CUTTING PLANES IN MIXED INTEGER PROGRAMMING: THEORY AND ALGORITHMS

A Thesis
Presented to
The Academic Faculty

by

Steve Tyber

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology
May 2013

CUTTING PLANES IN MIXED INTEGER PROGRAMMING: THEORY AND ALGORITHMS

Approved by:

Dr. Shabbir Ahmed, Co-Advisor
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. George Nemhauser, Co-Advisor
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Santanu Dey
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Santosh Vempala
School of Computer Science
Georgia Institute of Technology

Dr. Zonghao Gu
Chief Technical Officer
Gurobi Optimization

Date Approved: 30 January 2013

In loving memory of Kristen.

ACKNOWLEDGEMENTS

First and foremost, I owe a debt of gratitude to my advisors Shabbir Ahmed and George Nemhauser. When I was my toughest critic, they stuck with me. It has truly been a privilege to work with them these past few years. I must also extend a special thanks to Pinar Keskinocak, Gary Parker, and Brani Vidakovic for encouraging me to pursue a PhD. This has been a truly rewarding experience that I might have otherwise missed were it not for their marked enthusiasm.

My time at Georgia Tech would not be the same without my fellow PhD students: Doug Altner, Gustavo Angulo, Arash Asadi, Rodolfo Carvajal, Daniel Dadush, Akshay Gupte, Qie He, Chris Healey, Fatma Kılınç-Karzan, Diego Morán Ramiréz, Dimitri Papageorgiou, Pete Petersen, Feng Qiu, Sangho Shim, Dan Steffy, Kael Stilp, Juan Pablo Vielma, Xing Wang, and Linji Yang. Thank you for all the years of friendship and for all the knowledge and wisdom you have imparted to me.

Thanks to Jerry Carney and Frank Grange for the years of practical insight that have kept me grounded in reality.

I would also like to thank the many excellent professors I've had the pleasure to learn from both in and out of the classroom, and I would especially like to thank my committee for their time and dedication.

Last and certainly not least, thanks to my family who have been a constant source of support. They have been with me through the good times and the bad, and I know I couldn't have made it to this point without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Recent developments in mixed integer programming have highlighted the need for multi-row cuts. To this day, the performance of such cuts has typically fallen short of the single-row Gomory mixed integer cut. This disparity between the theoretical need and the practical shortcomings of multi-row cuts motivates the study of both the mixed integer cut and multi-row cuts. In this thesis, we build on the theoretical foundations of the mixed integer cut and develop techniques to derive multi-row cuts.

The first chapter introduces the mixed integer programming problem. In this chapter, we review the terminology and cover some basic results that find application throughout this thesis. Furthermore, we describe the practical solution of mixed integer programs, and in particular, we discuss the role of cutting planes and our contributions to this theory.

In Chapter 2, we investigate the Gomory mixed integer cut from the perspective of group polyhedra. In this setting, the mixed integer cut appears as a facet of the master cyclic group polyhedron. Our chief contribution is a characterization of the adjacent facets and the extreme points of the mixed integer cut. This provides insight into the families of cuts that may work well in conjunction with the mixed integer cut. We further provide extensions of these results under mappings between group polyhedra.

For the remainder of this thesis we explore a framework for deriving multi-row cuts. For this purpose, we favor the method of superadditive lifting. This technique is largely driven by our ability to construct superadditive under-approximations of a special value function known as the lifting function. We devote our effort to precisely this task.

Chapter 3 reviews the theory behind superadditive lifting and returns to the classical problem of lifted flow cover inequalities. For this specific example, the lifting function we wish to approximate is quite complicated. We overcome this difficulty by adopting an indirect method for proving the validity of a superadditive approximation. Finally, we adapt

the idea to high-dimensional lifting problems, where evaluating the exact lifting function often poses an immense challenge. Thus we open entirely unexplored problems to the powerful technique of lifting.

Next, in Chapter 4, we consider the computational aspects of constructing strong superadditive approximations. Our primary contribution is a finite algorithm that constructs non-dominated superadditive approximations. This can be used to build superadditive approximations on-the-fly to strengthen cuts derived during computation. Alternately, it can be used offline to guide the search for strong superadditive approximations through numerical examples.

We follow up in Chapter 5 by applying the ideas of Chapters 3 and 4 to high-dimensional lifting problems. By working out explicit examples, we are able to identify non-dominated superadditive approximations for high-dimensional lifting functions. These approximations strengthen existing families of cuts obtained from single-row relaxations. Lastly, we show via the stable set problem how the derivation of the lifting function and its superadditive approximation can be entirely embedded in the computation of cuts.

Finally, we conclude by identifying future avenues of research that arise as natural extensions of the work in this thesis.

CHAPTER I

INTRODUCTION

This thesis primarily concerns methods that can be applied to mixed integer programming. In this chapter, we give a brief overview of mixed integer programming, including both applications and techniques used in the practical solution of these challenging problems.

Our discussion is by no means exhaustive, but provides a sufficient background for the reader less familiar with mixed integer programming. The reader more familiar with mixed integer programming and cutting plane theory may wish to skip ahead to the final section where we describe our contributions and outline the remainder of the thesis.

1.1 Mixed Integer Programming

The *mixed integer programming* problem is defined by

$$\begin{aligned} z_{\mathrm{MIP}} = \max\ & cx + hy \\ \text{s.t. }\ & Ax + Gy \leq b \\ & (x, y) \in \mathbf{R}^p_+ \times \mathbf{Z}^q_+. \end{aligned} \tag{1}$$

The vectors x and y are typically referred to as the *decision vectors*. The data (A, G, b) define a collection of *constraints* that the decision vectors must satisfy. Throughout we shall assume that $A \in \mathbf{Q}^{m\times p}$, $G \in \mathbf{Q}^{m\times q}$, and $b \in \mathbf{Q}^m$. The vectors $c \in \mathbf{Q}^p$, and $h \in \mathbf{Q}^q$ define the *objective function* $cx + hy$ that we wish to maximize. We say a point $(x, y) \in \mathbf{R}^p_+ \times \mathbf{Z}^q_+$ is *feasible* if it satisfies $Ax + Gy \leq b$. If no such point exists, we say that (1) is *infeasible*.

Despite its fairly simple structure, mixed integer linear programming enjoys a wide variety of applications because of its ability to capture problem structure. Some of the classical examples and now a staple of any text include the diet problem [71], the cutting stock problem [31, 32], the traveling salesman problem [1, 24], the matching problem [27], the network flow problem [29, 30], and the lot-sizing problem [66, 74]. Of course, this barely scratches the surface, and there are many more problems that fit into the framework of mixed integer programming.

There are a number of variants of (1) commonly encountered. When $q = 0$, this problem is the well-known *linear programming* problem; when $p = 0$, this problem is the *integer programming* or *pure integer programming* problem; and when $y \in \{0,1\}^q$, this problem is the *mixed binary integer programming* problem.

Despite their apparent similarities, these problems differ vastly with respect to their complexity. Consider for a moment the linear programming problem:

$$z_{\mathrm{P}} = \max\left\{cx : Ax \leq b, x \in \mathbf{R}^p_+\right\} \tag{2}$$

Both infeasibility and optimality can be efficiently certified as demonstrated in the next two theorems.

Theorem 1.1.1 (Farkas' lemma)**.** *There exists some $x \geq 0$ such that $Ax \leq b$ if and only if $yb \geq 0$ for all $y \geq 0$ such that $yA \geq 0$.*

Theorem 1.1.2 (LP duality)**.** *The linear program*

$$z_{\mathrm{D}} = \min\left\{yb : yA \geq c, y \in R^m_+\right\} \tag{3}$$

satisfies $z_{\mathrm{D}} = z_{\mathrm{P}}$ whenever both systems are feasible.

The linear programs (2) and (3) are known respectively as the *primal* and the *dual.* To certify the infeasibility of a linear program, one simply needs to produce some $y \geq 0$ such that $yA \geq 0$ and $yb < 0$. Similarly to certify optimality, one simply needs to produce a dual feasible solution y attaining the same objective value.

Notably, linear programming is amenable to a number of solution techniques. The first of these techniques, and still widely used today, is the simplex method [23]. This algorithm has become a cornerstone of linear programming theory and is a standard topic in most texts [13,59,69]. Since then, polynomial time algorithms such as the ellipsoid method [47,48] and interior point methods [45] have been developed for the linear programming problem.

In contrast, the dual of (1) resides in the space of superadditive functions, and is far less accessible than the linear programming dual. It is still unknown whether there exists any certificate of polynomial size for declaring the optimality of a feasible solution of a mixed integer program. In fact, binary integer programming is among Karp's famous 21 $\mathcal{NP}$-Complete problems [46], which suggests that the mixed integer programming problem in general may be quite different from linear programming.

Nevertheless, there are approaches to establishing feasibility and optimality that leverage the theory of cutting planes in conjunction with other techniques. We describe the geometry of polyhedra and cutting planes in the next section and then give a high level overview of their role in solving mixed integer programs.

1.2 The Geometry of Polyhedra and Cutting Planes

To develop the idea of cutting planes, it will be useful to explore the geometry of polyhedra. A *polyhedron* is a set

$$P = \{x \in \mathbf{R}^n : a_i x \leq b_i, \;\; i = 1, \ldots r\}$$

for some finite r. A more geometric perspective is to regard P as the intersection of the r half-spaces defined by the inequalities $a_i x \leq b_i$.

The set P resides in some affine subspace (i.e. a translation of a linear subspace). The *dimension* of P is the minimum dimension of an affine subspace containing P. Equivalently, if P contains at most $d+1$ affinely independent points, then P has dimension d. If $d = n$, then P is said to be *full-dimensional*. By applying an appropriate affine transformation, we may assume without loss of generality that P is full-dimensional; thus, we take P to have dimension n.

Next, we briefly discuss valid inequalities. An inequality $\pi x \leq \pi_0$ is said to be *valid* for P if it is satisfied by all $x \in P$. From a geometric perspective, this implies that P is contained in the closed half-space defined by $\pi x \leq \pi_0$. In this regard, one may either consider inequalities from an algebraic perspective or a geometric perspective as the context requires.

Some inequalities may be *redundant* in the sense that their exclusion does not change P. The remaining inequalities in a minimal description of P are called *facet-defining*, and are unique up to scaling. Given a facet-defining inequality $a_i x \leq b_i$, the corresponding *facet* is the set

$$F = P \cap \{x \in \mathbf{R}^n : a_i x = b_i\}.$$

A graphical representation of the distinction between valid inequalities and facet-defining inequalities is given in Figure 1.

We sometimes use the term "facet" rather loosely to refer either to F or the underlying inequality $a_i x \leq b_i$. Given our assumptions, F is itself a polyhedron of dimension $n-1$. Closely related are *faces*, which are obtained as the intersection of multiple facets, and of particular importance are *minimal faces* (with respect to non-emptiness). For our purposes, minimal faces typically have dimension 0 and are referred to as *extreme points*.

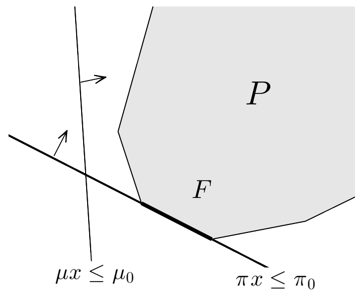

Figure 1: A valid inequality $\mu x \leq \mu_0$ and a facet-defining inequality $\pi x \leq \pi_0$ for an unbounded polyhedron P

Two important classes of polyhedra are *polytopes* and *polyhedral cones.* A polytope Q is defined as the convex hull of finitely many points: i.e.

$$Q = \text{conv}\left\{x^1, \ldots, x^s\right\} = \left\{ x \in \mathbf{R}^n : \begin{array}{l} x = \sum_{j=1}^{s} \lambda_j x^j \\ \sum_{j=1}^{s} \lambda_j = 1 \\ 0 \leq \lambda_j \leq 1, \quad j = 1, \ldots, s \end{array} \right\}.$$

Analogously, a polyhedral cone, C is defined as the conic hull of finitely many points. Thus

$$C = \text{cone}\left\{y^1, \ldots, y^t\right\} = \left\{ y \in \mathbf{R}^n : \begin{array}{l} y = \sum_{j=1}^{t} \mu_j y^j \\ y_j \geq 0, \quad j = 1, \ldots, s \end{array} \right\}.$$

If y^j is not in the conic hull of the remaining points, then y^j is called an *extreme ray.* As it turns out, every polyhedron can be described precisely in terms of polytopes and cones.

Theorem 1.2.1 (Decomposition for polyhedra)**.** *A set $P \subseteq \mathbf{R}^n$ is a polyhedron if and only if $P = Q + C$ for some polytope Q and some polyhedral cone C.*

Despite this nice geometric intuition, it is often more convenient to approach the facial structure of a polyhedron from an algebraic perspective. From this vantage, a valid inequality $a_i x \leq b_i$ is facet-defining if and only if there exist n affinely independent points $x^1, \ldots, x^n \in P$ satisfying $a_i x = b_i$. Furthermore, two facets $a_i x \leq b_i$ and $a_j x \leq b_j$ are *adjacent* if and only if they share $n-1$ affinely independent points. Extreme points therefore can be described as the solution of a system of equations

$$x_B = \{x : a_i x = b_i, \;\; i \in B\},$$

where the inequalities $a_i x \leq b_i$ define mutually adjacent facets. Thus adjacency is a useful property for identifying the extreme points of a polyhedron.

We now return to (1), and introduce the idea of cutting planes. The set

$$X = \left\{(x, y) \in \mathbf{R}^p_+ \times \mathbf{Z}^q_+ : Ax + Gy \leq b\right\}$$

is referred to as the *feasible region.* By the assumption that all data are rational, the *integer hull,* $P_{\text{I}} = \text{conv}(X)$, is itself a polyhedron. Conveniently,

$$z_{\text{MIP}} = \max\left\{cx + hy : (x, y) \in P_{\text{I}}\right\}.$$

Despite its simplicity, this result is a key component in the cutting plane paradigm. By optimizing over P_{I}, we have at our disposal all the machinery of linear programming.

Unfortunately, the set P_{I} is defined fairly abstractly, and is often not explicitly known. To this end, let $P_{\text{LP}} = \left\{(x, y) \in \mathbf{R}^p_+ \times \mathbf{R}^q_+ : Ax + Gy \leq b\right\}$. The maximization problem,

$$z_{\text{LP}} = \max\left\{cx + hy : (x, y) \in P_{\text{LP}}\right\}, \tag{4}$$

is known as the *linear programming relaxation* of (1).

Clearly $P_{\text{I}} \subseteq P_{\text{LP}}$, and typically this containment is strict. Thus, it is often the case that $z_{\text{MIP}} < z_{\text{LP}}$. Therefore, consider an extreme point solution (x^*, y^*) of (4). Either $y^* \in Z^q_+$, in which case $z_{\text{MIP}} = z_{\text{LP}}$, or $y^* \notin \mathbf{Z}^q_+$, in which case $(x^*, y^*) \notin P_{\text{I}}$. In the latter case, there exists an inequality $\pi x + \mu y \leq \pi_0$ that is valid for X (and hence P_{I}) such that $\pi x^* + \mu y^* > \pi_0$. Such an inequality is called a *cutting plane* or a *cut*. As seen in Figure 2, cuts need not define facets of the integer hull.

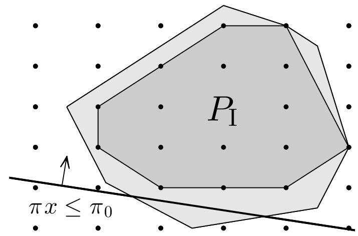

Figure 2: The integer hull, P_{I} and a cut $\pi_x \leq \pi_0$

1.3 Cutting Planes in the Solution of Mixed Integer Programs

The terminology of cutting planes is well-suited to its application in mixed integer programming. At a high level, the cutting plane scheme consists primarily of two steps. First the LP relaxation is solved producing some extreme point solution (x^*, y^*). If $y^* \in \mathbf{Z}_+^q$, then (x^*, y^*) is optimal. Otherwise, identify some cutting plane $\pi x + \mu y \leq \pi_0$, and introduce this constraint to the LP relaxation, cutting off the point (x^*, y^*). This process is repeated until a feasible solution is identified.

The earliest example of this approach appeared in the solution of a 49-city traveling salesman problem [24]. After employing some simple preprocessing reducing the problem to 42 cities, the authors are able to identify a tour with the addition of just seven subtour elimination constraints and two comb inequalities. Impressively, this feat was accomplished using only hand calculations.

Of course, haphazardly adding cutting planes would quickly render this approach hopelessly ineffectual. Therefore, one must seek well-behaved cutting planes to guarantee finite convergence.

Ideally, these cuts are facets of P_{I}. In some cases, such as matching [27, 28] and lot-sizing [9], a complete description of P_{I} is known. However, this is more often the exception than the rule. At best, we typically have a partial understanding of P_{I} that is manifested in the form of *cut classes* or *cut families*. These may either be derived from P_{I} itself or from some subsystem of $Ax + Gy \leq b$. Some well-known families of cuts are the subtour and comb inequalities for the traveling salesman problem [16, 24, 39]; the clique and odd-hole inequalities for set packing [63]; the cover inequalities for the knapsack problem [5,8,40,58]; the ℓ-S inequalities for lot-sizing [9]; and the flow cover inequalities for fixed-charge network flow [41, 64, 72].

The sufficiency of an incomplete description of P_{I} to characterize optimal solutions is highly dependent on the objective function; for example, Wagner-Whitin costs greatly simplify the facets needed to describe the optimal solution in many lot-sizing problems [65]. Relying on this behavior is tenuous at best. Therefore, many solvers come equipped with an arsenal of general-purpose cutting plane techniques that can be applied to any mixed integer program.

The potential of cutting planes for solving general problems was first demonstrated by Gomory in the setting of pure integer programming [34]. Here, he describes a procedure that correctly identifies a sequence of cutting planes that is guaranteed to produce an integer solution in finitely many iterations. Gomory later extended this result to mixed integer programming [33]. The driving force behind these finitely-convergent algorithms are the Gomory fractional cut and the mixed integer cut. Since then, a number of cutting planes have been developed for mixed binary integer programs such as lift-and-project cuts [6] and Lovász-Schrijver cuts [51] and for mixed integer programs such as the mixed integer rounding cuts [60] and split cuts [21].

Classifying cutting planes is not strictly limited to either problem specific cuts or generic cuts. One particularly successful avenue of research is aimed at characterizing "simple" mixed integer sets that possess some non-trivial structure, but can be used to derive valid inequalities for a wide array of problems. The mixing set and its variants have recently enjoyed considerable attention [17–19, 43, 73]. The so-called mixing inequalities have been used to expand existing families of cuts and have found further application in mixed integer programs derived from probabilistic constraints [49, 52].

Despite the broad selection of cuts at our disposal, pure cutting planes algorithms are often subject to numerical instability. At the heart of this issue, adding cuts alters which cuts can subsequently be derived. Gomory's original cutting plane algorithm overcomes this obstacle by precisely identifying which cuts to add, but in fact, utterly fails under different cut selection criteria [78].

An alternate approach to solving mixed integer programs is *branch-and-bound*. In the context of mixed integer programming, branch-and-bound operates by solving the LP relaxation to produce some optimal (x^*, y^*). If y^* is integral, then (x^*, y^*) is optimal for (1). Otherwise, there exists some $y_j^* \notin \mathbf{Z}$, and the problem can be split into two disjoint subproblems: one with $y_j \leq \lfloor y_j^* \rfloor$ and the other with $y_j \geq \lceil y_j^* \rceil$. Trivially, the maximum of these two subproblems will be a maximum for the original problem. This process is applied recursively until the resulting subproblem has an integral optimum, is infeasible, or can be eliminated using global bound data (typically by comparing the objective value of the LP relaxation with best known integral solution). This can be represented by a branch-and-bound tree as in Figure 3.

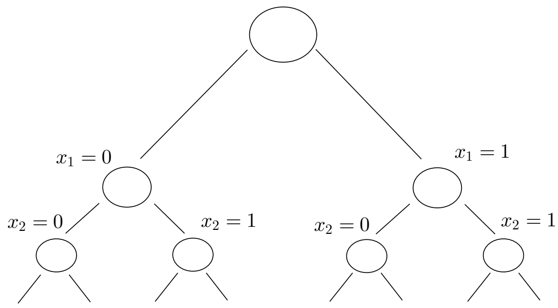

Figure 3: A typical branch-and-bound tree for a binary integer program

The theory behind a pure cutting plane approach tends to understate the computational challenges that inevitably arise. Certainly, early work identified the potential of cutting planes, but this was infrequently realized in practice. General purpose cuts were relegated to theory, but once again found life in the landmark work of Crowder, Johnson, and Padberg [22]. Their approach incorporated general cutting planes into a traditional branch-and-bound framework–a method now known as *branch-and-cut*–to solve binary integer programs. This idea was extended to mixed integer programming [7], and is now a major driver in the practical solution of mixed integer programs.

Branching strategies are no less important than cutting planes, and can vastly improve solution times. For example, branching can be made to accommodate column generation [10] in the solution of high-dimensional integer programs. Another clever branching strategy can break the symmetry that would otherwise hamstring a naïve branch-and-cut implementation [54,55,61,62]. Indeed, branching rules play just as integral in modern day solvers, and merit continued research.

1.4 Thesis Outline

For this dissertation, we focus on the role of cutting planes in mixed integer programming. Our goal is to improve the quality of cuts used in a branch-and-cut framework. Within this context, we wish to obtain cuts that produce better bounds earlier in the solution process. This in turn can greatly reduce the number of branching steps required to solve a mixed integer program, improving solve times.

In Chapter 2, we further develop the structure of corner polyhedra. These polyhedra provide a framework for understanding and deriving cuts for pure integer programs. Most notably, the mixed integer cut is one of many cuts that arise in this setting. We give a structural description of the mixed integer cut, specifically characterizing its adjacent facets. It is not unreasonable that these adjacent facets may work well in conjunction with the mixed integer cut.

We next explore the role of lifting in developing cuts in Chapter 3. In particular, we expand on previous work in superadditive lifting to show how the technique can be used to derive cuts without explicitly solving the lifting problem. We show this idea first in the context of fixed-charge network flow, and then provide a more general template for how this can be achieved to simplify the approximation of higher dimensional lifting functions. Using this approach, it is possible to derive cuts that might otherwise be prohibitively expensive to obtain.

In Chapter 4, we probe deeper into the structure of superadditive lifting functions. Our primary contribution is an algorithm that can construct strong superadditive approximations of piecewise linear functions used in the derivation of cuts. Subsequently, we also obtain a concise description of the "strongest" lifting functions. This work is the first of its kind and provides a tool that can either be used offline to assist in the derivation of new families of cuts, or can be used online to enable the application of superadditive lifting in a more general setting.

In Chapter 5, we show how the ideas from Chapters 3 and 4 can be applied to several classical problems in mixed integer programming. We first show how one can construct superadditive approximations for a fixed-charge flow set that is additionally constrained by a knapsack inequality. Next, we consider the intersection of multiple knapsacks, and show how traditional cover inequalities can incorporate additional problem structure. Lastly, we conclude by revisiting the odd-hole inequalities for the stable set polytope and demonstrate how one might modify the problem to accommodate superadditive lifting.

Finally, in Chapter 6, we revisit our main contributions and discuss some future work.

CHAPTER II

GOMORY'S GROUP PROBLEM REVISITED

The first topic we explore in cutting plane theory is Gomory's group problem (also known as the corner relaxation). The connection between integer programming and groups is embodied in the periodicity of solutions. All other things remaining fixed, as the right hand side of an integer program is increased, certain parts of the solution begin to repeat. This is most easily observed in the context of knapsack problems, but is entirely general.

Many of the recurring themes in mixed integer programming find their roots in the group problem. In many respects, the master cyclic group polyhedron is among the first simple integer sets to be studied extensively. For example, the mixed integer cut enjoys a very natural representation in this setting and can be extended using various group operations. Perhaps the most noteworthy of these extensions arises by using homomorphisms between groups. It is here that the idea of lifting, which we study extensively in subsequent chapters, finds its origin.

Similarly, recent work studying the relation between lattice-free convex sets and valid inequalities build off of ideas originating from the corner relaxation (see for example [11,12, 15]). Whereas the group problem we consider in this chapter is finite, the work in lattice-free convex sets is more closely connected to the infinite group problem [36, 37, 44]. This work establishes theoretical foundations for multi-row cuts that cannot be obtained by simple constraint aggregation.

In this chapter, we characterize the mixed integer cut as a facet of the master cyclic group polyhedron. This is a significant step in understanding one of the most widely used cuts in the practical solution of mixed integer programs, and it may shed light on possible cuts that tend to work well in conjunction with the mixed integer cut. We primarily build off of the work of Aráoz et al [3] and show that their tilted knapsack facets are the only non-trivial facets adjacent to the mixed integer cut. Lastly, we extend earlier results on various group mappings to provide extensions of our characterization under homomorphic lifting.

2.1 Algebraic Prerequisites

Before going into depth about the group problem, we take a moment to review some standard definitions and useful theorems from algebra in order to provide a clearer and more complete exposition for the reader less familiar with the subject. We will further provide numerical examples that we develop throughout our discussion to highlight our results. Although the main ideas presented here are common throughout any algebra text, we borrow much of our notation from [26].

Definition 2.1.1. Let the pair $(\mathcal{G}, \star)$ denote a set $\mathcal{G}$, and a binary operation $\star : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$. We say that $(\mathcal{G}, \star)$ is a *group* if it satisfies the following axioms:

(i) $(a \star b) \star c = a \star (b \star c)$ for all $a, b, c \in \mathcal{G}$ (i.e. $\star$ is associative),

(ii) there exists an element $e \in \mathcal{G}$ called the *identity* of $\mathcal{G}$, such that $a \star e = e \star a = a$ for all $a \in \mathcal{G}$,

(iii) for each $a \in \mathcal{G}$ there is an element $a^{-1} \in \mathcal{G}$ called the *inverse* of a such that $a \star a^{-1} = a^{-1} \star a = e$.

Further if $a \star b = b \star a$, then the group $(\mathcal{G}, \star)$ is called *abelian*.

For brevity, we will typically refer to the group by its underlying set $\mathcal{G}$ and we will suppress $\star$ entirely when the underlying operation is understood from context. When speaking of abelian groups we often replace $\star$ with $+$ and e with 0.

The groups that we study in this chapter will all be finite and abelian, so we restrict our examples accordingly.

Example 2.1.1. The following two examples most closely reflect the groups we study:

1. Let $\mathbf{Z}_m$ denote the set $\{0, \ldots, m-1\}$ under addition modulo m where m is some finite integer. It is easy to verify that $\mathbf{Z}_m$ is a group with $e = 0$ and $a^{-1} = m - a$ for $a \neq 0$.

2. Let $\Lambda = \{Az : z \in \mathbf{Z}^n\}$. Then Λ is a group under addition with $e = 0$ and for $a = Az$, $a^{-1} = -a = A(-z)$.

Let $g \in \mathcal{G}$, and define $g^0 = e$. For $k > 0$, let $g^k = g^{k-1} \cdot g$ and $g^{-k} = g^{-(k-1)} \cdot g^{-1}$. In the context of abelian groups, g^k is typically denoted kg.

The group $\mathbf{Z}_m$ in the above example represents a special kind of group called a *cyclic group* that will be the primary focus of our study of the group problem. A group $\mathcal{G}$ is called *cyclic* if there exists some element $g \in \mathcal{G}$ such that $\mathcal{G} = \left\{g^k : k \in \mathbf{Z}\right\}$.

When we get into the details of the group problem, we will need to use some elementary results from algebra. The following theorem will be explicitly used later:

Theorem 2.1.1. *If $|\mathcal{G}| = n$ is finite, then $g^n = e$ for all $g \in \mathcal{G}$.*

Beyond the groups themselves, we will also refer explicitly to well-behaved mappings between groups.

