \ldots, \tilde{g}^{k+1}$ corresponding to $\mathcal{P}_1, \ldots, \mathcal{P}_{k+1}$ such that for any $P \in \mathcal{P}_i$, there exists δ_P such that $\pi'(r) = \tilde{g}^i \cdot r + \delta_P$. However, note that we have not shown $\tilde{g}^1, \tilde{g}^2, \ldots, \tilde{g}^{k+1}$ to be all distinct.

6.3.2 Constructing a System of Linear Equations

As the next step in proving that $\pi = \pi'$, we construct a system of linear equations which is satisfied by both $\bar{g}^1, \ldots, \bar{g}^{k+1}$ and $\tilde{g}^1, \ldots, \tilde{g}^{k+1}$.

The system has two sets of constraints, the first of which follows from Theorem 4 and Lemma 36. The second set of constraints is more involved. Consider two adjacent cells $P, P' \in \mathcal{P}$ that contain a segment $[x, y] \subseteq \mathbb{R}^k$ in their intersection. Along the line segment $[x, y]$, the gradients of P and P' projected onto the line spanned by the vector $y - x$ must agree; the second set of constraints captures this observation. We will identify a set of vectors $r^1, \ldots, r^{k+1}$ such that every subset of k vectors is linearly independent and such that each vector r^i is contained in k cells of $\mathcal{P}$ with different gradients. We then use the segment $[0, r^i]$ to obtain linear equations involving the gradients of π and π'. The fact that every subset of k vectors is linearly independent will be crucial in ensuring the uniqueness of the system of equations.

Remark 1. *In the case $k = 2$, in the terminology of [53], these vectors would all be* directions *of the piecewise linear function π; see also the discussion in Section 6.7.*

To show the existence of such a set of vectors, we utilize the following classical lemma in combinatorial topology.

Lemma 42 (KKM [99, 2])**.** *Consider an n-simplex $\operatorname{conv}(u^j)_{j=1}^n$. Let $F_1, F_2, \ldots, F_n$ be closed sets such that for all $I \subseteq \{1, \ldots, n\}$, the face $\operatorname{conv}(u^j)_{j \in I}$ is contained in $\bigcup_{j \in I} F_j$. Then the intersection $\bigcap_{j=1}^n F_j$ is non-empty.*

Lemma 43. *There exist vectors $r^1, r^2, \dots, r^{k+1} \in \mathbb{R}^k$ with the following properties:*

(i) *For every $i, j, \ell \in \{1, \dots, k+1\}$ with j, ℓ different from i, the equations $r^i \cdot \bar{g}^j = r^i \cdot \bar{g}^\ell$ and $r^i \cdot \tilde{g}^j = r^i \cdot \tilde{g}^\ell$ hold.*

(ii) $\operatorname{cone}(r^i)_{i=1}^{k+1} = \mathbb{R}^k$.

Proof. We consider the neighborhood $B_\epsilon(0)$ of the origin given by the local finiteness assumption (see Definition 5). Let $F_i = \bigcup_{P \in \mathcal{P}_i} (P \cap \bar{B}_\epsilon(0))$, namely the set of points in the closed ball $\bar{B}_\epsilon(0)$ for which π has gradient $\bar{g}^i$. Since $\bar{B}_\epsilon(0)$ is compact, Proposition 3 says that only finitely many terms are non-empty in the union $\bigcup_{P \in \mathcal{P}_i} (P \cap \bar{B}_\epsilon(0))$. Moreover each term $P \cap \bar{B}_\epsilon(0)$ is closed as it is the intersection of a polyhedron with a closed ball. Thus, each F_i is a finite union of closed sets and therefore is closed. Our first goal is to show that, for each $i = 1, \dots, k+1$, there is a vector r^i which belongs to $\bigcap_{j \neq i} F_j$.

In order to better understand how the sets F_i intersect, we start by defining the set $H_i = \{\, r \in \mathbb{R}^k \mid \bar{g}^i \cdot r \le 0 \,\}$. The crucial property of this set is that the gradient of π at these points must be *different* from $\bar{g}^i$, at least around the origin.

Claim 4. *For every $i = 1, \dots, k+1$, the set F_i is disjoint with H_i.*

Proof. Suppose to the contrary that there exists $P \in \mathcal{P}_i$ and $r \in H_i \cap P \cap \bar{B}_\epsilon(0)$. Since $r \in \bar{B}_\epsilon(0)$, the entire segment $[0, r]$ is contained in P. Moreover, $\bar{g}^i \cdot r \le 0$ as $r \in H_i$. Thus $\pi(\lambda r) = \lambda \bar{g}^i \cdot r \le 0$ for all $0 \le \lambda \le 1$. Since π is piecewise linear with a locally finite cell complex and subadditive, Lemmas 31 and 32 show that π is Lipschitz continuous. Therefore, π satisfies the hypotheses of Lemma 34 and we conclude that π is not genuinely k-dimensional. This is a contradiction. □

For a subset $I \subseteq \{1, \dots, k+1\}$, define the cone $C_I = \bigcap_{i \notin I} H_i$ (for convenience of notation, we use C_j instead of $C_{\{j\}}$ for a singleton set). From the above claim, for all $i \notin I$ we have F_i disjoint with C_I; since $B_\epsilon(0) \subseteq \bigcup_{i=1,\dots,k+1} F_i$, we get that $C_I \cap B_\epsilon(0) \subseteq \bigcup_{i \in I} F_i$. Alternatively, the gradient of π in any point in $C_I \cap B_\epsilon(0)$ must be within the set $\{\bar{g}^i\}_{i \in I}$. We need the following technical property of the cones C_I.

Claim 5. *C_j is full-dimensional for all $j = 1, \dots, k+1$.*

Proof. Observe that the polar cone

$$(C_j)^\circ = \left\{\, \textstyle\sum_{i \neq j} \lambda_i \bar{g}^i \mid \lambda_i \ge 0 \,\right\}$$

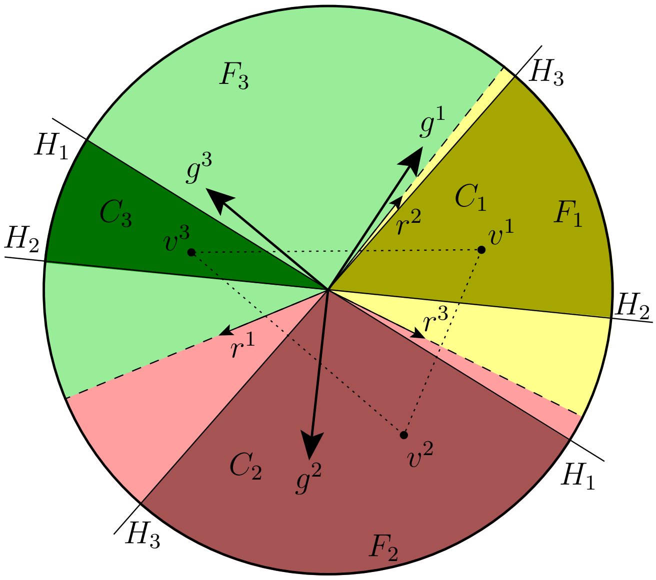

Figure 6.2: The geometry of the proof of Lemma 43. Each cone C_i (shaded in dark colors) is the intersection of the halfspaces H_j (defined by the gradients $\bar{g}^j$) for $j \neq i$. Near the origin (within the ball $B_\epsilon(0)$), each point of C_i lies in the set F_i of points where the function π has gradient $\bar{g}^i$ (shaded in light colors). Picking points v_i near the origin in the interior of C_i, we construct a simplex Δ with 0 in its interior. By applying the KKM Lemma to each of its facets Δ_i, we show the existence of the vectors r^i with the desired properties.

does not contain any lines because the set $\{\bar{g}^i\}_{i\neq j}$ is linearly independent by Lemma 29 and Lemma 35. Hence, C_j is full-dimensional. $\square$

In order to continue analyzing how the sets F_i intersect, it is useful to focus on a full-dimensional simplex $\text{conv}(v^j)_{j=1}^{k+1}$ around the origin. More precisely, Claim 5 allows us to pick $v^j \in \text{int}(C_j) \cap B_\epsilon(0)$ for every $j = 1, \dots, k+1$. Since $v^j \in \text{int}(C_j)$, we have $v^j \cdot \bar{g}^i < 0$ for all $i \neq j$. Then employing Lemma 30 with $a^i = \bar{g}^i$ and $b^i = v^i$, we deduce that $\text{cone}(v^i)_{i=1}^{k+1} = \mathbb{R}^k$. Therefore, $\Delta = \text{conv}(v^i)_{i=1}^{k+1}$ is indeed a full-dimensional simplex.

Since $\Delta \subseteq B_\epsilon(0) \subseteq \bigcup_{i=1,\dots,k+1} F_i$, the sets F_i form a closed cover of Δ, and in particular they form a closed cover of each facet $\Delta_i = \text{conv}(v^j)_{j\neq i}$. We will show that, for each $i = 1, \dots, k+1$, there is a point r^i in Δ_i which belongs to $\bigcap_{i\neq j} F_j$. For that, we apply the KKM Lemma (Lemma 42) to the simplex Δ_i.

To do so, we need to show that for every $I \subseteq \{1, \dots, k+1\} \setminus \{i\}$, the face $\text{conv}(v^j)_{j\in I}$ is contained in $\bigcup_{j\in I} F_j$. To see that this holds, consider an arbitrary subset $I \subseteq \{1, \dots, k+1\} \setminus \{i\}$. By definition, for every $j \in I$ we have $v^j \in \text{int}(C_j) \cap B_\epsilon(0) \subseteq C_I \cap B_\epsilon(0)$. Since $C_I \cap B_\epsilon(0)$ is convex, it follows that the entire face $\text{conv}(v^j)_{j\in I}$ belongs to $C_I \cap B_\epsilon(0)$. As mentioned previously, $C_I \cap B_\epsilon(0) \subseteq \bigcup_{j\in I} F_j$ and hence the face $\text{conv}(v^j)_{j\in I}$ is contained in $\bigcup_{j\in I} F_j$.

Therefore, for each $i = 1, \dots, k+1$, the KKM Lemma (Lemma 42) implies the existence of a point $r^i \in \Delta_i$ belonging to $\bigcap_{j\neq i} F_j$ as desired.

Now it is easy to see that $r^1, \dots, r^{k+1}$ satisfy property (i) as claimed. Fix $i \in \{1, \dots, k+1\}$. Consider $j \neq i$ and let $P \in \mathcal{P}_j$ contain r^i; notice that actually $r^i \in P \cap \Delta_i \subseteq P \cap B_\epsilon(0)$. Since $\mathcal{P} \cap B_\epsilon(0) = \mathcal{F} \cap B_\epsilon(0)$ for some polyhedral fan $\mathcal{F}$, P also contains the entire segment $[0, r]$. Since π is affine in P with gradient $\bar{g}^j$, it follows that $\pi(r^i) - \pi(0) = r^i \cdot \bar{g}^j$; this implies that for all $j, \ell \neq i$ we have $r^i \cdot \bar{g}^j = r^i \cdot \bar{g}^\ell$. Similarly, since π' is a piecewise linear function compatible with $\{\mathcal{P}_i\}_{i=1}^{k+1}$, again we have that $\pi'(r^i) - \pi'(0) = r^i \cdot \tilde{g}^j$ for all $j \neq i$, and hence $r^i \cdot \tilde{g}^j = r^i \cdot \tilde{g}^\ell$ for all $j, \ell \neq i$.

Finally, we prove that $r^1, \dots, r^{k+1}$ satisfy property (ii) as claimed. Because $r^i \in \bigcap_{j\neq i} F_j$, Claim 4 directly implies that $r^i \notin H_j$ for every $j \neq i$, namely $r^i \cdot \bar{g}^j > 0$ when $j \neq i$. Now using Lemma 30 with $a^i = -\bar{g}^i$ and $b^i = r^i$, we deduce that $\text{cone}(r^i)_{i=1}^{k+1} = \mathbb{R}^k$. This concludes the proof of Lemma 43. $\square$

We finally present the system of linear equations that we consider.

Corollary 1. *Consider vectors $a^1, a^2, \dots, a^{k+1} \in \mathbb{Z}^k - f$ such that* $\text{cone}(a^i)_{i=1}^{k+1} = \mathbb{R}^k$. *Also, let $r^1, r^2, \dots, r^{k+1}$ be the vectors given by Lemma 43. Then there exist $\mu_{ij} \in \mathbb{R}_+$,*

$i, j \in \{1, \dots, k+1\}$ with $\sum_{j=1}^{k+1} \mu_{ij} = 1$ for all $i \in \{1, \dots, k+1\}$ such that both $\tilde{g}^1, \dots, \tilde{g}^{k+1}$ and $\bar{g}^1, \dots, \bar{g}^{k+1}$ are solutions to the linear system

$$\begin{aligned} \textstyle\sum_{j=1}^{k+1} (\mu_{ij} a^i) \cdot g^j &= 1 && \textit{for all } i \in \{1, \dots, k+1\}, \\ r^i \cdot g^j - r^i \cdot g^\ell &= 0 && \textit{for all } i, j, \ell \in \{1, \dots, k+1\} \textit{ such that } i \neq j, \ell, \end{aligned} \tag{6.4}$$

with variables $g^1, \dots, g^{k+1} \in \mathbb{R}^k$.

Proof. Feasibility for the first set of constraints follows directly from the minimality of π and π', Theorem 4 and Lemma 36. Feasibility for the second set of constraints follows from Lemma 43 (i). □

We remark that we can always find vectors $a^1, a^2, \dots, a^{k+1} \in \mathbb{Z}^k - f$ such that $\text{cone}(a^i)_{i=1}^{k+1} = \mathbb{R}^k$, so the system above indeed exists.

6.3.3 Unique Solution of the Linear System

We now analyze the solution set of (6.4), which will be rewritten as a system of $k(k+1)$ linear equations for the $k+1$ gradient vectors, i.e., in $k(k+1)$ variables. We will show the gradients of π and π' coincide by demonstrating that this system either has no solutions or has a unique solution. Recall from linear algebra that, given a square matrix A and a vector b, if the augmented matrix $[b\ A]$ has full row rank, then the linear system $Ay = b$ either has no solutions or has a unique solution.

Proposition 5. *$\bar{g}^i = \tilde{g}^i$ for every $i = 1, \dots, k+1$.*

Proof. We wish to show that the system (6.4) either has no solution or a unique solution. We begin by rewriting the system in terms of some new variables. Since for any fixed i, the value of $r^i g^j$ must coincide for all $j = 1, \dots, k+1$, $i \neq j$, we use the variable z_i, $i = 1, \dots, k+1$ to denote this common value. We can rewrite the system (6.4) as

$$\begin{aligned} \textstyle\sum_{j=1}^{k+1} (\mu_{ij} a^i) \cdot g^j &= 1 && \text{for all } i = 1, \dots, k+1 \\ r^i \cdot g^j - z_i &= 0 && \text{for all } i, j = 1, \dots, k+1, \text{ such that } i \neq j. \end{aligned} \tag{6.5}$$

Note that there is a one-to-one mapping between solutions of (6.4) and (6.5). We now rearrange the variables and the constraints of (6.5) so that it can be represented as $Ay = b$, where

$$A = \begin{bmatrix} \Delta & O_{k+1\times k+1} \\ R & I' \end{bmatrix}, \quad y = [g^1, \ldots, g^{k+1}, z]^T, \quad \text{and} \quad b = [1, \ldots, 1, O_{1\times k} \ldots, O_{1\times k}]^T,$$

where Δ is a $(k+1) \times k(k+1)$-matrix, R is a $k(k+1) \times k(k+1)$-matrix, I' is a $k(k+1) \times (k+1)$-matrix and $O_{i\times j}$ is the $i \times j$-matrix with all zero entries. The i-th row of Δ, $i = 1, \ldots, k+1$, is given by $(\mu_{i1}a^i, \ldots, \mu_{i(k+1)}a^i)$ where a^i is written as a row vector. The matrix R has a block diagonal structure:

$$R = \begin{bmatrix} R_1 & & \\ & \ddots & \\ & & R_{k+1} \end{bmatrix},$$

where each R_i is a $k \times k$-matrix. For each $i = 1, \ldots, k+1$, the matrix R_i has rows r^j, $j \neq i$.

The matrix I' has entries corresponding to the coefficients on z, and will be written as

$$I' = \begin{bmatrix} -I_1 \\ \vdots \\ -I_{k+1} \end{bmatrix},$$

where I_i is a $k \times (k+1)$-matrix obtained from the $k \times k$ identity matrix with the 0 column inserted as the i-th column.

We now argue that the matrix $[b\ A]$ has full row rank. Since A is a $(k+1)^2 \times (k+1)^2$ square matrix, if $[b\ A]$ has full row rank, the system $Ay = b$ either has a unique solution or no solution.

We use one further trick to prove $[b\ A]$ has full row rank: we analyze the row rank of the matrix $[b\ D\ A]$, where D is a $(k+1)^2 \times k$-matrix of all zero entries. The rank of $[b\ D\ A]$ is the same as $[b\ A]$ and we now show that $[b\ D\ A]$ has full row rank. We now perform the block row and column operations on the matrix

$$[\, b \mid D \mid A\,] = \left[\begin{array}{c|c|ccc|c} 1 & O_{1\times k} & \mu_{11}a^1 & \cdots & \mu_{1(k+1)}a^1 & O_{1\times k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & O_{1\times k} & \mu_{(k+1)1}a^{k+1} & \cdots & \mu_{(k+1)(k+1)}a^{k+1} & O_{1\times k} \\ \hline O_{k\times 1} & O_{k\times k} & R_1 & & & -I_1 \\ \vdots & \vdots & & \ddots & & \vdots \\ O_{k\times 1} & O_{k\times k} & & & R_{k+1} & -I_{k+1} \end{array}\right].$$

First, add all the block columns of A corresponding to each $g^1, \dots, g^{k+1}$ to the block D, giving (recall that $\sum_{j=1}^{k+1} \mu_{ij} = 1$ for all $i \in \{1, \dots, k+1\}$)

$$\left[\begin{array}{c|c|ccc|c} 1 & a^1 & \mu_{11}a^1 & \dots & \mu_{1(k+1)}a^1 & O_{1\times k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & a^{k+1} & \mu_{(k+1)1}a^{k+1} & \dots & \mu_{(k+1)(k+1)}a^{k+1} & O_{1\times k} \\ \hline O_{k\times 1} & R_1 & R_1 & & & -I_1 \\ \vdots & \vdots & & \ddots & & \vdots \\ O_{k\times 1} & R_{k+1} & & & R_{k+1} & -I_{k+1} \end{array}\right].$$

Second, in the last matrix above, multiply the last block of $k+1$ columns (corresponding to the variables z_i) on the right by the matrix $\bar{R}$, which is the $(k+1) \times k$ matrix whose rows are the $k+1$ vectors r^i, for all $i = 1, \dots, k+1$. Note that $-I_i\bar{R} = -R_i$ for every $i = 1, \dots, k+1$. Hence, if we multiply the last block of columns with $\bar{R}$ and add to the second block of columns in the last matrix above, we obtain

$$\left[\begin{array}{c|c|ccc|c} 1 & a^1 & \mu_{11}a^1 & \dots & \mu_{1(k+1)}a^1 & O_{1\times k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & a^{k+1} & \mu_{(k+1)1}a^{k+1} & \dots & \mu_{(k+1)(k+1)}a^{k+1} & O_{1\times k} \\ \hline O_{k\times 1} & O_{k\times k} & R_1 & & & -I_1 \\ \vdots & \vdots & & \ddots & & \vdots \\ O_{k\times 1} & O_{k\times k} & & & R_{k+1} & -I_{k+1} \end{array}\right].$$

The final matrix has an upper triangular block structure. The blocks on the diagonal are

$$A' = \begin{bmatrix} 1 & a^1 \\ \vdots & \vdots \\ 1 & a^{k+1} \end{bmatrix}, \quad R_1, \dots, R_k, \quad \text{and} \quad \begin{bmatrix} R_{k+1} & -I_{k+1} \end{bmatrix}.$$

Each R_i has full row rank since every proper subset of $\{r^1, \dots, r^{k+1}\}$ is linearly independent by Lemma 29. Also, A' has full row rank because $a^1, \dots, a^{k+1}$ are affinely independent since $\text{cone}(a^i)_{i=1}^{k+1} = \mathbb{R}^k$. Hence, we have shown that $[b\ D\ A]$ has full row rank.