Definition 2.1.2. Let $\mathcal{G}$ and $\mathcal{H}$ be groups, and let $\varphi : \mathcal{G} \to \mathcal{H}$. We say that φ is a *homomorphism* if $\varphi(xy) = \varphi(x)\varphi(y)$. If $\mathcal{H} = \mathcal{G}$ and φ is bijective, then φ is more specifically called an *automorphism*.

We will assume that φ is surjective by restricting $\mathcal{H}$ to the image of $\mathcal{G}$. If φ is also injective, then the sets $\mathcal{G}$ and $\mathcal{H}$ are said to be *isomorphic*, denoted by $\mathcal{G} \cong \mathcal{H}$, and φ is called an *isomorphism*. As we will see in the next definition, homomorphisms naturally induce a partition of $\mathcal{G}$:

Definition 2.1.3. Let φ be a homomorphism from $\mathcal{G}$ onto $\mathcal{H}$, and $\mathcal{K} = \{g \in \mathcal{G} : \varphi(g) = e_{\mathcal{H}}\}$. $\mathcal{K}$ is referred to as the *kernel of* φ. For any $g \in \mathcal{G}$ the set $g\mathcal{K} = \{gk : k \in \mathcal{K}\}$ is called a *coset* of $\mathcal{G}$ and g is referred to as its *coset representative*.

Observe that all elements in $g\mathcal{K}$ map to the same value under φ. Furthermore, if $\varphi(g_1) = \varphi(g_2)$ then $g_2 \in g_1\mathcal{K}$. As $g \in g\mathcal{K}$ the cosets partition $\mathcal{G}$.

Definition 2.1.4. The set $\mathcal{G}/\mathcal{K} = \{g\mathcal{K} : g \in \mathcal{G}\}$ with $(g_1\mathcal{K})(g_2\mathcal{K}) = (g_1g_2)\mathcal{K}$ is called the *quotient group* of $\mathcal{G}$. Moreover, $\mathcal{G}/\mathcal{K} \cong \mathcal{H}$ (the isomorphism being $\vartheta(gK) = \varphi(g)$).

There is one key result about the cosets and the quotient group induced by a homomorphism that is explicitly used in this chapter.

Theorem 2.1.2. *Suppose that $\mathcal{G}$ is finite, and let $\varphi : \mathcal{G} \to \mathcal{H}$ be a homomorphism with kernel $\mathcal{K}$. Then $|g\mathcal{K}| = |\mathcal{K}|$ for all $g \in \mathcal{G}$ and $|\mathcal{H}||\mathcal{K}| = |\mathcal{G}|$.*

Example 2.1.2. The homomorphisms and automorphisms we consider later in this chapter closely resemble the following examples:

1. Let $\mathcal{G} = \mathbf{Z}$ and $\mathcal{H} = \mathbf{Z}_m$ under addition (taken modulo m in the latter set). Let $\varphi : \mathbf{Z} \to \mathbf{Z}_m$, map z to its residue modulo m. Then φ is a homomorphism, with kernel $\mathcal{K} = m\mathbf{Z} = \{mz : z \in \mathbf{Z}\}$. Further the cosets $h + \mathcal{K} = \{z \in Z : z = h + km\}$ partition $\mathbf{Z}$.

2. Let $\mathcal{G} = \mathbf{Z}_7$ and let $\varphi : \mathbf{Z}_7 \to \mathbf{Z}_7$ be defined by

$$\varphi(z) = 3z \pmod 7.$$

   Then φ is an automorphism; for if $\varphi(z_1) = \varphi(z_2)$ then $3(z_1 - z_2) = 7z_3$. Therefore $z_1 = z_2$ and hence φ is bijective.

Observe that in this first example $m\mathbf{Z}$ is a one-dimensional lattice. More generally consider the lattice Λ of Example 2.1.1 when A is restricted to be non-singular and integer valued. In this context, it is reasonable to define congruence modulo A:

Definition 2.1.5. For $z_1, z_2 \in \mathbf{Z}^n$,

$$z_1 = z_2 \pmod A \Leftrightarrow (z_1 - z_2) = Az$$

for some $z \in \mathbf{Z}^n$.

Analogous to the one-dimensional case, let $\mathcal{G} = \{z + \Lambda : z \in \mathbf{Z}^n\}$. Every element of $\mathbf{Z}^n$ naturally maps to its coset which is a translate of the lattice generated by A. The set $\mathcal{G}$ is finite and forms a group where addition is defined by adding the coset representatives. It is this group that will be most relevant in our discussion of the group problem.

2.2 Master Group Polyhedra

Now that we have highlighted the algebra underlying the group problem, we can introduce the group relaxation and master group polyhedra. Consider the integer program

$$z_{\mathrm{IP}} = \min \left\{ cx : Ax = b, x \in Z_+^n \right\},$$

where all data are assumed to be rational and A is $m \times n$ and b is $m \times 1$. Given a solution to the LP relaxation, we can partition A into basic and non-basic columns, hence we can rewrite z_{IP} as

$$z_{\mathrm{IP}} = \min \left\{ c_B x_B + c_N x_N : B x_B = b - N x_N, x_B \in Z_+^m, x_N \in Z_+^{n-m} \right\}.$$

Now observe that $x_B \in \mathbf{Z}^m$ if and only if $b - N x_N$ belongs to the lattice $\Lambda = \{Bz : z \in Z^m\}$. However, this condition on its own is insufficient to guarantee the non-negativity of x_B.

As B is invertible, this implies that $x_B = B^{-1}(b - N x_N)$. Therefore, if we relax the non-negativity of x_B, we can eliminate these variables entirely from the problem, yielding the *group relaxation*:

$$z_{\mathrm{GR}} = z_{\mathrm{LP}} + \min \left\{ \tilde{c}_N x_N : N x_N \equiv b \pmod{B},\ x_N \in \mathbf{Z}_+^{n-m} \right\} \tag{5}$$

where $z_{\mathrm{LP}} = c_B B^{-1} b$ denotes the objective of the LP relaxation and $\tilde{c}_N = (c_N - c_B B^{-1} N)$ denotes the reduced cost vector. The corresponding set in the original space of variables is often called the *corner polyhedron* and is depicted in Figure 4.

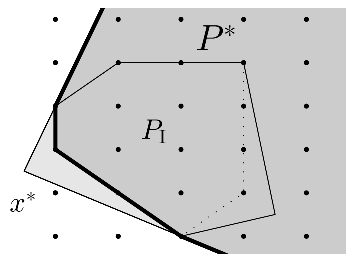

Figure 4: The corner polyhedron P^* obtained from the group relaxation for x^*

Gomory explores the properties of this relaxation and shows sufficient, although by no means necessary, conditions that guarantee the non-negativity of x_B and hence tightness of the group relaxation. However, these conditions are generally too restrictive to be of any practical use.

Instead, the utility of the group relaxation derives from its polyhedral structure. With this in mind we explore the underlying set, i.e.

$$X = \left\{ x \in \mathbf{Z}_+^{n-m} : Nx = b \pmod{B} \right\}.$$

Now let $\Lambda = \{Bz : z \in \mathbf{Z}^m\}$ and $\mathcal{G} = \{z + \Lambda : z \in Z^m\}$. Then every column of N naturally maps to an element of $\mathcal{G}$: namely its coset which is a translate of Λ. Define

$$\mathcal{N} = \{g \in \mathcal{G} \mid \exists j : N_j + \Lambda = g\}.$$

In plain terms $\mathcal{N}$ denotes the collection of cosets represented by the columns of N. Let $d = |\mathcal{N}|$. As not all $g \in \mathcal{N}$ may be uniquely represented by a column of N, it is quite possible that $d < n - m$. Let $g_0 = b + \Lambda$ and let

$$t(g) = \sum_{j : N_j + \Lambda = g} x_j.$$

At this point the algebraic structure of the set emerges:

$$X(\mathcal{G}, \mathcal{N}, g_0) = \left\{ t \in \mathbf{Z}_+^d : \sum_{g \in \mathcal{N}} g \cdot t(g) = g_0 \right\}, \tag{6}$$

where equality is now taken with respect to the group operation. We will refer to the set $P(\mathcal{G}, \mathcal{N}, g_0) = \text{conv}\{X(\mathcal{G}, \mathcal{N}, g_0)\}$ as the *group polyhedron.*

Valid inequalities for $P(\mathcal{G}, \mathcal{N}, g_0)$ map to valid inequalities of X by replacing $t(g)$ with an appropriate sum. The study of $P(\mathcal{G}, \mathcal{N}, g_0)$ is not itself transparent; hence it is more natural to consider the *master group polyhedron*:

$$P(\mathcal{G}, g_0) = \text{conv} \left\{ t \in \mathbf{Z}_+^{|\mathcal{G}^+|} : \sum_{g \in \mathcal{G}^+} g \cdot t(g) = g_0 \right\}, \tag{7}$$

where $\mathcal{G}^+ = \mathcal{G} \setminus \{0\}$. If $\mathcal{G}$ is a cyclic group, we refer to the above system as the *master cyclic group polyhedron.* Throughout, we shall assume that $g_0 \neq 0$ as this case has considerably less practical appeal.

We now give a concrete example with a one-row system to show an explicit derivation of this system.

Example 2.2.1. Let X be the integral system defined by the constraint

$$7x_1 + 13x_2 + 3x_3 - x_4 + 5x_5 + 11x_6 = 24.$$

Taking $B = 7$, the group relaxation is obtained by replacing equality with equivalence modulo 7. Thus we have

$$13x_2 + 3x_3 - x_4 + 5x_5 + 11x_6 = 24 \pmod 7.$$

Each column can be replaced with its residue modulo 7 which is an equivalent coset representative. Therefore, the above system is identical to

$$6x_2 + 3x_3 + 6x_4 + 5x_5 + 4x_6 = 3 \pmod 7,$$

and the corresponding group polyhedron is defined by the constraint

$$3t_3 + 4t_4 + 5t_5 + 6t_6 = 3 \pmod 7.$$

Therefore, the master cyclic group polyhedron for this system is

$$P(\mathcal{C}_7, 3) = \text{conv}\left\{t \in Z_+^6 : t_1 + 2t_2 + 3t_3 + 4t_4 + 5t_5 + 6t_6 = 3 \pmod 7\right\}.$$

We shall return to this example later as we explore in more detail the polyhedral structure of the set.

As suggested earlier, master group polyhedra provide a mechanism to obtain valid inequalities for specific instances of group polyhedra. The specific group polyhedron occurs as a face of its corresponding master group polyhedron:

$$P(\mathcal{G}, \mathcal{N}, g_0) = P(\mathcal{G}, g_0) \cap \{t : t(g) = 0, g \notin \mathcal{N}\}.$$

Thus, for any valid inequality $\pi t \geq \pi_0$ of the master polyhedron, a valid inequality of the group polyhedron can be obtained by including only the columns corresponding to $\mathcal{N}$.

Before getting into too much depth about valid inequalities, we take a moment to introduce a shorthand notation to describe an inequality. Namely,

$$(\pi, \pi_0) := \sum_{g \in \mathcal{G}^+} \pi(g)t(g) \geq \pi_0. \tag{8}$$

In the subsequent discussion, the dimension of π and π_0 are understood from context.

Returning to $P(\mathcal{G}, g_0)$, we are specifically interested in the inequalities (π, π_0) that define facets of this set. In his seminal work on corner polyhedra [35], Gomory fully characterizes the polar. This description has since been revisited and refined in the works of [70].

Theorem 2.2.1. *(π, π_0), $\pi_0 > 0$, is a non-trivial facet of $P(\mathcal{G}, g_0)$, $g_0 \neq 0$ if and only if it is a basic feasible solution to the system of equations and inequalities:*

$$\begin{aligned} &\pi(g_0) = \pi_0 \\ &\pi(g) + \pi(g_0 - g) = \pi_0, && g \in \mathcal{G}^+, g \neq g_0 \\ &\pi(g) + \pi(g') \geq \pi(g + g'), && g, g' \in \mathcal{G}^+ \\ &\pi(g) \geq 0, && g \in \mathcal{G}^+. \end{aligned} \tag{9}$$

Although the number of facets of $P(\mathcal{G}, g_0)$ typically grows exponentially with $|\mathcal{G}|$, its polar can still be expressed succinctly. This fact was exploited in [38] to investigate the strength of facets via shooting experiments. At a high level, a non-negative vector is randomly generated, and a linear program based on (9) is solved to identify the first facet hit along this direction. Despite the large number of facets of the polyhedra tested, a surprisingly small number of facets received the majority of hits, indicating their importance in characterizing $P(\mathcal{G}, g_0)$.

Typically these facets could be classified either as the mixed integer cut or some mapping of a mixed integer cut. It is precisely these facets and mappings that we study; hence, we will conclude this section by introducing the relevant theorems and returning to our example.

Theorem 2.2.2. *Let n be integral and $r \in \mathcal{C}_n^+$. The inequality $(\mu, 1)$ with*

$$\mu_i = \begin{cases} \frac{i}{r} & i \leq r \\ \frac{n-i}{n-r} & i > r \end{cases} \tag{10}$$

is facet-defining for $P(\mathcal{C}_n, r)$.

The inequality $(\mu, 1)$ is called the *mixed integer cut* and has played a central role throughout pure and mixed integer programming. Even though this inequality hardly resembles the mixed integer cut that most are familiar with, we show as an illustrative example that they are in fact the same.

Example 2.2.2. Consider the one row integer program where all data are rational and $b \notin \mathbf{Z}$:

$$X = \left\{ x \in \mathbf{Z}_+^p : \sum_{j \in J} a_j x_j = b \right\},$$

and let $P = \text{conv}(X)$. Let $f_j = a_j - \lfloor a_j \rfloor$ and $f_0 = b - \lfloor b \rfloor$. The mixed integer cut for this set is the inequality

$$\sum_{j \in J : f_j \leq f_0} f_j x_j + \frac{f_0}{1 - f_0} \sum_{j \in J : f_j > f_0} (1 - f_j) x_j \geq f_0.$$

Let N be chosen such that $Na_j \in \mathbf{Z}$ for all $j \in J$ and $Nb \in \mathbf{Z}$. Therefore, we can scale by N so that all coefficients are integral,

$$X = \left\{ x \in \mathbf{Z}_+^p : \sum_{j \in J} (\lfloor a_j \rfloor + f_j) \cdot N x_j = (\lfloor b \rfloor + f_0) \cdot N \right\}.$$

Now consider the group relaxation obtained by replacing the equality with equivalence modulo N. It follows that a_j maps to $f_j \cdot N$ for each $j \in J$ and b maps to $f_0 \cdot N$. The mixed integer cut for $P(\mathcal{C}_N, f_0 N)$ is given by

$$\mu_i = \begin{cases} \frac{i}{f_0 \cdot N} & i \leq f_0 \cdot N \\ \frac{N - i}{N - f_0 \cdot N} & i > f_0 \cdot N. \end{cases}$$

Therefore, to recover a valid inequality, we simply read off coefficients. If $f_j \leq a_j$ then x_j receives coefficient $\frac{f_j}{f_0}$; otherwise x_j receives coefficient $\frac{1-f_j}{1-f_0}$. So the two inequalities indeed coincide.

Using automorphisms and homomorphisms, it is possible to produce facets for $P(\mathcal{G}, g_0)$ from related polyhedra. Let $\varphi : \mathcal{G} \to \mathcal{G}$ be a non-trivial automorphism (i.e. there exist some g such that $\varphi(g) \neq g$).

Theorem 2.2.3. *If (π, π_0) is a facet of $P(\mathcal{G}, g_0)$, with components, $\pi(g)$, then (π', π_0) with components $\pi'(g) = \pi(\varphi^{-1}(g))$ is a facet of $P(\mathcal{G}, \varphi(g_0))$.*

Similarly, consider some homomorphism $\psi : \mathcal{G} \to \mathcal{H}$ such that $\psi(g_0) = h_0 \neq 0$.

Theorem 2.2.4. *Let (π, π_0) be a non-trivial facet of $P(\mathcal{H}, h_0)$. Then (π', π_0) is a facet of $P(\mathcal{G}, g_0)$ where $\pi'(g) = \pi(\psi(g))$ for all $g \in \mathcal{G} \setminus \mathcal{K}$, and $\pi'(k) = 0$ for all $k \in \mathcal{K}$.*

As described before, a homomorphism ψ induces a partition of $\mathcal{G}$ into cosets. In *homomorphic lifting* each element of a coset receives the same coefficient.

We conclude by giving a concrete example.

Example 2.2.3. Recall that

$$P(\mathcal{C}_7, 3) = t_1 + 2t_2 + 3t_3 + 4t_4 + 5t_5 + 6t_6 = 3.$$

The mixed integer cut for $P(\mathcal{C}_7, 3)$ is given by

$$\frac{1}{3} \cdot t_1 + \frac{2}{3} \cdot t_2 + \frac{3}{3} \cdot t_3 + \frac{3}{4} \cdot t_4 + \frac{2}{4} \cdot t_5 + \frac{1}{4} \cdot t_6 \geq 1.$$

Now consider the automorphism of φ of $\mathcal{C}_7$ that sends g to $4g$. This gives

g	0	1	2	3	4	5	6
$\varphi(g)$	0	4	1	5	2	6	3

Therefore, applying Theorem 2.2.3 to the mixed integer cut for $P(\mathcal{C}_7, 3)$ we obtain the inequality

$$\frac{2}{3} \cdot t_1 + \frac{3}{4} \cdot t_2 + \frac{1}{4} \cdot t_3 + \frac{1}{3} \cdot t_4 + \frac{3}{3} \cdot t_5 + \frac{2}{4} \cdot t_6 \geq 1.$$

as a facet of $P(\mathcal{C}_7, 5)$. Next we consider a homomorphism $\psi : \mathcal{C}_{14} \to \mathcal{C}_7$ such that $\psi(g) = g \pmod 7$. By applying Theorem 2.2.4 to the mixed integer cut for $P(\mathcal{C}_7, 3)$, we have that

$$\begin{aligned} &\frac{1}{3} \cdot t_1 + \frac{2}{3} \cdot t_2 + \frac{3}{3} \cdot t_3 + \frac{3}{4} \cdot t_4 + \frac{2}{4} \cdot t_5 + \frac{1}{4} \cdot t_6 \\ &\quad + \frac{1}{3} \cdot t_8 + \frac{2}{3} \cdot t_9 + \frac{3}{3} \cdot t_{10} + \frac{3}{4} \cdot t_{11} + \frac{2}{4} \cdot t_{12} + \frac{1}{4} \cdot t_{13} \geq 1. \end{aligned}$$

is a facet of $P(\mathcal{C}_{14}, 10)$ and $P(\mathcal{C}_{14}, 3)$.

The corresponding mixed integer cut and its related facets obtained through automorphisms and homomorphisms are depicted in Figures 5 and 6.

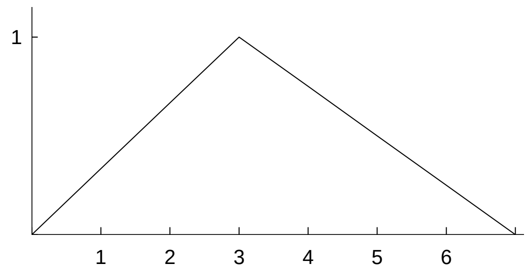

Figure 5: The mixed integer cut coefficients

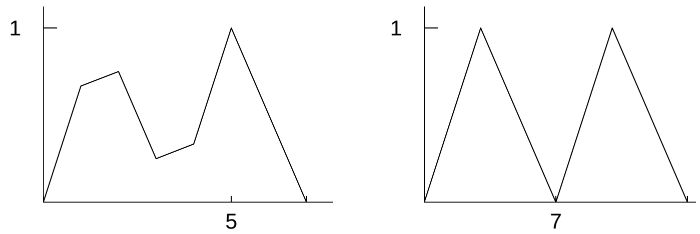

Figure 6: The mixed integer cut coefficients under automorphic mapping (left) and homomorphic lifting (right)

2.3 Master Knapsack Polyhedra and the Mixed Integer Cut

Our primary result in this section is a characterization of the extreme points and adjacent facets of the mixed integer cut. Along these lines, it will be easier to regard the mixed integer cut as the polytope:

$$P_{\text{MIC}}(n,r) = P(\mathcal{C}_n, r) \cap \{t : \mu t = 1\}. \tag{11}$$

In this setting, a facet (π, π_0) of $P(\mathcal{C}_n, r)$ is adjacent to $(\mu, 1)$ if (π, π_0) defines a facet of $P_{\text{MIC}}(n,r)$.

For our characterization, we will also need the *master knapsack polyhedron*:

$$P(K_m) = \left\{ x \in \mathbf{Z}_+^m : \sum_{i=1}^{m} i \cdot x_i = m \right\}. \tag{12}$$

As we will show, this set is intrinsically related to $P_{\text{MIC}}(n,r)$. Like the group problem, facets of the master knapsack problem can be described by a small system of inequalities [2]:

Theorem 2.3.1. *The facets (ρ, ρ_m) of $P(K_m)$ are extreme rays of the cone defined by*

$$\begin{aligned} \rho_i + \rho_j &\geq \rho_{i+j} \quad & 1 \leq i, j, i+j \leq m, \\ \rho_i + \rho_{m-i} &= \rho_m \quad & 1 \leq i \leq \left\lfloor \frac{m}{2} \right\rfloor. \end{aligned}$$

Using this characterization, the authors in [3] are able to derive a new class of non-trivial facets of $P(\mathcal{C}_n, r)$ from non-trivial facets of $P(K_r)$.

The next theorem describes a class of facets of $P(\mathcal{C}_n, r)$ called *tilted knapsack facets* introduced in [3].

Theorem 2.3.2. *Let (ρ, ρ_r) be a non-trivial facet of $P(K_r)$ such that $\rho \geq 0$, $\rho_i = 0$ for at least one i, and $\rho_r = 1$. Let*

$$\overline{\rho} = \left(\rho_1, \ldots, \rho_r = 1, \frac{n-r-1}{n-r}, \ldots, \frac{1}{n-r}\right).$$

Then there exists some $\alpha \in \mathbf{R}$ such that $(\pi, \pi_0) = (\overline{\rho} + \alpha\mu, 1 + \alpha)$ is a facet of $P(\mathcal{C}_n, r)$.

Details for computing the *tilting coefficient*, α, are contained in [3]. Although not stated explicitly in the original work, these facets can also be obtained from $P(K_{n-r})$.

Corollary 2.3.3. *Let (ρ, ρ_{n-r}) be a non-trivial facet of $P(K_{n-r})$ such that $\rho \geq 0$, $\rho_i = 0$ for at least one i, and $\rho_{n-r} = 1$. Let*

$$\overline{\rho} = \left(\frac{1}{r}, \ldots, \frac{r-1}{r}, 1 = \rho_{n-r}, \rho_{n-r-1}, \ldots, \rho_1\right).$$

Then there exists some $\alpha \in \mathbb{R}$ such that $(\pi, \pi_0) = (\overline{\rho} + \alpha\mu, 1 + \alpha)$ is a facet of $P(\mathcal{C}_n, r)$.

Proof. Observe that the mapping $\varphi : \mathcal{C}_n \to \mathcal{C}_n$ defined by $\varphi(i) = n - i$ for $i \in \{1, \ldots, n-1\}$ is an automorphism of $\mathcal{C}_n$. By Theorem 2.2.3 with φ and Theorem 2.3.2 applied to $P(K_{n-r})$ the result follows. □

The conditions placed on (ρ, ρ_r) are without loss of generality. For completeness we describe how any facet of $P(K_r)$ can be made to conform to these conditions.

Proposition 2.3.4. *Let (ρ, ρ_0) be a non-trivial facet of $P(K_m)$. Without loss of generality we may assume that $(\rho, \rho_0) \geq 0$, $\rho_0 = \rho_m = 1$. Moreover, we may assume there exists some $i \neq m$ such that $\rho_i = 0$.*

Proof. We may add the knapsack equation to (ρ, ρ_0) so that $(\rho, \rho_0) \geq 0$. Since $x = e_m$ is a feasible solution, $\rho_0 \leq \rho_m$. If $\rho_0 < \rho_m$, then $x_m = 0$ for all x satisfying (ρ, ρ_0) at equality, contradicting that the facet is non-trivial. Furthermore, if $\rho_0 = 0$ the inequality is implied by non-negativity constraints; thus $\rho_0 > 0$, and we may assume by scaling that $\rho_0 = 1$.

Now if for all $i, \rho_i \geq \frac{i}{m}$, we may subtract the knapsack equation scaled by $\frac{1}{m}$ to yield an inequality implied by the non-negativity constraints (since the right hand side is 0 and the left hand side is non-negative). Therefore, there must exist some $\rho_i < \frac{i}{m}$, so by subtracting the appropriately scaled knapsack equation, and rescaling, we may further assume that there exists some $i \neq m$ such that $\rho_i = 0$. □

Hence, every non-trivial facet of $P(K_r)$ maps to a corresponding facet of $P(\mathcal{C}_n, r)$. We will show these tilted knapsack facets are in fact facets of $P_{\text{MIC}}(n, r)$. Before doing so, we describe an operation we call *extending* a knapsack solution.

Proposition 2.3.5. *If $x \in P(K_r)$, $x = (x_1, \ldots, x_r)$, then $t = (x_1, \ldots, x_r, 0, \ldots, 0) \in P_{\text{MIC}}(n, r)$. If $x \in P(K_{n-r})$, $x = (x_1, \ldots, x_{n-r})$, then $t = (0, \ldots, 0, x_{n-r}, \ldots, x_1) \in P_{\text{MIC}}(n, r)$.*

Proof. For $x \in P(K_r)$, the result is trivial. So take $x \in P(K_{n-r})$. Since $P(K_{n-r})$ is convex and integral, we may assume that x is integral. Rewriting $i = n - (n - i)$ for $i = 1, \ldots, r$ and applying the assumption that x is an integer knapsack solution

$$\sum_{i=r}^{n-1} it_i = \sum_{i=r}^{n-1} nx_{n-i} - \sum_{i=r}^{n-1}(n-r)x_{n-i} = r \pmod{n},$$

and the proposition follows. □

Note that $P(\mathcal{C}_n, r)$ is full-dimensional since its recession cone is the non-negative orthant. As $(\mu, 1)$ defines a facet of $P(\mathcal{C}_n, r)$, it must have dimension $n - 2$.

Observation 2.3.6. *The dimension of $P_{\text{MIC}}(n, r)$ is $n - 2$.*

Therefore to prove that an inequality (π, π_0) defines a facet of $P_{\text{MIC}}(n, r)$, we can proceed by identifying $n - 2$ affinely independent points.

Moreover, $P(K_m)$ must have dimension $m - 1$. Indeed, its dimension cannot exceed $m - 1$ as it resides in a lower dimensional subspace. Hence by listing out the m affinely independent points $e_i + (m - i) \cdot e_1$ for $i = 2, \ldots, m$ and $m \cdot e_1$, we conclude that this upper bound is tight.

Observation 2.3.7. *The dimension of* $P(K_m)$ *is* $m-1$.

By this observation, every facet of $P(K_r)$, (respectively, $P(K_{n-r})$) must contain $r-1$ (respectively, $n-r-1$) affinely independent points.

Proposition 2.3.8. *The tilted knapsack facets are facets of* $P_{\text{MIC}}(n,r)$.

Proof. We argue for facets tilted from $P(K_r)$; an analogous argument proves the result for facets tilted from $P(K_{n-r})$.