Therefore the system (6.4) has either no solutions or has a unique solution. Since Corollary 1 shows that both $\tilde{g}^1, \dots, \tilde{g}^{k+1}$ and $\bar{g}^1, \dots, \bar{g}^{k+1}$ are solutions to (6.4), it follows that $\tilde{g}^i = \bar{g}^i$ for all $i = 1, \dots, k+1$. $\square$

6.3.4 Conclusion of the Proof

We finally put all the pieces together to complete the proof of Theorem 7. Recall that our main goal before beginning Subsection 6.3.1 was to show that $\pi = \pi'$. Since both π and π' are minimal, Theorem 4 guarantees that $\pi(0) = \pi'(0) = 0$. At the end of Subsection 6.3.1, we concluded that there exist $\tilde{g}^i$, $i = 1, \ldots, k$, and δ_P for all $P \in \mathcal{P}$ such that $\pi'(r) = \tilde{g}^i \cdot r + \delta_P$, where $\mathcal{P}$ is the polyhedral complex corresponding to π. Proposition 5 shows that $\bar{g}^i = \tilde{g}^i$ for all $i = 1, \ldots, k+1$. From Lemma 36, for every $r \in \mathbb{R}^k$ there exist $\mu_1, \mu_2, \ldots, \mu_{k+1}$ such that

$$\pi(r) = \pi(0) + \sum_{i=1}^{k+1} \mu_i(\bar{g}^i \cdot r) = \pi'(0) + \sum_{i=1}^{k} \mu_i(\tilde{g}^i \cdot r) = \pi'(r).$$

This proves that $\pi = \pi'$ and concludes the proof of Theorem 7.

6.4 Conclusion

As mentioned, the $(k+1)$-slope theorem we present here answers an open question of Gomory and Johnson from the 1970's and brings us one step closer to understanding the infinite group relaxation for arbitrary k. We do not currently know if or how our theorem can be generalized further. There are several directions that the assumptions could be weakened, for example, removing the locally finite criterion for piecewise linear function or allowing for discontinuous functions. A non–piecewise linear 2-slope extreme function was presented in [32], and nothing more is known about such functions. The issue of discontinuous extreme functions has also only begun to be addressed in [70]. Also, it is unknown if the definitions of facets and extreme functions coincide or if there do exist extreme functions which are not facets.

Another area of interest is generalizing the interval lemma. Although we applied the interval lemma here in k dimensions, we always used an interval connected to the origin when applying it. The interval lemma is stronger than just this one application, and we should try to understand how and when it can be used in higher dimensions. The only known generalization of the interval lemma for when $k = 2$ given in [69]. An interval lemma for general k would likely lead to some interesting results in higher dimensions.

Lastly, we are also interested to know how $(k+1)$-slope functions compare computationally. Studying this requires determining useful classes of $(k+1)$-slope functions in varying dimensions and generating cutting planes from them. We hope that such functions perform well in practice.

Appendix

6.5 Proof of the existence of minimal valid functions – Theorem 3

Proof of Theorem 3. Consider the non-empty set Σ of valid functions π' with $\pi' \leq \pi$ (the set is non-empty because $\pi \in \Sigma$). We now consider $(\Sigma, \leq)$ as a partially ordered set, where the partial order is imposed by the relation $\pi_1 \leq \pi_2$ for $\pi_1, \pi_2 \in \Sigma$. If we can show that every chain in $(\Sigma, \leq)$ has a lower bound, then applying Zorn's lemma we would conclude that Σ has a minimal element which will be the minimal function π' we are looking for.

Consider any chain $\mathcal{C}$ in $(\Sigma, \leq)$, i.e., for $\pi_1, \pi_2 \in \mathcal{C}$ either $\pi_1 \leq \pi_2$ or $\pi_2 \leq \pi_1$. Consider the function $\pi_{\mathcal{C}}$ defined as follows: $\pi_{\mathcal{C}}(r) = \inf_{\pi' \in \mathcal{C}} \pi'(r)$. We claim that $\pi_{\mathcal{C}} \in \Sigma$. We only need to verify that it is a valid function; it is clear that $\pi_{\mathcal{C}} \leq \pi$. Since $\pi' \geq 0$ for all $\pi' \in \mathcal{C}$, $\pi_{\mathcal{C}} \geq 0$.

Suppose to the contrary that there exists $s \geq 0$ with finite support such that $f + \sum_{r \in \mathbb{R}^k} r s_r \in \mathbb{Z}^k$, but $\sum_{r \in \mathbb{R}^k} \pi_{\mathcal{C}}(r) s_r < 1$. Let $\{r^1, \dots, r^n\}$ be the finite support of s, i.e., $s_r = 0$ for all $r \notin \{r^1, \dots, r^n\}$. Let $S = \max\{s_{r^1}, \dots, s_{r^n}\}$ and let $\epsilon = 1 - \sum_{r \in \mathbb{R}^k} \pi_{\mathcal{C}}(r) s_r > 0$. Since $\pi_{\mathcal{C}}(r^i) = \inf_{\pi' \in \mathcal{C}} \pi'(r^i)$, there exists $\pi_i \in \mathcal{C}, i = 1, \dots, n$ such that $\pi_i(r^i) \leq \pi_{\mathcal{C}}(r^i) + \frac{\epsilon}{2nS}$. Since $\mathcal{C}$ is a chain, there exists $i^* \in \{1, \dots, n\}$ such that $\pi_{i^*} \leq \pi_i$ for all $i \in \{1, \dots, n\}$. Hence, $\pi_{i^*}(r^i) \leq \pi_{\mathcal{C}}(r^i) + \frac{\epsilon}{2nS}$ for every $i \in \{1, \dots, n\}$. But then

$$\sum_{r \in \mathbb{R}^k} \pi_{i^*}(r) s_r \leq \sum_{r \in \mathbb{R}^k} \pi_{\mathcal{C}}(r) s_r + \sum_{i=1}^{n} \frac{\epsilon}{2nS} s_{r^i} \leq 1 - \epsilon + \frac{\epsilon}{2nS} nS < 1,$$

which shows that π_{i^*} is not a valid function, which is a contradiction because $\pi_{i^*} \in \Sigma$. $\square$

6.6 Facet Theorem

The next lemma shows that a weaker condition than that in the definition of a facet is enough to guarantee facetness.

Lemma 44. *Let π be minimal valid function. Suppose for every minimal valid function π^*, the condition holds that $S(\pi) \subseteq S(\pi^*)$ implies $\pi^* = \pi$. Then π is a facet.*

Proof. Consider any valid function π' (not necessarily minimal) such that $S(\pi) \subseteq S(\pi')$; we show that $\pi' = \pi$.

Suppose to the contrary that there exists $r_1 \in \mathbb{R}^k$ such that $\pi(r_1) \neq \pi'(r_1)$. We claim that actually there is r_2 such $\pi(r_2) > \pi'(r_2)$. To see this, first notice that the symmetry condition of π (via Theorem 4) guarantees that $\pi(r_1) + \pi(-f - r_1) = 1$. Moreover, it is clear that the solution $\bar{s}$ given by $\bar{s}_{r_1} = \bar{s}_{-f-r_1} = 1$ and $\bar{s}_r = 0$ otherwise is feasible; together, these observations imply that $\bar{s} \in S(\pi)$. Since $S(\pi) \subseteq S(\pi')$, we have that $\bar{s} \in S(\pi')$ and hence

$$\pi'(r_1) + \pi'(-f - r_1) = \sum_{r \in \mathbb{R}^k} \pi'(r)\bar{s}_r = 1 = \pi(r_1) + \pi(-f - r_1).$$

Since $\pi(r_1) \neq \pi'(r_1)$, it follows that either $\pi(r_1) > \pi'(r_1)$ or $\pi(-f - r_1) > \pi'(-f - r_1)$, and the claim holds.

Now consider a minimal valid function $\pi^* \leq \pi'$ (which exists by Theorem 3). Notice that $S(\pi') \subseteq S(\pi^*)$: for $\bar{s} \in S(\pi')$, using its validity we get $1 \leq \sum_{r \in \mathbb{R}^k} \pi^*(r)\bar{s}_r \leq \sum_{r \in \mathbb{R}^k} \pi'(r)\bar{s}_r = 1$, hence equality hold throughout and $\bar{s} \in S(\pi^*)$. Since $S(\pi) \subseteq S(\pi')$, we get that $S(\pi) \subseteq S(\pi^*)$. However, $\pi \neq \pi^*$, since there is r_2 such that $\pi(r_2) > \pi'(r_2) \geq \pi^*(r_2)$. This contradicts the assumptions on π, which concludes the proof. □

Proof of Theorem 8. By Lemma 44, all we need to show is that for every minimal valid function π', $S(\pi) \subseteq S(\pi')$ implies $\pi' = \pi$. We simply show that for every minimal valid function π', $S(\pi) \subseteq S(\pi')$ implies $E(\pi) \subseteq E(\pi')$.

So let π' be a minimal valid function with $S(\pi) \subseteq S(\pi')$. Consider any $(r_1, r_2) \in E(\pi)$, namely such that $\pi(r_1) + \pi(r_2) = \pi(r_1 + r_2)$. Notice that the solution $\bar{s}$ given by $\bar{s}_{r_1} = \bar{s}_{r_2} = \bar{s}_{-f-r_1-r_2} = 1$ and $\bar{s}_r = 0$ is feasible. Moreover, using symmetry condition of π we get that $\bar{s} \in S(\pi)$. Indeed,

$$\sum_{r \in \mathbb{R}^k} \pi(r)\bar{s}_r = \pi(r_1) + \pi(r_2) + \pi(-f - (r_1 + r_2)) = \pi(r_1 + r_2) + \pi(-f - (r_1 + r_2)) = 1.$$

Since $S(\pi) \subseteq S(\pi')$, the solution $\bar{s}$ also belongs to $S(\pi')$, and now the symmetry condition of π' gives

$$1 = \sum_{r \in \mathbb{R}^k} \pi'(r)\bar{s}_r = \pi'(r_1) + \pi'(r_2) + \pi'(-f - r_1 - r_2) = \pi'(r_1) + \pi'(r_2) + (1 - \pi'(r_1 + r_2)).$$

Thus, $\pi'(r_1)+\pi'(r_2) = \pi'(r_1+r_2)$ and $(r_1, r_2) \in E(\pi')$. This concludes the proof. $\square$

6.7 Proof that Theorem 7 generalizes Theorem 3 of [53]

We restate Theorem 3 of [53] here using our terminology. A *direction* of a piecewise linear function π with cell complex $\mathcal{P}$ is a linear space parallel to a one-dimensional element of $\mathcal{P}$, such as an edge.

Theorem 9 (Theorem 3 of [53]). *Let $\pi\colon \mathbb{R}^2 \to \mathbb{R}$ be a minimal valid function. If π is piecewise linear with a locally finite cell complex, has 3 slopes and has 3 directions, then π is extreme.*

Consider π satisfying the hypothesis of the above theorem; it suffices to show that π satisfies the hypothesis of Theorem 7. So let $\{a^i\}_{i=1}^3$ be the gradient set of π. Lemma 3.6 of [53] implies that $\text{cone}(a^i)_{i=1}^3 = \mathbb{R}^3$. The next lemma, which provides a partial converse to Lemma 35, shows that this property guarantees that π is genuinely 3-dimensional; this implies that π satisfies the hypothesis of Theorem 7 and concludes the proof.

Lemma 45. *Let $\theta\colon \mathbb{R}^k \to \mathbb{R}$ be a piecewise linear function with gradient set $\{a^i\}_{i\in I}$. If $\text{cone}(a^i)_{i\in I} = \mathbb{R}^k$, then θ is genuinely k-dimensional.*

Proof. By means of contradiction, suppose that θ is not genuinely k-dimensional. So consider a function $\phi\colon \mathbb{R}^{k-1} \to \mathbb{R}$ and a linear map $T\colon \mathbb{R}^k \to \mathbb{R}^{k-1}$ such that $\theta = \phi \circ T$. Notice that the kernel of T contains some non-zero vector, and let v be one such vector.

Let θ be a piecewise linear function with cell complex $\mathcal{P}$ and take a maximal cell $P \in \mathcal{P}$; we claim that $a^P \cdot v = 0$. Since P is full-dimensional, we can find $x, y \in P$ such that $y - x = \lambda v$ for some $\lambda \neq 0$. Since $T(v) = 0$, we have $\pi(y) = \pi(x+\lambda v) = \phi(T(x+\lambda v)) = \phi(T(x)) = \pi(x)$. Moreover, by definition of $\mathcal{P}$, we have that $\pi(r) = a^P \cdot r + \delta_P$ for all $r \in P$. Putting the two previous observations together, we get that $0 = \pi(y) - \pi(x) = a^P \cdot (y-x) = \lambda a^P \cdot v$. Since $\lambda \neq 0$, this implies that $a^P \cdot v = 0$.

However, since this holds for every $P \in \mathcal{P}$, it is clear that $\text{cone}(a^i)_{i\in I}$ belongs to the orthogonal complement of v, and hence does not equal $\mathbb{R}^k$. This contradicts the assumption on the vectors a^i and concludes the proof of the lemma. $\square$

6.8 Lifting of Cuts from Lattice-free Simplices

Let $S \subseteq \mathbb{R}^k$ be a simplex which contains f in its interior and which is lattice-free, namely S does not contain integer points in its interior. It is known that S can be expressed using linear inequalities in the form $\{x \in \mathbb{R}^k : a^i(x - f) \leq 1, i = 1, 2, \dots, k+1\}$, for some collection $\{a^i\}_{i=1}^{k+1}$ where $\text{cone}(a_i)_{i=1}^{k+1} = \mathbb{R}^k$. Let $\pi_S : \mathbb{R}^k \to \mathbb{R}$ be defined as follows:

$$\pi_S(r) = \min_{w \in \mathbb{Z}^k} \max_{i \in \{1,2,\dots,k+1\}} a^i(r + w). \tag{6.6}$$

Note that $\pi_S \geq 0$ since $\text{cone}(a_i)_{i=1}^{k+1} = \mathbb{R}^k$. The main observation is that π_S is a valid function for (IR) (see Corollary 7.6 in [44]).

In [33], the authors provide conditions on S, such that π_S gives a so-called *unique lifting* (see [45] and [29] for a discussion of unique liftings). It can be shown that for such special simplices, π_S is a minimal valid function for (IR) (see Theorem 7.2 in [44]). We claim that given such a simplex S, the function π_S satisfies all the hypothesis of Theorem 7, and therefore the latter can be used to show that π_S is a facet of (IR). We need to verify that π_S is a genuinely k-dimensional piecewise linear function with a locally finite cell complex and having at most $k + 1$ slopes. From equation (6.6) it follows that π_S is a piecewise linear function with $k + 1$ slopes (one slope for each a^i). Moreover, one can also show that π_S is genuinely k-dimensional because $\pi_S(r) = 0$ if and only if $r \in \mathbb{Z}^k$ (one needs to use the fact that $\text{cone}(a_i)_{i=1}^{k+1} = \mathbb{R}^k$ implies that $\max_{i \in \{1,2,\dots,k+1\}} a^i(r + w) = 0$ if and only if $r = -w$). Finally, using the fact that $\text{cone}(a_i)_{i=1}^{k+1} = \mathbb{R}^k$, one can show that the cell complex of π_S is locally finite: given any bounded region $B \subseteq \mathbb{R}^k$ the minimum in equation (6.6) can be replaced by minimum over all w's of bounded norm:

$$\forall r \in B, \quad \pi_S(r) = \min_{w \in \mathbb{Z}^k, \|w\| \leq U} \max_{i \in \{1,2,\dots,k+1\}} a^i(r + w), \tag{6.7}$$

for some sufficiently large scalar U. This can be used to prove the local finiteness. This concludes the proof of the claim.

CHAPTER 7

Lifting Gomory Cuts With Bounded Variables

Recently, Balas and Qualizza [22] introduced new cuts for mixed 0/1 programs called *lopsided cuts*. Their derivation is based on the Balas-Jeroslow modularization technique [17] and uses significantly the (upper) bounds present in the variables. This is of special interest given that current techniques for generating cuts (e.g. based on the corner relaxation) ignore the bounds on several variables; in fact, the role of that latter information is still not well understood.

In this chapter we provide a geometric derivation of these lopsided cuts and generalize it to an infinite family, which includes the GMI inequality. The first step in our approach is to see the Balas-Qualizza cuts as valid cuts for version of the mixed-integer infinite relaxation where the basic variable is upper bounded. Then we notice that these cuts are actually liftings (taking into account the upper bound on the basic variable) of the fractional Gomory function. For that, we use a variation of the "geometric lifting" technique introduced by in [45]. This perspective allow us to extend the construction to obtain the generalization of the Balas-Qualizza cuts.

Moreover, we also show that all the cuts in the family we obtain are "new": they are all extreme for the mixed-integer infinite relaxation with upper bounded basic variable. We also provide some preliminary computational results that unfortunately shows that these cuts decrease in importance as they move away from the GMI inequality, complementing the experimental results from [22].

Organization of the chapter. We start defining the upper bounded mixed-integer relaxation that we use and recall some definitions used in this chapter. In Section 7.2 we briefly introduce the modified geometric lifting perspective that we employ. In Section 7.3, we construct the family of cuts that is the main object of study of this chapter, while in Sections 7.4 and 7.5 we provide an analysis of their strength theoretically and empirically, respectively.

Acknowledgment. This chapter is joint work with Gérard Cornuéjols and Tamás Kis.

7.1 Preliminaries

We consider a linear equation where a bounded integer variable x is expressed in terms on nonnegative variables. By a change of variable, we may assume that $x \leq 1$. Consider the upper bounded system

$$\begin{aligned} x = f + \sum_{r \in \mathbb{R}} r s_r + \sum_{r \in \mathbb{R}} r y_r \\ x \in \{-\infty, \dots, 0, 1\} \\ s_r \geq 0 \\ y_r \in \mathbb{Z}_+ \\ (s, y) \text{ has finite support,} \end{aligned} \tag{7.1}$$

which is an upper bounded version of a mixed-integer infinite relaxation.

For simplicity of exposition, we focus on the case $0 < f < 1$. Note however that the approach that we present below can also be used when f is further from the bound. We will fix $f \in (0, 1)$ from now own. We use IP to denote the set of (x, s, y) feasible solutions for the above.

Let ψ_{GMI} denote the classical Gomory function for the coefficients of the continuous variables

$$\psi_{GMI}(r) = \begin{cases} -\frac{r}{f} & r < 0 \\ \frac{r}{1-f} & r \geq 0 \end{cases} .$$

Our goal is to lift ψ_{GMI} for the integral variables, namely find π such that

$$\sum_r \psi_{GMI}(r) s_r + \sum_r \pi(r) y_r \geq 1$$

is satisfied by all $(x, s, y) \in IP$.

7.2 Modified Geometric Lifting

Since integrality constraints on basic variables (i.e. on the left-hand side of the equation in 7.1) are more easily handled, the idea is to 'transfer' the integrality of y to a basic variable. For that, we first consider the extended system

$$\begin{aligned} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} + \sum_{r \in \mathbb{R}} \begin{pmatrix} r \\ 0 \end{pmatrix} s_r + \sum_{r \in \mathbb{R}} \begin{pmatrix} r \\ \ell(r) \end{pmatrix} y_r & \\ x \in \{-\infty, \dots, 0, 1\} & \\ z \in \mathbb{Z} & \\ s_r \geq 0 & \\ y_r \in \mathbb{Z}_+ & \\ (s, y) \text{ has finite support.} & \end{aligned} \tag{IP(ℓ)}$$

We use $IP(\ell)$ to denote the set of feasible solutions (x, z, s, y) for the above and observe that, when $\ell(r)$ is integral for all $r \in \mathbb{R}$, this extended system is equivalent to the original one.

Proposition 6. *$IP = \text{proj}_{x,y,s}\, IP(\ell)$ for all $\ell : \mathbb{R} \to \mathbb{Z}$.*

We assume $\ell : \mathbb{R} \to \mathbb{Z}$ in the remainder of the paper. Now we relax the integrality of the non-basic variables to obtain the system that we work with:

$$\begin{aligned} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} + \sum_{r \in \mathbb{R}} \begin{pmatrix} r \\ 0 \end{pmatrix} s_r + \sum_{r \in \mathbb{R}} \begin{pmatrix} r \\ \ell(r) \end{pmatrix} y_r & \\ x \in \{-\infty, \dots, 0, 1\} & \\ z \in \mathbb{Z} & \\ s_r \geq 0 & \\ y_r \geq 0 & \\ (s, y) \text{ has finite support.} & \end{aligned} \tag{Z(ℓ)}$$

Notice that this system partially captures the original integrality of y and hence (its projection onto the (x, y, s)-space) is tighter than the relaxation obtained by completely dropping the integrality of y from IP.

7.3 Wedge Inequalities

We now construct the family of cuts which are the object of study in this paper. Let $S = \{-\infty, \dots, 0, 1\} \times \mathbb{Z}$ and notice that every feasible solution (x, z, s, y) for $Z(\ell)$ satisfies $(x, z) \in S$. We say that a set $K \subseteq \mathbb{R}^2$ is *S-free* if K does not contain in its interior any point of S.

For $\alpha \in (0, 1]$, we consider the S-free convex set $K_\alpha = \{(x, z) : a^1 \cdot [(x, z) - (f, 0)] \leq 1, a^2 \cdot [(x, z) - (f, 0)] \leq 1\}$ with

$$a^1 = \left(-\frac{1}{f}, \frac{1}{f}\right) \qquad a^2 = \left(\frac{1}{1-f}, -\frac{1-\alpha(1-f)}{\alpha f(1-f)}\right).$$

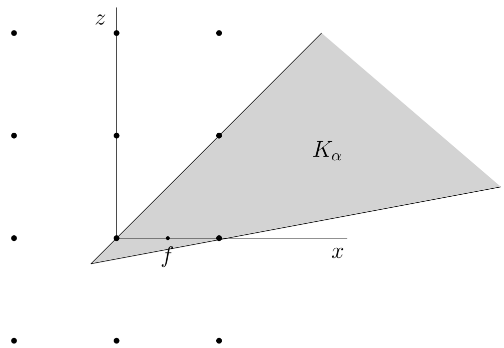

Figure 7.1: S-free convex set K_α. The slope of the lower facet of K_α is $\frac{\alpha f}{1-\alpha(1-f)}$.

The S-free convex set K_α contains $(f, 0)$ in its interior and therefore it can be used to derive an intersection cut for $Z(\ell)$. More specifically, from the theory of S-free cuts [30], we obtain the valid inequality

$$\sum_r \bar{\psi}(r) s_r + \sum_r \bar{\pi}^\ell_\alpha(r) y_r \geq 1 \tag{7.2}$$

with $\bar{\psi}(r) = \max\{a^1 \cdot (r, 0),\ a^2 \cdot (r, 0)\}$ and $\bar{\pi}^\ell_\alpha(r) = \max\{a^1 \cdot (r, \ell(r)),\ a^2 \cdot (r, \ell(r))\}$. We have the following explicit formula for the coefficients (see Figure 7.2):

$$\bar{\psi}(r) = \begin{cases} -\frac{r}{f} & r < 0 \\ \frac{r}{1-f} & r \geq 0 \end{cases}$$

$$\bar{\pi}^\ell_\alpha(r) = \begin{cases} \frac{-r+\ell(r)}{f} & \ell(r) > \alpha r \\ \frac{r}{1-f} - \frac{\ell(r)(1-\alpha(1-f))}{\alpha f(1-f)} & \ell(r) \leq \alpha r. \end{cases}$$

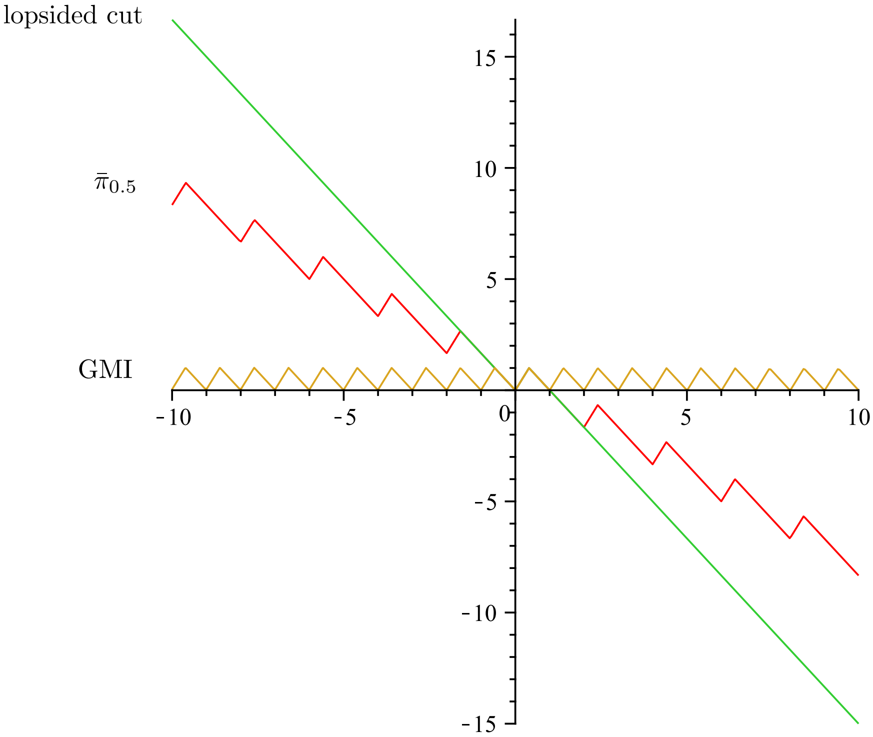

Figure 7.2: Graph of $\bar{\pi}_{0.5}$ and of functions corresponding to lopsided and GMI cuts.

Inequality (7.2) will be called *wedge inequality* in the remainder. Since $\bar{\psi} = \psi_{GMI}$, the function $(\bar{\psi}, \bar{\pi}^{\ell}_{\alpha})$ is a lifting of the Gomory function. Geometrically this is clear: the Gomory function ψ_{GMI} is obtained by considering the lattice-free set $[0,1]$ in $IP(f)$. For all α the function $\bar{\psi}$ is obtained from the set K_α along rays $(r, 0)$, and therefore the relevant part of K_α is $K_\alpha \cap \{z = 0\}$ which is the segment $[0,1] \times \{0\}$.

We remark that if x is a 0,1 variable, another cut, exploiting the lower bound of 0, can be obtained by substituting x by $1 - x$.

7.3.1 Optimizing ℓ

Notice that the cut above is valid for every $\ell : \mathbb{R} \to \mathbb{Z}$; we can choose the one which gives the 'best' coefficients, i.e. the smallest value of $\bar{\pi}^{\ell}_{\alpha}(r)$. Thankfully, each ray is associated to a different component of ℓ, so we can actually get the best coefficient for all the rays simultaneously.

Proposition 7. *For given r and α, the value of ℓ that minimizes $\bar{\pi}^{\ell}_{\alpha}(r)$ is $\ell_\alpha(r) = \lceil \alpha(r + f - 1) \rceil$. Furthermore $\bar{\pi}^{\ell_\alpha}_{\alpha}(r) = min\{\frac{-r+\lceil \alpha r \rceil}{f}, \frac{r}{1-f} - \frac{\lfloor \alpha r \rfloor (1-\alpha(1-f))}{\alpha f(1-f)}\}$.*

Proof. Fix $r \in \mathbb{R}$. Note that $\bar{\pi}^{\ell}_{\alpha}(r)$ as a function of $\ell(r)$ is a piece-wise linear function which is decreasing in the interval $(-\infty, \alpha r]$ and increasing in the interval $[\alpha r, \infty)$, hence with minimum at $\ell(r) = \alpha r$. Therefore, the minimum over all integer values of $\ell(r)$ is attained at either $\ell(r) = \lfloor \alpha r \rfloor$ or $\lceil \alpha r \rceil$. This shows the second part of the proposition.

To prove the first part, let $\tilde{\ell}$ be the (unique) value $\tilde{\ell} \leq \alpha r < \tilde{\ell} + 1$ such that $\bar{\pi}^{\tilde{\ell}}_{\alpha}(r) = \bar{\pi}^{\tilde{\ell}+1}_{\alpha}(r)$. A simple calculation gives $\tilde{\ell} = \alpha(r + f - 1)$. Since there is only one integer in the range $[\tilde{\ell}, \tilde{\ell} + 1)$, the optimum choice of the integer $\ell(r)$ is $\ell_\alpha(r) = \lceil \alpha(r + f - 1) \rceil$. □

This has the following geometric interpretation: The optimum choice of ℓ is such that the ray $(r, \ell(r))$ belongs to the strip $(f, 0) + R$ (see Figure 7.3). This is related to the region of best possible liftings for wedges as introduced in [45] Section 3.1.

To simplify the notation, let $\bar{\pi}_\alpha = \bar{\pi}^{\ell_\alpha}_{\alpha}$.

7.3.2 Limit Cases

Now we consider the extreme cases $\alpha = 1$ and $\alpha \to 0$.

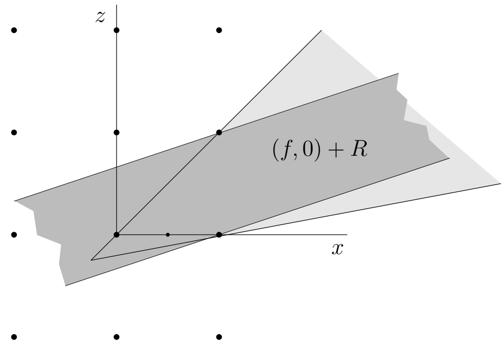

Figure 7.3: Region $(f,0)+R$ of best possible liftings.

Proposition 8. *For $\alpha = 1$, $\bar{\pi}_1(r) = min\{\frac{\lceil r\rceil - r}{f}, \frac{r-\lfloor r\rfloor}{1-f}\}$, i.e. we get the GMI cut.*

Now we focus on the case $\alpha \to 0$. Let

$$\bar{\pi}_0(r) = \begin{cases} \frac{-r}{f} & r < 0 \\ \frac{r}{1-f} & r \in [0, 1-f] \\ \frac{-r+1}{f} & r > 1-f. \end{cases}$$

Lemma 46. *Fix $r \in \mathbb{R}$. Then there exists $\alpha_0 > 0$ such that for any $0 < \alpha \le \alpha_0$, we have $\ell_\alpha(r) = 0$ if $r \le 1-f$ and $\ell_\alpha(r) = 1$ if $r > 1-f$. Hence, for all $0 < \alpha \le \alpha_0$ we have $\bar{\pi}_\alpha(r) = \bar{\pi}_0(r)$.*

The above lemma directly gives the behavior of $\bar{\pi}_\alpha$ with $\alpha \to 0$.

Proposition 9. *The pointwise limit $\lim_{\alpha\to 0} \bar{\pi}_\alpha$ equals $\bar{\pi}_0$, i.e. we get the lopsided cut of Balas and Qualizza [22].*

7.4 Strength of $(\bar{\psi}, \bar{\pi}_\alpha)$

The main goal of this section is to show that the function $(\bar{\psi}, \bar{\pi}_\alpha)$ is extreme for 7.1. In fact we will prove an even stronger result, namely that this function is extreme in the $0,1$ case. That is, we consider the more restricted system

$$\begin{aligned} x = f + \sum_{r\in\mathbb{R}} r s_r + \sum_{r\in\mathbb{R}} r y_r \\ x \in \{0,1\} \\ s_r \geq 0 \\ y_r \in \mathbb{Z}_+ \\ (s,y) \text{ has finite support.} \end{aligned} \tag{B(f)}$$

Since 7.1 is a relaxation of B(f), $(\bar{\psi}, \bar{\pi}_\alpha)$ is valid for the latter for all $\alpha \in [0,1]$.

We say that a valid function (ψ, π) is *extreme* for (B(f)) if there are no distinct valid functions (ψ_1, π_1) and (ψ_2, π_2) such that $\psi = \frac{\psi_1}{2} + \frac{\psi_2}{2}$ and $\pi = \frac{\pi_1}{2} + \frac{\pi_2}{2}$. The following theorem formally states the main result of this section.

Theorem 14. *For all $\alpha \in [0,1]$, the function $(\bar{\psi}, \bar{\pi}_\alpha)$ is extreme for* (B(f)).

In particular, this implies that we cannot improve the coefficients of the valid inequality for (7.1)

$$\sum_r \bar{\psi}(r) s_r + \sum_r \bar{\pi}_\alpha(r) y_r \geq 1$$

even if we use the additional information that $x \in \{0,1\}$. We start by showing something weaker, namely that this inequality is minimal.

7.4.1 Minimality

Just as for the infinite relaxation, we say that a valid function (ψ, π) for B(f) is *minimal* if there is no other valid function (ψ', π') such that $\psi' \leq \psi$ and $\pi' \leq \pi$. The following lower bound on valid functions is the main observation to prove that $(\bar{\psi}, \bar{\pi}_\alpha)$ is minimal.

Lemma 47. *Let (ψ, π) be a valid function for B(f). Then:*

1. $\psi \geq \bar{\psi}$
2. $\pi(r) + \pi(1 - f - r) \geq 1$ *for all* $r \in \mathbb{R}$.

Proof. Consider the first property. Notice that for $r > 0$, setting $\bar{s}_r = \frac{1-f}{r}$ and all other $\bar{s}_{r'}$'s equal to 0 gives a feasible solution for (B(f)); the validity of (ψ, π)

implies that $\psi(r)\bar{s}_r \geq 1$, or equivalently that $\psi(r) \geq \bar{\psi}(r)$. For $r < 0$, we employ the same reasoning to the solution given by $\bar{s}_r = \frac{-f}{r}$ and $\bar{s}_{r'} = 0$ for $r' \neq r$ to obtain $\psi(r) \geq \bar{\psi}(r)$. Finally, we use the solution $\bar{s}_0 = M > 0$, $\bar{s}_1 = 1 - f$ and $\bar{s}_{r'} = 0$ for $r' \notin \{0, 1\}$ to obtain that $\psi(0) \geq \frac{1-\psi(1)(1-f)}{M}$; taking $M \to \infty$ gives $\psi(0) \geq 0 = \bar{\psi}(0)$.

The second property is proved similarly by considering the feasible solution obtained by setting $\bar{y}_r = 1$, $\bar{y}_{1-f-r} = 1$ and every other component of $\bar{y}$ to 0. □

Lemma 48. *For all $\alpha \in [0, 1]$, the function $(\bar{\psi}, \bar{\pi}_\alpha)$ is minimal for* (B(f)).

Proof. Consider (ψ, π) such that $\psi \leq \bar{\psi}$ and $\pi \leq \bar{\pi}_\alpha$. The first part of Lemma 47 shows that $\psi = \bar{\psi}$. Moreover, elementary calculations show that for all r, $\bar{\pi}_\alpha(r) + \bar{\pi}_\alpha(1 - f - r) = 1$. This shows that $\pi = \bar{\pi}_\alpha$. Hence the pair $(\bar{\psi}, \bar{\pi}_\alpha)$ is minimal. □

7.4.2 Extremality

First we focus on proving Theorem 14 for the case $\alpha \in (0, 1]$. It is easy to check that the function $\bar{\pi}_\alpha$ is piecewise linear; furthermore, for $\alpha > 0$ its breakpoints occur at $\frac{k}{\alpha}$ and $\frac{k}{\alpha} + 1 - f$ for $k \in \mathbb{Z}$; notice that the first and the second set of breakpoints are respectively the local minima and maxima of $\bar{\pi}_\alpha$. Let $\ldots < x_{-2} < x_{-1} < x_0 = 0 < x_1 < x_2 < \ldots$ be this set of breakpoints; according to this definition, x_i is a local minimum for i even. It is easy to check that $\bar{\pi}_\alpha$ is *quasiperiodic*, namely for all $i \in \mathbb{Z}$ and $r \in [x_{2i-1}, x_{2i+1}]$ we have $\bar{\pi}_\alpha(r) = \bar{\pi}_\alpha(x_{2i}) + \bar{\pi}_\alpha(r - x_{2i})$.

In order to prove Theorem 14, consider valid functions (ψ_1, π_1) and (ψ_2, π_2) satisfying $\bar{\psi} = \frac{\psi_1}{2} + \frac{\psi_2}{2}$ and $\bar{\pi}_\alpha = \frac{\pi_1}{2} + \frac{\pi_2}{2}$; we show that $\bar{\psi} = \psi_1 = \psi_2$ and $\bar{\pi}_\alpha = \pi_1 = \pi_2$.

First notice that Lemma 47 implies $\psi_1 \geq \bar{\psi}$ and $\psi_2 \geq \bar{\psi}$; but then since $\frac{\psi_1}{2} + \frac{\psi_2}{2} = \psi$, it is clear that we must have the equality $\psi_1 = \psi_2 = \bar{\psi}$. Therefore, we only need to prove $\pi_1 = \pi_2 = \bar{\pi}_\alpha$. It is easy to see that $(\bar{\psi}, \pi_1)$ and $(\bar{\psi}, \pi_2)$ are minimal: if there were, say, a valid $\pi_1' \neq \pi_1$ such that $\pi_1' \leq \pi_1$, then $(\bar{\psi}, \frac{\pi_1'}{2} + \frac{\pi_2}{2})$ would be a valid function contradicting the minimality of $(\bar{\psi}, \bar{\pi}_\alpha)$. Now we evoke the following property about minimal valid functions.

Lemma 49. *Consider a minimal valid function (ψ, π) for* (B(f)). *Then π is subadditive, namely for all $r_1, r_2 \in \mathbb{R}$, $\pi(r_1 + r_2) \leq \pi(r_1) + \pi(r_2)$.*

After proving that $\pi(0) \geq 0$ as in the first part of Lemma 47, the proof of Lemma 5.2 given in [44] goes through to prove the above lemma; details are omitted.

Claim 6. *For $r \in [x_{-1}, x_1]$, $\pi_1(r) = \pi_2(r) = \bar{\psi}(r) = \bar{\pi}_\alpha(r)$.*

Proof. Fix $j \in \{1,2\}$. Using standard techniques (see [45]), one can show that since $(\bar{\psi}, \pi_j)$ is valid then $(\bar{\psi}, \min\{\bar{\psi}, \pi_j\})$ is also valid. The minimality of $(\bar{\psi}, \pi_j)$ then implies that $\pi_j \leq \min\{\bar{\psi}, \pi_j\} \leq \bar{\psi}$.