Let (π, π_0) be tilted from $(\rho, 1)$, and let $\overline{\rho}$ be as described in Theorem 2.3.2 and α be the corresponding tilting coefficient. Since $(\rho, 1)$ is a facet of $P(K_r)$, there exist $r-1$ affinely independent extreme points $x^1, \dots, x^{r-1}$ satisfying $(\rho, 1)$ at equality. As described in Proposition 2.3.5, these points may be extended to points $t^1, \dots, t^{r-1} \in P_{\text{MIC}}(n,r)$. Clearly this operation preserves affine independence. Moreover, for $i = 1, \dots, r-1$, $\mu t^i = 1$ and $\overline{\rho} t^i = \rho x = 1$, thus

$$\pi t^i = (\overline{\rho} + \alpha\mu) t^i = \overline{\rho} t^i + \alpha \cdot \mu t^i = 1 + \alpha = \pi_r.$$

Now consider $n-r$ affinely independent extreme points $y^1, \dots, y^{n-r}$ of $P(K_{n-r})$, and again as in Proposition 2.3.5, extend them to points $s^1, \dots, s^{n-r} \in P_{\text{MIC}}(n,r)$.

$$\pi s^i = (\overline{\rho} + \alpha\mu) s^i = \overline{\rho} s^i + \alpha \cdot \mu s^i = 1 + \alpha = \pi_r.$$

It is easily seen that $\{t^1, \dots, t^{r-1}\} \cap \{s^1, \dots, s^{n-r}\} = e_r$. Therefore we have produced $n-2$ affinely independent points, proving the claim. $\square$

Consider a tilted knapsack facet (π, π_0) arising from the facet $(\rho, 1)$ of $P(K_r)$ with tilting coefficient α. Letting μ' denote the first r coefficients of μ, the same facet of $P(K_r)$ is described by $(\gamma, 0) = (\rho, 1) - (\mu', 1)$. In particular letting,

$$(\bar{\gamma}, 0) = (\gamma_1, \dots, \gamma_r = 0, 0, \dots, 0),$$

it follows that $(\pi, \pi_0) = (\bar{\gamma}, 0) + (1+\alpha)(\mu, 1)$. The same applies to tilted knapsack facets arising from $P(K_{n-r})$.

Therefore we will think of tilted knapsack facets as arising from facets of the form $(\rho, 0)$, and by subtracting off the mixed integer cut we think of tilted knapsack facets in the form $(\bar{\rho}, 0)$.

We now prove our main result.

Theorem 2.3.9. *The convex hull of $P_{\mathrm{MIC}}(n,r)$ is given by the tilted knapsack facets and the non-negativity constraints.*

Proof. For convenience, say that $P(K_r)$ has non-trivial facets $(\rho^1, 0), \ldots, (\rho^M, 0)$ and that $P(K_{n-r})$ has non-trivial facets $(\gamma^1, 0), \ldots, (\gamma^N, 0)$. Let $(\bar{\rho}^i, 0)$ and $(\bar{\gamma}^i, 0)$ denote the tilted knapsack facets from $(\rho^i, 0)$ and $(\gamma^i, 0)$ respectively.

We shall show that the system

$$\begin{array}{lllll} \min & c \cdot t & & & \\ \text{s.t.} & \mu \cdot t & = & 1 & \\ & \bar{\rho}^i \cdot t & \geq & 0 & i = 1, \ldots M \\ & \bar{\gamma}^i \cdot t & \geq & 0 & i = 1, \ldots N \\ & t & \geq & 0 & \end{array} \tag{13}$$

attains an integer optimum that belongs to $P_{\mathrm{MIC}}(n,r)$ for every c.

Let

$$c' = (c_1, \ldots, c_r), \qquad c'' = (c_{n-1}, \ldots, c_r)$$

and

$$\mu' = \left(\frac{1}{r}, \ldots, \frac{r-1}{r}, 1\right), \qquad \mu'' = \left(\frac{1}{n-r}, \ldots, \frac{n-r-1}{n-r}, 1\right).$$

Consider the systems

$$\begin{array}{lllll} \min & c' \cdot x' & & & \\ \text{s.t.} & \mu' \cdot x' & = & 1 & \\ & \rho^i \cdot x' & \geq & 0 & i = 1, \ldots M \\ & x' & \geq & 0 & \end{array} \tag{14}$$

and

$$\begin{array}{lllll} \min & c'' \cdot x'' & & & \\ \text{s.t.} & \mu'' \cdot x'' & = & 1 & \\ & \gamma^i \cdot x'' & \geq & 0 & i = 1, \ldots N \\ & x'' & \geq & 0 & \end{array} \tag{15}$$

representing $P(K_r)$ and $P(K_{n-r})$ respectively. Since both systems are integral, the minima are obtained at integer extreme points x^* and x^{**} respectively. Now let t^* be obtained by extending the solution achieving the smaller objective value to a feasible point of $P_{\text{MIC}}(n,r)$. Indeed this t^* is feasible and integral; it remains to show that it is optimal.

We now consider the duals. The dual of (14) is given by

$$\max_{\lambda_1,\alpha'}\left(\lambda_1 : \lambda_1\mu' + \alpha'_1\rho^1 + \cdots + \alpha'_M\rho^M \leq c', \alpha' \geq 0\right), \tag{16}$$

and the dual of (15) is given by

$$\max_{\lambda_2,\beta'}\left(\lambda_2 : \lambda_2\mu'' + \beta'_1\gamma^1 + \cdots + \beta'_N\gamma^N \leq c'', \beta' \geq 0\right). \tag{17}$$

Lastly the dual of (13) is given by

$$\max_{\lambda,\alpha,\beta}\left(\lambda : \begin{array}{l}\lambda\mu + \alpha_1\bar{\rho}^1 + \cdots + \alpha_M\bar{\rho}^M + \beta_1\bar{\gamma}^1 + \cdots + \beta_N\bar{\gamma}^N \leq c\\ \alpha,\beta \geq 0\end{array}\right). \tag{18}$$

Let (λ_1^*, α^*) and (λ_2^*, β^*) attain the maxima in (16) and (17) respectively. Setting

$$\lambda^* = \min(\lambda_1^*, \lambda_2^*),$$

it easily follows from the zero pattern of (13) and non-negativity of μ that $(\lambda^*, \alpha^*, \beta^*)$ is feasible to (18). Moreover $\lambda^* = c \cdot t^*$, proving optimality. □

Further observe that $P_{\text{MIC}}(n,r)$ is pointed, and so from this same proof we get the following characterization of extreme points:

Corollary 2.3.10. *A point t is an extreme point of $P_{\text{MIC}}(n,r)$ if and only if it can be obtained by extending an extreme point of $P(K_r)$ or $P(K_{n-r})$.*

Therefore, we have fully characterized the polyhedral structure of $P_{\text{MIC}}(n,r)$. However, by applying a different proof technique, we can in fact generalize the previous corollary to all integer points.

Theorem 2.3.11. *$t \in P_{\text{MIC}}(n,r) \cap \mathbf{Z}^{n-1}$ if and only if t can be obtained by extending an integer solution of $P(K_r)$ or $P(K_{n-r})$.*

Proof. If $t = e_r$ the claim is obvious. So we suppose that $t_r = 0$. We shall show that if $t \in P_{\mathrm{MIC}}(n,r) \cap \mathbf{Z}^{n-1}$, $t_r = 0$, then either (I) $(t_1, \ldots, t_r) > 0$ or (II) $(t_r, \ldots, t_{n-1}) > 0$ but not both.

Since $t \in P(\mathcal{C}_n, r)$,

$$t_1 + \cdots + (r-1)t_{r-1} + rt_r + (r+1)t_{r+1} + \cdots + (n-1)t_{n-1} \equiv r \pmod{n}.$$

Thus there exists some $\beta \in \mathbf{Z}$ such that

$$t_1 + \cdots + (r-1)t_{r-1} + rt_r + (r+1)t_{r+1} + \cdots + (n-1)t_{n-1} = r + \beta n,$$

and since $r > 0$, we may rewrite this

$$\frac{1}{r}t_1 + \cdots + \frac{r-1}{r}t_{r-1} + t_r + \frac{r+1}{r}t_{r+1} + \cdots + \frac{n-1}{r}t_{n-1} = 1 + \beta \cdot \frac{n}{r}. \tag{19}$$

Now, $t \in P_{\mathrm{MIC}}(n,r)$ therefore

$$\frac{1}{r}t_1 + \cdots + \frac{r-1}{r}t_{r-1} + t_r + \frac{n-r-1}{n-r}t_{r+1} + \cdots + \frac{1}{n-r}t_{n-1} = 1$$

or

$$\frac{1}{r}t_1 + \cdots + \frac{r-1}{r}t_{r-1} + t_r = 1 - \left[\frac{n-r-1}{n-r}t_{r+1} + \cdots + \frac{1}{n-r}t_{n-1}\right]. \tag{20}$$

Substituting (20) into (19), we obtain

$$\begin{aligned}
& 1 - \left[\tfrac{n-r-1}{n-r}t_{r+1} + \cdots + \tfrac{1}{n-r}t_{n-1}\right] + \tfrac{r+1}{r}t_{r+1} + \cdots + \tfrac{n-1}{r}t_{n-1} = 1 + \beta \cdot \tfrac{n}{r} \\
\Rightarrow\quad & \left(\tfrac{r+1}{r} - \tfrac{n-r-1}{n-r}\right)t_{r+1} + \cdots + \left(\tfrac{n-1}{r} - \tfrac{1}{n-r}\right)t_{n-1} = \beta \cdot \tfrac{n}{r} \\
\Rightarrow\quad & \tfrac{n}{r} \cdot \tfrac{1}{n-r}t_{r+1} + \cdots + \tfrac{n}{r} \cdot \tfrac{n-r-1}{n-r}t_{n-1} = \beta \cdot \tfrac{n}{r} \\
\Rightarrow\quad & \tfrac{1}{n-r}t_{r+1} + \cdots + \tfrac{n-r-1}{n-r}t_{n-1} = \beta \\
\Rightarrow\quad & \underbrace{[t_{r+1} + \cdots + t_{n-1}]}_{(*)} - \underbrace{\left[\frac{n-r-1}{n-r}t_{r+1} + \cdots + \frac{1}{n-r}t_{n-1}\right]}_{(**)} = \beta.
\end{aligned}$$

Because t was assumed to be integral $(*)$ is necessarily integral. Suppose conversely that both (I) and (II) hold; by the assumption that $t_r = 0$ and because t is necessarily non-negative, the relation

$$\frac{n-r-1}{n-r}t_{r+1} + \cdots + \frac{1}{n-r}t_{n-1} = 1 - \left[\frac{1}{r}t_1 + \cdots + \frac{r-1}{r}t_{r-1} + t_r\right]$$

implies that $(**)$ must be fractional. But this contradicts that β is integral. Therefore (I) and (II) cannot simultaneously hold. $\square$

Example 2.3.1. In the following example, we use PORTA to explicitly determine the convex hulls of $P_{\text{MIC}}(7,3)$ and $P(K_3)$ and $P(K_4)$ to demonstrate our results in action. We have that

$$P(K_3) = \left\{ x \in \mathbf{R}^3_+ : \begin{array}{l} x_1 + 2x_2 + 3x_3 = 3 \\ x_1 - x_2 \geq 0 \end{array} \right\}$$

and

$$P(K_4) = \left\{ x \in \mathbf{R}^4 : \begin{array}{l} x_1 + 2x_2 + 3x_3 + 4x_4 = 4 \\ x_1 - x_3 \geq 0 \end{array} \right\}.$$

The convex hull of $P_{\text{MIC}}(7,3)$ is given by

$$P_{\text{MIC}}(7,3) = \left\{ t \in \mathbf{R}^6 : \begin{array}{l} \frac{1}{3} \cdot t_1 + \frac{2}{3} \cdot t_2 + \frac{3}{3} \cdot t_3 + \frac{3}{4} \cdot t_4 + \frac{2}{4} \cdot t_5 + \frac{1}{4} \cdot t_6 = 1 \\ t_1 - t_2 \geq 0 \\ t_6 - t_4 \geq 0 \end{array} \right\}.$$

For $P(\mathcal{C}_7, 3)$, the tilted knapsack facets given by the two inequalities are defined by

$$5x_1 + 3x_2 + 8x_3 + 6x_4 + 4x_5 + 2x_6 \geq 8$$

$$2x_1 + 4x_2 + 6x_3 + x_4 + 3x_5 + 5x_6 \geq 6,$$

which can easily be verified to arise by appropriately scaling and adding the mixed integer cut to the knapsack inequalities.

2.4 Extreme Points and Adjacency under Group Mappings

As described in Theorems 2.2.3 and 2.2.4, automorphisms and homomorphisms of groups can be used to generate facets for one master group polyhedron from facets of another. A natural question arising from these theorems is how extreme points and adjacent facets translate through such mappings. In this section, we depart from the setting of cyclic groups, and once again consider $P(\mathcal{G}, g_0)$ in full generality. We will always denote an automorphism by $\varphi : \mathcal{G} \to \mathcal{G}$ and a homomorphism by $\psi : \mathcal{G} \to \mathcal{H}$.

2.4.1 Automorphic Mappings

The case of automorphisms is handled very naturally. Indeed, if t satisfies (π, π_0) at equality, then s with components $s(g) = t(\varphi^{-1}(g))$ satisfies (π', π_0) at equality, and since φ is an automorphism of $\mathcal{G}$,

$$\varphi(g_0) = \varphi\left(\sum_{g \in \mathcal{G}^+} g \cdot t(g)\right) = \sum_{g \in G^+} \varphi(g) \cdot t(g) = \sum_{g \in G^+} g \cdot s(g);$$

and thus s necessarily satisfies the group equation for $P(\mathcal{G}, \varphi(g_0))$. As an obvious consequence, a point t lies on the facet (π, π_0) of $P(\mathcal{G}, g_0)$ if and only if the corresponding point s lies on the facet (π', π_0) of $P(G, \varphi(g_0))$.

Hence we obtain the following:

Proposition 2.4.1. *If (π, π_0) and (γ, γ_0) are facets of $P(\mathcal{G}, g_0)$, then (π, π_0) and (γ, γ_0) are adjacent if and only if (π', π_0) and (γ', γ_0) are adjacent facets of $P(\mathcal{G}, \varphi(g_0))$, where $\pi'(g) = \pi(\varphi^{-1}(g))$ and $\gamma'(g) = \gamma(\varphi^{-1}(g))$.*

Proof. Since (γ, γ_0) and (π, π_0) define adjacent facets, there exist affinely independent points $t^1, \ldots, t^{(|\mathcal{G}|-2)}$ satisfying both at equality. By the previous remarks, we may define points $s^1, \ldots, s^{(|\mathcal{G}|-2)}$ satisfying both (π', π_0) and (γ', γ_0) at equality. Since these are all defined by applying a fixed permutation to the indices of $t^1, \ldots, t^{(|\mathcal{G}|-2)}$, affine independence is preserved. □

As an immediate application of this proposition we can extend our characterization of $P_{\text{MIC}}(n, r)$ to accommodate facets obtained from an automorphic mapping of the mixed integer cut. Letting $(\mu', 1)$ denote the permuted mixed integer cut, we have the following theorems.

Theorem 2.4.2. *The non-trivial facets of $P(\mathcal{C}_n, \varphi(r))$ adjacent to $(\mu', 1)$ are exactly those obtained by applying φ to tilted knapsack facets.*

Likewise, we can describe the integer points (including the extreme points) that satisfy $(\mu', 1)$ at equality.

Theorem 2.4.3. *An integer point $t \in P(\mathcal{C}_n, \varphi(r))$ satisfies $(\mu', 1)$ at equality if and only if t is obtained by extending a knapsack solution of $P(K_r)$ or $P(K_{n-r})$ and applying φ to the indices of t.*

2.4.2 Homomorphic Lifting

The next task will be to describe how adjacency and integer points are preserved under homomorphic lifting. Unfortunately, the extension of Theorem 2.2.4 is not nearly as transparent as the extension of Theorem 2.2.3.

To begin, we prove the following proposition showing that the intersection of two non-trivial facets is not contained on the face described by a non-negativity constraint.

Proposition 2.4.4. *Let (π, π_0) and (γ, γ_0) be adjacent non-trivial facets in $P(\mathcal{H}, h_0)$, $h_0 \neq 0$. Then the affine subspace*

$$T = P(\mathcal{H}, h_0) \cap \left\{ t \in \mathbf{R}^{|\mathcal{H}|-1} : \pi t = \pi_0, \gamma t = \gamma_0 \right\}$$

does not lie in the hyperplane $H(h) = \{t \in \mathbf{R}^{|\mathcal{H}|-1} : t(h) = 0\}$ for any $h \in \mathcal{H}_+$.

Proof. By Theorem 2.2.1 every non-trivial facet (π, π_0) of $P(\mathcal{H}, h_0)$ satisfies

$$\pi(h) + \pi(h_0 - h) = \pi(h_0) = \pi_0.$$

In particular for all $h \in \mathcal{H}_+ \setminus h_0$, the point $t = e_h + e_{h_0 - h}$ belongs to T, and has $t(h) > 0$. Similarly, the point $t = e_{h_0}$ belongs to T and has $t(h_0) > 0$. □

Despite its simplicity, this proposition provides a useful tool for producing affinely independent points. We now show as an extension of Theorem 2.2.4 that homomorphic lifting preserves the adjacency of facets.

Proposition 2.4.5. *Let (π, π_0) and (γ, γ_0) be adjacent non-trivial facets of $P(\mathcal{H}, h_0)$, and let (π', π_0) and (γ', γ_0) be facets of $P(\mathcal{G}, g_0)$ obtained by homomorphic lifting using the homomorphism ψ. Then (π', π_0) and (γ', γ_0) are adjacent.*

Proof. Let $\mathcal{K} = \ker(\psi)$. Let ϑ be a function selecting one element from each coset of $\mathcal{G}/\mathcal{K}$ distinct from $\mathcal{K}$. Denote by $\vartheta(\mathcal{H})$ the set of coset representatives chosen by ϑ.

Since (π, π_0) and (γ, γ_0) are adjacent there exist $|\mathcal{H}| - 2$ affinely independent points $t^1, \ldots, t^{|\mathcal{H}|-2}$ in $P(\mathcal{H}, h_0)$ satisfying (π, π_0) and (γ, γ_0) at equality. As (π', π_0) and (γ', γ_0) are obtained by homomorphic lifting, $g_0 \notin \mathcal{K}$; thus $\psi(g_0) = h_0 \neq 0$ so by Proposition 2.4.4, for all $h \in \mathcal{H}_+$, there exists an $i \in \{1, \ldots, |\mathcal{H}| - 2\}$ such that $t^i(h) > 0$.

Using these points, we will construct $|\mathcal{G}| - 2$ affinely independent points belonging to $P(\mathcal{G}, g_0)$ that satisfy both (π', π_0) and (γ', γ_0) at equality. Before going into specifics, we give a high level description of how we will proceed. Given a point t^i, we will use ϑ to map each element of $\mathcal{H}$ to an element of $\mathcal{G}$; by construction, this assignment is unique. Observe that

$$\psi\left(\sum_{h \in \mathcal{H}} \vartheta(h) s^i(\vartheta(h))\right) = h_0;$$

therefore

$$\sum_{h \in \mathcal{H}} \vartheta(h) s^i(\vartheta(h)) = g_0 + k$$

for some $k \in \mathcal{K}$, i.e. it is in the same coset as g_0. So by setting $s^i(-k) = 1$, the point is made to satisfy the group equation. Next if $s^i(\vartheta(h)) > 0$, then we can interchange $\vartheta(h)$ with any element of g belonging to the same coset by modifying which element of the kernel $\mathcal{K}$ we include in our solution. Finally, we generate affinely independent points from each kernel element of $\mathcal{K}$.

Taking this high level description, we now explicitly construct $|\mathcal{G}|-2$ affinely independent points:

(i) Set $\mathcal{N} = \mathcal{H}^+$

(ii) For $i = \{1, \ldots, |\mathcal{H}| - 2\}$

Set $\mathcal{N}(i) = \{h \in \mathcal{H}^+ : t^i(h) > 0\} \cap \mathcal{N}$

Define s^i as follows:

$$\begin{aligned}
&s^i(\vartheta(h)) = t^i(h), && \forall h \in \mathcal{H} \\
&s^i(g) = 0, && g \in \mathcal{G} \setminus (\mathcal{K} \cup \vartheta(\mathcal{H})) \\
&s^i(k) = 1, && k = g_0 - \textstyle\sum_{g \in \mathcal{G}^+ \setminus \mathcal{K}} s^i(g) \cdot g \\
&s^i(k) = 0, && k \neq g_0 - \textstyle\sum_{g \in \mathcal{G}^+ \setminus \mathcal{K}} s^i(g) \cdot g
\end{aligned}$$

For each $h' \in \mathcal{N}(i)$, $k' \in \mathcal{K}^+$, define the point $s_{k',h'}$ as follows:

$$\begin{aligned}
&s_{k',h'}(\vartheta(h')) + k') = t^i(h') \\
&s_{k',h'}(\vartheta(h')) = 0 \\
&s_{k',h'}(g) = s^i(g), && g \in \mathcal{G}^+ \setminus \mathcal{K}, g \neq \vartheta(h'), g \neq \vartheta(h') + k' \\
&s_{k',h'}(k) = 1, && k = g_0 - \textstyle\sum_{g \in \mathcal{G}^+ \setminus \mathcal{K}} s^i_{k',h'}(g) \cdot g \\
&s_{k',h'}(k) = 0, && k \neq g_0 - \textstyle\sum_{g \in \mathcal{G}^+ \setminus \mathcal{K}} s^i_{k',h'}(g) \cdot g
\end{aligned}$$

Set $\mathcal{N} = \mathcal{N} \setminus \mathcal{N}(i)$

(iii) For each $k \in \mathcal{K}^+$, define s_k by $s_k = s^1 + |\mathcal{G}|e_k$

By construction these points satisfy (π', π_0) and (γ', γ_0) at equality. By the above remarks, it is clear that the points constructed in (ii) are feasible to $P(\mathcal{G}, g_0)$. Furthermore, from Theorem 2.1.1, $|\mathcal{G}|k = 0$; hence the points constructed in (iii) still satisfy the group equation. We first verify that the above procedure produces $|\mathcal{G}| - 2$ points.

Note that we have the $|\mathcal{H}| - 2$ points $s^1, \ldots, s^{|\mathcal{H}|-2}$. By Proposition 2.4.4, we obtain $(|\mathcal{H}| - 1)(|\mathcal{K}| - 1)$ points $s_{k,h}$ for $k \in \mathcal{K}^+$ and $h \in \mathcal{H}^+$, and lastly, we obtain $|\mathcal{K}| - 1$ points, s_k for $k \in \mathcal{K}^+$. Using the identity $|\mathcal{G}| = |\mathcal{K}||\mathcal{H}|$, it immediately follows that we have $|\mathcal{G}| - 2$ points.

Lastly, we prove that these points are affinely independent. Suppose to the contrary that these points are affinely dependent, then there exist coefficients α, β, and ζ such that

$$\begin{aligned}
\sum_{i=1}^{|\mathcal{H}|-2} \alpha^i s^i + \sum_{h \in \mathcal{H}^+, k \in \mathcal{K}^+} \beta_{h,k} s_{h,k} + \sum_{k \in K_+} \zeta_k s_k = 0, \\
\sum_{i=1}^{|\mathcal{H}|-2} \alpha^i + \sum_{h \in \mathcal{H}^+, k \in \mathcal{K}^+} \beta_{h,k} + \sum_{k \in K_+} \zeta_k = 1.
\end{aligned} \tag{21}$$

Now observe that $s_{k,h}$ is the only element with $s(\vartheta(h) + k) > 0$. Therefore $\beta_{k,h} = 0$ for all $k \in \mathcal{K}^+, h \in \mathcal{H}^+$. Suppose there exists some α and ζ satisfying (21). Then

$$\left(\alpha^1 + \sum_{k \in K^+} \zeta_k \right) s^1(\vartheta(h)) + \sum_{i > 2} \alpha^i s^i(\vartheta(h)) = 0,$$

for all $h \in \mathcal{H}$. But this implies that

$$\left(\alpha^1 + \sum_{k\in K^+} \zeta_k\right) t^1 + \sum_{i=2}^{|\mathcal{H}|-2} \alpha^i t^i = 0,$$

contradicting that $t^1, \ldots, t^{|\mathcal{H}|-2}$ are affinely independent. □

Next we show the converse of this proposition, proving that adjacency of lifted facets occurs only if the original facets were adjacent.

Theorem 2.4.6. *Let (π, π_0) and (γ, γ_0) be non-trivial facets of $P(\mathcal{H}, h_0)$, and let (π', π_0) and (γ', γ_0) be facets of $P(\mathcal{G}, g_0)$ obtained by homomorphic lifting using the homomorphism ψ. Then (π', π_0) and (γ', γ_0) are adjacent if and only if (π, π_0) and (γ, γ_0) are adjacent.*

Proof. We already have the forward implication by Proposition 2.4.5, so it remains to show that if (π', π_0) and (γ', γ_0) are adjacent, then (π, π_0) and (γ, γ_0) are adjacent.

Consider the affine space,

$$T = P(\mathcal{H}, h_0) \cap \left\{ t \in \mathbf{R}^{|\mathcal{H}|-1} : \pi t = \pi_0, \gamma t = \gamma_0 \right\},$$

and suppose that T contains at most $|\mathcal{H}| - r$ affinely independent points. Following the construction of affinely independent points in Proposition 2.4.5 and by noting Proposition 2.4.4, we can construct at least $|\mathcal{G}| - r$ affinely independent points satisfying both lifted inequalities.

Now we take a closer look at combinations of certain columns. First note that for any $k \in \mathcal{K}^+$,

$$|\mathcal{G}| e_k = s^1 - s^k.$$

Because the sum of the coefficients is 0, we may add and subtract out elements of $\mathcal{K}^+$ at our leisure when considering affine combinations. With this in mind, we may freely disregard elements of the kernel for the remaining points. Similarly, for any $g \in \vartheta(h) + \mathcal{K}$, there exists some i such that

$$s_{k,h} - s^i = s^i(\vartheta(h)) \left[e_{\vartheta(h)+k} - e_{\vartheta(h)} \right],$$

so we can distribute weight among the elements of a given coset as we choose.

Thus take some point s that satisfies (π', π_0) and (γ', γ_0) at equality. From this point, construct a point t by setting

$$t(h) = \sum_{g \in \mathcal{G}:\psi(g)=h} s(g).$$

Trivially t satisfies both (π, π_0) and (γ, γ_0) at equality, so it can be expressed as an affine combination:

$$\sum_{i=1}^{|\mathcal{H}|-r} \alpha^i t^i = t.$$

Therefore by applying weights α^i to the lifted points s^i and using the previous observation, it follows that s can be expressed as an affine combinations of the points we constructed. Thus there can be no more than $|\mathcal{G}| - r$ affinely independent points satisfying the lifted inequalities at equality. □

Now consider $\mathcal{G} = \mathcal{C}_{n'}$, $g_0 = r'$, a homomorphism $\psi : \mathcal{C}_{n'} \to \mathcal{C}_n$, $\psi(r') = r \neq 0$, and let $(\mu', 1)$ be obtained by applying homomorphic lifting to $(\mu, 1)$. Similarly by applying Theorem 2.4.6, we know that the only facets lifted under ψ that are adjacent to $(\mu', 1)$ come from tilted knapsack facets. Stated precisely:

Theorem 2.4.7. *Let (π', π_0) be obtained by homomorphic lifting using ψ applied to (π, π_0). Then (π', π_0) is adjacent to $(\mu', 1)$ if and only if (π, π_0) is a tilted knapsack facet.*

Moreover, for the integer points we obtain the following:

Theorem 2.4.8. *If an integer point $s \in P(\mathcal{C}_{n'}, r')$ satisfies $(\mu', 1)$ at equality. Then the point t defined by the mapping*

$$t_i = \sum_{j:\psi(j)=i} s_j$$

is an integer point of $P(\mathcal{C}_n, r)$ and satisfies $(\mu, 1)$ at equality. In particular it is obtained from extending a knapsack solution of $P(K_r)$ or $P(K_{n-r})$.

Although we are able to produce a nice characterization of which lifted facets are adjacent to the lifted mixed integer cut, it is not in general true that all adjacent facets come from homomorphic lifting. We conclude this section with an example that suggests that the problem of identifying the remaining adjacent facets is still open.

Example 2.4.1. Consider $P(\mathcal{C}_{14}, 10)$. Let $\psi : \mathcal{C}_{14} \to \mathcal{C}_7$ be defined by $\psi(g) = g \pmod 7$. Recall from Example 2.2.3 that the lifted mixed integer cut is given by

$$\frac{1}{3} \cdot t_1 + \frac{2}{3} \cdot t_2 + \frac{3}{3} \cdot t_3 + \frac{3}{4} \cdot t_4 + \frac{2}{4} \cdot t_5 + \frac{1}{4} \cdot t_6$$
$$+ \frac{1}{3} \cdot t_8 + \frac{2}{3} \cdot t_9 + \frac{3}{3} \cdot t_{10} + \frac{3}{4} \cdot t_{11} + \frac{2}{4} \cdot t_{12} + \frac{1}{4} \cdot t_{13} \geq 1.$$

By using Theorem 2.4.8 it is easy to list out the extreme points of the lifted mixed integer cut to use in PORTA. Solving for the convex hull yields 12 non-trivial facets. Noting Example 2.3.1, $P_{\text{MIC}}(7, 3)$ only has 2 non-trivial facets. Therefore, most of these facets cannot be described by homomorphic lifting.

It is quite easy to see, for example, that the facet

$$8x_2 + 12x_3 + 9x_4 + 6x_5 + 3x_6 + 4x_8 + 8x_9 + 12x_{10} + 9x_{11} + 6x_{12} + 3x_{13} \leq 12,$$

cannot arise by applying homomorphic lifting to either of the tilted knapsack inequalities as the coefficients do not exhibit the proper periodic structure.

2.5 Closing Remarks

Several questions remain for both the group polyhedron and knapsack polytope. One worthy avenue of research is to expand the existing library of knapsack facets, which in turn will provide even more information about the mixed integer cut. This cut has been one of the most successful tools in the solution of mixed integer programs, and it is reasonable that a greater understanding of its polyhedral structure will yield similarly effective cuts. To this end, the mixed integer cut can also be represented in the infinite group setting. It may be possible to generalize the notion of adjacency to this setting and develop a family of valid inequalities that work well in conjunction with the mixed integer cut.

Another interesting problem is to obtain non-trivial necessary and sufficient conditions to describe the extreme points of the master knapsack polytope and the master group polyhedron. A natural idea was considered for the group polyhedron in terms of irreducibility. This condition is necessary for all vertices, but insufficient. One might hope that this condition becomes sufficient for the master knapsack polytope; however, it again fails.

Lastly, a closer inspection will reveal that in homomorphic lifting we gain no information about the kernel of our homomorphism. If we consider the lifted mixed integer cut as a polyhedron, it is no longer sufficient to characterize its extreme points in terms of two related knapsacks. As the last example demonstrated, lifted tilted knapsack facets are not the only adjacent non-trivial facets of the lifted mixed integer cut. One might address whether there exists a family of facets that when added to the lifted tilted knapsack facets completely characterizes the adjacent facets of the lifted mixed integer cut.

CHAPTER III

TECHNIQUES FOR SUPERADDITIVE LIFTING

In conjunction with branching methods, cutting planes have played a crucial role in the practical solution of mixed integer programs. In this chapter, we will explore a powerful technique called lifting that has produced some of the most effective cutting planes implemented in modern-day mixed integer program solvers.

At a high level, the lifting problem takes a valid inequality for a lower-dimensional set and *lifts* it to a valid inequality for a higher-dimensional set. We have already seen this idea in the previous chapter, and indeed, lifting traces its roots to homomorphic lifting introduced by Gomory in the context of corner polyhedra [35]. The idea reappears in Padberg's treatment of the set packing polytope [63] and was subsequently generalized in the works of Wolsey [75], Zemel [79], and Balas and Zemel [8].

Lifting is usually carried out by fixing variables and sequentially reintroducing them to the problem. Performing exact lifting can unfortunately be quite cumbersome because it depends on this sequence; however, this difficulty may be remedied through the use of superadditivity. The connection between lifting and superadditivity was first recognized by Wolsey [76] in the context of binary integer programs. Wolsey shows that if the lifting function is superadditive, then lifting does not depend on the sequence that variables are reintroduced. This result was further extended by Gu et al. [40, 42] to mixed binary integer programs and by Atamtürk [4] to general mixed integer programs.

In the study of cutting planes, lifting is one of the more frequently used tools to derive and strengthen cuts. Padberg uses lifting to strengthen odd-hole inequalities for the set packing polytope. Similarly, Balas and Zemel lift minimal cover inequalities for the knapsack polytope. Crowder et al. [22] use lifting as a key ingredient to strengthen the general purpose cuts they used in their landmark work on branch-and-cut. Gu, Nemhauser, and Savelsbergh [41] apply superadditive lifting to derive lifted flow cover inequalities for the fixed-charge network flow problem. In each of these works, the application of lifting quite often produces all-around better cuts that dominate their unlifted counterparts both theoretically and empirically.

Lifting has also found applications to more general problem structures. One fairly demonstrative example of the power of lifting is highlighted by the work of de Farias et al. [25]. The authors derive cuts for a continuous knapsack set subject to a cardinality constraint; however, by using lifting, they are able to obtain these cuts without explicitly introducing auxiliary variables. Superadditive lifting is also used to obtain facets of the knapsack with a single continuous variable [53] and the related dynamic knapsack set [50]. Many mixed binary integer programs can be mapped to these sets to obtain other well-known classes of cuts. Lastly, Richard et al. study the lifting problem, specifically with respect to the continuous variables, of a mixed knapsack polytope [67, 68]. This model can be applied to rows of a simplex tableau as an alternative to the standard Gomory mixed integer cut [56].

The standard application of superadditive lifting typically requires that an exact lifting function be computed and then approximated with a superadditive approximation. For certain well-behaved problems, this is often possible, but for more complicated problems it may not be reasonable to compute the exact lifting function. This is particularly true of higher-dimensional lifting functions.

In this chapter, we develop techniques to accommodate superadditive lifting that do not explicitly require the exact lifting function. Our approach is primarily aimed at simplifying the approximation of higher-dimensional lifting functions, but the idea is by no means exclusive to higher-dimensional problems. The framework we provide is guaranteed to produce approximations that are at least as good as and can easily dominate those obtained from the one-row relaxations. Our techniques help expand the applications of lifting, and open previously inaccessible problems to this powerful approach.

We begin this section by defining the lifting problem and review several important results for superadditive lifting. Next, we explore a variant of lifted flow cover inequalities and show how a superadditive approximation can be obtained even though an exact lifting function is quite difficult to characterize. Finally, we show a more general approach, essentially surrogate constraints for lifting functions, that can be used to simplify the approximation of higher-dimensional lifting functions.

3.1 Lifting and Superadditivity

Intuitively, the sets involved in the lifting procedure must be related. Rather than speak in generalities, we begin by introducing the main concepts in an example for the knapsack polytope and then proceed to make these ideas more general.

The results and proofs in this section essentially come from [40] and borrow from [4] to extend to the setting of general mixed integer programs.

Example 3.1.1. Consider the binary knapsack problem defined by

$$\begin{aligned} X &= \left\{ y \in \{0,1\}^5 : 5y_1 + 5y_2 + 4y_3 + 2y_4 + 2y_5 \leq 10 \right\} \\ &= \left\{ y \in \{0,1\}^5 : 5y_1 + 5y_2 + 4y_3 \leq 10 - 2y_4 - 2y_5 \right\}. \end{aligned}$$

The set $C = \{1, 2, 3\}$ is a *cover* in the sense that $5 + 5 + 4 > 10$. Moreover, C is minimal with this property; in particular, no subset of elements of C exceeds the knapsack capacity.

Consider a restriction of the problem obtained by setting $y_4 = 0$ and $y_5 = 0$. This restricted system is defined by

$$X' = \left\{ y \in \{0,1\}^3 : 5y_1 + 5y_2 + 4y_3 \leq 10 \right\}$$

Observe that since C is a cover, no solution can have $y_1 = y_2 = y_3 = 1$. Therefore we obtain the valid inequality

$$y_1 + y_2 + y_3 \leq 2. \tag{22}$$

It is easy to see that this inequality is facet-defining for $\operatorname{conv}(X')$. Indeed, the three affinely independent points

$$y^1 = (1, 1, 0), \quad y^2 = (1, 0, 1), \quad y^3 = (0, 1, 1)$$

all belong to X' and satisfy (22) at equality. Our goal will be to reintroduce y_4 and y_5 to obtain a valid (and hopefully facet-defining) inequality for $\operatorname{conv}(X)$. First reintroducing y_4, we seek a coefficient α_4 such that

$$y_1 + y_2 + y_3 + \alpha_4 y_4 \leq 2$$

is valid for the system

$$X'' = \left\{ y \in \{0,1\}^4 : 5y_1 + 5y_2 + 4y_3 + \leq 10 - 2y_4 \right\}.$$

We can rewrite this as a small integer programming problem:

$$\alpha_4 y_4 \leq \min \left(2 - [y_1 + y_2 + y_3] : \begin{array}{l} 5y_1 + 5y_2 + 4y_3 \leq 10 - 2y_4, \\ y \in \{0,1\}^4 \end{array} \right).$$

The expression on the right is an example of a *lifting function*. We will visit this in more detail later.

When $y_4 = 0$, we get the inequality $0 \leq 0$, yielding no additional information about α_4. On the other hand, if $y_4 = 1$, then we obtain the inequality $\alpha_4 \leq 1$. Thus we take $\alpha_4 = 1$ to obtain the inequality

$$y_1 + y_2 + y_3 + y_4 \leq 2.$$

It is easily verified that this inequality is facet-defining for $\text{conv}(X'')$. Proceeding as before we solve the integer program

$$\alpha_5 y_5 \leq \min \left(2 - [y_1 + y_2 + y_3 + y_4] : \begin{array}{l} 5y_1 + 5y_2 + 4y_3 + 2y_4 \leq 10 - 2y_5, \\ y \in \{0,1\}^5. \end{array} \right)$$

When $y_5 = 1$ we get the inequality $\alpha_5 \leq 0$; thus we set $\alpha_5 = 0$. The final inequality we obtain is given by

$$y_1 + y_2 + y_3 + y_4 \leq 2.$$

This can further be verified to be facet-defining for $\text{conv}(X)$.

Observe that in this example we could have chosen to reintroduce y_5 before y_4. The process would proceed identically; however it would yield the facet-defining inequality

$$y_1 + y_2 + y_3 + y_5 \leq 2.$$

This highlights an obstacle that is inherent to the lifting procedure: namely, that it is *sequential* and requires the solution of multiple integer programs. However, through the use of superadditivity, we can often overcome this challenge.

We first describe the lifting procedure in greater detail. Consider the mixed integer set

$$X = \left\{ x \in \mathbf{R}^n_+ : \begin{array}{l} \sum_{j \in N} a_j x_j \leq d, \\ x_j \in \mathbf{Z}, \;\; j \in J \end{array} \right\},$$

where $a_j,\ d \in \mathbf{Q}^m$. Let the sets N_i for $i = 0, \ldots, r$ partition N and have cardinality n_i. Further, let $J_i = N_i \cap J$. Our goal is to derive valid inequalities for X by considering simpler subsystems obtained by fixing variables in N_i for $i = 1, \ldots, r$.

For $i > 0$, fix $x_j = b_j$ for all $j \in N_i$. This produces a sequence of restrictions:

$$X^i = \left\{ x \in \mathbf{R}^{n_0+\cdots+n_i}_+ : \begin{array}{l} \sum_{0 \leq k < i} \sum_{j \in N_k} a_j x_j \leq d^i - \sum_{j \in N_i} a_j x_j, \\ x_j \in \mathbf{Z}, \;\; j \in \bigcup_{0 \leq k \leq i} J_k \end{array} \right\}, \tag{23}$$

where $d^i = d - \sum_{k>i} \sum_{j \in N_k} a_j b_j$. Noting that

$$d^i = d - \sum_{k>0} \sum_{j \in N_k} a_j b_j + \sum_{1 \leq k \leq i} \sum_{j \in N_k} a_j b_j = d^0 + \sum_{1 \leq k \leq i} \sum_{j \in N_k} a_j b_j,$$

we can rewrite (23).

$$X^i = \left\{ x \in \mathbf{R}^{n_0+\cdots+n_i}_+ : \begin{array}{l} \sum_{j \in N_0} a_j x_j + \sum_{1 \leq k < i} \sum_{j \in N_k} a_j (x_j - b_j) \leq d^0 - \sum_{j \in N_i} a_j (x_j - b_j), \\ x_j \in \mathbf{Z}, \;\; j \in \bigcup_{0 \leq k \leq i} J_k \end{array} \right\}.$$

Beginning with a valid inequality for X^0, we simultaneously reintroduce variables in N_1 to construct a valid inequality for X^1. Next, we reintroduce the variables in N_2 and so forth until we recover a valid inequality for X.

Suppose that we start with a valid inequality for X^0,

$$\sum_{j \in N_0} \alpha_j x_j \leq \alpha_0. \tag{24}$$

In the first iteration, we must appropriately choose coefficients α_j for $j \in N_1$ such that

$$\sum_{j \in N_1} \alpha_j (x_j - b_j) \leq \alpha_0 - \sum_{j \in N_0} \alpha_j x_j. \tag{25}$$

We can condition both sides of (25) on $\sum_{j\in N_1} a_j(x_j - b_j) = z$. To this end, we define two functions that will enable us to choose valid coefficients α_j. For the left hand side, we introduce the function

$$\begin{aligned} h_i(z) = \max \quad & \sum_{j\in N_i} \alpha_j(x_j - b_j) \\ \text{s.t.} \quad & \sum_{j\in N_i} a_j(x_j - b_j) = z \\ & x_j \in \mathbf{Z},\;\; j \in J_i \\ & x \in \mathbf{R}_+^{n_i} \end{aligned} \tag{26}$$

If this maximization is infeasible for a given choice of z, then $h_i(z) = -\infty$.

The second function we introduce deserves special emphasis as it will receive considerable attention throughout this chapter.

Definition 3.1.1. The *lifting function* f is defined by

$$\begin{aligned} f(z) = f_1(z) = \min \quad & \alpha_0 - \sum_{j\in N_0} \alpha_j x_j \\ \text{s.t.} \quad & \sum_{j\in N_0} a_j x_j \le d^0 - z \\ & x_j \in \mathbf{Z},\;\; j \in J_0 \\ & x \in \mathbf{R}_+^{n_0}. \end{aligned} \tag{27}$$

If the system is infeasible for a certain choice of z then $f(z) = +\infty$.

As a matter of notation, we will typically let $\mathbf{D}$ denote the domain of the lifting function. By construction, we have the following proposition:

Proposition 3.1.1. *The inequality (25) is valid for X^1 for any choice of α_j such that $h_1(z) \le f(z)$.*

Fixing $x_j = b_j$ for some $b_j > 0$ is generally risky as valid lifting coefficients are not guaranteed to exist. We show a small example on an integer knapsack problem where we encounter this difficulty:

Example 3.1.2. Let $X = \left\{x \in \mathbf{Z}_+^2 : 3x_1 + x_2 \le 6\right\}$. Fix $x_2 = 1$, and consider the cover inequality for the resulting system: i.e. $x_1 \le 1$. The function h is defined by

$$h(z) = \max\left\{\alpha_2(x_2 - 1) : x_2 - 1 = z\right\}.$$

Thus $h(-1) = -\alpha_2$ and $h(1) = \alpha_2$. Next define the lifting function

$$f(z) = \min\left\{1 - x_1 : 3x_1 \leq 5 - z, x_1 \in \mathbf{Z}_+\right\}.$$

Observe that $f(-1) = -1$ and $f(1) = 0$. However, this implies $-\alpha_2 \leq -1$ and $\alpha_2 \leq 0$, but this is impossible. Therefore there cannot exist any valid lifting coefficients.

Despite this hazard, we introduce the general theory of superadditive lifting fixing the variables at arbitrary values, and we presuppose the existence of valid lifting coefficients. This is a big assumption, and warrants some discussion about when valid lifting coefficients are guaranteed to exist.

Proposition 3.1.2. *If $b_j = 0$ for all $j \in N_1$, then there exist some choice of lifting coefficients satisfying Proposition 3.1.1.*

Proof. If $X^1 = \emptyset$, then any choice of α_j is trivially valid. Therefore assume that X^1 is non-empty and let $P^1 = \text{conv}\{X\}$. By the decomposition theorem for polyhedra, $P^1 = Q^1 + C^1$ for some polytope Q^1 and finitely generated cone C^1. Let

$$Q^1 = \text{conv}\left\{x^1, \ldots, x^p\right\},$$

$$C^1 = \text{cone}\left\{y^1, \ldots, y^r\right\}.$$

Therefore, for any $x \in P^1$, we can express x as follows:

$$x = \sum_{i=1}^{p} \lambda_i x^i + \sum_{i=1}^{r} \mu_i y^i,$$

with $\sum_{i=1}^{p} \lambda_i = 1$ and $\lambda_i \in [0,1]$ for all i, and $\mu_i \geq 0$ for all i.

Suppose that α satisfies the following the following inequalities:

$$\begin{aligned}
\sum_{j \in N_1} \alpha_j x_j^i &\leq \alpha_0 - \sum_{j \in N_0} \alpha_j x_j^i, \qquad i = 1, \ldots, p \\
\sum_{j \in N_1} \alpha_j y_j^i &\leq - \sum_{j \in N_0} \alpha_j y_j^i, \qquad i = 1, \ldots, r.
\end{aligned} \tag{28}$$

Then it clearly follows that $\alpha x \leq \alpha_0$ defines a valid inequality of P^1. Observe α_j is fixed for $j \in N_0$; therefore, the right hand side of (28) consists of scalars. Now if $x_j^i = 0$ for all

$j \in N_1$, then it follows by validity that $\alpha_0 - \sum_{j \in N_0} \alpha_j x_j^i \geq 0$. Likewise if $y_j^i = 0$ for all $j \in N_1$, then $-\sum_{j \in N_0} \alpha_j y_j^i \geq 0$. In particular, these constraints are redundant, so we shall assume that (28) does not contain any such rows.

We now claim that a valid choice of α_j for $j \in N_1$ is guaranteed to exist. Indeed, we can rewrite (28) as $G\alpha \leq h$. By the Farkas lemma, this system has a solution if and only if $sh \geq 0$ for all $s \geq 0$ such that $sG = 0$. As $G \geq 0$ with no zero rows, the only $s \geq 0$ satisfying $sG = 0$ is $s = 0$. Trivially, this must satisfy $sh \geq 0$, therefore a valid choice of α exists. □

Another important theorem provides conditions under which a lifted inequality is facet-defining. For a vector $x \in \mathbf{R}^{n_1}$ let $t = x - b$, i.e. $t_j = x_j - b_j$. A valid lifting is called *maximal* if there exist exist n_1 linearly vectors $t^1, \dots t^{n_1}$ such that $h_1(z) = f(z)$.

Theorem 3.1.3. *If* $\text{conv}(X^0)$ *and* $\text{conv}(X^1)$ *are full dimensional, (24) is facet-defining for* $\text{conv}(X^0)$*, and* $\alpha_0 \neq 0$*, then (25) defines a facet of* X^1 *if and only if the lifting is maximal.*

As we reintroduce variables, we determine valid choices of the coefficients α_j. Accordingly, we update the lifting function:

$$
\begin{aligned}
f_i(z) = \min \;& \alpha_0 - \sum_{j \in N_0} \alpha_j x_j - \sum_{0<k<i} \sum_{j \in N_k} \alpha_j (x_j - b_j) \\
\text{s.t.} \;& \sum_{j \in N_0} a_j x_j + \sum_{0 \leq k < i} \sum_{j \in N_k} a_j (x_j - b_j) \leq d^0 - z \\
& x_j \in \mathbf{Z}, \;\; j \in \bigcup_{0 \leq k < i} J_k \\
& x \in \mathbf{R}_+^{n_0 + \cdots + n_{i-1}}.
\end{aligned}
\tag{29}
$$

Although, this computation is usually prohibitively expensive, by repeated application of Proposition 3.1.1, we can eventually construct a valid inequality of X:

$$
\sum_{j \in N_0} \alpha_j x_j + \sum_{1 \leq k \leq r} \sum_{j \in N_k} a_j (x_j - b_j) \leq d. \tag{30}
$$

As a natural consequence of (29), we obtain the following proposition:

Proposition 3.1.4. $f_1 \geq f_2 \geq \cdots \geq f_r$.

Although X^0 and the valid inequality (24) may permit us to efficiently compute f, this structure is quickly lost upon the reintroduction of variables, and it becomes harder and harder to compute f_i. Therefore we seek tools to overcome this difficulty, and fortunately we find refuge in superadditivity.

Definition 3.1.2. A function $g : \mathbf{D} \to \mathbf{R}$ is *superadditive* if

$$g(u) + g(v) \leq g(u + v)$$

for all u, v, $u + v \in \mathbf{D}$.

Gu et al. [40,42], show under moderate restrictions that if f is superadditive, then $f_i = f$ for all f. This idea is further generalized by Atamtürk [4] to accommodate an even greater variety of problems. In general, f will not be superadditive. Gu proposed to remedy this by the introduction of *superadditive valid lifting functions*

Definition 3.1.3. A superadditive function $g \leq f$ is called a *superadditive valid lifting function* for f.

If for some choice of α_j, $h_i(z) \leq g(z)$ for all i, and g is a superadditive valid lifting function, then (30) is valid for X. Therefore, by sacrificing some accuracy, we are able to avoid the computational effort needed to recompute f.

Intuitively, some approximations may be better than others. Hence Gu proposes two reasonable indicators of the quality of g.

Definition 3.1.4. A superadditive valid lifting function g is *non-dominated* if there does not exist any superadditive valid lifting function $g' \geq g$ such that $g'(z) > g'(z)$ for some z.

Definition 3.1.5. Define the set

$$E = \{z \in \mathbf{D} : f_i(z) = f(z), \text{for all } i,\, N_i \text{ and lifting orders}\}.$$

A superadditive valid lifting function g is *maximal* if $g(z) = f(z)$ for all $z \in E$.

This definition of maximality is difficult to work with, so we must develop a much more concrete characterization of this property. Therefore, we adapt a proof of Gu's that relied on more stringent assumptions. In doing so, we incorporate the ideas from Atamtürk, but are less restrictive in our choice of N_i. To condense our notation we let $t_j = x_j$ for $j \in N_0$ and $t_j = x_j - b_j$ for $j \in N \setminus N_0$. Therefore the i-th lifting problem becomes

$$\begin{aligned} f_i(z) = \min \quad & \alpha_0 - \sum_{0 \le k < i} \sum_{j \in N_k} \alpha_j t_j \\ \text{s.t.} \quad & \sum_{j \in N_0} a_j t_j + \sum_{0 \le k < i} \sum_{j \in N_k} a_j t_j \le d^0 - z \\ & t_j \in \mathbf{Z}, \quad j \in \bigcup_{0 \le k < i} J_k \\ & t_j + b_j \in \mathbf{R}_+, \quad j \in \bigcup_{0 \le k < i} N_k. \end{aligned} \tag{31}$$

Proposition 3.1.5. *$f_i(z) \ge \min_{u, z+u \in \mathbf{D}} \{f_\ell(z+u) - f_\ell(u)\}$, for $\ell < i$.*

Proof. For brevity, we will omit bound and integrality constraints from the descriptions of the lifting functions. Let t^* be an optimum solution to (31) for some $z \in \mathbf{D}$. Let $u^* = \sum_{\ell \le k < i} \sum_{j \in N_k} a_j t_j^*$.

$$\begin{aligned} f_i(z) = \min \quad & \alpha_0 - \sum_{0 \le k < i} \sum_{j \in N_k} \alpha_j t_j \\ \text{s.t.} \quad & \sum_{j \in N_0} a_j t_j + \sum_{0 \le k < i} \sum_{j \in N_k} a_j t_j \le d^0 - z. \end{aligned}$$

Splitting the sums,

$$\begin{aligned} f_i(z) = \min \quad & \alpha_0 - \sum_{0 \le k < \ell} \sum_{j \in N_k} \alpha_j t_j - \sum_{\ell \le k < i} \sum_{j \in N_k} \alpha_j t_j^* \\ \text{s.t.} \quad & \sum_{j \in N_0} a_j t_j + \sum_{0 \le k < \ell} \sum_{j \in N_k} a_j t_j \le d^0 - z - u^*. \end{aligned}$$

Note, that by fixing t_j for $j \in \bigcup_{\ell \le k < i} N_j$, we do not change the optimum. From the definition of f_ℓ, we thus have

$$f_i(z) = f_\ell(z + u^*) - \sum_{\ell \le k < i} \sum_{j \in N_k} \alpha_j t_j^*.$$

By assumption,

$$\sum_{0 \le k < \ell} \sum_{j \in N_k} \alpha_j t_j + \sum_{\ell \le k < i} \sum_{j \in N_k} \alpha_j t_j \le \alpha_0,$$

is valid for (31). In particular, from Proposition 3.1.1

$$\sum_{\ell \leq k < i} \sum_{j \in N_k} \alpha_j t_j^* \leq f_\ell\left(u^*\right).$$

Putting this all together, $f_i(z) \geq f_\ell\left(z + u^*\right) - f_\ell(u^*) \geq \min_{u, z+u \in \mathbf{D}} \{f_\ell(z+u) - f_\ell(u)\}$. $\square$

This proposition allows us to clearly characterize maximality.

Proposition 3.1.6. *For fixed $u \in \mathbf{D}$, $u \in E$ if and only if $f(u) \leq f(u+v) - f(v)$.*

Proof. Suppose that $f(u) \leq f(u+v) - f(v)$ for all $u+v, v \in \mathbf{D}$. By Proposition 3.1.4 $f(u) \geq f_i(u)$. By Proposition 3.1.5, $f_i(u) \geq f(u+v) - f(v) \geq f(u)$. Thus $f_i(u) = f(u)$, proving that $u \in E$.

Conversely suppose that there exist some v such that $f(u) > f(u+v) - f(v)$. Then suppose that $h_1(v) = f(v)$. Let t^* be an optimal solution to $f(u+v)$. Setting $t_j = t_j^*$ for $j \in N_0$, and choosing $t_j \in N_1$ such that

$$\sum_{j \in N_1} \alpha_j t_j = f(v), \qquad \sum_{j \in N_1} a_j t_j = v,$$

the objective value is at most $f(u+v) - f(v) < f(u)$. Therefore $f_2(u) < f(u)$, so $u \notin E$. $\square$

It is a simple corollary that if f is superadditive then $f_i = f$ since this implies $E = \mathbf{D}$. We use this proposition to redefine maximality in a much more tractable form:

Proposition 3.1.7. *A superadditive valid approximate lifting function is maximal if and only if $g(u) = f(u)$ for all u such that $f(u) + f(v) \leq f(u+v)$ for all $v, u+v \in \mathbf{D}$.*

We conclude by showing that under mild assumptions there always exists a maximal superadditive valid approximate lifting function. Let $\mathbf{D}$ satisfy the property that for all $u, v, w \in \mathbf{D}$ such that $u+v+w \in \mathbf{D}$, $u+v, u+w, v+w \in \mathbf{D}$. Gu imposed a similar restriction by considering $\mathbf{D} = \mathbf{R}$ or $\mathbf{D} = [0, d]$. Let

$$\gamma(u) = \inf_{v, u+v \in \mathbf{D}} \{f(u+v) - f(v)\}. \tag{32}$$

This function was introduced in Wolsey in the context of binary integer programs [76], but has since been revisited by Gu for binary mixed integer programs [40]. In general, we cannot replace the infimum with a minimum when $\mathbf{D}$ is not finite and f has discontinuities.