However, notice that for $r \in [x_{-1}, x_1]$, $\bar{\psi}(r) = \bar{\pi}_\alpha(r)$. As before, for $r \in [x_{-1}, x_1]$ the fact that $\bar{\psi}(r) = \pi_1(r) + \pi_2(r)$ together with the previous paragraph implies $\pi_1(r) = \pi_2(r) = \bar{\psi}(r)$. The result then follows. □

Claim 7. *For $r \in [x_{2i-1}, x_{2i+1}]$ and $j = 1, 2$, we have*

$$\pi_j(r) - \pi_j(x_{2i}) = \bar{\pi}_\alpha(r) - \bar{\pi}_\alpha(x_{2i}).$$

Proof. Since π_j is minimal, Lemma 49 gives that $\pi_j(r) \leq \pi_j(x_{2i}) + \pi_j(r - x_{2i})$. But since $r - x_{2i} \in [x_{-1}, x_1]$, it follows from Claim 6 and the quasiperiodicity of $\bar{\pi}_\alpha$ that $\pi_j(r - x_{2i}) = \bar{\pi}_\alpha(r - x_{2i}) = \bar{\pi}_\alpha(r) - \bar{\pi}_\alpha(x_{2i})$; this gives that $\pi_j(r) - \pi_j(x_{2i}) \leq \bar{\pi}_\alpha(r) - \bar{\pi}_\alpha(x_{2i})$ for $j = 1, 2$. Adding these inequalities for $j = 1, 2$, dividing by 2 and using the fact that $\bar{\pi}_\alpha = \frac{\pi_1}{2} + \frac{\pi_2}{2}$, we again obtain that we must have the equality $\pi_j(r) - \pi_j(x_{2i}) = \bar{\pi}_\alpha(r) - \bar{\pi}_\alpha(x_{2i})$ for $j = 1, 2$. This concludes the proof. □

Now take $i \in \mathbb{N}$, and $r \in [x_{2i-1}, x_{2i+1}]$; we will show that $\pi_1(r) = \pi_2(r) = \bar{\pi}_\alpha(r)$. For that, simply write $\pi_j(r) = (\pi_j(r) - \pi_j(x_{2i})) + \sum_{k=1}^{2i} (\pi_j(x_k) - \pi_j(x_{k-1})) + \pi_j(x_0)$. Applying Claim 7 to each parenthesized expression and Claim 6 to the last term, we obtain that $\pi_j(r) = \bar{\pi}_\alpha(r)$, giving the desired result. The case $i \in -\mathbb{N}$ can be handled analogously, which then proves Theorem 14 when $\alpha \in (0, 1]$.

The case $\alpha = 0$ needs to be handled separately. As before, we still have $\psi_1 = \psi_2 = \bar{\psi}$. Moreover, as in Claim 6, the fact that $\bar{\pi}_0(r) = \bar{\psi}(r)$ for all $r \leq 1 - f$ implies that $\pi_1(r) = \pi_2(r) = \bar{\pi}_0(r)$ for all $r \leq 1 - f$. For $r > 1 - f$, we claim that we have an equality analogous to Claim 7, namely that $\pi_j(r) - \pi_j(1 - f) = \bar{\pi}_0(r) - \bar{\pi}_0(1 - f)$ for $j = 1, 2$. To see this, first notice that for $r > 1 - f$ we have $\bar{\pi}_0(r) = \bar{\pi}_0(1-f) - \bar{\pi}_0(1-f-r)$. Then using the subadditivity of π_j and the fact that $1-f-r \leq 1-f$, we get $\pi_j(1-f) \leq \pi_j(1-f-r) + \pi_j(r) = \bar{\pi}_0(1-f) - \bar{\pi}_0(r) + \pi_j(r)$, and the claim follows as in Claim 7. Since $\pi_j(1 - f) = \bar{\pi}_0(1 - f)$, this equation gives that $\pi_j(r) = \bar{\pi}_0(r)$ for all $r > 1 - f$ and $j = 1, 2$. This concludes the proof of Theorem 14.

7.4.3 Split Cuts

Let LP denote the linear relaxation of formulation 7.1, i.e. it is obtained from 7.1 by replacing the integrality conditions $x \in \{-\infty, \ldots, 0, 1\}$, $y_r \in \mathbb{Z}_+$ by $x \leq 1$, $y_r \geq 0$.

We say that an inequality $\sum_r \psi(r)s_r + \sum_r \pi(r)y_r \geq 1$ is a *split* inequality if it is satisfied by all (x, s, y) in $LP \cap (\{ax + \sum_r b_r y_r \leq c\} \cup \{ax + \sum_r b_r y_r \geq c+1\})$, for some $a, c \in \mathbb{Z}, b_r \in \mathbb{Z}$ for all $r \in \mathbb{R}$.

Notice that the split disjunction $\{ax+\sum_r b_r y_r \leq c\} \cup \{ax+\sum_r b_r y_r \geq c+1\}$ is valid (i.e., it contains all feasible solutions) even when we relax the upper bound on x in 7.1, namely when we simply take x to be an integer. Thus, split inequalities do not use the information of the upper bound on x in the choice of the disjunction used to generate it. What we show next is that, although the disjunction $\{(x, z, s, y) : (x, y) \notin \operatorname{int} K_\alpha\}$ employed to generate the wedge inequality uses the upper bound on x, this inequality is dominated by a split inequality.

Theorem 15. *For all $\alpha \in (0, 1]$ and all $\ell : \mathbb{R} \to \mathbb{Z}$, the wedge inequality (7.2) is a split inequality.*

Proof. To prove the theorem, we show that the wedge inequality (7.2) is satisfied by all (x, s, y) in $LP \cap (\{x \leq \sum_r \ell(r)y_r\} \cup \{x - 1 \geq \sum_r \ell(r)y_r\})$.

Let $\beta = \frac{\alpha f}{1-\alpha(1-f)}$. Notice that $0 < \beta \leq 1$ since $0 < \alpha \leq 1$.

Claim 1 If $(\bar{x}, \bar{s}, \bar{y}) \in LP \cap (\{x \leq \sum_r \ell(r)y_r\} \cup \{x - 1 \geq \sum_r \ell(r)y_r\})$, then $(\bar{x}, \bar{s}, \bar{y}) \in LP \cap (\{x \leq \sum_r \ell(r)y_r\} \cup \{\beta(x-1) \geq \sum_r \ell(r)y_r\})$.

Proof. If $(\bar{x}, \bar{s}, \bar{y}) \in LP \cap \{x \leq \sum_r \ell(r)y_r\}$, then the claim holds. Now suppose $(\bar{x}, \bar{s}, \bar{y}) \in LP \cap \{x - 1 \geq \sum_r \ell(r)y_r\}$. Since $0 < \beta \leq 1$, and $x \leq 1$ is a constraint of LP, we have $(1-\beta)(\bar{x}-1) \leq 0$. Hence, $\sum_r \ell(r)\bar{y}_r \leq (\bar{x}-1)-(1-\beta)(\bar{x}-1) = \beta(\bar{x}-1)$ and we are done. □

To finish the proof of the theorem, observe that (7.2) is valid for $LP \cap (\{x \leq \sum_r \ell(r)y_r\} \cup \{\beta(x-1) \geq \sum_r \ell(r)y_r\})$. To see this, consider the function $\psi : \mathbb{R}^2 \to \mathbb{R}$ defined as $\psi(r, \ell) = \max\{a^1 \cdot (r, \ell), a^2 \cdot (r, \ell)\}$. It is easy to verify that ψ is convex, and positively homogeneous, and thus it is subadditive, i.e., $\psi(r, \ell) + \psi(r', \ell') \geq \psi(r + r', \ell + \ell')$. Moreover, $\psi(x - f, z) = 1$ on the boundary of the set K_α. By convexity of ψ and K_α, we get that $\psi(x - f, z) \geq 1$ if $(x, z) \notin \operatorname{int} K_\alpha$, i.e. $x \leq z$ or $\beta(x-1) \geq z$. Using the equations for (x, z) in the definition of $LP(\ell)$, we derive the following valid inequality for all (s, y) such that (x, z) satisfies $x \leq z$ or $\beta(x-1) \geq z$:

$$1 \leq \psi(x - f, z) = \psi(\sum_r (r, 0)s_r + \sum_r (r, \ell(r))y_r) \leq \sum_r \psi(r, 0)s_r + \sum_r \psi(r, \ell(r))y_r.$$

Substituting $\sum_r \ell(r)y_r$ for z into the inequalities $x \leq z$ or $\beta(x - 1) \geq z$ gives the desired result. □

We obtain as a corollary a result for the cuts $(\bar{\psi}, \bar{\pi}_\alpha)$, including the case $\alpha = 0$.

Corollary 5. *For all $\alpha \in [0,1]$, the inequality generated by $(\bar{\psi}, \bar{\pi}_\alpha)$ is a split inequality.*

Proof. It suffices to prove for $\alpha = 0$. Consider any feasible solution (x,s,y) in $LP \cap (\{(x,s,y) : x \leq \sum_{r>1-f} y_r\} \cup \{(x,s,y) : x \geq 1 + \sum_{r>1-f} y_r\})$. Since s and y have finite support, Lemma 46 implies that there exists $\alpha_0 > 0$ such that: (i) $\ell_{\alpha_0}(r)$ equals 0 if $r \leq 1-f$ and equals 1 if $r > 1-f$; (ii) $\bar{\pi}_{\alpha_0}(r) = \bar{\pi}_0(r)$ for all r such that $s_r \neq 0$ or $y_r \neq 0$. Property (i) and the proof of Theorem 15 give that (x,s,y) satisfies the inequality defined by $(\bar{\psi}, \bar{\pi}_{\alpha_0})$. Property (ii) then shows that (x,s,y) also satisfies the inequality defined by $(\bar{\psi}, \bar{\pi}_0)$. This concludes the proof. □

7.5 Computations

We performed computational tests to assess the practical consequences of Theorem 15 of Section 4.3. We have selected 63 instances from MIPLIB 2010 with binary and continuous variables only. We have generated wedge cuts $(\bar{\psi}, \bar{\pi}_\alpha)$ with $\alpha \in \{0, 0.5, 0.8, 0.85, 0.9, 0.95, 1\}$ for the most fractional binary variables in the optimal basis of the LP relaxation. For the experiments we have used the LP solver of FICO Xpress with no preprocessing at all. For each instance, and each fixed $\alpha < 1$, we generated one round of 50 + 50 wedge cuts for strengthening the LP relaxation (50 for each one of the two orientations of the cone K_α, cf. Section 3). For $\alpha = 1$, when wedge cuts are equivalent to GMI cuts by Proposition 2, we have generated one round of 50 cuts. Hence, for each instance we have a total of 7 runs of our cutting plane algorithm, and in each run we have added cuts with a single fixed α parameter only. Let $LB(I)$ and $LB^+_\alpha(I)$ denote the optimum value of the LP relaxation, and that after adding one round of cuts of parameter α, respectively, on instance I. For each instance I, and parameter α, we have computed the quantity $A_\alpha(I) = (LB^+_\alpha(I) - LB(I))/|LB(I)+1|$, where the denominator is perturbed by 1 to handle those cases with $LB(I) = 0$. Let $\bar{A}_\alpha$ denote the average of the $A_\alpha(I)$ values over the 63 instances for each α. Table 7.1 depicts the averages.

Observe that as α tends to 1 (i.e., the cuts approach GMI cuts), the improvement over the LP optimum strictly increases. Our findings complement those of Balas and Qualizza [22] which showed that in practice lopsided cuts do not improve much over GMI cuts from the tableau, although they are occasionally stronger. On average, they tend to be weaker. What these results indicate is that it does not pay to use

α	0	0.5	0.8	0.85	0.9	0.95	1 (GMI)
$\bar{A}_\alpha$	0.1305	0.131	0.136	0.139	0.141	0.144	0.146

Table 7.1: Average improvement over the LP optimum for increasing α values.

bounds on the basic variables when generating GMI cuts from tableau rows. Instead it is preferable to generate the standard GMI cut and to simply use the bounds directly in the formulation. This seems counterintuitive but it is in agreement with the proof of Theorem 15.

CHAPTER 8

Characterization of the Split Closure by Cut Generating Functions via Geometric Lifting

In this chapter we further explore properties and characterizations of split cuts. For that, we focus on the generalized corner relaxations defined in Section 2.3.

The backbone of this work is a description of split cuts for this relaxation from the perspective of cut-generating functions. This description establishes a connection between these split cuts and (a generalization of) the k-cuts introduced by Cornuéjols, Li and Vandenbussche [50]. To obtain this result, we further explore the geometric lifting idea employed in the previous chapter, illustrating its flexibility as a tool for analyzing cuts.

In this chapter we also present some implications of this result. First, we show that every split cut for a corner relaxation is the restriction of a split cut for the mixed-integer infinite relaxation; this further indicates the universality of the latter. Moreover, we use our characterization of split cuts to construct a pure-integer set (actually a pure-integer corner relaxation) that has an arbitrarily weak split closure with respect to the blow up measure introduced in Section 3.1; this result gives a pure-integer counterpart of the mixed-integer example given by Basu et al. in [27].

Organization of the chapter. We start by stating more formally our results and discussing their connections with related work. In Section 8.2 we introduce the definitions and tools used during this chapter. Section 8.3 then proves our characterization of split cuts, while Section 8.4 illustrates some implications of this result.

Acknowledgment. This chapter is joint work with Amitabh Basu.

8.1 Introduction

In the context of corner relaxations (in particular in the *group problem* setting [87, 88, 89]) a very important perspective is to study cuts as *cut-generating functions* [42, 54]. Informally, cut-generating functions allow one to compute the coefficient of a variable in a given cut using only information pertaining to this variable. Two main advantages of cut generating functions is that they usually allow a computationally efficient way of generating cuts in practice, and also possess special structure that allows one to better understand the properties of these cuts [87, 88, 89, 53, 36].

On the other hand, most cuts, split cuts included, are naturally defined from a geometric perspective (e.g. disjunctive cuts and intersection cuts). In fact, the connection between the geometric and functional perspectives on cuts has been an important tool to improve our understanding of the properties and relationship between cuts.

Despite the importance of split cuts, to the best of our knowledge, no explicit description of split closures of corner relaxations in terms of cut generating functions exists in the literature. The backbone of this work is precisely to give such a description. In fact, for 1-row corner relaxations, we establish a tight connection between split cuts and the k-*cuts* defined by Cornuéjols, Li and Vandenbussche [50]; for n-row corner relaxations, we naturally extend the definition of the latter to the so-called α-*cuts*. Informally, an α-cut for an n-row corner relaxation is obtained by taking a permissible integer vector $\alpha \in \mathbb{Z}^n$, aggregating the n rows of the problem using the α_i's as multipliers and then employing the GMI function to the resulting equality (see Section 8.2 for proper definitions).

Theorem 16. *Consider a generalized corner relaxation $\mathcal{C} = \mathcal{C}(f, R, Q)$. Then the α-cuts are exactly the split cuts for $\mathcal{C}$. More precisely:*

- *Every α-cut is a split cut for $\mathcal{C}$.*
- *Every split cut for $\mathcal{C}$ is dominated by an α-cut.*

We remark that this result is similar in spirit to the equivalence of split cuts, GMI cuts and MIR cuts (see for instance [107, 49, 56]). One might be able to recover Theorem 16 for *finite-dimensional* corner relaxations by carefully inspecting these available proofs (notice, for instance, that the multipliers α that we use form a subset of the *integral* vectors $\mathbb{Z}^n$). However, our proof, which unlike these works does not rely on linear programming duality, naturally handles the case of infinite-dimensional relaxations.

Another goal of this work is to explore the power of geometric lifting as a tool for analyzing the structure of cuts. This idea was originally introduced by Conforti, Cornuéjols and Zambelli [45] and was already used in a modified way in Chapter 7 to understand cuts for a version of the corner relaxation with extra upper/lower bounds on the variables. It is the use of this perspective that allow us to handle the infinite dimensional case seamlessly, since it allows one to translate all the work to a finite set of basic variables.

As a consequence of the generality of the above characterization, we can establish a connection between split cuts for finite-dimensional corner relaxations and split cuts for the mixed-integer infinite relaxation $\mathcal{C}(f, \mathbb{R}^n, \mathbb{R}^n)$. This result illustrates the universality of the latter: not only the mixed-integer infinite relaxation encodes all corner polyhedra, but it also preserves the structure of split cuts.

Theorem 17. *Consider a generalized corner relaxation $\mathcal{C}(f, R, Q)$.*

- *If $\mathcal{C}(f, R, Q) \neq \emptyset$ then every split cut for it is the restriction of a split cut for the infinite relaxation $\mathcal{C}(f, \mathbb{R}^n, \mathbb{R}^n)$.*
- *If $\mathcal{C}(f, R, Q) = \emptyset$ then every split cut for it is dominated by a restriction of a split cut for the infinite relaxation $\mathcal{C}(f, \mathbb{R}^n, \mathbb{R}^n)$.*

In addition, using Theorem 16 we can make the following observation about the split closure. Recall that Cook, Kannan and Schrijver constructed an example of a *mixed-integer* program with infinite split rank [47]. Reinforcing the potential weakness of the split closure, Basu et al. [27] constructed mixed-integer programs (actually projected corner relaxations) whose split closures provide an arbitrarily weak approximation to their integer hull (with respect to the *blow up* measure). On the other hand, it is known that every *pure-integer* program has finite split rank [41, 112]. However, our next result shows that in the pure-integer setting, we can still find examples where the split closure provides an arbitrarily weak approximation of the integer hull.

Theorem 18. *For every rational $\epsilon > 0$, there is a pure-integer corner relaxation $\mathcal{C}_\epsilon \subseteq \mathbb{Z}^3$ whose split closure $\mathcal{SC}_\epsilon$ has the property that*

$$\frac{\min\{y_1 + y_2 : (x,y) \in \mathcal{C}_\epsilon\}}{\min\{y_1 + y_2 : (x,y) \in \mathcal{SC}_\epsilon\}} \geq \frac{1}{12\epsilon}.$$

Corollary 6. *For a rational $\epsilon > 0$ let $\mathcal{C}_\epsilon$ and $\mathcal{SC}_\epsilon$ be as in Theorem 18. Let $\overline{\mathcal{C}}_\epsilon$ and $\overline{\mathcal{SC}}_\epsilon$ be, respectively, the projections of $\text{conv}(\mathcal{C}_\epsilon)$ and $\mathcal{SC}$ onto the y-space. Then*

$$\inf\{\alpha : \alpha\overline{\mathcal{C}}_\epsilon \supseteq \overline{\mathcal{SC}}_\epsilon\} \geq \frac{1}{12\epsilon}.$$

(Recall that for a set X we define $\alpha X = \{x/\alpha : x \in X\}$.)

8.2 Preliminaries

8.2.1 Generalized Corner Relaxation

Valid cuts. Recall from Section 2.3.1 the definition of a valid function for a generalized corner relaxation $\mathcal{C}(f, R, Q)$. Also with slight overload in notation, given sets $R' \supseteq R$ and $Q' \supseteq Q$ and functions $\psi \in \mathbb{R}^{R'}$ and $\pi \in \mathbb{R}^{Q'}$, we say that (ψ, π) is a valid cut/function for $\mathcal{C}(f, R, Q)$ if the restriction $(\psi|_R, \pi|_Q)$ is valid for it. Notice that if (ψ, π) is valid for $\mathcal{C}(f, R', Q')$, then it is valid for the restriction $\mathcal{C}(f, R, Q)$.

GMI and k-cuts. Given a real number $a \in \mathbb{R}$, let $[a]$ denote its fractional part $a - \lfloor a \rfloor$. Then given $f \in \mathbb{R} \setminus \mathbb{Z}$, the *GMI function* $(\psi^f_{\text{GMI}}, \pi^f_{\text{GMI}})$ is defined as

$$\psi^f_{\text{GMI}}(r) = \max\left\{\frac{r}{1 - [f]}, -\frac{r}{[f]}\right\} \tag{8.1}$$

$$\pi^f_{\text{GMI}}(q) = \max\left\{\frac{q - \lfloor q + [f] \rfloor}{1 - [f]}, -\frac{q - \lfloor q + [f] \rfloor}{[f]}\right\}. \tag{8.2}$$

The GMI function is valid for $\mathcal{C}(f, \mathbb{R}, \mathbb{R})$.

For $f \in [0,1]^n \setminus \mathbb{Z}^n$, define the sets $\mathbb{Z}_f = \{w \in \mathbb{Z}^n : wf \notin \mathbb{Z}\}$ and $\mathbb{Z}^+_f = \mathbb{Z}_f \cap \{w : \sum_{i=1}^n w_i \geq 0\}$. Given $\alpha \in \mathbb{Z}^+_f$, the *$\alpha$-cut function* $(\psi^f_\alpha, \pi^f_\alpha)$ is defined as

$$\psi^f_\alpha(r) = \psi^{\alpha f}_{\text{GMI}}(\alpha r) \tag{8.3}$$

$$\pi^f_\alpha(q) = \pi^{\alpha f}_{\text{GMI}}(\alpha q). \tag{8.4}$$

In the case $n = 1$ this definition is a cut-generating function view of the *k-cuts* introduced in [50].