Proposition 3.1.8. *If* $\mathbf{D}$ *satisfies the property that for all* $u, v, w \in \mathbf{D}$ *such that* $u+v+w \in \mathbf{D}$*,* $u+v, u+w, v+w \in \mathbf{D}$*, then* γ *is a maximal superadditive valid approximate lifting function.*

Proof. Observe that by the validity of (24), $f(0) \geq 0$. Therefore $\gamma(u) \leq f(u) - f(0) \leq f(u)$. Hence $\gamma \leq f$.

To show that γ is superadditive consider $\gamma(u+v)$. For any $w \in \mathbf{D}$ such that $u+v+w \in \mathbf{D}$, observe that

$$\gamma(u) \leq f(u+w) - f(w),$$

$$\gamma(v) \leq f(u+v+w) - f(u+w).$$

Therefore,

$$\gamma(u) + \gamma(v) \leq f(u+v+w) - f(w).$$

As $\gamma(u) + \gamma(v)$ is a lower bound for this quantity,

$$\gamma(u) + \gamma(v) \leq \inf_{w, u+v+w \in \mathbf{D}} \{f(u+v+w) - f(w)\} = \gamma(u+v).$$

So γ is superadditive as desired.

Lastly, we show that γ is maximal. If $f(0) > 0$, then $E = \emptyset$ as $f(u) + f(0) > f(u)$ for all u. Otherwise, if $u \in E$, then $f(u) + f(v) \leq f(u+v)$ for all $v, u+v \in \mathbf{D}$. Therefore $f(u) \leq f(u+v) - f(v)$, so $f(u) \leq \gamma(u)$. Hence equality holds by the validity of γ. □

While γ is maximal, it is typically dominated as we show in this small example.

Example 3.1.3. Let f be defined by

$$f(z) = \begin{cases} 0 & 0 \leq z \leq 1 \\ 1 & 1 < z \leq 4. \end{cases}$$

Such a lifting function might arise from a knapsack cover inequality. Clearly f is not superadditive as $f(2) + f(2) = 2 > f(4)$.

For all $u \in [0,1]$, $f(u+v) \geq f(v) = f(u) + f(v)$, so $\gamma(u) = 0$. If $u \in (1,3)$, then $f(4-u) = 1$. Since $f(4) = 1$, this gives $\gamma(u) = f(4) - f(4-u) = 0$. Finally if $u \in [3,4]$, then $v \in [0,1]$, so $f(u+v) - f(v) = 1$. Thus $\gamma(u) = 1$. Hence

$$\gamma(z) = \begin{cases} 0 & 0 \leq z < 3 \\ 1 & 3 \leq z \leq 4. \end{cases}$$

However, this is clearly dominated by the function g defined by

$$g(z) = \begin{cases} 0 & 0 \leq z \leq 2 \\ 1 & 2 < z \leq 4, \end{cases}$$

which is easily verified to be a superadditive valid approximate lifting function.

In the next section, we will show by example that it is possible to construct superadditive valid approximate lifting functions even without a closed form solution of the lifting function. We follow this with a more general framework that permits us to construct approximate lifting functions for higher-dimensional lifting functions.

3.2 A New Family of Lifted Flow Cover Inequalities

In this section we demonstrate that it is possible to construct non-trivial superadditive valid approximate lifting functions even without a closed form description of the original lifting function. We build on the work of Gu et al. [41] to derive a new family of lifted flow cover inequalities.

The *single node fixed-charge network flow problem* is defined by the system

$$X = \left\{ (x,y) \in \mathbf{R}^n \times \mathbf{Z}^n : \begin{array}{ll} \sum_{j \in N^+} x_j - \sum_{j \in N^-} x_j \leq b & \\ 0 \leq x_j \leq u_j y_j, & \forall j \in N \\ y_j \in \{0,1\}, & \forall j \in N \end{array} \right\}. \tag{33}$$

This set has been extensively studied [64, 72, 77], and its facets have proved quite useful in the practical solution of mixed integer programs.

The sets N^- and N^+ represent the collection of capacitated inflow and outflow arcs, each carrying some fixed cost, and the constant b is some exogenous flow. The constraint $\sum_{j \in N^+} x_j - \sum_{j \in N^-} x_j \leq b$ is often called the *flow balance constraint* and restricts the outgoing flow to be less than the incoming flow. This can be depicted graphically as in Figure 7.

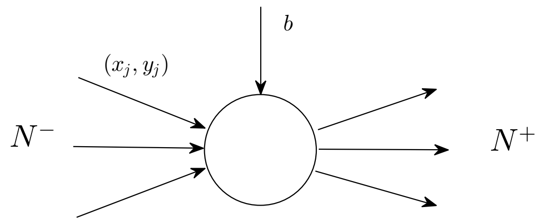

Figure 7: Single node flow set

The key structure of the facets that we study is known as a *flow cover.*

Definition 3.2.1. A set $C^+ \cup C^- \subseteq N^+ \cup N^-$ is a *flow cover* if

$$\sum_{j \in C^+} u_j - \sum_{j \in C^-} u_j = b + \lambda, \tag{34}$$

with $\lambda > 0$.

For a given flow cover, fix $(x_j, y_j) = (0, 0)$ for $j \in N^+ \setminus C^+$; fix $(x_j, y_j) = (u_j, 1)$ for $j \in C^-$; and fix $(x_j, y_j) = (0, 0)$ for $j \in N^- \setminus C^-$. Letting $b' = b + \sum_{j \in C^-} u_j$, the restricted system is the set

$$X^0 = \left\{ (x, y) \in \mathbf{R}^{|C^+|} \times \mathbf{Z}^{|C^+|} : \begin{array}{ll} \sum_{j \in C^+} x_j \leq b' & \\ 0 \leq x_j \leq u_j y_j, & \forall j \in C^+ \\ y_j \in \{0, 1\}, & \forall j \in C^+ \end{array} \right\}. \tag{35}$$

Let $P^0 = \text{conv}(X^0)$. The *flow cover inequality* for P^0 is

$$\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \leq b', \tag{36}$$

where $C^{++} = \{j \in C^+ : u_j > \lambda\}$. For notational convenience, let $C^{++} = \{1, \ldots, r\}$ with $u_1 \geq \cdots \geq u_r$, and define $U_j = U_{j-1} + u_j$ for $j = 1, \ldots, r$ with $U_0 = 0$. The following result is well known:

Proposition 3.2.1. *The flow cover inequality is facet-defining for P^0 whenever $C^{++} \neq \emptyset$.*

The lifting function associated with the flow cover inequality is given by

$$f(z) = \left\{ b' - \left[\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \right] : \begin{array}{ll} \sum_{j \in C^+} x_j \le b' - z & \\ 0 \le x_j \le u_j y_j, & \forall j \in C^+ \\ y_j \in \{0,1\}, & \forall j \in C^+ \end{array} \right\}. \quad (37)$$

This has a compact closed form description given in the next theorem due to Gu [41].

Theorem 3.2.2.

$$f(z) = \begin{cases} -\lambda & z \le -\lambda \\ z & -\lambda \le z \le 0 \\ k\lambda & U_k \le z \le U_{k+1} - \lambda, \quad (k = 0, \ldots, r-1) \\ z - U_k + k\lambda & U_k - \lambda \le z \le U_k, \quad (k = 1, \ldots, r-1) \\ z - U_r + r\lambda & U_r - \lambda \le z \le b'. \end{cases} \quad (38)$$

It is easy to verify that f is superadditive when z is restricted to be either negative or positive; however, the function f is not itself superadditive over **R**. One can construct a superadditive valid lifting function

$$g(z) = \begin{cases} k\lambda & ku_1 \le z \le (k+1)u_1 - \lambda, \quad k = 0, \pm 1, \pm 2, \ldots \\ z - ku_1 + k\lambda & ku_1 - \lambda \le z \le ku_1, \quad k = 0, \pm 1, \pm 2, \ldots \end{cases}$$

This function is maximal and non-dominated, but tends to produce computationally less effective cuts than the *lifted simple generalized flow cover inequalities*

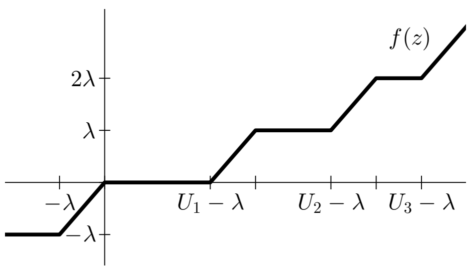

Figure 8: Flow cover lifting function

As its name suggests, these cuts are obtained by lifting the *simple generalized flow cover inequality*:

$$\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) - \sum_{j \in L^-} \lambda y_j - \sum_{j \in L^{--}} x_j \leq b',$$

where $L^- = \{j \in N^- \setminus C^- : u_j > \lambda\}$ and $L^{--} = \{j \in N^- \setminus C^- : u_j \leq \lambda\}$. This inequality arises from (36) by reintroducing the variables in $N^- \setminus C^-$. Thus, the lifted simple generalized flow cover inequalities are obtained by first performing exact lifting on a well-behaved set of variables and then performing approximate lifting on the remaining variables.

Given the success of these inequalities, a promising avenue is to develop cuts obtained by reversing the lifting order: first introducing variables in $N^+ \setminus C^+$ and C^- using exact lifting, and then approximately lifting in variables in $N^- \setminus C^-$. For the remainder of this section, we will develop the theoretical foundations needed to obtain a new family of cuts by doing precisely this.

3.2.1 Computing an Alternate Lifting Function

After reintroducing the variables in $(N^+ \setminus C^+) \cup C^-$, we obtain the set

$$X^1 = \left\{ (x, y) \in \mathbf{R}^q \times \mathbf{Z}^q : \begin{array}{ll} \sum_{j \in N^+} x_j - \sum_{j \in C^-} x_j \leq b & \\ 0 \leq x_j \leq u_j y_j, & \forall j \in N^+ \cup C^- \\ y_j \in \{0, 1\}, & \forall j \in N^+ \cup C^- \end{array} \right\} \tag{39}$$

where $q = |N^+| + |C^-|$. Under mild assumptions on the arc capacities u_j (see for example [64]), $\text{conv}(X^0)$ and $\text{conv}(X^1)$ are full dimensional; therefore, we assume that these conditions are satisfied.

Reintroducing the variables in $N^+ \setminus C^+$ and C^- via exact lifting, we obtain the inequality

$$\begin{aligned} &\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \\ &\quad + \sum_{j \in N^+ \setminus C^+} (\alpha_j x_j - \beta_j y_j) \leq b' - \sum_{j \in C^-} f(u_j)(1 - y_j), \end{aligned} \tag{40}$$

where $(\alpha_j, \beta_j) = (0, 0)$ if $U_k < u_j \leq U_{k+1} - \lambda$ for some k and $(\alpha_j, \beta_j) = (1, U_k - k\lambda)$ if $U_k - \lambda < u_j \leq U_k$ for some k. If $U_r - \lambda < u_j$, then $(\alpha_j, \beta_j) = (1, U_r - r\lambda)$.

One can verify that if we reintroduce these variables into the original flow cover inequality with these coefficients, there exist two linearly independent points satisfying the $h_i(z) = f(z)$ at equality. For $j \in N^+ \setminus C^+$, if $(\alpha_j, \beta_j) = (1, U_k - k\lambda)$, these points are precisely $(x_j, y_j) = (u_j, 1)$ and $(x_j, y_j) = (u_j - \epsilon, 1)$ for $\epsilon > 0$ sufficiently small. Otherwise, for $j \in C^-$, these points are $(x_j, y_j) = (u_j - \epsilon, 1)$ and $(x_j, y_j) = (0, 0)$. In particular, by Theorem 3.1.3, (40) must be facet-defining for $\mathrm{conv}(X^1)$.

Our goal is to lift in the variables in $N^- \setminus C^-$. In doing so, we calculate the lifting function:

$$\begin{aligned} g'(z) = \min \quad & b' - \left[\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \right. \\ & \left. + \sum_{j \in N^+ \setminus C^+} (\alpha_j x_j - \beta_j y_j) + \sum_{j \in C^-} f(u_j)(1 - y_j) \right] \\ \text{s.t.} \quad & \sum_{j \in N^+} x_j + \sum_{j \in C^-} (u_j - x_j) \le b' - z \\ & 0 \le x_j \le u_j y_j, \; j \in N^+ \cup C^- \\ & y_j \in \{0, 1\}, \; j \in N^+ \cup C^- \end{aligned} \tag{41}$$

To prevent the negative signs from becoming too cumbersome, we will let $g(z) = g'(-z)$, so that the domain is non-negative. Desirable properties of the lifting function, namely superadditivity, are not affected by this change.

The exact lifting function is complicated, and may not be superadditive as we show in the following example.

Example 3.2.1. Let $u_1 = u_2 = u_3 = 4$ and let $b' = 10$. Suppose that we reintroduce one variable in C^- with $u_4 = 9$. Therefore, (40) takes on the form:

$$x_1 + x_2 + x_3 + 2(1 - y_1) + 2(1 - y_2) + 2(1 - y_3) \le 10 - 4(1 - y_4).$$

A solution to (41) either has $(x_4, y_4) = (9, 1)$ or $(x_4, y_4) = (0, 0)$. Thus, the lifting function can be expressed as the minimum of two functions, and is given in Table 1.

Table 1: Explicit calculation of the lifting function g

$g(z)$	z	$g(z)$	z
$-z$	$0 \le z \le 2$	-4	$7 \le z \le 9$
-2	$2 \le z \le 5$	$5 - z$	$9 \le z \le 11$
$3 - z$	$5 \le z \le 7$	-6	$z \ge 11$

Now observe that $g(5) = -2$, therefore $g(5) + g(5) = -4 \ge g(10) = -5$, so g is not superadditive.

As there are more elements in $(N^+ \setminus C^+) \cup C^-$, the lifting function becomes even more complicated. For convenience, let $(N^+ \setminus C^+) \cup C^- = \{j_1, \dots, j_p\}$ with $u_{j_1} \ge \dots \ge u_{j_p}$. We now establish some simple observations that help characterize the structure of optimal solutions to the lifting function:

Observation 3.2.3. *In any optimal solution, if $y_j = 1$ for $j \in C^-$, then without loss of generality $x_j = u_j$.*

Proof. If $y_j = 1$ and $x_j < u_j$ for $j \in C^-$, then increasing x_j neither changes the objective nor makes the resulting solution infeasible. □

Observation 3.2.4. *If $(\alpha_j, \beta_j) = (0, 0)$ for some $j \in (N^+ \setminus C^+)$, then without loss of generality $(x_j, y_j) = (0, 0)$. Similarly if $f(u_j) = 0$ for $j \in C^-$, without loss of generality $(x_j, y_j) = (u_j, 1)$ in all solutions.*

Proof. If this condition does not hold for some $j \in N^+ \setminus C^+$, then we can obtain an equivalent solution setting $(x_j, y_j) = (0, 0)$. Similarly if $f(u_j) = 0$ for some $j \in C^-$ then we can set $(x_j, y_j) = (u_j, 1)$ to produce a solution with the same objective. □

Therefore, for the purposes of lifting, we can assume that all variable pairs in $N^+ \setminus C^+$ receive non-zero coefficients and that $f(u_j) > 0$ for all $j \in C^-$.

Observation 3.2.5. *Without loss of generality any optimal solution has at most one x_j with $0 < x_j < u_j y_j$.*

Proof. If not then there exist i and j such that $0 < x_i < u_i$ and $0 < x_j < u_j$. This implies that $i, j \in N^+$, so we can increase x_i and decrease x_j by $\delta = \min(u_i - x_i, x_j)$ to preserve feasibility and optimality. □

Observation 3.2.6. *In any optimal solution to the lifting function, if $x_j > 0$ for some $j \in C^{++}$, then $x_j \geq u_j - \lambda$. Otherwise, $(x_j, y_j) = (0, 0)$.*

Proof. Clearly, if $0 < x_j < u_j - \lambda$ in some solution, the solution obtained by setting $(x_j, y_j) = (0, 0)$ is feasible and dominates the current solution. □

Observation 3.2.7. *In any optimal solution to the lifting function, if $x_i > 0$ for some $i \in C^+$, then there exists an optimal solution with $x_j > 0$ for all $j > i$, $j \in C^+$.*

Proof. Suppose that $x_i > 0$ and $(x_j, y_j) = (0, 0)$ for some $j > i$. By setting $(x_i, y_i) = (0, 0)$, and $(x_j, y_j) = (\min(u_j, x_i), 1)$, the change in the objective function is $[x_i + (u_j - \lambda)] - [\min(x_i, u_j) + (u_i - \lambda)] \leq 0$. □

Proposition 3.2.8. *Without loss of generality, if $y_j = 1$ for some $j \in (N^+ \setminus C^+)$, then $x_j = u_j$.*

Proof. Consider some solution to the lifting function such that $y_{j^*} = 1$ and $x_{j^*} < u_{j^*}$ for some $j^* \in N^+ \setminus C^+$. We can obtain an equivalent solution by moving flow from x_j to x_{j^*} for $j \in C^+$. In this solution, either $x_{j^*} = u_{j^*}$ or $x_j = 0$ for all $j \in C^+$.

Suppose now that $x_{j^*} < u_{j^*}$ and $x_j = 0$ for all $j \in C^+$. Since this solution is feasible, $b' + z - x_{j^*} \geq 0$; hence $b' + z \geq x_{j^*}$. Setting $(x_{j^*}, y_{j^*}) = (0, 0)$, the new capacity is at least x_{j^*}. Let $x_{j^*} = U_k + \delta$ for $0 \leq \delta < u_{k+1}$. We assume $k < r$ (the other case being handled similarly). Setting $(x_j, y_j) = (u_j, 1)$ for $j = 1, \ldots, k$ and $x_{k+1} = \delta$, $y_{k+1} = 1$. The change in the objective value is at most

$$[U_{k+1} - (k+1)\lambda] - \beta_{j^*}.$$

Since $x_{j^*} > U_k$, $\beta_{j^*} \geq U_{k+1} - (k+1)\lambda$, so the objective value decreases. □

In effect, this proposition implies that arcs belonging to $N^+ \setminus C^+$ and C^- behave the same with respect to optimizing (41). In particular, when introducing any element from

$N^+ \setminus C^+$, $\alpha_j x_j - \beta_j y_j = f(u_j)$. Therefore, we will say that an element $j \in (N^+ \setminus C^+) \cup C^-$ is *active* if $(x_j, y_j) = (u_j, 1)$ for $j \in N^+ \setminus C^+$ or $(x_j, y_j) = (0, 0)$ for $j \in C^-$, and we need not distinguish which set the element comes from.

The next proposition asserts that in any optimal solution, whenever the set of active elements is non-empty, it is minimal with respect to covering z.

Proposition 3.2.9. *Suppose that j is active for all $j \in S \subseteq (N^+ \setminus C^+) \cup C^-$. Then without loss of generality there does not exist some $\emptyset \subsetneq S' \subsetneq S$ such that $\sum_{j \in S'} u_j > z$.*

Proof. Suppose that in an optimal solution some such set S' exists. Let $j^* \in S \setminus S'$. Since the solution is feasible

$$b' + z - \sum_{j \in S} u_j \geq 0.$$

Since $u_j > 0$ and $S' \subsetneq S$,

$$b' + z - \sum_{j \in S'} u_j \geq u_{j^*}.$$

For convenience, let $u_{j^*} = U_k + \delta$ for $0 \leq \delta < u_{k+1}$. By the assumption that $\sum_{j \in S'} u_j > z$, the current solution must have $(x_j, y_j) = (0, 0)$ for $j = 1, \ldots, k+1$. Assume that $k < r$ (again, we can argue similarly for $k \geq r$).

Setting $(x_{j^*}, y_{j^*}) = (0, 0)$ if $j^* \in N^+ \setminus C^+$ or $(x_{j^*}, y_{j^*}) = (u_{j^*}, 1)$ if $j^* \in C^-$, we increase capacity by u_{j^*} and decrease the objective function by $f(u_{j^*})$. If $0 \leq \delta \leq u_{k+1} - \lambda$, we set $(x_j, y_j) = (u_j, 1)$ for $j = 1, \ldots, k$, $x_{k+1} = 0, y_{k+1} = 0$, increasing the objective by $k\lambda = f(u_{j^*})$. Otherwise, if $u_{k+1} - \lambda < \delta < u_{k+1}$, we set $(x_j, y_j) = (u_j, 1)$ for $j = 1, \ldots, k$, $(x_{k+1}, y_{k+1}) = (\delta, 1)$, increasing the objective by $k\lambda + \delta - (u_{k+1} - \lambda) = f(u_{j^*})$. □

Using this proposition, we can characterize g in terms of several related lifting functions. This will in turn provide a starting point for deriving a superadditive approximation of g.

Define the functions g_i for $i = 1, \ldots, p$ by

$$
\begin{aligned}
g_i(z) = \min \quad & b' - \left[\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \right. \\
& \left. + \sum_{j \in N^+ \setminus C^+} (\alpha_j x_j - \beta_j y_j) + \sum_{j \in C^-} f(u_j)(1 - y_j) \right] \\
\text{s.t.} \quad & \sum_{j \in N^+} x_j + \sum_{j \in C^-} (u_j - x_j) \leq b' - z \\
& 0 \leq x_j \leq u_j y_j, \; j \in N^+ \cup C^- \\
& y_j \in \{0, 1\}, \; j \in N^+ \cup C^- \\
& (x_{j_i}, y_{j_i}) = (u_{j_i}, 1), \; j_i \in N^+ \setminus C^+ \\
& (x_{j_i}, y_{j_i}) = (0, 0), \; j_i \in C^-
\end{aligned}
\tag{42}
$$

corresponding to when we force j_i to be active. We are interested in this function restricted to $z \in [0, u_{j_i}]$, and so by Proposition 3.2.9 this reduces to

$$
\begin{aligned}
g_i(z) = -f(u_{j_i}) + \min \quad & b' - \left[\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \right] \\
\text{s.t.} \quad & \sum_{j \in C^+} x_j \leq b' - u_{j_i} + z \\
& 0 \leq x_j \leq u_j y_j, \; j \in C^+ \\
& y_j \in \{0, 1\}, \; j \in C^+.
\end{aligned}
\tag{43}
$$

In particular we have $g_i(z) = -f(u_{j_i}) + f(u_{j_i} - z)$. We will also define a function $g_{p+1}(z) = f(-z)$ with z restricted to be non-negative.

Proposition 3.2.10. *Suppose that $j_{k_1}, \ldots, j_{k_q}$ are active in a solution to the lifting function, with $j_{k_1} \geq \cdots \geq j_{k_q}$. If $z \leq \sum_{i=1}^{q} u_{j_k}$, then*

$$
g(z) = -\sum_{i=1}^{q-1} f(u_{j_{k_i}}) + g_{k_q} \left(z - \sum_{i=1}^{q-1} u_{j_{k_i}} \right).
$$

Otherwise,

$$
g(z) = -\sum_{i=1}^{q} f(u_{j_{k_i}}) + g_{p+1} \left(z - \sum_{i=1}^{q} u_{j_{k_i}} \right).
$$

Proof. The proof follows easily by writing the restricted problem:

$$
\begin{aligned}
g(z) = \min \quad & b' - \Bigg[\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) \\
& + \sum_{j \in N^+ \setminus C^+} (\alpha_j x_j - \beta_j y_j) + \sum_{j \in C^-} f(u_j)(1 - y_j) \Bigg] \\
\text{s.t.} \quad & \sum_{j \in N^+} x_j + \sum_{j \in C^-} (u_j - x_j) \le b' + z \\
& 0 \le x_j \le u_j y_j, \; j \in N^+ \cup C^- \\
& y_j \in \{0,1\}, \; j \in N^+ \cup C^- \\
& (x_j, y_j) = (u_{j_i}, 1), \; j_i \in \{j_{k_1}, \dots j_{k_q}\} \cap N^+ \setminus C^+ \\
& (x_j, y_j) = (0, 0), \; j_i \in \{j_{k_1}, \dots j_{k_q}\} \cap C^-.
\end{aligned}
$$

Again, simplifying yields

$$
\begin{aligned}
g(z) = -\sum_{i=1}^{q} f(u_{j_{k_i}}) + \min \quad & b' - \Bigg[\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} [u_j - \lambda](1 - y_j) \Bigg] \\
\text{s.t.} \quad & \sum_{j \in C^+} x_j \le b' - \sum_{i=1}^{q} u_{j_{k_i}} + z \\
& 0 \le x_j \le u_j y_j, \; j \in C^+ \\
& y_j \in \{0,1\}, \; j \in C^+
\end{aligned}
$$

Substituting in g_{k_q} or g_{p+1} as necessary, the statement of the proposition follows. □

This proposition plays a key role in the special case that we consider next as it will allow us to completely describe the lifting function.

3.2.2 A Special Case

Now assume that $u_{j_i} = U_{\phi(i)}$ for $j_i \in (N^+ \setminus C^+) \cup C^-$. Observe that $f(u_{j_i}) = f(U_{\phi(i)}) = \phi(i)\lambda$. Despite the slight modification, we show that this special case is considerably more tractable than the general case.

Proposition 3.2.11. *If $u_{j_i} = U_{\phi(i)}$, then g_i is superadditive.*

Proof. Applying the definitions of f and g_i, (38) and (43) respectively, we can easily evaluate g_i. Explicitly, this is given by

$$g_i(z) = \begin{cases} \left[U_{\phi(i)} - U_{\phi(i)-j} - j\lambda\right] - z & \begin{array}{l} U_{\phi(i)} - U_{\phi(i)-j} \le z \le U_{\phi(i)} - U_{\phi(i)-j} + \lambda, \\ (j = 0, \ldots, \phi(i) - 1) \end{array} \\ -(j+1)\lambda & \begin{array}{l} U_{\phi(i)} - U_{\phi(i)-j} + \lambda \le z \le U_{\phi(i)} - U_{\phi(i)-(j+1)}, \\ (j = 0, \ldots, \phi(i) - 1) \end{array} \end{cases}$$

when $i \le p$. Further

$$g_{p+1}(z) = \begin{cases} -z & 0 \le z \le \lambda \\ -\lambda & z \ge \lambda. \end{cases}$$

That g_i is non-increasing is clear. It is represented graphically in Figure 9.

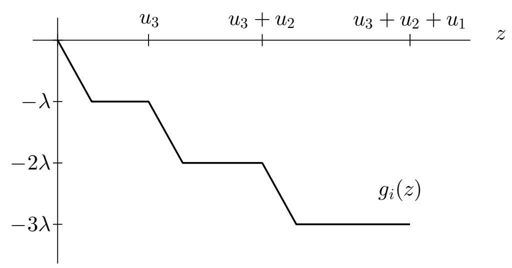

Figure 9: Plot of g_i

We now show that g_i is superadditive for $1 \le i \le p$. Consider $g(u+v) - [g(u) + g(v)]$. If the slope of g to the right of u is 0, then for some small $\epsilon > 0$,

$$g(u+v+\epsilon) - [g(u+\epsilon) + g(v)] \le g(u+v) - [g(u) + g(v)],$$

$$g(u+v) - [g(u+\epsilon) + g(v-\epsilon)] \le g(u+v) - [g(u) + g(v)],$$

as $g(u+v+\epsilon) \le g(u+v)$ and $g(v-\epsilon) \ge g(v)$. On the other hand, if the slope of g to the left of u is -1, then

$$g(u+v-\epsilon) - [g(u-\epsilon) + g(v)] \le g(u+v) - [g(u) + g(v)],$$

$$g(u+v) - [g(u-\epsilon) + g(v+\epsilon)] \le g(u+v) - [g(u) + g(v)],$$

since $g(u+v-\epsilon) \le g(u+v) + \epsilon$ and $g(v+\epsilon) \ge g(v) - \epsilon$. Therefore, we may assume that u and v are breakpoints of g where the slope changes from 0 to -1. Or u is such a breakpoint, and $v = d - u$.