Split cuts. Consider an n-dimensional corner relaxation $\mathcal{C}(f,R,Q)$. Given $\alpha \in \mathbb{Z}_f$ and $\beta : Q \to \mathbb{Z}$, we define the split disjunction

$$D(\alpha,\beta,f) \triangleq \left\{(x,s,y) : \alpha x + \sum_{q\in Q} \beta(q)y(q) \le \lfloor \alpha f \rfloor \right\} \cup \left\{(x,s,y) : \alpha x + \sum_{q\in Q} \beta(q)y(q) \ge \lceil \alpha f \rceil \right\}.$$

A cut (ψ,π) is a *split cut* for $\mathcal{C}(f,R,Q)$ with respect to the disjunction $D(\alpha,\beta,f)$ if it is satisfied by all points in the set $\mathcal{C}_{LP}(f,R,Q) \cap D(\alpha,\beta,f)$.[1]

Notice that this definition holds for the case where R and/or Q are infinite, thus generalizing the standard definition of split cuts to this setting; this further generalizes the definition used in Chapter 7.

8.2.2 Continuous Corner Relaxation

It will be convenient to define a slightly modified version of the continuous relaxation $\mathcal{C}(f,R,\emptyset)$ which allows up to two copies of vectors in R. For that, define the continuous corner relaxation $\mathcal{CC}(f,R,Q)$:

$$\begin{aligned} & x = f + \sum_{r\in R} r\cdot s(r) + \sum_{q\in Q} q\cdot y(q) \\ & x \in \mathbb{Z}^n \\ & s \in \mathbb{R}_+^{\{R\}}, y \in \mathbb{R}_+^{\{Q\}} \end{aligned} \qquad (\mathcal{CC}(f,R,Q))$$

Valid cuts and lattice-free cuts. Recall the notion of lattice-free cuts from Section 2.3.4; we adapt it slightly for our setting.

Given a convex lattice-free set $S \subseteq \mathbb{R}^n$ containing f in its interior, we define the function (ψ_S,π_S) by setting

$$\psi_S = \pi_S = \gamma_{S-f} \qquad (8.5)$$

where γ is the gauge or Minkowski functional from Definition 2. The function (ψ_S,π_S) is valid for $\mathcal{CC}(f,\mathbb{R}^n,\mathbb{R}^n)$. We call these *lattice-free* cuts.

[1]Notice that we only consider α in $\mathbb{Z}_f$, and not in $\mathbb{Z}^n$, because if $\alpha \in \mathbb{Z}^n \setminus \mathbb{Z}_f$ we have $\mathcal{C}_{LP}(f,R,Q) \cap D(\alpha,\beta,f) = \mathcal{C}_{LP}(f,R,Q)$.

Split cuts. As in the previous section, we define split cuts for $\mathcal{CC}(f, R, Q)$. For that, given $\alpha \in \mathbb{Z}_f$, define the split disjunction

$$\bar{D}(\alpha, f) = \{(x, s, y) : \alpha x \leq \lfloor \alpha f \rfloor\} \cup \{(x, s, y) : \alpha x \geq \lceil \alpha f \rceil\}.$$

The cut (ψ, π) is a *split cut* for $\mathcal{CC}(f, R, Q)$ with respect to the disjunction $\bar{D}(\alpha, f)$ if it is satisfied by all points in $\mathcal{CC}_{LP}(f, R, Q) \cap \bar{D}(\alpha, f)$.

Given an integer vector $\alpha \in \mathbb{Z}^n$, we define the *lattice-free split* $S(\alpha, f)$ as

$$S(\alpha, f) = \{x \in \mathbb{R}^n : \lfloor \alpha f \rfloor \leq \alpha x \leq \lceil \alpha f \rceil\}.$$

The set $S(\alpha, f)$ is convex and lattice-free, and notice that whenever $\alpha \in \mathbb{Z}_f$ it contains f in its interior; in this case, $(\psi_{S(\alpha,f)}, \pi_{S(\alpha,f)})$ is a valid cut for $\mathcal{CC}(f, R, Q)$. We call this cut a *lattice-free split cut.*

8.2.3 Equivalence of Split Cuts and Lattice-free Split Cuts

We now show that the notion of a split cut and lattice-free split cut are equivalent (see Theorem 19).

Theorem 19.

$$\mathcal{CC}_{LP}(f, R, Q) \cap (\psi_{S(\alpha,f)}, \pi_{S(\alpha,f)}) = \text{conv}\left(\mathcal{CC}_{LP}(f, R, Q) \cap \bar{D}(\alpha, f)\right).$$

In other words, for a continuous corner relaxation, every split cut is a lattice-free split cut and vice versa.

Proof.

Claim 8. *Let $S(\alpha, f)$ be a lattice-free split for some $\alpha \in \mathbb{Z}_f$. Let r^* be such that $\psi_{S(\alpha,f)}(r^*) = \pi_{S(\alpha,f)}(r^*) = 0$. Then a point $\tilde{x} \in \mathbb{R}^n$ belongs to $\text{int}(S(\alpha, f))$ if and only if $\tilde{x} + \lambda r^*$ belongs to $\text{int}(S(\alpha, f))$ for all $\lambda \in \mathbb{R}$.*

Proof. By definition $S(\alpha, f) = \{x : \lfloor \alpha f \rfloor \leq \alpha x \leq \lceil \alpha f \rceil\}$. By definition of the Minkowski functional, $\psi_{S(\alpha,f)}(r^*) = 0$ implies that r^* is in the recession cone of $S(\alpha, f)$. Thus, $\alpha r^* = 0$. Then $\tilde{x}$ belongs to $\text{int}(S(\alpha, f))$ iff $\lfloor \alpha f \rfloor < \alpha \tilde{x} < \lceil \alpha f \rceil$, which happens iff $\lfloor \alpha f \rfloor < \alpha(\tilde{x} + \lambda r^*) < \lceil \alpha f \rceil$ for all $\lambda \in \mathbb{R}$, which is equivalent to $(\tilde{x} + \lambda r^*) \in \text{int}(S(\alpha, f))$ for all $\lambda \in \mathbb{R}$. □

For ease of notation, we denote $\mathcal{CC}_{LP} = \mathcal{CC}_{LP}(f, R, Q)$ and $S = S(\alpha, f)$ in the remainder of this proof.

($\supseteq$) Since $\mathcal{CC}_{LP} \cap (\psi_S, \pi_S)$ is convex, it suffices to show that $\mathcal{CC}_{LP} \cap \bar{D}(\alpha, f) \subseteq \mathcal{CC}_{LP} \cap (\psi_S, \pi_S)$; so it suffices to show that $\mathcal{CC}_{LP} \cap \bar{D}(\alpha, f)$ satisfies (ψ_S, π_S).

Take $(\bar{x}, \bar{s}, \bar{y}) \in \mathcal{CC}_{LP} \cap \bar{D}(\alpha, f)$. Suppose by contradiction that

$$\sum_{r \in R} \bar{s}(r)\psi_S(r) + \sum_{q \in Q} \bar{y}(q)\pi_S(q) < 1. \tag{8.6}$$

Define $R' = \{r \in R : \psi_S(r) > 0\}$ and $Q' = \{q \in Q : \pi_S(q) > 0\}$. Also define $\lambda_r = \bar{s}(r)\psi_S(r)$ and $\mu_q = \bar{y}(q)\pi_S(q)$, and let $\nu = \sum_{r \in R'} \lambda_r + \sum_{q \in Q'} \mu_q$; notice that $\nu \in [0, 1)$ because of (8.6).

Notice that we can write $\bar{x}$ as

$$\bar{x} = f + \sum_{r \in R'} \lambda_r \left(\frac{r}{\psi_S(r)} \right) + \sum_{q \in Q'} \mu_q \left(\frac{q}{\pi_S(q)} \right) + \Delta',$$

where the leftover Δ' equals $\sum_{r \in R \setminus R'} \bar{s}(r)r + \sum_{q \in Q \setminus Q'} \bar{y}(q)q$. Rewriting a bit more, we have

$$\bar{x} = (1 - \nu)f + \nu \left[\sum_{r \in R'} \frac{\lambda_r}{\nu} \left(f + \frac{r}{\psi_S(r)} \right) + \sum_{q \in Q'} \frac{\mu_q}{\nu} \left(f + \frac{q}{\pi_S(q)} \right) \right] + \Delta'. \tag{8.7}$$

Now notice that for all $r \in R'$, the vector $f + \frac{r}{\psi_S(r)}$ belongs to S, and for all $q \in Q'$ the vector $f + \frac{q}{\pi_S(q)}$ belongs to S. Since the bracketed term in the above displayed expression is a convex combination of these vectors, it also belongs to S. Moreover, f belongs to the *interior* of S, and hence (going back to equation (8.7)) we have that $\bar{x} - \Delta'$ belongs to the interior of S.

To conclude the proof of this part, we can apply the Claim 8 repeatedly using each of the terms in the definition of Δ' to obtain that $\bar{x} = (\bar{x} - \Delta') + \Delta'$ also belongs to the interior of S (recall that the sum defining Δ' is a finite sum becaue $\bar{s}$ and $\bar{y}$ have finite support). But this implies that the solution $(\bar{x}, \bar{s}, \bar{y})$ does not belong to $\mathcal{CC}_{LP} \cap \bar{D}(\alpha, f)$, which raises a contradiction.

($\subseteq$) Take a point $(\bar{x}, \bar{s}, \bar{y}) \in \mathcal{CC}_{LP} \cap (\psi_S, \pi_S)$; we show that it belongs to $\text{conv}(\mathcal{CC}_{LP} \cap \bar{D}(\alpha, f))$.

Define R', Q', λ_r, μ_q and ν as before; now notice that since $(\bar{x}, \bar{s}, \bar{y})$ satisfies (ψ_S, π_S), we have $\nu = \sum_{r \in R'} \lambda_r + \sum_{q \in Q'} \mu_q = \sum_{r \in R'} \bar{s}(r)\psi_S(r) + \sum_{q \in Q'} \bar{y}(q)\pi_S(q) =$

$\sum_{r \in R} \bar{s}(r)\psi_S(r) + \sum_{q \in Q} \bar{y}(q)\pi_S(q) \geq 1$ (the third equality follows from the fact that $\psi_S(r) = 0$ for all $r \in R \setminus R'$ and $\pi_S(r) = 0$ for all $r \in Q \setminus Q'$, and the inequality follows from the fact that $(\bar{x}, \bar{s}, \bar{y})$ satisfies (ψ_S, π_S)). For every $r \in R'$, define the point

$$U^r = \left(f + \frac{\nu r}{\psi_S(r)}, \left(\frac{\nu}{\psi_S(r)}\right) \cdot 1_r, 0\right)$$

where 1_r denotes the indicator function of r. Let $u^r = f + \frac{\nu r}{\psi_S(r)}$ be its first component; similarly, for every $q \in Q'$ define

$$V^q = \left(f + \frac{\nu q}{\pi_S(q)}, 0, \left(\frac{\nu}{\pi_S(q)}\right) \cdot 1_q\right)$$

and let $v^q = f + \frac{\nu q}{\pi_S(q)}$ be its first component.

Notice that all U^r and V^q belong to $\mathcal{CC}_{LP}$; moreover, using the definition of ψ_S and π_S and the fact that $\nu \geq 1$, we have that $u^r \notin \text{int}(S)$ and $v^q \notin \text{int}(S)$ for all $r \in R'$ and $q \in Q'$. Therefore, the points U^r and V^q belong to $\mathcal{CC}_{LP} \cap \bar{D}(\alpha, f)$.

Now define

$$\Delta = \left(\sum_{r \in R \setminus R'} \bar{s}(r) r + \sum_{q \in Q \setminus Q'} \bar{y}(q) q, \quad \bar{s}|_{R \setminus R'}, \quad \bar{y}|_{Q \setminus Q'}\right).$$

Again using Claim 8 repeatedly, we get that $U^r + \Delta \in \mathcal{CC}_{LP} \cap \bar{D}(\alpha, f)$ and $V^q + \Delta \in \mathcal{CC}_{LP} \cap \bar{D}(\alpha, f)$.

Now notice that we can write $(\bar{x}, \bar{s}, \bar{y})$ as follows:

$$(\bar{x}, \bar{s}, \bar{y}) = \sum_{r \in R'} \frac{\lambda_r}{\nu}(U^r + \Delta) + \sum_{q \in Q'} \frac{\mu_q}{\nu}(V^q + \Delta). \tag{8.8}$$

It then follows that $(\bar{x}, \bar{s}, \bar{y})$ belongs to $\text{conv}(\mathcal{CC}_{LP} \cap \bar{D}(\alpha, f))$ as desired (notice that the λ_r's and μ_q's are finitely supported, because they come from $\bar{s}$ and $\bar{y}$).

□

Since we will work mostly with lattice-free split cuts, to simplify the notation define $\bar{S}(\alpha, f) = S(\alpha, f) - f$ as the "centered" version of $S(\alpha, f)$, so that $\psi_{S(\alpha,f)} = \pi_{S(\alpha,f)} = \gamma_{\bar{S}(\alpha,f)}$.

8.2.4 Geometric Lifting

The idea behind lifting of cuts is to first obtain a (weaker) cut by ignoring some of the integrality of the variables and then strengthen it by incorporating back some of the discarded information. In our context, we can formally define lifting as follows: given a cut (ψ, π) valid for $\mathcal{CC}(f, R, Q)$, we say that a cut (ψ', π') valid for $\mathcal{C}(f, R, Q)$ is a *lifting* of (ψ, π) if $\psi' = \psi$ and $\pi' \leq \pi$. For example, the GMI function introduced in Section 8.2.1 can be seen as a lifting of a lattice-free cuts, a fact that will be very important for our developments.

Fact 1 ([45]). *The GMI function* $(\psi^f_{GMI}, \pi^f_{GMI})$ *is a lifting of the lattice-free cut* $(\psi_{S(1,f)}, \pi_{S(1,f)})$. *More precisely,*

$$\psi^f_{GMI}(r) = \psi_{S(1,f)}(r), \qquad \pi^f_{GMI}(q) = \min_{w \in \mathbb{Z}} \pi_{S(1,f)}(q + w).$$

Conforti, Cornuéjols and Zambelli considered a geometric approach to lifting [45]. There they focus on lifting/improving one coefficient at a time; for that, they choose different lattice-free bodies for each coefficient. One drawback is that one needs to show that the cut obtained is valid.

We take a different perspective on geometric lifting, which is the same used in Chapter 7: we lift all coefficients simultaneously in an independent but coordinated manner, guaranteeing that the cut we end up with is always valid. The idea is to relax the integrality of all the non-basic variables in $\mathcal{C}(f, R, Q)$ and capture all of them in an aggregated way by introducing an additional equality to the system. More precisely, given an integral function $\ell : Q \to \mathbb{Z}$ we consider the program

$$\begin{aligned} \tbinom{x}{x_{n+1}} &= \tbinom{f}{0} + \textstyle\sum_{r \in R} \tbinom{r}{0} \cdot s\tbinom{r}{0} + \textstyle\sum_{q \in Q} \tbinom{q}{\ell(q)} \cdot y\tbinom{q}{\ell(q)} \\ x &\in \mathbb{Z}^{n+1} \\ z &\in \mathbb{Z} \\ s &\in \mathbb{R}^{\{R^0\}}_+, q \in \mathbb{R}^{\{Q^\ell\}}_+ \end{aligned} \qquad (\mathcal{CC}(f^0, R^0, Q^\ell))$$

where we define $f^0 = (f, 0)$, $R^0 = \{(r, 0) : r \in R\}$ and $Q^\ell = \{(q, \ell(q)) : q \in Q\}$.

Intuitively, $\mathcal{CC}(f^0, R^0, Q^\ell)$ is a relaxation of $\mathcal{C}(f, R, Q)$ and hence valid cuts for the former should be valid for the latter. Indeed, it is easy to check that given a lattice-free set $S \subseteq \mathbb{R}^{n+1}$ containing $(f, 0)$ in its interior, the cut (ψ^+_S, π^+_S) given by

$$\psi^+_S(r) = \gamma_{S-f}\begin{pmatrix} r \\ 0 \end{pmatrix} \quad \forall r \in R, \quad \pi^+_{S,\ell}(q) = \gamma_{S-f}\begin{pmatrix} q \\ \ell(q) \end{pmatrix} \quad \forall q \in Q$$

is valid for $\mathcal{C}(f, R, Q)$. Moreover, keeping the set $S \subseteq \mathbb{R}^{n+1}$ fixed, we can also look for the strongest possible cuts we can obtain by choosing different ℓ's; more precisely, the function $(\tilde{\psi}_S, \tilde{\pi}_S)$ given by

$$\tilde{\psi}_S(r) = \psi_S^+(r) = \gamma_{S-f}\begin{pmatrix} r \\ 0 \end{pmatrix} \quad \forall r \in R, \tag{8.9}$$

$$\tilde{\pi}_S(q) = \inf\{\pi_{S,\ell}^+(q) \mid \ell : Q \to \mathbb{Z}\} = \inf\left\{\gamma_{S-f}\begin{pmatrix} q \\ \ell(q) \end{pmatrix} : \ell(q) \in \mathbb{Z}\right\} \quad \forall q \in Q \tag{8.10}$$

is valid for $\mathcal{C}(f, R, Q)$.

Lemma 50. *Let $S = S(\alpha, f) \subseteq \mathbb{R}^n$ be a lattice-free split set for some $\alpha \in \mathbb{Z}_f$. Define $\pi^*(r) = \inf\{\gamma_{S-f}(r + w) : w \in \mathbb{Z}^n\}$. Then, there exists $w^* \in \mathbb{Z}^n$ such that $\pi^*(r) = \gamma_{S-f}(r + w^*)$.*

Proof. It can be verified that

$$\gamma_{S-f}(r) = \max\left\{\frac{\alpha r}{\lceil \alpha f \rceil}, \frac{-\alpha r}{\alpha f - \lfloor \alpha f \rfloor}\right\} \triangleq \max\left\{c_1 \alpha r, -c_2 \alpha r\right\},$$

see for instance [45].

For any $w \in \mathbb{Z}^n$, there exists $\lambda \in \mathbb{R}$ and $v \in \mathbb{R}^n$ such that $w = \lambda\alpha + v$ and $\alpha v = 0$. Thus, $\gamma_{S-f}(r + w) = \max\{c_1 \alpha r + c_1 \lambda \|\alpha\|^2, -c_2 \alpha r - c_2 \lambda \|\alpha\|^2\}$. Let $\Lambda = \{\lambda \in \mathbb{R} : \exists v \in \mathbb{R}^n \text{ such that } \lambda\alpha + v \in \mathbb{Z}^n, \alpha v = 0\}$ be the projection of the lattice $\mathbb{Z}^n$ onto the subspace spanned by α; since α is rational, we have that Λ is a lattice [24]. With this definition, we have

$$\pi^*(r) = \inf\{\gamma_{S-f}(r + w) : w \in \mathbb{Z}^n\} = \inf\left\{\max\{c_1 \alpha r + c_1 \lambda \|\alpha\|^2, -c_2 \alpha r - c_2 \lambda \|\alpha\|^2\} : \lambda \in \Lambda\right\}. \tag{8.11}$$

Now $\omega(\lambda) \triangleq \max\{c_1 \alpha r + c_1 \lambda \|\alpha\|^2, -c_2 \alpha r - c_2 \lambda \|\alpha\|^2\}$ is a piecewise linear convex function in λ with two pieces, and Λ is a lattice. Moreover, $\omega(\lambda) \geq 0$ for all $\lambda \in \mathbb{R}$. Thus, the infimum in the right-hand side of (8.11) is actually achieved. This proves the lemma. □

8.3 Characterization of Split Cuts

8.3.1 Split Cuts as Lifted Lattice-free-Split Cuts

In this section we show that every split cut for $\mathcal{C}(f, R, Q)$ can be obtained via geometric lifting from a cut derived from a lattice-free split set.