For ease of notation let $\phi(i) = \ell$. Let $z_k = U_\ell - U_{\ell - t_k}$ for $k = 1, 2$. By the ordering imposed on C^+, $u_\ell \geq u_{\ell+m}$, for all $m \geq 0$. Thus

$$\begin{aligned} U_\ell - U_{\ell-(t_1+t_2)} &= (u_\ell + \cdots + u_{\ell-t_1+1}) + (u_{\ell-t_1} + \cdots + u_{\ell-t_1-t_2+1}) \\ &\geq (U_\ell - U_{\ell-t_1}) + (U_\ell - U_{\ell-t_2}) \\ &= z_1 + z_2. \end{aligned}$$

Therefore, $g(z_1 + z_2) \geq g(U_\ell - U_{\ell-(t_1+t_2)}) = -(t_1 + t_2)\lambda = (-t_1\lambda) + (-t_2\lambda) = g(z_1) + g(z_2)$.

Lastly we test when $z_1 = U_\ell - U_{\ell-t}$ and $z_2 = U_\ell - z_1 = U_{\ell-t}$. It follows that $U_{\ell-t} \geq U_\ell - U_t$, so $g(z_2) \leq -(\ell - t)\lambda$, and $g(z_1) + g(z_2) \leq -\ell\lambda = g(U_\ell)$. □

Proposition 3.2.12. *If $u_{j_i} = U_{\phi(i)}$ for $i = 1, \ldots, p$, then $g_{i_1}(z) \leq g_{i_2}(z)$ if $i_1 < i_2$ and g_{i_1} and g_{i_2} are both defined at z.*

Proof. We show the result for i and $i + 1$, and the more general claim follows trivially by induction. If $0 \leq z \leq \lambda$, then the assertion is clear, so assume that $\lambda \leq z$. For $\ell \geq k$, $U_k - U_{k-j} = u_k + u_{k-1} + \cdots + u_{k-j+1} \geq u_\ell + u_{\ell-1} + \cdots + u_{\ell-j+1} = U_\ell - U_{\ell-j}$. Now write

$$z = \lambda + \left(U_{\phi(i+1)} - U_{\phi(i+1)-j} + \delta\right),$$

with $0 \leq \delta < u_{\phi(i+1)-j}$. If $0 \leq \delta \leq u_{\phi(i+1)-j} - \lambda$, then $g_{i+1}(z) = -(j+1)\lambda$. In this case set

$$z' = \lambda + \left(U_{\phi(i)} - U_{\phi(i)-j}\right).$$

Thus $g_i(z') = -(j+1)\lambda$. Since $U_{\phi(i)} \geq U_{\phi(i+1)}$, it follows that $z' \leq z$, and since g_i and g_{i+1} are both non-increasing, necessarily $g_i(z) \leq g_{i+1}(z)$. Otherwise, if $u_{\phi(i+1)-j} - \lambda < \delta < u_{\phi(i+1)-j}$, let $\epsilon = \delta - [u_{\phi(i+1)-j} - \lambda]$. Setting

$$z' = \lambda + \left(U_{\phi(i)} - U_{\phi(i)-j}\right) + (u_{\phi(i)-j} - \lambda + \epsilon).$$

Again $g_i(z') = g_{i+1}(z)$ and $z' \leq z$, thus showing $g_i(z) \leq g_{i+1}(z)$. □

The next two theorems are central to identifying the new class of lifted flow cover inequalities. We will first evaluate $g(z)$ and then show that it is superadditive. For convenience let $V_i = V_{i-1} + U_{\phi(i)}$ with $V_0 = 0$.

Theorem 3.2.13. *Suppose that* $u_{j_i} = U_{\phi(i)}$ *for* $i = 1, \ldots, p$. *Then*

$$g(z) = \begin{cases} -\left[\sum_{i=1}^{k} \phi(i)\right] \cdot \lambda + f\left(V_{k-1} - z\right) & V_{k-1} \le z \le V_k,\ (k = 1, \ldots, p) \\ -\left[\sum_{i=1}^{p} \phi(i)\right] \cdot \lambda + f\left(V_p - z\right) & V_p \le z \end{cases} \tag{44}$$

Proof. Consider a solution to the lifting problem, and suppose that $j_{k_1} < \cdots < j_{k_q}$ are active in this solution. Applying Proposition 3.2.10, it follows that

$$g(z) = -\sum_{i=1}^{q-1} f(u_{j_{k_i}}) + g_{k_q}\left(z - \sum_{i=1}^{q-1} u_{j_{k_i}}\right).$$

We claim that $j_{k_i} = j_i$ for $i = 1, \ldots, q$. Suppose not and let ℓ denote the smallest ℓ such that $j_{k_\ell} > j_\ell$. Necessarily $j_\ell < j_{k_q}$. Now consider the solution with $j_{k_1}, \ldots, j_{k_{q-1}}$ and j_ℓ active. Applying the same argument as Proposition 3.2.10, the objective function is

$$-\sum_{i=1}^{q-1} f(u_{j_{k_i}}) + g_\ell\left(z - \sum_{i=1}^{q-1} u_{j_{k_i}}\right).$$

However, by Proposition 3.2.12,

$$g_\ell\left(z - \sum_{i=1}^{q-1} u_{j_{k_i}}\right) \le g_{k_q}\left(z - \sum_{i=1}^{q-1} u_{j_{k_i}}\right),$$

so there exists a solution at least as good as the current solution that has $j_{k_i} = j_i$.

Since $g_{p+1} \ge g_i$ for all i it follows that we only use g_{p+1} once $z \ge V_p$. Applying the definitions of g_i, we obtain the function described above. □

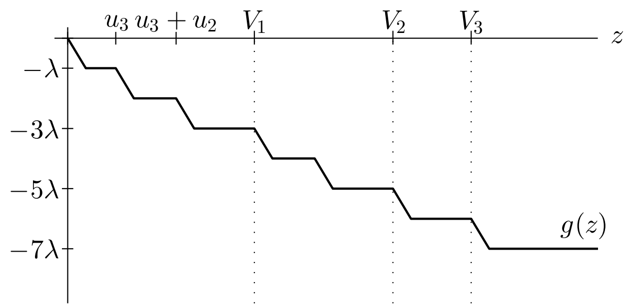

Figure 10: Plot of g when $u_{j_i} = U_{\phi(i)}$

Figure 10 gives an example of g. In the next theorem we show that g is superadditive. For a moment, consider the superadditivity condition:

$$g(z_1 + z_2) - [g(z_1) + g(z_2)] \ge 0$$

for all $z_1, z_2, z_1 + z_2 \in \mathbf{D}$. We further know that g is decreasing and $\mathbf{D} = \mathbf{R}_+$. Rather than work directly with this condition, we begin with two points z_1 and z_2 and compute $g(z_1) + g(z_2)$. We then identify a point z^* satisfying

$$g(z^*) \geq g(z_1) + g(z_2), \tag{45}$$

and show that $z^* \geq z_1 + z_2$. As g is decreasing, this implies that

$$g(z_1) + g(z_2) \leq g(z^*) \leq g(z_1 + z_2),$$

thereby showing superadditivity. As we shall see, this allows us to greatly simplify and even avoid much of the case analysis that tends to accompany superadditivity proofs.

Theorem 3.2.14. *Suppose that $u_{j_i} = U_{\phi(i)}$ for $i = 1, \ldots, p$. Then $g(z)$ is superadditive.*

Proof. As in the proof of Proposition 3.2.11, we only need to test breakpoints at which the slope of g changes from 0 to -1. We partition $\mathbf{D}$ into intervals of the form $\mathbf{D}_i = [V_{i-1}, V_i]$ for $i = 1, \ldots, p$ and $\mathbf{D}_{p+1} = [V_p, +\infty)$. For $1 \leq i \leq p$, each interval further partitions into $\phi(i)$ subintervals,

$$\mathbf{D}_i^k = V_{i-1} + [U_{\phi(i)} - U_{\phi(i)-k+1}, U_{\phi(i)} - U_{\phi(i)-k}], \quad k = 1, \ldots, \phi(i).$$

The subinterval D_i^k has length $u_{\phi(i)-k+1}$. Moreover, let $\mathbf{D}_{p+1}^1 = [V_p, +\infty)$.

Observe that the right endpoints of these subintervals correspond precisely with the previously described breakpoints. If z is one of these special breakpoints, then $g(z) = -N(z) \cdot \lambda$, where $N(z)$ denotes the total number of subintervals $\mathbf{D}_i^k = [\ell_i^k, u_i^k]$ such that $u_i^k \leq z$. Now let $z_1 = V_{s_1} + U_{\phi(s_1+1)} - U_{\phi(s_1+1)-t_1}$ and $z_2 = V_{s_2} + U_{\phi(s_2+1)} - U_{\phi(s_2+1)-t_2}$.

We claim that we may assume that $t_1 = 0$. Suppose that $t_1, t_2 > 0$. Without loss of generality,

$$\left|\mathbf{D}_{s_1+1}^{t_1}\right| \geq \left|\mathbf{D}_{s_2+1}^{t_2}\right|;$$

thus, we set $t_1' = t_1 + 1$ and $t_2' = t_2 + 1$. Updating z_1 and z_2, we have

$$z_1' = z_1 + \left|\mathbf{D}_{s_1+1}^{t_1+1}\right| \geq z_1 + \left|\mathbf{D}_{s_1+1}^{t_1}\right|$$

$$z_2' = z_2 - \left|\mathbf{D}_{s_2+1}^{t_2}\right| \geq z_2 - \left|\mathbf{D}_{s_1+1}^{t_1}\right|.$$

In particular $z_1' + z_2' \geq z_1 + z_2$ and $g(z_1') = g(z_1)$ and $g(z_2') = g(z_2)$. By the sorting on the intervals, it follows that we can repeatedly perform this update until either $t_1 = \phi(s_1 + 1)$ or $t_2 = 0$. By appropriately incrementing s_1 or interchanging the roles of s_1 and s_2, the claim follows.

If there are no more than $N(z_2)$ subintervals following V_{s_1}, then

$$g(z_1 + z_2) \geq -\left(N(z_1) + N(z_2)\right)\lambda = g(z_1) + g(z_2).$$

Therefore we shall assume that there are at least $N(z_2)$ subintervals after V_{s_1}. To prove superadditivity, we show that the total length of the $N(z_2)$ subintervals after V_{s_1} is greater than the total length of the first $N(z_2)$ subintervals.

First note that $V_{s_1} + V_{s_2} \geq V_{s_1+s_2}$, and thus the intervals $\mathbf{D}_{s_1+1}, \ldots, \mathbf{D}_{s_1+s_2}$ are entirely covered. Consider specifically the subintervals $\mathbf{D}_{s_1+i}^{\phi(s_1+i)-j+1}$ and $\mathbf{D}_i^{\phi(i)-j+1}$. By construction

$$\left|\mathbf{D}_{s_1+i}^{\phi(s_1+i)-j+1}\right| = \left|\mathbf{D}_i^{\phi(i)-j+1}\right|.$$

So for $i = 1, \ldots, s_2$ and $j = 1, \ldots, \phi(s_1 + i)$, we pair $\mathbf{D}_{s_1+i}^{\phi(s_1+i)-j+1}$ with $\mathbf{D}_i^{\phi(i)-j+1}$.

Thus, we must show that the remaining t unpaired subintervals in $[0, z_2]$ are shorter than the first t unpaired subintervals in $[V_{s_1+s_2}, \infty)$. This demonstrates the existence of a suitable z^* as in (45) with $g(z^*) = g(z_1) + g(z_2)$ and $z^* \geq z_1 + z_2$.

Now some of the subintervals in $\mathbf{D}_1, \ldots, \mathbf{D}_{s_2}$ may be unpaired. However, any unpaired interval $\mathbf{D}_i^j$ must satisfy

$$\left|\mathbf{D}_i^j\right| \leq \left|\mathbf{D}_{s_1+i}^1\right| \leq \left|\mathbf{D}_{s_1+s_2+1}^1\right|.$$

In particular, these unpaired subintervals in $[0, V_{s_2}]$ are all shorter than any unpaired subinterval in $[V_{s_1+s_2}, \infty)$.

If $t_2 = 0$, then we are done, so assume $t_2 > 0$. Now we consider the intervals $\mathbf{D}_{s_2+1}$ and $\mathbf{D}_{s_1+s_2+1}$. Let t' denote the number of the remaining t subintervals in $\mathbf{D}_{s_1+s_2+1}$. If $t' < t_2$, then $t' = \phi(s_1 + s_2 + 1)$. Thus

$$\left|\mathbf{D}_{s_2+1}^{t_2}\right| \leq \left|\mathbf{D}_{s_2+1}^{\phi(s_2+1)}\right| = \left|\mathbf{D}_{s_1+s_2+1}^{\phi(s_1+s_2+1)}\right|.$$

Thus we pair $\mathbf{D}_{s_2+1}^{t_2-j+1}$ with $\mathbf{D}_{s_1+s_2+1}^{\phi(s_1+s_2+1)-j+1}$ for $j = 1, \ldots, \phi(s_1 + s_2 + 1)$. It similarly follows that the remaining $t - t'$ unpaired subintervals in $[0, z_2]$ are shorter than the first

$t - t'$ unpaired subintervals in $[V_{s_1+s_2+1}, \infty)$. If $t' \geq t_2$, then

$$\left|\mathbf{D}^j_{s_2+1}\right| \leq \left|\mathbf{D}^j_{s_1+s_2+1}\right|,$$

and we pair $\mathbf{D}^j_{s_2+1}$ with $\mathbf{D}^j_{s_1+s_2+1}$. This concludes our pairing argument.

Thus we have shown that the maximum z^* such that $g(z^*) = -\left(N(z_1) + N(z_2)\right)\lambda$ is at least $z_1 + z_2$. Since g is decreasing, this implies that $g(z_1) + g(z_2) \leq g(z_1 + z_2)$. □

It is also possible to extend these results to when $u_{j_i} > U_r$. The proofs follow almost identically with the addition of an interval $\mathbf{D}^0_i = [V_i, V_i + u_i - U_r]$ whenever $u_i > U_r$. The same sort of pairing argument suffices for the intervals $\mathbf{D}^j_i$ for $j > 0$, and by noting that $\mathbf{D}^0_i$ always at least as long as $\mathbf{D}^0_{i+1}$.

3.2.3 Obtaining a Superadditive Approximation

The case when $u_j = U_{\phi(i)}$ is unlikely to hold. However, we can still use this special case to construct a valid superadditive approximate lifting function even without a closed form description of the exact lifting function. To show this we, will reformulate our problem and relax certain constraints.

Define the following sets:

$$\begin{aligned}
L^+_1 &= \{j \in N^+ \setminus C^+ \mid \exists k(j) : U_{k(j)} - \lambda < u_j \leq U_{k(j)}\},\\
L^+_2 &= \{j \in N^+ \setminus C^+ \mid U_r < u_j\},\\
C^-_1 &= \{j \in C^- \mid \exists k(j) : U_{k(j)} < u_j \leq U_{k(j)+1} - \lambda\},\\
C^-_2 &= \{j \in C^- \mid \exists k(j) : U_{k(j)} - \lambda < u_j \leq U_{k(j)}\},\\
C^-_3 &= \{j \in C^- \mid U_r < u_j\}.
\end{aligned}$$

We now describe a reformulation obtained in essence by *splitting* certain variables and increasing variable upper bounds. For $j \in L^+_1$, we relax the variable upper bound:

$$0 \leq x_j \leq u_j y_j \quad \mapsto \quad 0 \leq x_j \leq U_{k(j)} y_j.$$

For each $j \in C_1^-$, we split x_j by adding the set of constraints:

$$\begin{aligned} & & x_j &= x_j^1 + x_j^2, \\ (x_j, y_j), & & 0 &\le x_j^1 \le U_{k(j)} y_j^1, \\ 0 \le x_j^1 \le u_j y_j^1, & \quad \mapsto & 0 &\le x_j^2 \le \left(u_j - U_{k(j)}\right) y_j^2, \\ y_j \in \{0,1\} & & y_j^1 &= y_j^2, \\ & & y_j^1, y_j^2 &\in \{0,1\}. \end{aligned}$$

Lastly, for $j \in C_2^-$, we split x_j as follows:

$$\begin{aligned} & & x_j &= x_j^1 - x_j^2, \\ (x_j, y_j), & & \left(U_{k(j)} - u_j\right) y_j^1 &\le x_j^1 \le U_{k(j)} y_j^1, \\ 0 \le x_j \le u_j y_j, & \quad \mapsto & \left(U_{k(j)} - u_j\right) y_j^2 &\le x_j^2 \le \left(U_{k(j)} - u_j\right) y_j^2, \\ y_j \in \{0,1\} & & y_j^1 &= y_j^2, \\ & & y_j^1, y_j^2 &\in \{0,1\}. \end{aligned}$$

The remaining variables in L_2^+ and C_3^- remain unchanged. Therefore, we can rewrite (33) in terms of these new variables. We omit this intermediate step, and proceed to the system obtained by eliminating the constraints $y_j^1 = y_j^2$ and replacing the variable lower bounds with 0 for $j \in C_2^-$. As the variable lower bounds are positive, this produces the following relaxation of X:

$$\left[\sum_{j \in N^+} x_j + \sum_{j \in C_2^-} x_j^2\right] - \left[\sum_{j \in N^- \setminus C_1^- \cup C_2^-} x_j + \sum_{j \in C_1^-} \left(x_j^1 + x_j^2\right) + \sum_{j \in C_2^-} x_j^1\right] \le b$$

$$\begin{aligned}
& x_j = x_j^1 + x_j^2 && \forall j \in C_1^- \\
& x_j = x_j^1 - x_j^2 && \forall j \in C_2^- \\
& 0 \le x_j \le u_j y_j && \forall j \in L_2^+ \cup C_3^- \\
& 0 \le x_j \le U_{k(j)} y_j && j \in L_1^+ \\
& 0 \le x_j^1 \le U_{k(j)} y_j^1 && j \in C_1^- \\
& 0 \le x_j^2 \le (u_j - U_{k(j)}) y_j^2 && j \in C_1^- \\
& 0 \le x_j^1 \le U_{k(j)} y_j^1 && j \in C_2^- \\
& 0 \le x_j^2 \le (U_{k(j)} - u_j) y_j^2 && j \in C_2^- \\
& y_j \in \{0, 1\} && \forall j \in N.
\end{aligned} \tag{46}$$

We now must define a flow cover, $S^+ \cup S^-$. To avoid ambiguity we will identify elements in the flow cover by their continuous variables. Hence we define

$$\begin{aligned}
S^+ &= \left\{x_j : j \in C^+\right\} \cup \left\{x_j^2 : j \in C_2^-\right\} \\
S^- &= \left\{x_j^1 : j \in C_1^- \cup C_2^-\right\} \cup \left\{x_j^2 : j \in C_2^-\right\} \cup \left\{x_j : j \in C_3^-\right\}
\end{aligned}$$

Let $b' = b + \sum_{j \in C^-} u_j$. Fixing variables as we did before, we obtain the flow cover inequality

$$\sum_{j \in C^+} x_j + \sum_{j \in C^{++}} (u_j - \lambda)(1 - y_j) + \sum_{j \in C_2^-} x_j^2 \le b' + \sum_{j \in C_2^-} \left(U_{k(j)} - u_j\right).$$

We have used here that $U_{k(j)} - u_j < \lambda$ for all $j \in C_2^-$. For compactness let $b'' = b' + \sum_{j \in C_2^-} \left(U_{k(j)} - u_j\right)$.

Next we lift in variables in $N^+ \setminus S^+ \cup S^-$ as in (40). For $x_j \in L_2^+ \cup C_3^-$, the lifting coefficients are unchanged; for $j \in L_1^+$, the lifting coefficients likewise remain the same; for $j \in C_1^-$, $f(u_j) = f\left(U_{k(j)}\right)$ and $u_j - U_{k(j)} < u_1 - \lambda$, so $f\left(u_j - U_{k(j)}\right) = 0$; and lastly for $j \in C_2^-$, $f\left(U_{k(j)}\right) = f(u_j) + \left[U_{k(j)} - u_j\right]$.

Recovering the variables in $N^+ \setminus S^+ \cup S^-$, we arrive at the inequality

$$\sum_{j\in C^+} x_j + \sum_{j\in C_2^-} x_j^2 + \sum_{j\in C^{++}} (u_j - \lambda)(1-y_j) + \sum_{j\in N^+\setminus C^+} \left[x_j - \left(U_{k(j)} - k(j)\lambda\right) y_j\right]$$
$$\leq b'' - \sum_{j\in C_1^-\cup C_3^-} f(u_j)(1-y_j^1) - \sum_{j\in C_2^-} [f\left(U_{k(j)}\right)(1-y_j^1).$$

Using the lifting function (44) from Theorem 3.2.13, we can perform sequence independent lifting to obtain a new valid inequality:

$$\sum_{j\in C^+} x_j + \sum_{j\in C_2^-} x_j^2 x_j + \sum_{j\in C^{++}} (u_j - \lambda)(1-y_j) + \sum_{j\in N^+\setminus C^+} \left[x_j - \left(U_{k(j)} - k(j)\lambda\right) y_j\right]$$
$$\leq b'' - \sum_{j\in C_1^-\cup C_3^-} f(u_j)(1-y_j^1) - \sum_{j\in C_2^-} f\left(U_{k(j)}\right)(1-y_j^1) + \sum_{j\in N^-\setminus C^-} \left[\pi_j x_j + \mu_j y_j\right].$$

where π_j and μ_j represent valid lifting coefficients for (x_j, y_j). Noting that $x_j^2 = (U_{k(j)} - u_j)y_j^2$ for $j \in C_2^-$ and identifying y_j^1 and y_j^2 as necessary, we obtain the inequality

$$\sum_{j\in C^+} x_j + \sum_{j\in C^{++}} (u_j - \lambda)(1-y_j) + \sum_{j\in N^+\setminus C^+} \left[x_j - \left(U_{k(j)} - k(j)\lambda\right) y_j\right]$$
$$\leq b' - \sum_{j\in C^-} f(u_j)(1-y_j) + \sum_{j\in N^-\setminus C^-} \left[\pi_j x_j + \mu_j y_j\right].$$

In particular, none of the coefficients on the variables in C^- and $N^+ \setminus C^+$ have changed. Thus we have demonstrated a superadditive valid approximate lifting function.

Theorem 3.2.15. *Let $\hat{g}(z)$ be the approximation of g obtained by replacing u_j with $U_{k(j)}$ for all $j \in (N^+ \setminus C^+) \cup C^-$. Then $\hat{g}$ is a superadditive valid approximate lifting function for g.*

This theorem is significant in two regards. First and most immediately, it establishes the existence of a new family of lifted flow covers obtained through superadditive lifting. Beyond this, it suggests the potential of superadditive lifting even in the absence of an exact description of the lifting function.

It is unlikely that $\hat{g}$ is non-dominated. Letting $\nu = \lim_{z\to+\infty} g(z)$, we will typically have that $\lim_{z\to+\infty} \hat{g}(z) < \nu$. As $\hat{g}$ is negative and superadditive and ν is negative, the function

$$\tilde{g}(z) = \max\left(\hat{g}(z), \nu\right)$$

is superadditive, valid, and dominates $\hat{g}$. However, the initial approximation $\hat{g}$ is itself non-trivial, and uses much of the structure of g without explicitly computing g. In the next section, we provide a more general framework for constructing superadditive valid approximate lifting functions in the absence of a closed form description of the lifting function. We close out this section, by revisiting Example 3.2.1.

Example 3.2.2. In Example 3.2.1 we encountered a small fixed-charge flow problem with $u_1 = u_2 = u_3 = 4$ and $b' = 10$. We reintroduced a single variable x_4 from C^- with $u_4 = 9$. Because $u_4 \neq U_j$, the resulting lifting function was not superadditive.

Applying the previous reformulation, we replace x_4 and its corresponding variable upper bound constraints with

$$\begin{gathered} x_4 = x_4^1 + x_4^2, \\ 0 \leq x_4^1 \leq 8y_4^1, \\ 0 \leq x_4^2 \leq 1y_4^2, \\ y_4^1 = y_4^2, \\ y_4^1, y_4^2 \in \{0,1\}. \end{gathered}$$

Next we relax the constraint $y_4^1 = y_4^2$.

The new flow cover is given by $C^+ = \{x_1, x_2, x_3\}$ and $C^- = \left\{x_4^1, x_4^2\right\}$. Noting that $b'' = b' = 10$, the lifted flow cover inequality is given by:

$$x_1 + x_2 + x_3 + 2(1 - y_1) + 2(1 - y_2) + 2(1 - y_3) \leq 10 - 4(1 - y_4^1).$$

As $y_4^1 = y_4$, this corresponds precisely with the original flow cover inequality. On the other hand, because $u_4^1 = 8$, the corresponding lifting function is as follows in Table 2:

Table 2: Approximation $\hat{g}$ of the lifting function g

$\hat{g}(z)$	z	$\hat{g}(z)$	z
$-z$	$0 \leq z \leq 2$	-4	$6 \leq z \leq 8$
-2	$2 \leq z \leq 4$	$4 - z$	$8 \leq z \leq 10$
$2 - z$	$4 \leq z \leq 6$	-6	$z \geq 10$

This approximation can easily be verified to be superadditive and is compared against g in Figure 11.

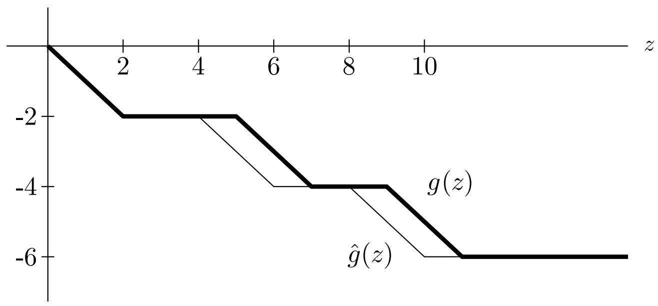

Figure 11: Comparison of g and $\hat{g}$

3.3 Approximating High-Dimensional Lifting Functions

We now shift our focus to lifting functions in higher dimensions. As a motivating example, we show that even for simple integer programs, it can be advantageous to perform lifting in higher dimensions:

Example 3.3.1. Consider the two-dimensional knapsack problem defined by the following constraints:

$$4x_1 + 4x_2 + 4x_3 + 2x_4 \leq 10,$$

$$3x_1 + 2x_2 + 3x_3 + 10x_4 \leq 12,$$

$$x_j \in \{0, 1\}.$$

The set $\{1, 2, 3\}$ defines a minimal cover of the first knapsack, but not the second. The associated minimal cover inequality is

$$x_1 + x_2 + x_3 \leq 2.$$

We fix $x_4 = 0$. If we ignore the second knapsack, the lifting problem is

$$\min \left\{ 2 - [x_1 + x_2 + x_3] : \begin{array}{l} 4x_1 + 4x_2 + 4x_3 \leq 10 - z, \\ x_j \in \{0, 1\} \end{array} \right\}.$$

Therefore, when we reintroduce x_4, $z = 2$, and there exists an optimal solution with $\alpha_4 = 0$.

On the other hand, if we consider the second knapsack, then the lifting problem is given by

$$\min \left\{ 2 - [x_1 + x_2 + x_3] : \begin{array}{r} 4x_1 + 4x_2 + 4x_3 \leq 10 - z_1, \\ 3x_1 + 2x_2 + 3x_3 \leq 12 - z_2, \\ x_j \in \{0, 1\} \end{array} \right\}.$$

So when we reintroduce x_4, we have $(z_1, z_2) = (2, 10)$. This implies that $x_1 = x_3 = 0$ in an optimal solution. Therefore, $\alpha_4 = 1$, and we obtain the inequality

$$x_1 + x_2 + x_3 + x_4 \leq 2,$$

which dominates the inequality obtained from the first knapsack alone. Furthermore, it is easy to see that this inequality is not valid for either of the individual knapsacks.

Lifting has proved an incredibly effective tool in developing inequalities that incorporate multiple constraints. However, the challenge of performing lifting in higher dimensions has limited its practical application. In this section, we describe a simple tool that facilitates the application of superadditive lifting even without full knowledge of the lifting function.

Consider the system

$$X = \left\{ x \in S : \begin{array}{l} A^0 x \leq b^0 \\ A^i x \leq b^i, \quad i = 1, \ldots, k \end{array} \right\}.$$

The set S captures integrality, variable bounds, and other such information not contained in the constraint matrix. Let $P = \mathrm{conv}(X)$, and let $\alpha x \leq \alpha_0$ be a valid inequality for P. Consider the lifting function

$$\begin{aligned} f\left(z^0, z^1, \ldots, z^k\right) = \min \quad & \alpha_0 - \alpha x \\ \text{s.t.} \quad & A^0 x \leq b^0 - z^0 \\ & A^i x \leq b^i - z^i, \quad i = 1, \ldots, k \\ & x \in S. \end{aligned} \tag{47}$$

Assume that $(z^0, z^1, \ldots, z^k) \in \mathbf{D}^0 \times \mathbf{D}^1 \times \cdots \times \mathbf{D}^k$, and let $\mu^i x \leq \mu^i_0$ be some inequality.

Define the function

$$\Phi^i\left(z^i\right) = \min \quad \mu_0^i - \mu^i x$$
$$\text{s.t.} \quad A^i x \leq b^i - z^i \tag{48}$$
$$x \in S.$$

Using this function we are able to construct a surrogate constraint for the lifting function. We describe this in two simple propositions:

Proposition 3.3.1. *Let $\phi^i \leq \Phi^i$. Then for all $z^i \in \mathbf{D}^i$, the inequality*

$$\mu^i x \leq \mu_0^i - \phi^i(z^i),$$

is valid for the set $\operatorname{conv}\left\{x \in S : A^i x \leq b^i - z^i\right\}$ *for all $z^i \in \mathbf{D}^i$.*

Proof. By construction, $\Phi(z^i) \leq \mu_0^i - \mu^i x$ for all $x \in \operatorname{conv}\left\{x \in S : A^i x \leq b^i - z^i\right\}$. As $\phi^i(z^i) \leq \Phi^i(z^i)$, the proposition immediately follows. □

Proposition 3.3.2. *If $\phi^i(z^i) \leq \Phi^i(z^i)$ for $i = 1, \ldots, k$, then the function*

$$g\left(z^0, z^1, \ldots, z^k\right) = \min \quad \alpha_0 - \alpha x$$
$$s.t. \quad A^0 x \leq b^0 - z^0 \tag{49}$$
$$\mu^i x \leq \mu_0^i - \phi^i(z^i), \quad i = 1, \ldots, k$$
$$x \in S.$$

is a valid approximate lifting function.

Proof. Let x^* minimize (47), then x^* satisfies $A^i x^* \leq b^i - z^i$. By applying Proposition 3.3.1, it follows that $\mu^i x^* \leq \mu_0^i - \phi^i(z^i)$. □

This framework still allows for the elimination of constraints by replacing them with the trivial inequality $0x \leq 0$. Let $\mathbf{E}^i \supseteq \left\{\phi^i(z^i) : z^i \in \mathbf{D}^i\right\}$. Define the function

$$\hat{g}\left(z^0, v^1, \ldots, v^k\right) = \min \quad \alpha_0 - \alpha x$$
$$\text{s.t.} \quad A^0 x \leq b^0 - z^0 \tag{50}$$
$$\mu^i x \leq \mu_0^i - v^i, \quad i = 1, \ldots, k$$
$$x \in S,$$

where $v^i \in \mathbf{E}^i$. Trivially $g\left(z^0, z^1, \ldots, z^k\right) = \hat{g}\left(z^0, \phi^1(z^1), \ldots, \phi^k(z^k)\right)$. The function $\hat{g}$ may be much easier to deal with than g. For example if $\mathbf{D}^i$ is continuous but $\mathbf{E}^i$ is finite, we may have a much easier time evaluating $\hat{g}$.

Under certain conditions we are able to use $\hat{g}$ to construct a superadditive valid approximate lifting function for f without explicitly calculating f.

Theorem 3.3.3. *Let $\hat{h}$ be a superadditive valid approximate lifting function for $\hat{g}$. If $\hat{h}$ is non-decreasing and $\phi^i \leq \Phi^i$ is superadditive for $i = 1, \ldots, k$, then the function h defined by*

$$h\left(z^0, z^1, \ldots, z^k\right) = \hat{h}\left(z^0, \phi^1(z^1), \ldots, \phi^k(z^k)\right)$$

is a superadditive valid approximate lifting function for f.

Proof. Let $u, v, u+v \in \mathbf{D}$. Note that $u^i, v^i, u^i + v^i \in \mathbf{D}^i$, so that $\phi^i(u^i), \phi^i(v^i), \phi^i(u^i + v^i) \in \mathbf{E}^i$. By definition

$$h\left(u^0, u^1, \ldots, u^k\right) = \hat{h}\left(u^0, \phi^1(u^1), \ldots, \phi^k(u^k)\right),$$

and

$$h\left(v^0, v^1, \ldots, v^k\right) = \hat{h}\left(v^0, \phi^1(v^1), \ldots, \phi^k(v^k)\right).$$

As ϕ^i is superadditive $\phi^i(u^i) + \phi^i(v^i) \leq \phi^i(u^i + v^i)$. Therefore, because $\hat{h}$ is increasing and superadditive,

$$\begin{aligned}
&\hat{h}\left(u^0 + v^0, \phi^1(u^1 + v^1), \ldots, \phi^k(u^k + v^k)\right) \\
&\qquad \geq \hat{h}\left(u^0 + v^0, \phi^1(u^1) + \phi^1(v^1), \ldots, \phi^k(u^k) + \phi^k(v^k)\right) \\
&\qquad \geq \hat{h}\left(u^0, \phi^1(u^1), \ldots, \phi^k(u^k)\right) + \hat{h}\left(v^0, \phi^1(v^1), \ldots, \phi^k(v^k)\right).
\end{aligned}$$

Therefore $h(u^0, u^1, \ldots, u^k) + h(v^0, v^1, \ldots, v^k) \leq h(u^0 + v^0, u^1 + v^1, \ldots, u^k + v^k)$. Validity is an immediate consequence of Proposition 3.3.2. □

The assumption that $\hat{h}$ is non-decreasing is not restrictive as the lifting function $\hat{g}$ is guaranteed to be non-decreasing. Likewise, Φ^i is non-decreasing, and we will assume henceforth that ϕ^i is also non-decreasing. Observe that Φ^i is itself a lifting function. In particular, this suggests that we can recursively use this approach to construct a superadditive approximation ϕ^i of Φ^i.

The choice of μ^i has been left open. In Chapter 5, we show how a properly chosen μ^i and μ_0 can facilitate the derivation of superadditive lifting functions for several different problems. In the meantime, we apply this idea to the knapsack intersection in Example 3.3.1.

Example 3.3.2. In Example 3.3.1, recall we examined the intersection of knapsacks

$$\begin{aligned} &4x_1 + 4x_2 + 4x_3 + 2x_4 \leq 10, \\ &3x_1 + 2x_2 + 3x_3 + 10x_4 \leq 12, \\ &x_j \in \{0, 1\}. \end{aligned}$$

By lifting we were able to obtain the facet-defining inequality $x_1 + x_2 + x_3 + x_4 \leq 2$. If we replace the second knapsack in the lifting problem with a cardinality constraint, we obtain the system

$$\begin{aligned} &4x_1 + 4x_2 + 4x_3 \leq 10 - z_1, \\ &x_1 + x_2 + x_3 \leq 3 - \Phi(z_2), \\ &x_j \in \{0, 1\}. \end{aligned}$$

The function Φ is given by the integer program

$$\Phi(z_2) = \min \left\{ 3 - x_1 + x_2 + x_3 : \begin{array}{l} 3x_1 + 2x_2 + 3x_3 \leq 12 - z_2, \\ x_1, x_2, x_3 \in \{0, 1\} \end{array} \right\},$$

and evaluates explicitly to

$$\Phi(z_2) = \begin{cases} 0 & 0 \leq z_2 \leq 4 \\ 1 & 4 < z_2 \leq 7 \\ 2 & 7 < z_2 \leq 10 \\ 3 & 10 < z_2 \leq 12. \end{cases}$$

As $\Phi(10) = 2$, it follows that this relaxation still gives $\alpha_4 = 1$.

3.4 Closing Remarks

In this chapter, we demonstrated that superadditive lifting can be applied without explicitly computing the lifting function. This deviates from the typical framework of first deriving an exact lifting function and then identifying a valid superadditive approximation.

For the lifted flow cover inequality, an exact description of the second lifting function is quite cumbersome. Nevertheless, we are able to avoid this obstacle by considering a special case of this problem and reformulating the original problem to conform to this case. The original cut coefficients remain the same, but we gain access to the machinery of superadditive lifting. We are able to build on this idea and provide a general framework for constructing valid superadditive approximations by relaxing constraints in the lifting function.

Our approach to deriving approximate lifting functions indirectly resolves the question of validity through the use of relaxation and reformulation. This idea has not yet been extensively studied and seems like a promising avenue to overcoming the challenges posed by dimension in the lifting problem.

There are still many questions to consider: perhaps most importantly, identifying conditions that guarantee the non-dominance of an approximation derived using Theorem 3.3.3. The non-dominance of $\hat{h}$ is a clearly necessary condition for the non-dominance of h, but it is not sufficient. Of course, it may still be difficult to construct a non-dominated $\hat{h}$. We begin to address this question in the next chapter and show how it is possible to construct non-dominated approximations of some simple (although still quite non-trivial) lifting functions.

CHAPTER IV

SUPERADDITIVE APPROXIMATIONS OF LIFTING FUNCTIONS

In the previous chapter we discussed some general techniques that can be used to facilitate the application of superadditive lifting to higher-dimensional lifting functions. Despite the collection of work utilizing superadditive lifting, there has been very little done to address how appropriate superadditive approximations can be obtained. Past work has typically followed the framework of first deriving a closed form of the lifting function in question, and then proving that some approximation possesses desirable properties.

We seek a more constructive approach to producing superadditive approximations. This idea has received some treatment in the context of knapsack problems with disjoint cardinality constraints [80, 81], but is still problem specific. Rather than begin with a closed form description of f, we assume that we obtain f in a fairly generic form. For example, we may obtain f as the output of a dynamic program.

In this chapter, we first describe the structural properties of the lifting function and restrict our analysis to a class of functions fitting these properties. We then show that under these conditions we can test superadditivity and non-dominance in polynomial time. Next, we describe a slight modification of the superadditive approximation γ (see (32) from Chapter 2) and explore its properties. We then show that it is always possible to construct a superadditive approximation of f in polynomial time. Finally, we give an algorithm that produces a non-dominated superadditive approximation in finite time.

4.1 Structure of the Lifting Function

The problem of even testing superadditivity is completely hopeless on general functions. Consider a function f on $[0, 1]$ and suppose that the function is represented by an oracle. Any algorithm to test superadditivity sends some value z to the oracle, and the oracle returns $f(z)$. Now let f be superadditive, and suppose that an algorithm queries the oracle finitely many times and correctly declares f superadditive. Then there exists a finite k such that the algorithm sends $z_1, \ldots, z_k$ to the oracle and the oracle returns $f(z_1), \ldots, f(z_k)$. Let the function $\hat{f}(z)$ be constructing by setting $\hat{f}(z^*) = +\infty$ for $z^* \in (0,1) \setminus \{z_1, \ldots, z_k\}$ and $\hat{f}(z) = f(z)$ otherwise. The algorithm behaves no differently and will incorrectly declare $\hat{f}$ superadditive. If the algorithm is restricted to be finite but can additionally use randomness, then it will still declare $\hat{f}$ superadditive with probability 1.

Fortunately, our choice of f is not completely general. Specifically it arises as the solution of a mixed integer program:

$$\min \left\{ cx + gy : \begin{array}{c} Ax + Gy = d \\ (x,y) \in \mathbf{R}^p_+ \times \mathbf{Z}^q_+ \end{array} \right\}.$$

More generally, we can define the *value function.*

Definition 4.1.1. The *value function* of a mixed integer program is

$$\Theta(z) = \min \left\{ cx + gy : \begin{array}{c} Ax + Gy = z \\ (x,y) \in \mathbf{R}^p_+ \times \mathbf{Z}^q_+ \end{array} \right\} \tag{51}$$

The properties of Θ were established by Blair and Jeroslow [14]:

Theorem 4.1.1. *The value function, $\Theta(z)$, is piecewise linear and lower semicontinuous.*

By a simple transformation, we can apply this result to characterize the lifting function:

Corollary 4.1.2. *The lifting function $f(z)$ is non-decreasing, piecewise linear, and lower semicontinuous.*

Proof. Rewrite the lifting function

$$\begin{aligned} f(z) = \min \quad & \pi_0 t - \pi x \\ \text{s.t.} \quad & td^0 - Ax - Is = z \\ & t = 1 \\ & x \in \mathbf{R}^p_+ \times \mathbf{Z}^q_+ \\ & (s,t) \in \mathbf{R}^m_+ \times \mathbf{R}_+. \end{aligned}$$

It is a simple exercise to show that this coincides with the original lifting function. By Theorem 4.1.1, f is therefore piecewise linear and lower semicontinuous. On the other hand, if $(x^*, s^*, 1)$ is an optimal solution to $f(z)$, then for $z' \geq z$, $(x^*, s^* + (z' - z), 1)$ is feasible to $f(z')$. Therefore f is non-decreasing. $\square$

Hence, we can always assume that f is non-decreasing, piecewise linear, and lower semicontinuous. Additionally, we shall assume that $\mathbf{D}$ takes on a specific form: $\mathbf{D} = \mathbf{D}_0 \times \mathbf{D}_1$, with $\mathbf{D}_0 = [0, d_0]$ and either $\mathbf{D}_1 = \{0\}$ or

$$\mathbf{D}_1 = \{0, \ldots, d_1\} \times \cdots \times \{0, \ldots, d_m\}$$

with $d_i > 0$ for $i = 1, \ldots, m$. Note that these restrictions on $\mathbf{D}$ allow us to consider one dimensional lifting functions and lifting functions over a discrete domain. In the first case $\mathbf{D}_1 = \{0\}$ and in the latter case $\mathbf{D}_0 = \{0\}$.

For each $y \in \mathbf{D}_1$, we consider the function f_y defined on $\mathbf{D}_0$ by $f_y(z) = f(z, y)$. By Corollary 4.1.2, this function is itself piecewise linear and lower semicontinuous. We assume that f_y has finitely many breakpoints for all $y \in \mathbf{D}_1$. We relax the restriction that f is non-decreasing, and assume instead that f_y is non-decreasing for each $y \in \mathbf{D}_1$.

4.2 Efficiently Testable Conditions

In this section we show that validity, superadditivity, and non-dominance can be efficiently tested. As an important corollary we show that non-dominance implies maximality. Let f and g be two piecewise linear and lower semicontinuous functions defined over $\mathbf{D} = \mathbf{D}_0 \times \mathbf{D}_1$. Assume that for all $y \in \mathbf{D}_1$, f_y and g_y are both non-decreasing. Further, we assume throughout that $g(0) = 0$.

Definition 4.2.1. We say that a point $z \in \mathbf{D}_1$ is a *breakpoint* of f_y if $z = 0$ or $z = d_0$, f_y is discontinuous at z, or f_y changes slope at z.

The main result in this section is the classification of test points in Theorem 4.2.3. Theorem 4.2.7 is also significant in that it defines necessary and sufficient conditions for non-dominance (given our assumptions) that can be tested efficiently using these test points. These results subsequently play an integral role in the algorithmic construction of non-dominated approximations.

Let W_y and V_y denote the breakpoints of f_y and g_y respectively. Define

$$W = \bigcup_{y \in D_1} \bigcup_{w \in W_y} (w, y),$$

$$V = \bigcup_{y \in D_1} \bigcup_{v \in V_y} (V, y).$$

Hence, W and V represent the collection of breakpoints of f and g. By efficiently testable, we mean that it takes polynomially many steps with respect to $|W|$ and $|V|$ to certify some property of g.

Define the function $\bar{f}_y$ by

$$\bar{f}_y(z) = \lim_{z' \downarrow z} f_y(z'), \tag{52}$$

and let $\bar{f}_y(d_0) = f_y(d_0)$. Because f_y is non-decreasing and lower semicontinuous, $\bar{f}_y \geq f_y$. If $w \in W_y$ is a discontinuity then $\bar{f}_y(w) > f_y(w)$ otherwise the two functions coincide. We define the function $\bar{f}$ by $\bar{f}(z, y) = \bar{f}_y(z)$ and analogously define $\bar{g}$ and $\bar{g}_y$.

Suppose that $W_y = \{w_1, \ldots, w_t\}$ with $0 = w_1 < \cdots < w_t = d_0$. We can write f_y

$$f_y(z) = \begin{cases} f_y(0) & z = 0 \\ \bar{f}_y(w_j) + \lambda_j(z - w_j) & w_j < z \leq w_{j+1}, \quad (j = 1, \ldots, t-1) \end{cases} \tag{53}$$

where

$$\lambda_j = \frac{f_y(w_{j+1}) - \bar{f}_y(w_j)}{w_{j+1} - w_j}.$$

We can similarly express g_y. Therefore, we assume that f and g are encoded precisely by their breakpoints and at most two values at these breakpoints.

Proposition 4.2.1. *$f_y(z)$ and $\bar{f}_y(z)$ can be computed in time $\mathcal{O}(\log(|W_y|))$.*

Proof. We assume that the breakpoints of $f_y(z)$ are stored as a sorted list. Then using binary search, we can identify whether $z \in W_y$. Otherwise, we identify $w_1, w_2 \in W_y$ such that $w_1 < z < w_2$, and there does not exist a $w \in W_y$ such that $w_1 < w < w_2$. By interpolating f_y between these two points, we can evaluate $f_y(z)$.

As f and $\bar{f}$ only differ at breakpoints, the same result immediately applies to $\bar{f}$. □

Throughout we will ignore this log factor in our analysis, and instead focus on the number of steps our algorithms require.

4.2.1 Testing Validity

Validity is the easiest property of g to test. We say that g is *valid* if $g \leq f$. The validity of g can be efficiently verified by considering the breakpoints f and g.

Theorem 4.2.2. *g is valid if and only if $g(u) \leq f(u)$ and $\bar{g}(u) \leq \bar{f}(u)$ for all $u \in W \cup V$.*

Proof. Necessity is trivial; hence we prove sufficiency. Fix y and consider $T_y = W_y \cup V_y = \{t_1, \ldots, t_r\}$ with $t_1 < \cdots < t_r$.

Now consider some arbitrary $t \in \mathbf{D}_0$. If $t \in T_y$, then by assumption $g_y(t) \leq f_y(t)$ and $\bar{g}_y(t) \leq \bar{f}_y(t)$. Otherwise, t is not a breakpoint of g_y or f_y, so $\bar{g}_y(t) = g_y(t)$ and $\bar{f}_y(t) = f_y(t)$. Let $t_i < t < t_{i+1}$. Then by piecewise linearity and lower semicontinuity,

$$\begin{aligned} f_y(t) &= \frac{t_2 - t}{t_2 - t_1} \cdot \bar{f}_y(t_1) + \frac{t - t_1}{t_2 - t_1} \cdot f_y(t_2) \\ &\geq \frac{t_2 - t}{t_2 - t_1} \cdot \bar{g}_y(t_1) + \frac{t - t_1}{t_2 - t_1} \cdot g_y(t_2) \\ &= g(t). \end{aligned}$$

Iterating over all y, sufficiency follows. □

Therefore, we only need to test $|W| + |V|$ points of f and g to certify the validity of g. As we proceed to describing superadditivity and non-dominance, we will always assume that g is valid.

4.2.2 Testing Superadditivity

We now show that we can efficiently test superadditivity by considering only a small number of points. In particular, we are interested in which points minimize the quantity

$$f(u + v) - [f(u) + f(v)]. \tag{54}$$

By definition, f is superadditive if $f(u) + f(v) \leq f(u + v)$ for all $u, v, u + v \in \mathbf{D}$. Hence $f(u + v) - [f(u) + f(v)] \geq 0$, for all $u, v, u + v \in \mathbf{D}$: i.e.

$$\inf_{u,v,u+v \in \mathbf{D}} \{f(u + v) - [f(u) + f(v)]\} \geq 0. \tag{55}$$

Therefore we seek a succinct description of this infimum. We consider a slight generalization of (55) replacing $f(u)$ and $f(v)$ with $g(u)$ and $g(v)$. First fix y_1 and y_2, so instead we consider the problem of determining

$$\zeta_{y_1,y_2} = \inf_{z_1,z_2,z_1+z_2 \in \mathbf{D}_0} \left\{ f_{y_1+y_2}(z_1+z_2) - [g_{y_1}(z_1) + g_{y_2}(z_2)] \right\}. \tag{56}$$

We can obtain the overall infimum by taking the minimum of (56) over all y_1 and y_2 such that $y_1 + y_2 \in \mathbf{D}_1$

For this task, we will define four different but closely related quantities:

$$\begin{aligned}
\zeta_1 &= \min_{\substack{v_1 \in V_{y_1}, v_2 \in V_{y_2} \\ v_1+v_2 < d_0}} \left\{ \bar{f}_{y_1+y_2}(v_1+v_2) - [\bar{g}_{y_1}(v_1) + \bar{g}_{y_2}(v_2)] \right\} \\
\zeta_2 &= \min_{\substack{w \in W_{y_1+y_2}, v_1 \in V_{y_1} \\ w - v_1 > 0}} \left\{ f_{y_1+y_2}(w) - [\bar{g}_{y_1}(v_1) + g_{y_2}(w - v_1)] \right\} \\
\zeta_3 &= \min_{\substack{w \in W_{y_1+y_2}, v_2 \in V_{y_2} \\ w - v_2 > 0}} \left\{ f_{y_1+y_2}(w) - [g_{y_1}(w - v_2) + \bar{g}_{y_2}(v_2)] \right\} \\
\zeta_4 &= f_{y_1+y_2}(0) - [g_{y_1}(0) + g_{y_2}(0)]
\end{aligned} \tag{57}$$

The minimum is appropriate here as we are considering only a finite number of points. As it turns out, the infimum of (56) is one of these values.

Theorem 4.2.3. $\zeta_{y_1,y_2} = \min\{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$

Proof. Let $\zeta' = \min\{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$. We will first show that $\zeta' \leq \zeta_{y_1,y_2}$, and then we will demonstrate that ζ' is in fact an infimum.

For some arbitrary $z_1, z_2, z_1 + z_2 \in \mathbf{D}_0$ consider

$$\zeta(z_1, z_2) = f_{y_1+y_2}(z_1 + z_2) - [g_{y_1}(z_1) + g_{y_2}(z_2)]. \tag{58}$$

If $z_1 = z_2 = z_1 + z_2 = 0$ then this quantity is precisely equal to ζ_4. So assume that at least one of z_1 and z_2 is strictly positive.

The remainder of the proof is driven by a simple idea depicted in Figure 12. If any two of the points u, v, or $u + v$ are not breakpoints, then fixing the third, there is always a direction that decreases the gap ζ.

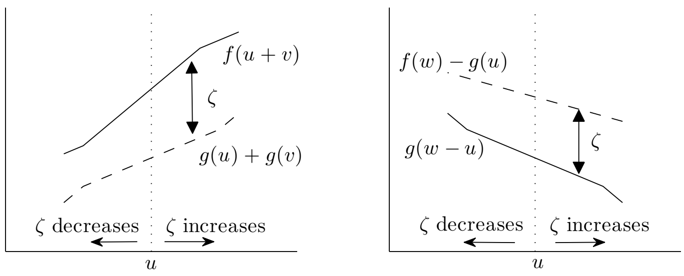

Figure 12: $f(u+v)$ and $g(u)+g(v)$ with v fixed (left), and $g(w-u)$ and $f(w)-g(u)$ with w fixed (right)

First suppose that neither z_1 nor z_2 are breakpoints. Let λ_1 and λ_2 denote the slopes of g_{y_1} at z_1 and g_{y_2} at z_2. Without loss of generality, we may assume that $\lambda_1 \leq \lambda_2$. Let

$$v_1 = \max\{v \in V_{y_1} : v < z_1\},$$

$$v_2 = \min\{v \in V_{y_2} : v > z_2\}.$$

Then $z_1 = v_1 + \epsilon_1$ and $z_2 = v_2 - \epsilon_2$ for some $\epsilon_1, \epsilon_2 > 0$. Let $\epsilon' = \min\{\epsilon_1, \epsilon_2\}$. For small $0 < \epsilon < \epsilon'$,

$$\begin{aligned} g_{y_1}(z_1 - \epsilon) + g_{y_2}(z_2 + \epsilon) &= g_{y_1}(z_1) + g_{y_2}(z_2) + (\lambda_2 - \lambda_1)\epsilon \\ &\geq g_{y_1}(z_1) + g_{y_2}(z_2). \end{aligned}$$

There are two possibilities to consider: either $\epsilon_1 = \epsilon_2$, or $\epsilon_1 \neq \epsilon_2$.

First, suppose that $\epsilon_1 = \epsilon_2$. By taking the limit as we increase ϵ to ϵ', we conclude that

$$\zeta(z_1, z_2) \geq f_{y_1+y_2}(v_1 + v_2) - [\bar{g}_{y_1}(v_1) + g_{y_2}(v_2)].$$

If $v_1 + v_2 \in W_{y_1+y_2}$, then this implies that $\zeta(z_1, z_2) \geq \zeta_2$. Otherwise, $v_1 + v_2$ is not a breakpoint of $f_{y_1+y_2}$, and therefore $\bar{f}_{y_1+y_2}(v_1 + v_2) = f_{y_1+y_2}(v_1 + v_2)$. Furthermore, this assumption implies that $v_1 + v_2 < d_0$. Thus,

$$\begin{aligned} &\bar{f}_{y_1+y_2}(v_1 + v_2) - [\bar{g}_{y_1}(v_1) + \bar{g}_{y_1}(v_1)] \\ &\qquad = \lim_{\epsilon \downarrow 0} f_{y_1+y_2}(v_1 + v_2 + \epsilon) - [\bar{g}_{y_1}(v_1) + g_{y_1}(v_1 + \epsilon)] \\ &\qquad \leq f_{y_1+y_2}(v_1 + v_2) - [\bar{g}_{y_1}(v_1) + g_{y_1}(v_1)]. \end{aligned}$$

Note that the last inequality follows from the continuity of $f_{y_1+y_2}$ at $z_1 + z_2$ and lower semicontinuity of g_{y_2}. Therefore, we conclude in this case that $\zeta(z_1, z_2) \geq \zeta_1$.

It remains to consider when $\epsilon_1 \neq \epsilon_2$. We only consider $\epsilon_1 < \epsilon_2$; the other case follows similarly. Let $z_2' = z_2 + \epsilon_1$. By assumption z_2' is not a breakpoint. If $v_1 + z_2'$ is a breakpoint of $f_{y_1+y_2}$, then we again have that $\zeta(z_1, z_2) \geq \zeta_2$. So we shall assume that neither z_2' nor $v_1 + z_2'$ are breakpoints.