As hinted in the previous section, the LP relaxations $\mathcal{CC}_{LP}(f^0, R^0, Q^\ell)$ and $\mathcal{C}_{LP}(f, R, Q)$ are isomorphic. More explicitly, define the functional $\Gamma_\ell : \mathcal{C}_{LP}(f, R, Q) \to \mathcal{CC}_{LP}(f^0, R^0, Q^\ell)$ which maps a solution (x, s, y) into (x', s', y') by setting

$$\begin{aligned} x'_i &= x_i, \quad i = 1, 2, \ldots, n \\ x'_{n+1} &= \sum_{q \in Q} \ell(q) y(q) \\ s'(r, 0) &= s(r) \quad \text{for all } r \in R \\ y'(q, \ell(q)) &= y(q) \quad \text{for all } q \in Q. \end{aligned}$$

It is easy to see that Γ is bijective.

Fact 2. *The tuple (x, s, y) is a solution to $\mathcal{C}_{LP}(f, R, Q)$ iff $\Gamma(x, s, y)$ is a solution to $\mathcal{CC}_{LP}(f^0, R^0, Q^\ell)$.*

This isomorphism naturally induces an isomorphism in the cuts between these programs: define the functional $\Gamma^\circ_\ell : \mathbb{R}^R \times \mathbb{R}^Q \to \mathbb{R}^{R^0} \times \mathbb{R}^{Q^\ell}$ which maps (ψ, π) into (ψ^1, π^1) by setting

$$\psi^1(r, 0) = \psi(r) \quad \forall r \in R, \qquad \pi^1(q, \ell(q)) = \pi(q) \quad \forall q \in Q.$$

This map is also a bijection, and we have

Fact 3. *The tuple (x, s, y) satisfies (ψ, π) if and only if $\Gamma_\ell(x, s, y)$ satisfies $\Gamma^\circ_\ell(\psi, \pi)$.*

Fact 3 implies

Lemma 51. *Let (ψ, π) and (ψ', π') be valid cuts for $\mathcal{C}(f, R, Q)$. Then*

$$\mathcal{C}_{LP}(f, R, Q) \cap (\psi, \pi) \subseteq \mathcal{C}_{LP}(f, R, Q) \cap (\psi', \pi')$$

iff

$$\mathcal{CC}_{LP}(f^0, R^0, Q^\ell) \cap \Gamma^\circ_\ell(\psi, \pi) \subseteq \mathcal{CC}_{LP}(f^0, R^0, Q^\ell) \cap \Gamma^\circ_\ell(\psi', \pi').$$

In other words, (ψ, π) dominates (ψ', π') with respect to $\mathcal{C}_{LP}(f, R, Q)$ iff the cut $\Gamma^\circ_\ell(\psi, \pi)$ dominates $\Gamma^\circ_\ell(\psi', \pi')$ with respect to $\mathcal{CC}_{LP}(f^0, R^0, Q^\ell)$.

Moreover,

Fact 4. *Let $S \subseteq \mathbb{R}^{n+1}$ be a lattice-free set containing $(f, 0)$ in its interior. Then for every integer ℓ, $(\psi_S^+, \pi_{S,\ell}^+) = (\Gamma_\ell^\circ)^{-1}(\psi_S, \pi_S)$.*

The heart of the argument for the main result of this section is the following lemma, which establishes the equivalence between split cuts for $\mathcal{C}(f, R, Q)$ and $\mathcal{CC}(f^0, R^0, Q^\ell)$; the idea is that we can simulate any split disjunction for the former by setting the value of ℓ appropriately and using a disjunction on the latter program which utilizes the new constraint $x_{n+1} = \sum_{q \in Q} \ell(q) \cdot y(q)$.

Lemma 52. *The cut (ψ, π) is a split cut for $\mathcal{C}(f, R, Q)$ with respect to the disjunction $D(\alpha, \beta, f)$ iff $\Gamma_\beta^\circ(\psi, \pi)$ is a split cut for $\mathcal{CC}(f^0, R^0, Q^\beta)$ with respect to the disjunction $\bar{D}((\alpha, 1), (f, 0))$.*

Proof. Observe that $(x, s, y) \in D(\alpha, \beta, f)$ if and only if $\Gamma_\beta(x, s, y) \in \bar{D}((\alpha, 1), (f, 0))$. From Fact 2, $(x, s, y) \in \mathcal{C}_{LP}(f, R, Q)$ if and only if $\Gamma_\beta(x, s, y) \in \mathcal{CC}_{LP}(f^0, R^0, Q^\beta)$. Thus, $(x, s, y) \in \mathcal{C}_{LP}(f, R, Q) \cap D(\alpha, \beta, f)$ if and only if $\Gamma_\beta(x, s, y) \in \mathcal{CC}_{LP}(f^0, R^0, Q^\beta) \cap \bar{D}((\alpha, 1), (f, 0))$. By Fact 3, (ψ, π) is valid for $\mathcal{C}_{LP}(f, R, Q) \cap D(\alpha, \beta, f)$ if and only if $\Gamma_\beta^\circ(\psi, \pi)$ is valid for $\mathcal{CC}_{LP}(f^0, R^0, Q^\beta) \cap \bar{D}((\alpha, 1), (f, 0))$, proving the result. ☐

Now we are ready to prove the main result of this section (recall the definition of $(\tilde{\psi}_S, \tilde{\pi}_S)$ from equations (8.9) and (8.10)).

Lemma 53. *A valid cut (ψ, π) for $\mathcal{C}(f, R, Q)$ is a split cut iff there is $\alpha \in \mathbb{Z}_f$ such that (ψ, π) is dominated by the cut $(\tilde{\psi}_S, \tilde{\pi}_S)$ with respect to $\mathcal{C}_{LP}(f, R, Q)$, where $S = S((\alpha, 1), (f, 0))$.*

Proof. ($\Rightarrow$) Suppose that (ψ, π) is a split cut for $\mathcal{C}(f, R, Q)$ with respect to the disjunction $D(\alpha, \beta, f)$. By Lemma 52 we get that $(\psi', \pi') = \Gamma_\beta^\circ(\psi, \pi)$ is a split cut for $\mathcal{CC}(f^0, R^0, Q^\ell)$ with respect to the disjunction $\bar{D}((\alpha, 1), (f, 0))$. From Theorem 19, we get that (ψ', π') is dominated by (ψ_S, π_S) with respect to $\mathcal{CC}_{LP}(f^0, R^0, Q^\ell)$. From Lemma 51 we then have that (ψ, π) is dominated by $(\Gamma_\beta^\circ)^{-1}(\psi_S, \pi_S) = (\psi_S^+, \pi_{S,\beta}^+)$ with respect to $\mathcal{C}_{LP}(f, R, Q)$, were the equality follows from Fact 4. Since $(\tilde{\psi}_S, \tilde{\pi}_S)$ is at least as strong as the latter cut, we have the first part of the lemma.

($\Leftarrow$) By Lemma 50, $(\tilde{\psi}_S, \tilde{\pi}_S) = (\psi_S^+, \pi_{S,\ell}^+)$ for some $\ell : Q \to \mathbb{R}$. Thus, it suffices to show that for all $\ell : Q \to \mathbb{Z}$ the cut $(\psi_S^+, \pi_{S,\ell}^+)$ is a split cut for $\mathcal{C}(f, R, Q)$. First notice that Theorem 19 implies that (ψ_S, π_S) is a split cut for $\mathcal{CC}(f^0, R^0, Q^\ell)$. Fact 4 says $(\psi_S^+, \pi_{S,\ell}^+) = (\Gamma_\ell^\circ)^{-1}(\psi_S, \pi_S)$; Lemma 52 then gives that $(\psi_{S,\ell}^+, \pi_S^+)$ is a split cut for $\mathcal{C}(f, R, Q)$, thus concluding the proof of the lemma. ☐

Remark 2 (Lifting is complete for splits)**.** *Consider a split set $S = S((\alpha, 1), (f, 0))$ and let $S' = \{x \in \mathbb{R}^n : (x, 0) \in S\}$ be the lower-dimensional set corresponding to the slice of S along the plane $x_{n+1} = 0$. Notice that $\tilde{\psi}_S(r) = \psi_S\binom{r}{0} = \psi_{S'}(r)$ for all $r \in R$ and similarly $\tilde{\pi}_S(q) \leq \pi_{S'}(q)$. The above lemma then shows that every split cut for $\mathcal{C}(f, R, Q)$ is (dominated by) a lifting of a lattice-free split cut for the relaxation $\mathcal{CC}(f, R, Q)$ coming from a (possibly non-maximal) split set $S' \subseteq \mathbb{R}^n$.*

8.3.2 Linear Transformation of Lattice-free Cuts

An important tool to connect the GMI and α-cuts with lattice-free split cuts for $\mathcal{CC}(f^0, R^0, Q^\ell)$ is to understand how the latter change when we apply a linear transformation to the split sets used to generate them.

The first observation follows directly from the definition of γ.

Claim 9. *Consider a convex set $S \subseteq \mathbb{R}^n$ with the origin in its interior. Then for every linear transformation $A : \mathbb{R}^m \to \mathbb{R}^n$ we have*

$$\gamma_S(Ar) = \gamma_{A^{-1}(S)}(r) \quad \forall r \in \mathbb{R}^m.$$

Proof. Using linearity of A we get that

$$\begin{aligned}\gamma_S(Ar) &= \inf\left\{\frac{1}{\lambda} > 0 : \lambda(Ar) \in S\right\} = \inf\left\{\frac{1}{\lambda} > 0 : A(\lambda r) \in S\right\} \\ &= \inf\left\{\frac{1}{\lambda} > 0 : \lambda r \in A^{-1}(S)\right\} = \gamma_{A^{-1}(S)}(r).\end{aligned}$$

□

Now we see how the description of a centered split $\bar{S}(u, f)$ changes when we apply a linear transformation to this set. To make this formal, given a linear map $A : \mathbb{R}^m \to \mathbb{R}^n$, we use $A^t : \mathbb{R}^n \to \mathbb{R}^m$ to denote the *adjoint* of A (i.e., the unique linear map satisfying $A(u) \cdot v = u \cdot A^t(v)$ for all $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$); recall that (fixing orthonormal bases for $\mathbb{R}^m$ and $\mathbb{R}^n$), the matrix corresponding to A^t is just the transpose of the matrix corresponding to A.

Lemma 54. *Consider a centered split set $\bar{S}(u, f)$ with $u \in \mathbb{Z}^n$ and $f \in \mathbb{R}^n$, and a linear transformation $A : \mathbb{R}^m \to \mathbb{R}^n$. Then for every vector $f' \in A^{-1}(f)$, we have*

$$A^{-1}(\bar{S}(u, f)) = \bar{S}(A^t(u), f').$$

Proof. Assume that $A^{-1}(f)$ is non-empty, otherwise there is nothing to prove, and consider $f' \in A^{-1}(f)$. Recall that $\bar{S}(u,f) = S(u,f) - f = \{x : \lfloor uf \rfloor \leq u(x+f) \leq \lceil uf \rceil\}$. Then we have

$$\begin{aligned} A^{-1}(\bar{S}(u,f)) &= \{x \in \mathbb{R}^m : \lfloor uf \rfloor \leq u \cdot (A(x) + f) \leq \lceil uf \rceil\} \\ &= \{x \in \mathbb{R}^m : \lfloor u \cdot A(f') \rfloor \leq u \cdot A(x + f') \leq \lceil u \cdot A(f') \rceil\} \\ &= \{x \in \mathbb{R}^m : \lfloor A^t(u) \cdot f' \rfloor \leq A^t(u) \cdot (x + f') \leq \lceil A^t(u) \cdot f' \rceil\} \\ &= \bar{S}(A^t(u), f'). \end{aligned}$$

□

Employing Claim 9 and Lemma 54, we directly get the following.

Corollary 7. *Consider a linear transformation $A : \mathbb{R}^m \to \mathbb{R}^n$. Then for every centered split set $\bar{S}(u,f) \subseteq \mathbb{R}^n$ and vector $f' \in A^{-1}(f)$, we have*

$$\gamma_{\bar{S}(u,f)}(Ar) = \gamma_{\bar{S}(A^t(u),f')}(r) \quad \forall r \in \mathbb{R}^n.$$

The next corollary states that we can use the geometric lifting to simulate the "trivial lifting" of a cut given by an interval, which is very handy given that we can construct the GMI function this way (recall Fact 1).

Corollary 8. *For every centered split set $\bar{S}(u,f) \subseteq \mathbb{R}^1$, we have*

$$\gamma_{\bar{S}(u,f)}(r + w) = \gamma_{\bar{S}((u,u),(f,0))}\begin{pmatrix} r \\ w \end{pmatrix} \qquad \forall r, w \in \mathbb{R}.$$

Proof. Employ the previous corollary with the linear map $(x_1, x_2) \mapsto x_1 + x_2$, whose adjoint is given by $x_1 \mapsto (x_1, x_1)$, and with $f' = (f, 0)$. □

In combination with the above result, the next corollary is useful to connect the definition of α-cuts with our geometric lifting.

Corollary 9. *Consider $f \in \mathbb{R}^n$ and an integral vector $\alpha \in \mathbb{Z}_f$. Then*

$$\gamma_{\bar{S}((1,1),(\alpha f,0))}\begin{pmatrix} \alpha q \\ w \end{pmatrix} = \gamma_{\bar{S}((\alpha,1),(f,0))}\begin{pmatrix} q \\ w \end{pmatrix} \qquad \forall q \in \mathbb{R}^n, w \in \mathbb{R}.$$

Proof. Employ Corollary 7 with the linear map $(x_1, x_2, \ldots, x_{n+1}) \mapsto (\sum_{i=1}^n \alpha_i x_i, x_{n+1})$, whose adjoint is given by $(x_1, x_2) \mapsto (\alpha_1 x_1, \alpha_2 x_1, \ldots, \alpha_n x_1, x_2)$, and with $f' = (f, 0)$. □

8.3.3 α-cuts as Lifted Lattice-free Split Cuts

Now we use the tools from the previous section to show that α-cuts for $\mathcal{C}(f, R, Q)$ are liftings of lattice-free split cuts for $\mathcal{CC}(f, R, Q)$. Notice the similarity of the statement below and Lemma 53, which shows that every split cut for $\mathcal{C}(f, R, Q)$ is a lifting of a lattice-free split cut for $\mathcal{CC}(f, R, Q)$; in fact, it is through these liftings of lattice-free split cuts that we can show the equivalence of α-cuts and split cuts for $\mathcal{C}(f, R, Q)$ in the sequel. 5

Lemma 55. *For every $\alpha \in \mathbb{Z}_f^+$, the α-cut can be obtained using the geometric lifting procedure. More precisely, defining $S = S((\alpha, 1), (f, 0))$ we have*

$$(\psi_\alpha^f, \pi_\alpha^f) = (\tilde{\psi}_S, \tilde{\pi}_S).$$

Proof. Recall that from the definition of α-cuts and Fact 1 that $\psi_\alpha^f(r) = \psi_{\text{GMI}}^{\alpha f}(\alpha r) = \gamma_{\bar{S}(1,\alpha f)}(\alpha r)$. Then using Corollary 8 with $u = 1$ and $w = 0$ and Corollary 9 with $w = 0$, we get that

$$\psi_\alpha^f(r) = \gamma_{\bar{S}(1,\alpha f)}(\alpha r) \stackrel{\text{Cor 8}}{=} \gamma_{\bar{S}((1,1),(\alpha f,0))}\begin{pmatrix}\alpha r\\ 0\end{pmatrix} \stackrel{\text{Cor 9}}{=} \gamma_{\bar{S}((\alpha,1),(f,0))}\begin{pmatrix}r\\ 0\end{pmatrix} = \tilde{\psi}_{S((\alpha,1),(f,0))}(r).$$

Similarly, recall that

$$\pi_\alpha^f(q) = \pi_{\text{GMI}}^{\alpha f}(\alpha q) = \min_{w\in\mathbb{Z}} \psi_{\text{GMI}}^{\alpha f}(\alpha q + w) = \min_{w\in\mathbb{Z}} \gamma_{\bar{S}(1,\alpha f)}(\alpha q + w).$$

Then employing Corollaries 8 and 9 we get that

$$\begin{aligned}\pi_\alpha^f(q) &= \min_{w\in\mathbb{Z}} \gamma_{\bar{S}(1,\alpha f)}(\alpha q + w) \stackrel{\text{Cor 8}}{=} \min_{w\in\mathbb{Z}} \gamma_{\bar{S}((1,1),(\alpha f,0))}\begin{pmatrix}\alpha q\\ w\end{pmatrix}\\ &\stackrel{\text{Cor 9}}{=} \min_{w\in\mathbb{Z}} \gamma_{\bar{S}((\alpha,1),(f,0))}\begin{pmatrix}q\\ w\end{pmatrix} = \tilde{\pi}_{S((\alpha,1),(f,0))}(q),\end{aligned}$$

concluding the proof of the lemma. □

8.3.4 Concluding the Proof of Theorem 16

Notice that Lemmas 53 and 55 almost give us Theorem 16: the difference is that in the former the characterization of split cuts uses vectors α in $\mathbb{Z}_f$, while the latter uses vectors α in the smaller set $\mathbb{Z}_f^+$; hence, at this point, there can potentially be split cuts for $\mathcal{C}(f, R, Q)$ which are not dominated by α-cuts. To show that such α's are redundant, we start with the following simple observation about the symmetry of split sets.

Lemma 56. *Consider a centered split set $\bar{S}(u,f)$, for $u, f \in \mathbb{R}^n$. Then $\bar{S}(u,f) = \bar{S}(-u,f)$.*

Proof. Notice that for every real number $a \in \mathbb{R}$ we have $-\lfloor a \rfloor = \lceil -a \rceil$. Then by definition of a split set

$$\begin{aligned} S(u,f) &= \{x : \lfloor uf \rfloor \le ux \le \lceil uf \rceil\} = \{x : -\lfloor uf \rfloor \ge -ux \ge -\lceil uf \rceil\} \\ &= \{x : \lceil -uf \rceil \ge -ux \ge \lfloor -uf \rfloor\} = S(-u,f). \end{aligned}$$

By definition of centered split sets, we get $\bar{S}(u,f) = S(u,f) - f = S(-u,f) - f = \bar{S}(-u,f)$, concluding the proof. $\square$

As a consequence, we can get the following invariance of lifted lattice-free split cuts.

Lemma 57. *Consider $f \in [0,1]^n \setminus \mathbb{Z}^n$ and $\alpha \in \mathbb{Z}_f$. Define $SP = S((\alpha,1),(f,0))$ and $SM = S((-\alpha,1),(f,0))$. Then*

$$(\tilde{\psi}^+_{SP}, \tilde{\pi}^+_{SP}) = (\tilde{\psi}^+_{SM}, \tilde{\pi}^+_{SM}).$$

Proof. To prove that $\tilde{\psi}^+_{SP} = \tilde{\psi}^+_{SM}$ we can use Corollary 7 with the self-adjoint linear map $(x_1,x_2) \mapsto (x_1,-x_2)$:

$$\tilde{\psi}^+_{SP}(r) = \gamma_{\bar{S}((\alpha,1),(f,0))}\begin{pmatrix} r \\ 0 \end{pmatrix} \overset{\text{Cor 7}}{=} \gamma_{\bar{S}((\alpha,-1),(f,0))}\begin{pmatrix} r \\ 0 \end{pmatrix} \overset{\text{Lemma 56}}{=} \gamma_{\bar{S}((-\alpha,1),(f,0))}\begin{pmatrix} r \\ 0 \end{pmatrix} = \tilde{\psi}^+_{SM}(r).$$

Similarly, to prove that $\tilde{\pi}^+_{SP} = \tilde{\pi}^+_{SM}$ we have that for every $\ell : Q \to \mathbb{R}$

$$\begin{aligned} \pi^+_{SP,\ell}(q) &= \gamma_{\bar{S}((\alpha,1),(f,0))}\begin{pmatrix} q \\ \ell(q) \end{pmatrix} \overset{\text{Cor 7}}{=} \gamma_{\bar{S}((\alpha,-1),(f,0))}\begin{pmatrix} q \\ -\ell(q) \end{pmatrix} \\ &\overset{\text{Lemma 56}}{=} \gamma_{\bar{S}((-\alpha,1),(f,0))}\begin{pmatrix} q \\ -\ell(q) \end{pmatrix} = \pi^+_{SM,-\ell}(q). \end{aligned}$$

Then using the symmetry in the possible choices of ℓ,

$$\tilde{\pi}^+_{SP}(q) = \min_{\ell} \pi^+_{SP,\ell}(q) = \min_{\ell} \pi^+_{SM,-\ell}(q) = \tilde{\pi}^+_{SM}(q),$$

concluding the proof of the lemma. $\square$

Proof of Theorem 16. First we prove that every α-cut is a split cut for $\mathcal{C}(f,R,Q)$. Consider some $\bar{\alpha} \in \mathbb{Z}^+_f$ and the cut $(\psi^f_{\bar{\alpha}}, \pi^f_{\bar{\alpha}})$. From Lemma 55 we have that $(\psi^f_{\bar{\alpha}}, \pi^f_{\bar{\alpha}})$ equals $(\tilde{\psi}_S, \tilde{\pi}_S)$ for $S = S((\bar{\alpha},1),(f,0))$. The validity of the latter for $\mathcal{C}(f,R,Q)$ then implies the validity of the former, and using the "if" part of Lemma 53 concludes the proof of this part.

For the second part, consider a split cut (ψ,π) for $\mathcal{C}(f,R,Q)$. From Lemma 53 there exists α in $\mathbb{Z}_f$ such that (ψ,π) is dominated by $(\tilde{\psi}_{S((\alpha,1),(f,0))}, \tilde{\pi}_{S((\alpha,1),(f,0))})$ with respect to $\mathcal{C}_{LP}(f,R,Q)$.

Set $\bar{\alpha}$ as either α or $-\alpha$, such that $\bar{\alpha} \in \mathbb{Z}^+_f$. Then employing Lemmas 57 and 55, respectively, we get

$$(\tilde{\psi}_{S((\alpha,1),(f,0))}, \tilde{\pi}_{S((\alpha,1),(f,0))}) = (\tilde{\psi}_{S((\bar{\alpha},1),(f,0))}, \tilde{\pi}_{S((\bar{\alpha},1),(f,0))}) = (\psi^f_{\bar{\alpha}}, \pi^f_{\bar{\alpha}}).$$

Therefore, the α-cut $(\psi^f_{\bar{\alpha}}, \pi^f_{\bar{\alpha}})$ dominates the split cut (ψ,π) with respect to $\mathcal{C}_{LP}(f,R,Q)$, as desired. □

8.4 Applications

8.4.1 Universality of Infinite Relaxation with Respect to Split Closure

We now use Theorem 16 to prove Theorem 17. Before that we need the following technical lemma that gives a more natural characterization of the domination relationship between valid cuts.

Lemma 58. *Consider a non-empty corner relaxation $\mathcal{C} = \mathcal{C}(f,R,Q)$. Let (ψ,π) and (ψ',π') be valid cuts for $\mathcal{C}$. Then (ψ,π) dominates (ψ',π') with respect to $\mathcal{C}_{LP}(f,R,Q)$ iff $(\psi,\pi) \le (\psi',\pi')$.*

Proof. The "if" part is trivial, so we prove only the "only if" part. So assume (ψ,π) dominates (ψ',π') with respect to $\mathcal{C}_{LP}(f,R,Q)$ and, by contradiction, assume that $(\psi,\pi) \not\le (\psi',\pi')$. We consider the case where $\psi \not\le \psi'$, the other one is analogous. So let $\bar{r} \in R$ be such that $\psi(\bar{r}) > \psi'(\bar{r})$.

First consider the case where $\psi'(\bar{r}) \ge 0$, and hence $\psi(\bar{r}) > 0$. Then we construct the solution

$$(\bar{x}, \bar{s}, \bar{y}) = \left(f + \frac{1}{\psi(\bar{r})}\bar{r}, \quad \frac{\bar{r}}{\psi(\bar{r})}, \quad 0 \right),$$

where $\;_{\bar{r}} : R \to \mathbb{R}$ is the function taking value $\;_{\bar{r}}(\bar{r}) = 1$ and $\;_{\bar{r}}(r) = 0$ for all other r. Notice that $(\bar{x}, \bar{s}, \bar{y})$ belongs to $\mathcal{C}_{LP}(f, R, Q) \cap (\psi, \pi)$ but not to $\mathcal{C}_{LP}(f, R, Q) \cap (\psi', \pi')$. This contradicts the assumption that (ψ, π) dominates (ψ', π') with respect to $\mathcal{C}_{LP}(f, R, Q)$, which concludes the proof in this case.

Now consider the case where $\psi'(\bar{r}) < 0$. Let $(\bar{x}, \bar{s}, \bar{y})$ be a feasible solution for $\mathcal{C}(f, R, Q)$; in particular we have $(\bar{x}, \bar{s}, \bar{y}) \in \mathcal{C}_{LP}(f, R, Q) \cap (\psi, \pi)$. Now let $\lambda > 0$ be large enough and consider the solution

$$(x', s', y') = (\bar{x} + \lambda \bar{r}, \bar{s} + \lambda \;_{\bar{r}}, \bar{y}).$$

Since valid functions are non-negative by assumption (and in particular $\psi(\bar{r}) \geq 0$), notice that we still have $(x', s', y') \in \mathcal{C}_{LP}(f, R, Q) \cap (\psi, \pi)$. However, since $\psi'(\bar{r}) < 0$, setting λ large enough gives that (x', s', y') does not belong to $\mathcal{C}_{LP}(f, R, Q) \cap (\psi', \pi')$, again reaching a contradiction. This concludes the proof of the lemma. □

Proof of Theorem 17. The second statement of the theorem follows directly from Theorem 16, so we prove the first statement. Let (ψ, π) be a split cut for $\mathcal{C}(f, R, Q) \neq \emptyset$. The second part of Theorem 16 guarantees that there is $\bar{\alpha} \in \mathbb{Z}_f^+$ such that $(\psi_{\bar{\alpha}}^f|_R, \pi_{\bar{\alpha}}^f|_Q)$ dominates (ψ, π) with respect to $\mathcal{C}(f, R, Q)$. Since the latter cut is non-negative, we can employ Lemma 58 above to get that $(\psi_{\bar{\alpha}}^f|_R, \pi_{\bar{\alpha}}^f|_Q) \leq (\psi, \pi)$.

Then define the cut (ψ', π') for $\mathcal{C}(f, \mathbb{R}, \mathbb{R})$ such that $(\psi'|_R, \pi'|_Q) = (\psi, \pi)$ and $(\psi'|_{\mathbb{R} \setminus R}, \pi'|_{\mathbb{R} \setminus Q}) = (\psi_{\bar{\alpha}}^f|_{\mathbb{R} \setminus R}, \pi_{\bar{\alpha}}^f|_{\mathbb{R} \setminus Q})$. We then have that: (i) (ψ', π') is a split cut for $\mathcal{C}(f, R, Q)$, because $(\psi', \pi') \geq (\psi_{\bar{\alpha}}^f, \pi_{\bar{\alpha}}^f)$ and $(\psi_{\bar{\alpha}}^f, \pi_{\bar{\alpha}}^f)$ is a split cut for $\mathcal{C}(f, R, Q)$ (by the first part of Theorem 16); (ii) (ψ, π) is a restriction of (ψ', π'). This concludes the proof of the theorem. □

8.4.2 Pure Integer Program with Weak Split Closure

In this section we prove Theorem 18 by exhibiting a pure integer corner relaxation with a weak split closure; throughout, we will only work with corner relaxations $\mathcal{C}(f, R, Q)$ where both R and Q are finite sets of rational vectors.

Recall that the split closure of a corner relaxation $\mathcal{C}(f, R, Q)$, denoted by $\mathcal{SC}(f, R, Q)$, is the set of all points in $\mathcal{C}_{LP}(f, R, Q)$ that satisfy all split cuts for $\mathcal{C}(f, R, Q)$. Using Theorem 16 we can describe this closure as

$$\mathcal{SC}(f, R, Q) = \mathcal{C}_{LP}(f, R, Q) \cap \left(\bigcap_{\alpha \in \mathbb{Z}_f} (\psi_\alpha^f, \pi_\alpha^f) \right). \tag{8.12}$$

Bad example. Now we present the family of integer corner relaxations with a weak split closure. Consider the following family parametrized by a rational number $\epsilon > 0$:

$$x = \frac{1}{2} + \left(\frac{1}{2} + \frac{\epsilon}{2}\right) y_1 + \left(\frac{1}{2} + \epsilon\right) y_2 \qquad (\mathcal{C}_\epsilon)$$
$$y_1, y_2 \geq 0$$
$$x, y_1, y_2 \in \mathbb{Z}.$$

To make things more clear, we define $f = 1/2$, $q^1 = (\frac{1}{2} + \frac{\epsilon}{2})$ and $q^2 = (\frac{1}{2} + \epsilon)$; so formally $\mathcal{C}_\epsilon = \mathcal{C}(f, \emptyset, \{q^1, q^2\})$.

Let $\mathcal{SC}_\epsilon$ denote the split closure of $\mathcal{C}_\epsilon$; we claim that there is a gap of $1/12\epsilon$ between these two sets when minimizing $y_1 + y_2$, i.e., they satisfy Theorem 18. The intuition behind this construction is the following. The equivalence in Theorem 16, via equation (8.12) allow us to focus solely on understanding the effect of α-cuts on the above program. Since the latter only has integrally constrained non-basic variables, this means focusing on the functions π^f_α. Observing the behavior of the function π^f_α, we see that it essentially has "high" value (close to 1) for inputs close to 1/2; while this is not exactly true for large $\alpha \in \mathbb{Z}^f$, our particular choice of q^1 and q^2 guarantees that *one of* $\pi^f_\alpha(q^1)$ or $\pi^f_\alpha(q^2)$ is reasonably large for every choice of $\alpha \in \mathbb{Z}_f$.

Lemma 59. *For every $\epsilon > 0$, we have that $\max\left\{\pi^f_\alpha(q^1), \pi^f_\alpha(q^2)\right\} \geq 1/3$ for all $\alpha \in \mathbb{Z}^f$.*

Proof. By the definition of $\mathbb{Z}_f$, notice that for every $\alpha \in \mathbb{Z}_f$ we have that the fractional part $[\alpha f]$ equals f. Thus, using the definition of π^f_α given by equations (8.2) and (8.4), we get that π^f_α takes the form

$$\pi^f_\alpha(q) = \begin{cases} 2[\alpha q] & , \text{if } [\alpha q] \leq \frac{1}{2} \\ 2 - 2[\alpha q] & , \text{if } [\alpha q] > \frac{1}{2} \end{cases} \qquad (8.13)$$

The next claim gives some control on the behavior of the fractional part $[\alpha q]$ that appears in the above expression.

Claim 10. *For every $\alpha \in \mathbb{Z}_f$ either $[\alpha q^1]$ or $[\alpha q^2]$ lies in the interval* $[1/6, 5/6]$.

Proof. Take $\alpha \in \mathbb{Z}_f$. Since $[\alpha f] = [f]$, we have that $[\alpha q^1] = [1/2 + k/(2\epsilon)] = [1/2 + [k/(2\epsilon)]]$. Using the latter, it is easy to verify that $[\alpha q^1] \in [1/6, 5/6]$ if and only if $[k\epsilon/2] \notin (1/3, 2/3)$. Similarly, we have that $[\alpha q^2] \in [1/6, 5/6]$ if and only if $[k\epsilon] \notin (1/3, 2/3)$.

To conclude the proof, it suffices to show that $[\alpha\epsilon/2] \in (1/3, 2/3)$ implies $[\alpha\epsilon] \notin (1/3, 2/3)$. For that notice that when $[\alpha\epsilon/2] \in (1/3, 1/2]$, this implies $[\alpha\epsilon] \in (2/3, 1]$ and hence $[\alpha\epsilon] \notin (1/3, 2/3)$. On the other hand, when $[\alpha\epsilon/2] \in [1/2, 2/3)$ we have $[\alpha\epsilon] \in [0, 1/3)$, again reaching the conclusion $[\alpha\epsilon] \notin (1/3, 2/3)$. This concludes the proof of the claim. □

Then take $\alpha \in \mathbb{Z}_f$ and using the above claim let $i \in \{1, 2\}$ be such that $[\alpha q^i] \in [1/6, 5/6]$. If $[\alpha q^i] \in [1/6, 1/2]$, then employing equation (8.13) we get that $\pi^f_\alpha(q^i) = 2[\alpha q^i] \geq 1/3$; if $[\alpha q^i] \in [1/2, 5/6]$, the same equation gives $\pi^f_\alpha(q^i) = 2 - 2[\alpha q^i] \geq 2 - 10/6 = 1/3$. This concludes the proof of the lemma. □

Using the above lemma, we see that the point $(\bar{x}, \bar{y}_1, \bar{y}_2)$ given by $\bar{y}_1 = \bar{y}_2 = 3$ and $\bar{x} = f + q^1\bar{y}_1 + q^2\bar{y}_2$ belongs to the linear relaxation of $(\mathcal{C}_\epsilon)$ and satisfies all α-cuts for it; therefore, this point belongs to the split closure of $\mathcal{SC}_\epsilon$. This directly gives the following.

Lemma 60. *The optimal value of minimizing $y_1 + y_2$ over the split closure $\mathcal{SC}_\epsilon$ is at most* 6*.*

Now in order to show the weakness of the split cuts, we show that the optimal value of minimizing $y_1 + y_2$ over the whole of $\mathcal{C}_\epsilon$ is much larger.

Lemma 61. *The optimal value of minimizing $y_1 + y_2$ over $\mathcal{C}_\epsilon$ is at least* $1/(2\epsilon)$*.*

Proof. We can rewrite the equation in $(\mathcal{C}_\epsilon)$ as

$$x = \frac{1}{2}(1 + y_1 + y_2) + \epsilon\left(\frac{y_1}{2} + y_2\right).$$

Since for every solution in $\mathcal{C}_\epsilon$ we have $y_1, y_2 \in \mathbb{Z}_+$, this implies that $\frac{1}{2}(1 + y_1 + y_2) \in \mathbb{Z}_+/2$; since in such solution we also have $x \in \mathbb{Z}$, this implies that $\epsilon(y_1/2+y_2) \in \mathbb{Z}_+/2$. Moreover, we need to have one of y_1 or y_2 strictly positive (and the other non-negative), so the previous observation actually implies that $\epsilon(y_1/2 + y_2) \geq 1/2$ is satisfied by all feasible solutions. Such solutions then satisfy the weaker inequality $y_1 + y_2 \geq 1/(2\epsilon)$; this concludes the proof of the lemma. □

Now we are ready to prove Theorem 18 and Corollary 6.

Proof of Theorem 18. Follows directly from Lemmas 60 and 61. □

Proof of Corollary 6. Recall that $\overline{\mathcal{C}}_\epsilon$ and $\overline{\mathcal{SC}}_\epsilon$ are respectively the projections of $\text{conv}(\mathcal{C}_\epsilon)$ and $\mathcal{SC}_\epsilon$ onto the y-space. It is clear that we have

$$\frac{\min\{y_1 + y_2 : y \in \overline{\mathcal{C}}_\epsilon\}}{\min\{y_1 + y_2 : y \in \overline{\mathcal{SC}}_\epsilon\}} \geq \frac{1}{12\epsilon} \tag{8.14}$$

just as in Theorem 18.

A closed, convex set $X \subseteq \mathbb{R}^n_+$ is said to be of *blocking type* if for every $x \in X$, $y \geq x$ implies $y \in X$. It is known that $\overline{\mathcal{C}}_\epsilon$ is of blocking type [44]; moreover, using the description of $\mathcal{SC}_\epsilon$ given in (8.12), and the fact that the $(\psi^f_\alpha, \pi^f_\alpha)$'s are non-negative, we also have that $\overline{\mathcal{SC}}_\epsilon$ is of blocking type. Finally, since $\mathcal{SC}$ is a relaxation of $\text{conv}(\mathcal{C})$, we have $\overline{\mathcal{C}} \subseteq \overline{\mathcal{SC}}$.

Now Basu et al. [35] showed that, for two sets $A \subseteq B \subseteq \mathbb{R}^n$ of blocking type, the blow up measure is equivalent to the worst gap over all non-negative directions; in our case translates to the following:

$$\inf\{\alpha : \alpha\overline{\mathcal{C}}_\epsilon \supseteq \overline{\mathcal{SC}}_\epsilon\} = \sup_{c \in \mathbb{R}^2_+} \left\{ \frac{\min\{cy : y \in \overline{\mathcal{C}}_\epsilon\}}{\min\{cy : y \in \overline{\mathcal{SC}}_\epsilon\}} \right\}.$$

The proof of Corollary 6 then follows from (8.14). □

Part III

Final Remarks

CHAPTER 9

Final Remarks

In this thesis, we undertook a principled study of the strength of general-purpose cutting planes, giving a better understanding of the relationship between the different families of cuts available and analyzing the properties and limitations of our current methods for deriving cuts. However, there are still many open avenues in these directions, as pointed out at the end of previous chapters. Here, we briefly mention two interesting possibilities for further expanding our knowledge on cutting planes.

9.1 Constrained Corner/Infinite Relaxation

Although the corner and related relaxations are still arguably the the most important tools to derive and analyze cuts, there has been renewed interest in considering stronger relaxations that bring back some of the non-negativity constraints lost in their derivation (see discussion in Section 3.2.2). As examples, we have the works [66, 31, 82], which focus on the continuous relaxation. However, it seems that we still do not have a satisfactory understanding of how to make use of both *integrality of non-basic* and *bounds on basic variables*, namely in a model such as B(f) presented in Chapter 7 and repeated here for convenience:

$$
\begin{aligned}
x = f + \sum_{r \in \mathbb{R}} r s_r + \sum_{r \in \mathbb{R}} r y_r \\
x \in \{0, 1\} \\
s_r \geq 0 \\
y_r \in \mathbb{Z}_+ \\
(s, y) \text{ has finite support.}
\end{aligned}
\tag{9.1}
$$

One proposal in this direction, is to start by understanding how the split cuts for (9.1) look like. The results from Chapter 7 give a starting point for this program. Even though the computational results from the cuts presented there did not show a significant improvement over GMI cuts, the hope is that we can still obtain stronger splits by directly making use of the bounds on the basic variables.

Actually, it might be convenient to consider a finite-dimensional version of the model (for simplicity, we include here only the integer variables):

$$
\begin{aligned}
x = f + \sum_{j=1}^{n} r^j y_j \\
x \in \{0, 1\} \\
y \in \mathbb{Z}^n_+.
\end{aligned}
\tag{9.2}
$$

We present two more concrete suggestions for approaching this problem. One is to analyze the split cuts coming from all of the corner relaxations of (9.2), expressing these cuts back in the y-space. For each of the corner relaxations, we can then try to use the insights on split cuts obtained in Chapter 8. Another possibility is to use directly the geometric lifting idea considered in Chapters 7 and 8, considering a system with extra rows (and extra basic variables) which capture the integrality of the y variables (see $\mathrm{Z}(\ell)$ in Chapter 7). Notice that, in particular, we can add n new rows and obtain an equivalent system (after it is projected onto the original variables) which only has integrality on the basic variables:

$$\begin{aligned}
x &= f + \sum_{j=1}^{n} r^j y_j \\
z_j &= y_j \quad j = 1, 2, \ldots, n \\
x &\in \{0, 1\} \\
y &\in \mathbb{R}^n_+ \\
z &\in \mathbb{Z}^n.
\end{aligned}$$

In this case, we are essentially back in the setting of the continuous relaxation, where we can look at $(\{0,1\} \times \mathbb{Z}^n)$-free sets in the (x, z)-space. The hope here is that we do not actually need to add all n extra rows in order to obtain new information about split cuts, as it happened in Chapter 7.

Question 1. *Understand split cuts for* (9.1) *or* (9.2)*. For the former, in particular, obtain a characterization of the valid functions corresponding to split cuts.*

Another even more concrete question comes from the following observation: as we made explicit in Chapter 7, the cuts obtained for (9.1) are valid for a relaxation where x is only upper bounded by 1, and hence do not use the information that x is also non-negative. The question is if we can use both upper and lower bounds on x simultaneously.

Question 2. *Understand if it is possible to obtain a valid function* (ψ, π) *for* (9.1) *which is not valid for the set 7.1 presented in Chapter 7.*

9.2 Geometric Perspective on Cuts for Non-linear MIPs

There has been renewed interest in understanding mixed-integer programs that contain non-linear constraints. For instance, Atamtürk and Narayanan [6] consider *second-order conic mixed-integer programs* of the form

$$\begin{aligned}
\min\ & cx + ry \\
& \|A^i x + G^i y - b^i\| \leq d^i x + f^i y - h_i, \quad i = 1, 2, \ldots, k \\
& x \in \mathbb{Z}^n, y \in \mathbb{R}^p,
\end{aligned}$$

where A^i, G^i are rational matrices and b, c, r, d^i, f^i are rational vectors. In order to generate valid cuts for such problems, the authors adopt the simple set approach from Section 2.4.2, studying the facial structure of sets of the form

$$\{x \in \mathbb{Z}^n_+, y \in \mathbb{R}^p_+, t \in \mathbb{R} : t \geq |ax + gy - b|\}.$$

The computational experiments reported show that these cuts close a significant part of the integrality gap.

This approach was also extended by Masihabadi, Sanjeevi and Kianfar [105]. We remark that the possibility of generating cuts for non-linear programs using simple sets had already been observed in [92], although there the authors still consider a linear simple set and incorporate non-linearity in the embedding step.

Given that all these results are obtained from the simple set perspective on cut generation, it is natural to ask if we can interpret these cuts in a more geometric way, for example via the disjunctive or intersection cuts perspective. Not only this would give a better understanding of cuts for non-linear MIPs, but also could help leveraging the known results from linear MIPs to the non-linear setting.

Question 3. *Can we obtain the conic MIR inequalities of [9] using a disjunctive or intersection cuts perspective?*

Bibliography

[1] J. Aczél. *Lectures on Functional Equations and Their Applications.* Academic Press, New York, 1966.

[2] R. P. Agarwal, M. Meehan, and D. O'Regan. *Fixed Point Theory and Applications.* Cambridge University Press, 2001.

[3] K. Andersen, G. Cornuéjols, and Y. Li. Split closure and intersection cuts. *Mathematical Programming*, 102(3):457–493, 2005.

[4] K. Andersen, Q. Louveaux, R. Weismantel, and L. Wolsey. Inequalities from two rows of a simplex tableau. In *IPCO*, volume 4513, pages 1–15, 2007.

[5] D. L. Applegate, R. E. Bixby, V. Chvátal, and W. J. Cook. *The Traveling Salesman Problem: A Computational Study.* Princeton University Press, 2007.

[6] A. Atamtürk. On the facets of the mixed-integer knapsack polyhedron. *Mathematical Programming*, 98:145–175, 2003.

[7] A. Atamtürk and O. Günlük. Mingling: mixed-integer rounding with bounds. *Mathematical Programming*, 123:315–338, 2010.

[8] A. Atamtürk and K. Kianfar. n-step mingling inequalities: new facets for the mixed-integer knapsack set. *Mathemaical Programming*, 132(1-2):79–98, 2012.

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

[10] G. Averkov, C. Wagner, and R. Weismantel. Maximal lattice-free polyhedra: Finiteness and an explicit description in dimension three. *Mathematics of Operations Research*, 36(4):721–742, 2011.

[11] E. Balas. Intersection cuts-a new type of cutting planes for integer programming. *Operations Research*, 19(1):19–39, 1971.

[12] E. Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.

[13] E. Balas. Disjunctive programming. In M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, editors, *50 Years of Integer Programming 1958-2008*, pages 283–340. Springer Berlin Heidelberg, 2010.

[14] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 01 programs. *Mathematical Programming*, 58:295–324, 1993.

[15] E. Balas, S. Ceria, and G. Cornuéjols. Mixed 0-1 programming by lift-and-project in a branch-and-cut framework. *Management Science*, 42:1229–1246, 1996.

[16] E. Balas, S. Ceria, G. Cornuéjols, and N. Natraj. Gomory cuts revisited. *Operations Research Letters*, 19(1):1 – 9, 1996.

[17] E. Balas and R. G. Jeroslow. Strengthening cuts for mixed integer programs. *European Journal of Operational Research*, 4(4):224 – 234, 1980.

[18] E. Balas and F. Margot. Generalized intersection cuts and a new cut generating paradigm. *Mathematical Programming*, pages 1–17.

[19] E. Balas, F. Margot, and S. Nadarajah. A polynomial procedure for generating generalized intersection cuts.

[20] E. Balas and M. Perregaard. A precise correspondence between lift-and-project cuts, simple disjunctive cuts, and mixed integer gomory cuts for 0-1 programming. *Mathematical Programming*, 94(2-3):221–245, 2003.

[21] E. Balas and A. Qualizza. Monoidal cut strengthening revisited. *Discrete Optimization*, 9(1):40–49, 2012.

[22] E. Balas and A. Qualizza. Monoidal cut strengthening revisited. *Discrete Optimization*, 9(1):40 – 49, 2012.

[23] E. Balas and A. Saxena. Optimizing over the split closure. *Mathematical Programming*, 113(2):219–240, 2008.

[24] A. Barvinok. *A Course in Convexity*, volume 54 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, Rhode Island, 2002.

[25] A. Basu. *Corner Polyhedra and Maximal Lattice-free Convex Sets : A Geometric Approach to Cutting Planes.* PhD thesis, Carnegie Mellon University, 2010.

[26] A. Basu, P. Bonami, G. Cornuéjols, and F. Margot. Experiments with two-row cuts from degenerate tableaux. *INFORMS Journal on Computing*, 23(4):578–590, 2011.

[27] A. Basu, P. Bonami, G. Cornuéjols, and F. Margot. On the relative strength of split, triangle and quadrilateral cuts. *Mathematical Programming*, 126(2):281–314, 2011.

[28] A. Basu, M. Campêlo, M. Conforti, G. Cornuéjols, and G. Zambelli. On lifting integer variables in minimal inequalities. In *Proc. IPCO 2010*, June 2010.

[29] A. Basu, M. Campêlo, M. Conforti, G. Cornuéjols, and G. Zambelli. Unique lifting of integer variables in minimal inequalities. *Mathematical Programming*, pages 1–16, 2012.

[30] A. Basu, M. Conforti, G. Cornuéjols, and G. Zambelli. Maximal lattice-free convex sets in linear subspaces. *Mathematics of Operations Research*, 35:704–720, 2010.

[31] A. Basu, M. Conforti, G. Cornuéjols, and G. Zambelli. Minimal inequalities for an infinite relaxation of integer programs. *SIAM Journal on Discrete Mathematics*, 24:158–168, February 2010.

[32] A. Basu, M. Conforti, G. Cornuéjols, and G. Zambelli. A counterexample to a conjecture of Gomory and Johnson. *Mathematical Programming*, 133(1-2):25–38, 2012.

[33] A. Basu, G. Cornuéjols, and M. Köppe. Unique minimal liftings for simplicial polytopes. *Mathematics of Operations Research*, 37:346–355, 2012.

[34] A. Basu, G. Cornuéjols, and F. Margot. Intersection cuts with infinite split rank. *Mathematics of Operations Research*, 37(1):21–40, Feb. 2012.

[35] A. Basu, G. Cornuéjols, and M. Molinaro. A probabilistic analysis of the strength of the split and triangle closures. In *IPCO*, pages 27–38, 2011.

[36] A. Basu, R. Hildebrand, M. Köppe, and M. Molinaro. A (k+1)-Slope Theorem for the k-Dimensional Infinite Group Relaxation. *ArXiv e-prints*, 2011.

[37] D. Bertsimas and J. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, 1st edition, 1997.

[38] R. E. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling. Mixed integer programming: A progress report. In *The Sharpest Cut*, pages 309–325. MPS-SIAM Series on Optimization, Philadelphia, PA, 2004.

[39] P. Bonami. On optimizing over lift-and-project closures. *Mathematical Programming Computation*, 4(2):151–179, 2012.

[40] V. Borozan and G. Cornuéjols. Minimal valid inequalities for integer constraints. *Mathematics of Operations Research*, 34(3):538–546, 2009.

[41] V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4(4):305 – 337, 1973.

[42] M. Conforti, G. Cornuéjols, A. Daniilidis, C. Lemaréchal, and J. Malick. Cut-generating functions and s-free sets. In *IPCO*, 2013.

[43] M. Conforti, G. Cornuéjols, and G. Zambelli. Equivalence between intersection cuts and the corner polyhedron. *Operations Research Letters*, 38(3):153–155, 2010.

[44] M. Conforti, G. Cornuéjols, and G. Zambelli. Corner polyhedra and intersection cuts. *Surveys in Operations Research and Management Science*, 16:105–120, 2011.

[45] M. Conforti, G. Cornuéjols, and G. Zambelli. A geometric perspective on lifting. *Operations Research*, 59:569–577, 2011.

[46] W. Cook. `http://www.tsp.gatech.edu/`.

[47] W. J. Cook, R. Kannan, and A. Schrijver. Chvátal closures for mixed integer programming problems. *Mathematical Programming*, 47:155–174, 1990.

[48] G. Cornuéjols. Revival of the gomory cuts in the 1990's. *Annals of Operations Research*, 149:63–66, 2007.

[49] G. Cornuéjols and Y. Li. On the rank of mixed 0,1 polyhedra. *Mathematical Programming*, 91(2):391–397, 2002.

[50] G. Cornuéjols, Y. Li, and D. Vandenbussche. K-cuts: A variation of Gomory mixed integer cuts from the LP tableau. *INFORMS Journal on Computing*, 15(4):385–396, Dec. 2003.

[51] G. Cornuéjols and F. Margot. On the facets of mixed integer programs with two integer variables and two constraints. *Mathematical Programming*, 120:429–456, 2009. DOI: 10.1007/s10107-008-0221-1.

[52] G. Cornuéjols, C. Michini, and G. Nannicini. How tight is the corner relaxation? insights gained from the stable set problem. *Discrete Optimization*, 9(2):109 – 121, 2012.

[53] G. Cornuéjols and M. Molinaro. A 3-slope theorem for the infinite relaxation in the plane. *Mathematical Programming*, pages 1–23, 2012.

[54] G. Cornuéjols, L. Wolsey, and S. Yildiz. Sufficiency of cut-generating functions. 2013.

[55] S. Dash. Mixed integer rounding cuts and master group polyhedra. In V. Chvátal, editor, *Combinatorial Optimization - Methods and Applications*, volume 31 of *NATO Science for Peace and Security Series - D: Information and Communication Security*, pages 1–32. IOS Press, 2011.

[56] S. Dash, S. Dey, and O. Günlük. Two dimensional lattice-free cuts and asymmetric disjunctions for mixed-integer polyhedra. *Mathematical Programming.* To appear.

[57] S. Dash, S. S. Dey, and O. Günlük. On mixed-integer sets with two integer variables. *Operations Research Letters*, 39(5):305–309, 2011.

[58] S. Dash, N. B. Dobbs, O. Günlük, T. J. Nowicki, and G. M. Swirszcz. Lattice-free sets, multi-branch split disjunctions, and mixed-integer programming. *Mathematical Programming.* To appear.

[59] S. Dash and M. Goycoolea. A heuristic to separate rank-1 GMI cuts. *Mathematical Programming Computation*, 2:231–257, 2010.

[60] S. Dash and O. Günlük. Valid inequalities based on simple mixed-integer sets. *Mathematical Programming*, 105:29–53, 2006.

[61] S. Dash and O. Günlük. Valid inequalities based on the interpolation procedure. *Mathematical Programming*, 106(1):111–136, May 2006.

[62] S. Dash and O. Günlük. On the strength of Gomory mixed-integer cuts as group cuts. *Mathematical Programming*, 115(2):387–407, June 2008.

[63] S. Dash and O. Günlük. On t-branch split cuts for mixed-integer programs. *Mathematical Programming*, pages 1–9, 2012.

[64] S. Dash, O. Günlük, and A. Lodi. Mir closures of polyhedral sets. *Mathematical Programming*, 121(1):33–60, 2010.

[65] S. Dash, O. Günlük, and J. P. Vielma. Computational experiments with cross and crooked cross cuts. 2011.

[66] S. Dey and L. Wolsey. Constrained infinite group relaxations of MIPs. *SIAM Journal on Optimization*, 20(6):2890–2912, 2010.

[67] S. S. Dey, A. Lodi, A. Tramontani, and L. A. Wolsey. Experiments with two row tableau cuts. In *IPCO*, pages 424–437, 2010.

[68] S. S. Dey and Q. Louveaux. Split rank of triangle and quadrilateral inequalities. *Mathematics of Operations Research*, 36(3):432–461, Aug. 2011.

[69] S. S. Dey and J.-P. P. Richard. Facets of two-dimensional infinite group problems. *Mathematics of Operations Research*, 33(1):140–166, 2008.

[70] S. S. Dey, J.-P. P. Richard, Y. Li, and L. A. Miller. On the extreme inequalities of infinite group problems. *Mathematical Programming*, 121(1):145–170, June 2009.

[71] S. S. Dey and L. A. Wolsey. Composite lifting of group inequalities and an application to two-row mixing inequalities. *Discrete Optimization*, 7(4):256 – 268, 2010.

[72] S. S. Dey and L. A. Wolsey. Two row mixed-integer cuts via lifting. *Mathematical Programming*, 124:143–174, 2010.

[73] D. Dubhashi and A. Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms.* Cambridge University Press, New York, NY, USA, 2009.

[74] F. Eisenbrand. Integer programming and algorithmic geometry of numbers. In M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, editors, *50 Years of Integer Programming 1958-2008*, pages 505–559. Springer Berlin Heidelberg, 2010.

[75] D. G. Espinoza. Computing with multi-row Gomory cuts. *Operations Research Letters*, 38(2):115 – 120, 2010.

[76] L. Evans. Cyclic group and knapsack facets with applications to cutting planes. PhD Thesis.

[77] M. Fischetti and A. Lodi. Optimizing over the first Chvátal closure. *Mathematical Programming*, 110(1):3–20, 2007.

[78] M. Fischetti, A. Lodi, and A. Tramontani. On the separation of disjunctive cuts. *Mathematical Programming*, 128(1):205–230, 2011.

[79] M. Fischetti and M. Monaci. How tight is the corner relaxation? *Discrete Optimization*, 5(2):262–269, 2008.

[80] M. Fischetti and D. Salvagnin. A relax-and-cut framework for Gomory's mixed integer cuts. *Mathematical Programming Computation*, 3(2), 2011.

[81] M. Fischetti and C. Saturni. Mixed-integer cuts from cyclic groups. *Mathematical Programming*, 109(1):27–53, 2007.

[82] R. Fukasawa and O. Günlük. Strengthening lattice-free cuts using non-negativity. *Discrete Optimization*, 8(2):229–245, 2011.

[83] M. X. Goemans. Worst-case comparison of valid inequalities for the TSP. *Mathematical Programming*, 69:335–349, 1995.

[84] R. Gomory. An algorithm for integer solutions to linear programs. In *Recent Advances in Mathematical Programming*, pages 269–302. McGraw-Hill, New York, 1963.

[85] R. E. Gomory. On the relation between integer and noninteger solutions to linear programs. *Proceedings of the National Academy of Sciences*, 53(2):260–265, 1965.

[86] R. E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and its Applications*, 2(4):451 – 558, 1969.

[87] R. E. Gomory and E. L. Johnson. Some continuous functions related to corner polyhedra, I. *Mathematical Programming*, 3:23–85, 1972. 10.1007/BF01585008.

[88] R. E. Gomory and E. L. Johnson. Some continuous functions related to corner polyhedra, II. *Mathematical Programming*, 3:359–389, 1972. 10.1007/BF01585008.

[89] R. E. Gomory and E. L. Johnson. T-space and cutting planes. *Mathematical Programming*, 96:341–375, 2003. 10.1007/s10107-003-0389-3.

[90] R. E. Gomory, E. L. Johnson, and L. Evans. Corner polyhedra and their connection with cutting planes. *Mathematical Programming*, 96(2):321–339, 2003.

[91] P. M. Gruber and C. G. Lekkerkerker. *Geometry of Numbers*. North-Holland, 1987.

[92] O. Günlük and Y. Pochet. Mixing mixed-integer inequalities. *Mathematical Programming*, 90(3):429–457, 2001.

[93] Q. He, S. Ahmed, and G. L. Nemhauser. A probabilistic comparison of split and type 1 triangle cuts for two-row mixed-integer programs. *SIAM Journal on Optimization*, 21(3):617–632, 2011.

[94] E. Hille and R. Phillips. *Functional Analysis and Semi-Groups*. American Mathematical Society, 1957.

[95] C. A. J. Hurkens. Blowing up convex sets in the plane. *Linear Algebra and its Applications*, 134:121–128, 1990.

[96] 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, 2010.

[97] R. Karp. Reducibility among combinatorial problems. In R. Miller and J. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[98] K. Kianfar and Y. Fathi. Generalized mixed integer rounding inequalities: facets for infinite group polyhedra. *Mathematical Programming*, 120(2):313–346, May 2009.

[99] B. Knaster, C. Kuratowski, and S. Mazurkiewicz. Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe. *Fundamenta Mathematicae*, 14:132–137, 1929.

[100] J. Lenstra, H. W. Integer programming with a fixed number of variables. *Mathematics of Operations Research*, 8(4):pp. 538–548, 1983.

[101] Y. Li and J.-P. P. Richard. Cook, Kannan and Schrijver's example revisited. *Discrete Optimization*, 5:724–734, 2008.

[102] A. Lodi. Mixed integer programming computation. In M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, editors, *50 Years of Integer Programming 1958-2008*, pages 619–646. Springer Berlin Heidelberg, 2010.

[103] J. A. D. Loera, R. Hemmecke, S. Onn, and R. Weismantel. N-fold integer programming. *Discrete Optimization*, 5(2):231–241, 2008.

[104] H. Marchand and L. A. Wolsey. Aggregation and mixed integer rounding to solve MIPs. *Operations Research*, 49(3):pp. 363–371, 2001.

[105] S. Masihabadi, S. Sanjeevi, and K. Kianfar. n-step conic mixed integer rounding inequalities. http://www.optimization-online.org/DB_HTML/2011/11/3251.html.

[106] D. Morán and S. Dey. On maximal S-free convex sets. *SIAM Journal on Discrete Mathematics*, 25(1):379–393, 2011.

[107] G. Nemhauser and L. A. Wolsey. A recursive procedure to generate all cuts for 0-1 mixed integer programs. *Mathematical Programming*, 46:379–390, 1990.

[108] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. Wiley-Interscience, New York, NY, USA, 1988.

[109] A. D. Pia, C. Wagner, and R. Weismantel. A probabilistic comparison of the strength of split, triangle, and quadrilateral cuts. *Operations Research Letters*, 39(4):234 – 240, 2011.

[110] P. Pudlk. Lower bounds for resolution and cutting plane proofs and monotone computations. *The Journal of Symbolic Logic*, 62(3):pp. 981–998, 1997.

[111] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, 1993.

[112] A. Schrijver. On cutting planes. *Annals of Discrete Mathematics*, 9:291–296, 1980.

[113] A. Schrijver. *Theory of Linear and Integer Programming.* John Wiley & Sons, Inc., New York, NY, USA, 1986.

[114] L. A. Wolsey. *Integer Programming.* Wiley, 1998.

[115] G. M. Ziegler. *Lectures on Polytopes.* Springer, 1995.