Let μ_1 and μ_2 denote the slopes of $f_{y_1+y_2}$ at $v_1 + z_2'$ and g_{y_2} at z_2'. Either $\mu_1 \geq \mu_2$ or $\mu_1 < \mu_2$. For brevity, we shall only consider the former possibility, noting that the latter case can be handled using the same techniques. Let

$$
\begin{aligned}
w &= \min\left\{w' \in W_{y_1+y_2} : w' < v_1 + z_2'\right\}, \\
v_2 &= \min\left\{v \in V_{y_2} : v < z_2'\right\}.
\end{aligned}
$$

Thus, $v_1 + z_2' = w + \epsilon_1$ and $z_2' = v_2 + \epsilon_2$ for some $\epsilon > 0$. Again define $\epsilon' = \min\{\epsilon_1, \epsilon_2\}$. For $0 < \epsilon < \epsilon'$,

$$
\begin{aligned}
&f_{y_1+y_2}(v_1 + z_2' - \epsilon) - \left[\bar{g}_{y_1}(v_1) + g_{y_2}(z_2' - \epsilon)\right] \\
&\qquad = f_{y_1+y_2}(v_1 + z_2') - \left[\bar{g}_{y_1}(v_1) + g_{y_2}(z_2')\right] - (\mu_1 - \mu_2)\epsilon \\
&\qquad \leq f_{y_1+y_2}(v_1 + z_2') - \left[\bar{g}_{y_1}(v_1) + g_{y_2}(z_2')\right] \leq \zeta(z_1, z_2)
\end{aligned}
$$

We consider two possibilities: either $\epsilon_1 \geq \epsilon_2$ or $\epsilon_1 < \epsilon_2$. If $\epsilon_1 \geq \epsilon_2$, then by taking the limit as we increase ϵ to ϵ', we have

$$
\zeta(z_1, z_2) \geq \bar{f}_{y_1+y_2}(v_1 + v_2) - [\bar{g}_{y_1}(v_1) + \bar{g}_{y_2}(v_2)] \geq \zeta_1.
$$

So assume that $\epsilon_1 < \epsilon_2$. Again taking the limit, we conclude

$$
\zeta(z_1, z_2) \geq \bar{f}_{y_1+y_2}(v_1 + v_2) - [\bar{g}_{y_1}(v_1) + \bar{g}_{y_2}(w - v_1)] \geq \zeta_2.
$$

The remaining cases are handled analogously to conclude that $\zeta' \leq \zeta_{y_1,y_2}$.

Now we must show that ζ' is an infimum. By scaling f and g appropriately, we may assume that all slopes are bounded by 0 and 1. Now let $\zeta'' = \zeta' + \epsilon$ for $\epsilon > 0$. We must show that there exists some z_1 and z_2 such that $\zeta(z_1, z_2) < \zeta''$. If $\zeta' = \zeta_4$, then this is clear, so we consider when $\zeta' = \zeta_1$ and $\zeta' = \zeta_2$ (ζ_3 being handled identically).

If $\zeta' = \zeta_1$, then let v_1 and v_2 satisfy

$$
\zeta' = \bar{f}_{y_1+y_2}(v_1 + v_2) - [\bar{g}_{y_1}(v_1) + \bar{g}_{y_2}(v_2)].
$$

Then by our assumption about the slopes of f and g,

$$\begin{aligned} f_{y_1+y_2}(v_1+v_2+2\epsilon') - \left[g_{y_1}(v_1+\epsilon') + g_{y_2}(v_2+\epsilon')\right] & \\ \geq \bar{f}_{y_1+y_2}(v_1+v_2) - [\bar{g}_{y_1}(v_1) + \bar{g}_{y_2}(v_2)] + 2\epsilon' & \\ = \zeta' + 2\epsilon'. & \end{aligned}$$

Thus taking $\epsilon' < \epsilon/2$, we have shown that ζ'' cannot be a lower bound.

Otherwise, if $\zeta' = \zeta_2$, then let w and v satisfy

$$\zeta' = f_{y_1+y_2}(w) - [\bar{g}_{y_1}(v) + \bar{g}_{y_2}(w-v)]\,.$$

Similarly, we can apply our assumption about the slopes to conclude

$$\begin{aligned} f_{y_1+y_2}(w) - \left[g_{y_1}(v+\epsilon') + g_{y_2}(w-v-\epsilon')\right] & \\ \geq f_{y_1+y_2}(w) - [\bar{g}_{y_1}(v) + g_{y_2}(w-v)] + \epsilon' & \\ = \zeta' + \epsilon''. & \end{aligned}$$

So restricting $\epsilon' < \epsilon$, we again have shown that ζ'' is not a valid lower bound. Therefore, ζ' is an infimum. □

This theorem has an immediate consequence of allowing us to test in polynomial time the superadditivity of a function.

Corollary 4.2.4. *Superadditivity for f can be tested in $\mathcal{O}(|W|^2)$ time.*

Proof. The functions f and g in Theorem 4.2.3 were arbitrary. Therefore we can replace g with f. To verify (55), we simply compute

$$\zeta = \min_{y_1,y_2,y_1+y_2\in D_1} \zeta_{y_1,y_2}.$$

f is superadditive if and only if $\zeta \geq 0$. As we only need to test pairs of points, to verify superadditivity, we test at most $\mathcal{O}(|W|^2)$ points. □

4.2.3 Testing Non-Dominance

The task of certifying non-dominance seems a greater challenge. We will show, however, that for our class of functions, it is just as tractable as testing superadditivity. Throughout this section, we assume that the validity and superadditivity of g have already been established.

We first show a sufficient condition for non-dominance and provide a simple extension of Theorem 4.2.3. Finally we describe a procedure for testing non-dominance in polynomial time.

Proposition 4.2.5. *Suppose that $g \leq f$ is superadditive. If for every $u \in \mathbf{D}$ there exists some $v \in \mathbf{D}$ such that $g(u) + g(v) = f(u+v)$, then g is non-dominated.*

Proof. Suppose to the contrary that there exists some superadditive $g' \leq f$ such that $g' \geq g$ and $g'(u) > g(u)$ for some u. Let $v \in \mathbf{D}$ satisfy $g(u) + g(v) = f(u+v)$. Then

$$g'(u) + g'(v) > g(u) + g(v) = f(u+v) \geq g'(u+v),$$

but this contradicts that g' is superadditive. □

The sufficiency of this condition is general. It does not depend on any other properties of f, g or $\mathbf{D}$. We will show that under our restrictions, this condition is also necessary. However, before proceeding we dispel the notion that necessity holds in more general settings:

Example 4.2.1. Consider the function $f : \mathbf{D} \to \mathbf{R}$ where $\mathbf{D} = \{-3, -2, -1, 0, +1, +2, +3\}$ and an associated approximation g given in Table 3:

Table 3: Non-dominated approximation with $\Delta(u) > 0$

u	$+3$	$+2$	$+1$	0	-1	-2	-3
$f(u)$	$+4$	$+4$	$+4$	0	-2	-6	-6
$g(u)$	$+4$	$+4$	0	0	-4	-6	-8

Validity of g is evident and superadditivity is a simple exercise. Now we examine g in more detail at $u = -1$ and $u = -3$. Note that the remaining u either satisfy $g(u) = f(u)$ or there exists some v such that $g(u) + g(v) = f(u+v)$; therefore we cannot increase g at those points. Clearly

$$g(+2) + g(-1) = +4 - 4 = 0 = g(+1).$$

So we cannot increase $g(-1)$ without decreasing $g(+2)$. Similarly,

$$g(+2) + g(-3) = +4 - 8 = -4 = g(-1).$$

However, we showed that we cannot increase $g(-1)$, so we cannot increase $g(-3)$ without decreasing $g(+2)$. In particular, g must be non-dominated.

On the other hand, as the Table 4 demonstrates, there does not exist any u such that $g(u) + g(-1) = f(u-1)$:

Table 4: Failure of $\Delta(u)$ to capture non-dominance

u	$+3$	$+2$	$+1$	0	-1	-2
$f(u-1)$	$+4$	$+4$	$+4$	0	-2	-6
$g(u) + g(-1)$	0	0	-4	-4	-8	-10

Therefore, the non-dominance condition is not necessary in general. □

Returning to our setting, for all $u \in \mathbf{D}$, we can define the function

$$\Delta(u) = \inf_{v, u+v \in \mathbf{D}} \{ f(u+v) - [g(u) + g(v)] \}. \tag{59}$$

As before, we can greatly simplify the computation of $\Delta(u)$. Letting

$$\begin{aligned} \Delta_1(u) &= \min_{w \in W : w-u \in \mathbf{D}} f(w) - [g(u) + g(w-u)] \\ \Delta_2(u) &= \min_{\substack{v \in V : u+v \in \mathbf{D} \\ u_1+v_1 < d_0}} \bar{f}(u+v) - [g(u) + \bar{g}(v)], \end{aligned} \tag{60}$$

we can characterize $\Delta(u)$:

Proposition 4.2.6. $\Delta(u) = \min \{ \Delta_1(u), \Delta_2(u) \}$.

Proof. Consider $f(u+v) - [g(u) + g(v)]$, and suppose that neither $u+v \in W$ nor $v \in V$. Let $u = (z_1, y_1)$ and $v = (z_2, y_2)$. Then let λ_1 and λ_2 denote the slope of $f_{y_1+y_2}$ at $z_1 + z_2$ and g_{y_2} at z_2. Further, let $w' < z_1 + z_2 < w''$ and $v' < z_2 < v''$, for breakpoints $w', w'' \in W_{y_1+y_2}$ and $v', v'' \in V_{y_2}$.

If $\lambda_1 \geq \lambda_2$, then for sufficiently small $\epsilon > 0$,

$$\begin{aligned} f_{y_1+y_2}(z_1 + z_2 - \epsilon) - &[g_{y_1}(z_1) + g_{y_2}(z_2 - \epsilon)] \\ &= f_{y_1+y_2}(z_1 + z_2) - [g_{y_1}(z_1) + g_{y_2}(z_2)] - (\lambda_1 - \lambda_2)\epsilon \\ &\geq f_{y_1+y_2}(z_1 + z_2) - [g_{y_1}(z_1) + g_{y_2}(z_2)] \, . \end{aligned}$$

Letting $\epsilon' = \min(z_1 + z_2 - w', z_2 - v')$, we consider the limit

$$\lim_{\epsilon \uparrow \epsilon'} f_{y_1+y_2}(z_1 + z_2 - \epsilon) - [g_{y_1}(z_1) + g_{y_2}(z_2 - \epsilon)] \, .$$

By construction if $\epsilon' = z_2 - v'$, then $f(u + v) - [g(u) + g(v)] \geq \Delta_2(u)$, and if $\epsilon' < z_2 - v'$ then $f(u + v) - [g(u) + g(v)] \geq \Delta_1(u)$.

On the other hand, if $\lambda_1 < \lambda_2$, then we instead increase z_2. Again we stop when we hit a breakpoint and similarly show that $f(u + v) - [g(u) + g(v)] \geq \Delta(u)$. In particular, we have shown that $\Delta(u) \geq \min(\Delta_1(u), \Delta_2(u))$. The proof that this value is in fact equal to the infimum follows similarly to the proof of Theorem 4.2.3, and is omitted for brevity. $\square$

Define $u^*_y = \inf \{u' \in \mathbf{D}_0 : \Delta(u, y) = 0, \ \forall u \in \mathbf{D}_0, u > u'\}$. We can therefore recast our non-dominance condition as follows: $u^*_y = 0$ and $\Delta(0, y) = 0$ for all $y \in \mathbf{D}_1$. Indeed, as $\Delta(u)$ is an infimum, the condition that $\Delta(u) = 0$ implies that we cannot increase $g(u)$ without there existing some v such that $g(u) + g(v) > f(u + v)$.

For now we simply state the theorem, and defer its proof until Section 4.5 where it will follow immediately from the algorithm to construct a non-dominated approximation.

Theorem 4.2.7. *Suppose f_y and g_y are non-decreasing, piecewise linear, and lower semi-continuous. If $g \leq f$ is superadditive, then g is non-dominated if and only if $\Delta(u) = 0$ for all $u \in \mathbf{D}$.*

This theorem yields an immediate corollary that shows non-dominance implies maximality whenever f, g, and $\mathbf{D}$ fit our specified conditions:

Corollary 4.2.8. *If $g \leq f$ is a non-dominated superadditive approximation, then g is maximal.*

Proof. As g is non-dominated, for every $u \in \mathbf{D}$, $\Delta(u) = 0$. Moreover, for any $u \in \mathbf{E}$, $f(u+v) - [f(u) + f(v)] \geq 0$ for all $v, u+v \in \mathbf{D}$. If $g(u) < f(u)$, then

$$f(u+v) - [g(u) + g(v)] \geq f(u+v) - [f(u) + f(v)] + [f(u) - g(u)] > 0.$$

Thus $\Delta(u) > 0$ contradicting that g is non-dominated. □

Now to prove a function is non-dominated, it suffices to compute u_y^* for all $y \in \mathbf{D}_1$. If $u_y^* > 0$ for any y, then the function cannot be non-dominated. If $u_y^* = 0$ for all y, then we also test whether $\Delta(0, y) = 0$. As a trivial consequence of Proposition 4.2.6 we can efficiently compute $\Delta(0, y)$, so we must show that we can compute u_y^* efficiently.

Theorem 4.2.9. *In polynomial time, we can identify u_y^*. In particular, we can test non-dominance in polynomial time.*

Proof. We will prove this theorem by describing an algorithm that computes u_y^* efficiently. Initialize the algorithm by setting $\bar{u} = (d_0, y)$. We successively update $\bar{u}$ to compute u_y^*.

For notational convenience let $\bar{u} = (\bar{z}, y)$, and assume inductively that $u_y^* \leq \bar{z}$. Moreover, we may assume that $\bar{z} > 0$; otherwise $u_y^* = 0$. Set $\Delta_{\min} = f(\bar{u}) - g(\bar{u})$. Define sets

$$W(u) = \{w \in W : w - u \in \mathbf{D}\},$$

$$V(u) = \{v \in V : u + v \in \mathbf{D}\}.$$

Constructing these sets for $u = \bar{u}$ takes at most $\mathcal{O}(|W| + |V|)$ steps, but it can be considerably accelerated if they have been computed in a previous iteration.

For each $w \in W(\bar{u})$,

$$\Delta(\bar{u} : w) = f(w) - [g(\bar{u}) + g(w - u)].$$

Update $\Delta_{\min} = \min\{\Delta_{\min}, \Delta(\bar{u} : w)\}$. Similarly for each $v \in V(\bar{u})$, let $v = (\tilde{z}, \tilde{y})$. If $\tilde{z} + \bar{z} < d_0$, then compute

$$\Delta(\bar{u} : v) = \bar{f}(\bar{u} + v) - [g(\bar{u}) + \bar{g}(v)].$$

Update $\Delta_{\min} = \min\{\Delta_{\min}, \Delta(\bar{u} : v)\}$. As $|W(\bar{u})| + |V(\bar{u})| \leq |W| + |V|$, this step takes a linear number of iterations.

If $\Delta_{\min} > 0$, then we have shown that $\Delta(\bar{u}) > 0$, and therefore by our inductive hypothesis, $u_y^* = \bar{z}$. Otherwise, if $\Delta_{\min} = 0$, then we must test whether we can update $\bar{u}$. Let λ denote the slope of g_y to the left of $\bar{z}$, and let $z' < \bar{z}$ denote the breakpoint of g_y preceding $\bar{z}$. Therefore, for all $z \in (z', \bar{z})$,

$$g_y(z) = g_y(\bar{z}) - \lambda(\bar{z} - z).$$

Now we must identify a new $\bar{u}$ if appropriate.

Let $\nu_0 = 0$. For each $w = (\hat{z}, \hat{y}) \in W(\bar{u})$, compute

$$\bar{\Delta}(\bar{u} : w) = f(w) - [g(\bar{u}) + \bar{g}(w - u)] .$$

If $\bar{\Delta}(\bar{u} : w) = 0$, then let $z'' \in V_{\hat{y}-y}$ denote the smallest breakpoint of $g_{\hat{y}-y}$ strictly greater than $\hat{z} - \bar{z}$. Letting μ_w denote the slope of $g_{\hat{y}-y}$ to the right of $\hat{z}$,

$$g_{\hat{y}-y}(z) = \bar{g}_{\hat{y}-y}(\hat{z} - \bar{z}) + \mu_w \left(z - (\hat{z} - \bar{z})\right),$$

for all $z \in (\hat{z} - \bar{z}, z'']$. If $\mu_w = \lambda$, then set $\nu_0 = \max\{\nu_0, z'' - (\hat{z} - \bar{z})\}$.

For $v = (\tilde{z}, \tilde{y}) \in V(\bar{u})$, compute

$$\bar{\Delta}(\bar{u} : v) = f(\bar{u} + v) - [g(\bar{u}) + \bar{g}(v)] .$$

If $\bar{\Delta}(\bar{u} : v) = 0$, then let $z'' \in W_{y+\tilde{y}}$ denote the largest breakpoint of $f_{y+\tilde{y}}$ that is strictly smaller than $\bar{z} + \tilde{z}$. Again letting μ_v denote the slope of $f_{y+\tilde{y}}$ to the left of $\bar{z} + \tilde{z}$,

$$f_{y+\tilde{y}}(z) = f_{y+\tilde{y}}(\bar{z} + \tilde{z}) - \mu_v(\bar{z} + \tilde{z} - z),$$

for all $z \in (z'', \bar{z} + \tilde{z}]$. If $\mu_v = \lambda$, then set $\nu_0 = \max(\nu_0, \bar{z} + \tilde{z} - z'')$.

We set $\nu = \min\{\nu_0, \bar{z} - z'\}$. If $\nu = 0$, then we claim that $u_y^* = \bar{z}$; otherwise, we claim that we can set $\bar{u} = (\bar{z} - \nu, y)$ (shown in Figure 13).

First suppose that $\nu = 0$. As before, observe that for small $\epsilon > 0$, then $\tilde{u} = (\bar{z} - \epsilon, y)$, satisfies $W(\tilde{u}) = W(\bar{u})$ and $V(\tilde{u}) = V(\bar{u})$. For $w \in W(\bar{u})$,

$$f(w) - [g(\tilde{u}) + g(w - \tilde{u})] = f(w) - [g(\bar{u}) + \bar{g}(w - \bar{u})] + (\lambda - \mu_w)\epsilon.$$

Again we consider $f(w) - [g(\bar{u}) + \bar{g}(w - \bar{u})]$. If this quantity is equal to 0, then $\lambda > \mu_w$ by the validity of g. Otherwise, it exceeds some small positive value Δ'. We may assume

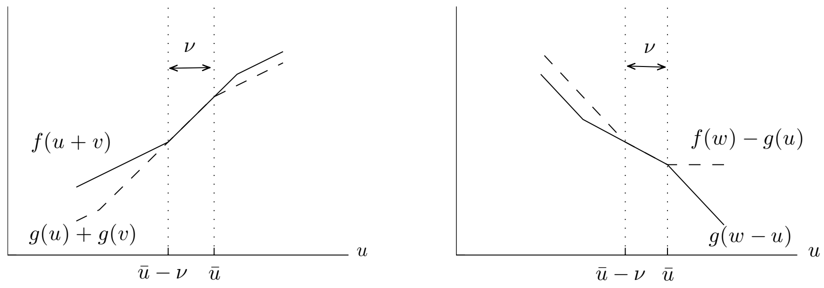

Figure 13: Possible updates for $\bar{u}$ with (a) v fixed or (b) w fixed.

by scaling that all slopes are between 0 and 1; hence we choose $\epsilon < \Delta'$. In particular, this implies that $\Delta(\tilde{u} : w) > 0$. A similar argument applies to $v \in V(\bar{u})$.

Next suppose that $\nu > 0$. Then there exists either some $w \in W(\bar{u})$ or some $v \in V(\bar{u})$ such that $\bar{\Delta}(\bar{u} : w) = 0$ and $\mu_w = \lambda$ or $\bar{\Delta}(\bar{u} : v) = 0$ and $\mu_v = \lambda$. Assume that for some such $w = (\hat{z}, \hat{y}) \in W$, $z'' - (\hat{z} - \bar{z}) \geq \nu$ with z'' defined as before. By construction, for all $0 < \epsilon < \nu$,

$$f(w) - [g(\tilde{u}) + g(w - \tilde{u})] = f(w) - [g(\bar{u}) + \bar{g}(w - \bar{u})] + (\lambda - \mu_w)\epsilon,$$

where $\tilde{u} = (\bar{z} - \epsilon, y)$. Therefore, $\Delta(\tilde{u}) = 0$. We proceed analogously for $v \in V(\bar{u})$.

We have shown that our update is valid in the sense that $\bar{z} - \nu$ is a valid upper bound for u^*_y. Further we have shown that we can perform the update in $\mathcal{O}(|W| + |V|)$ steps. Now observe that the updated $\bar{u}$ will either belong to V_y or is equal to $w - v$ for some $w \in W$ and $v \in V$. Thus we must perform at most $\mathcal{O}(|W||V| + |V|)$ iterations. □

4.3 A Modified Superadditive Approximation

There are many different superadditive approximations that we can construct. For example, let $\eta = \min_{z \in \mathbf{D}} f(z)$, then setting $g(z) = \min(\eta, 0)$ is a valid, albeit fairly trivial superadditive approximation of f. Of course, such an approximation would be entirely ill-suited for the purposes of superadditive lifting.

We shall instead revisit the function γ studied in [40, 76]:

$$\gamma(u) = \inf_{v, u+v \in \mathbf{D}} \{f(u + v) - f(v)\}. \tag{61}$$

For our particular choice of $\mathbf{D}$, γ is valid, superadditive, and maximal. Unfortunately, the function γ is not guaranteed to be lower semicontinuous even when f is lower semicontinuous (as in Example 3.1.3). Therefore we impose a slight modification on γ, and instead operate with the function α defined by

$$\alpha(u) = \inf_{v, u+v \in \mathbf{D}} \left\{ f(u+v) - \bar{f}(v) \right\}. \tag{62}$$

We will begin by establishing the properties of α. We will then show how α can be computed incrementally to produce a strengthened approximation. Finally, we will show that α can be constructed in polynomial time.

4.3.1 Properties of α

Here we show that α is computationally easier to handle than γ and still possesses many nice properties. First we will show that the infimum of (62) is in fact attained by some v such that either $v \in W$ or $u + v \in W$. As before, for each $y \in \mathbf{D}_1$ we can define a function α_y by $\alpha_y(z) = \alpha(z, y)$.

Proposition 4.3.1. *The infimum of (62) is attained by some v such that either $v \in W$ or $u + v \in W$.*

Proof. Let $u = (z_1, y_1)$ and let $v = (z_2, y_2)$. Suppose that neither v nor $u+v$ is a breakpoint of f. Then let λ denote the slope of $f_{y_1+y_2}$ at $z_1 + z_2$ and let μ denote the slope of f_{y_2} at z_2.

If $\lambda \geq \mu$, then we decrease z_2. Let $z' \in W_{y_2}$ denote the largest breakpoint strictly less than z_2, and let $z'' \in W_{y_1+y_2}$ be the largest breakpoint strictly less than $z_1 + z_2$. Choose $\epsilon' = \min(z_2 - z', z_1 + z_2 - z'')$. Taking the limit,

$$\begin{aligned} f_{y_1+y_2}(z_1 + z_2) - f_{y_2}(z_2) &\geq \lim_{\epsilon \uparrow \epsilon'} f_{y_1+y_2}(z_1 + z_2 - \epsilon) - f_{y_2}(z_2 - \epsilon) \\ &= \bar{f}_{y_1+y_2}(z_1 + z_2') - \bar{f}_{y_2}(z_2') \\ &\geq f_{y_1+y_2}(z_1 + z_2') - \bar{f}_{y_2}(z_2'), \end{aligned}$$

where $z_2' = z_2 - \epsilon'$. We proceed similarly if $\lambda < \mu$. Noting the last inequality, it follows by construction that the infimum is actually attained. $\square$

Now that we know the infimum in α is attained, we can easily prove that α is indeed superadditive.

Proposition 4.3.2. *α is superadditive.*

Proof. Consider $u, v, u+v \in \mathbf{D}$. Then by construction

$$\begin{aligned}\alpha(u+v) &= f(u+v+w) - \bar{f}(w)\\ &\geq f(u+v+w) - \bar{f}(v+w) + f(v+w) - \bar{f}(w)\\ &\geq \alpha(u) + \alpha(v).\end{aligned}$$

The first inequality follows from the relation $\bar{f}(z) \geq f(z)$, and the latter follows by the definition of α. $\square$

Next we show that the function α is well-behaved with respect to our conditions.

Proposition 4.3.3. *For all $y \in \mathbf{D}_1$, α_y is non-decreasing, piecewise linear, and lower semicontinuous.*

Proof. We first show that α_y is non-decreasing. Let $u = (z_1, y)$ and $v = (z_2, y)$ be chosen such that $z_1 \leq z_2$. Let

$$\alpha_y(z_2) = f_{y+y'}(z_2 + z') - \bar{f}_{y'}(z').$$

Then $f_{y+y'}(z_1 + z') \leq f_{y+y'}(z_2 + z')$; thus $\alpha_y(z_1) \leq f_{y+y'}(z_2 + z') - f_{y'}(z') = \alpha_y(z_2)$.

Next we show that α_y is piecewise linear. Let $w = (z', y') \in W$. Define the function $\alpha'_{y,w}$

$$\alpha'_{y,w}(z) = \begin{cases} f_{y'}(z') - \bar{f}_{y'-y}(z'-z) & 0 \leq z \leq z' \\ +\infty & z' < z \leq d_0 \\ +\infty & y' - y \notin \mathbf{D}_1. \end{cases}$$

Likewise, define the function $\alpha''_{y,w}$

$$\alpha''_{y,w}(z) = \begin{cases} f_{y+y'}(z+z') - \bar{f}_{y'}(z') & 0 \leq z \leq d_0 - z' \\ +\infty & d_0 - z' < z \leq d_0 \\ +\infty & y + y' \notin \mathbf{D}_1. \end{cases}$$

By Proposition 4.3.1, $\alpha_y(z)$ can be computed by taking the minimum of these functions. Therefore, α_y is the minimum of finitely many piecewise linear functions, and is thus piecewise linear.

Lastly we show that α_y is lower semicontinuous. Fix some $\bar{z} \in \mathbf{D}_0$, and consider the limit of $\bar{z} - \epsilon$ as ϵ decreases to 0. Now let

$$W'_y(z) = \{w \in W : w - (z, y) \in \mathbf{D}\} = \{w \in W : w \geq (z, y)\},$$

$$W''_y(z) = \{w \in W : w + (z, y) \in \mathbf{D}\} = \left\{(z', y') \in W : z' + z \leq d_0, y + y' \in \mathbf{D}_1\right\}.$$

For $\epsilon > 0$ sufficiently small $W'_y(\bar{z} - \epsilon) = W'_y(\bar{z})$ and $W''_y(\bar{z} - \epsilon) = W''_y(\bar{z})$.

Now assume by scaling that no slope of f is greater than 1. Thus for all $(z', y') \in W'_y(\bar{z})$ and $\epsilon > 0$ small,

$$\begin{aligned} f_{y'}(z') - f_{y'-y}(z' - \bar{z} + \epsilon) &\geq f_{y'}(z') - f_{y'-y}(z' - \bar{z}) - \epsilon \\ &\geq \alpha_y(\bar{z}) - \epsilon. \end{aligned}$$

Similarly for $(z', y') \in W''_y(\bar{z})$,

$$\begin{aligned} f_{y'+y}(\bar{z} + z' - \epsilon) - f_{y'}(z') &\geq f_{y'+y}(\bar{z} + z') - f_{y'}(z') - \epsilon \\ &\geq \alpha_y(\bar{z}) - \epsilon. \end{aligned}$$

In particular, by taking the minimum over all such terms, we conclude that $\alpha_y(\bar{z} - \epsilon) \geq \alpha_y(\bar{z}) - \epsilon$ for small $\epsilon > 0$. Therefore, $\lim_{\epsilon \downarrow 0} \alpha_y(\bar{z} - \epsilon) = \alpha_y(\bar{z})$, proving lower semicontinuity. □

Next we consider the slopes that the function α_y can attain. In particular, these slopes must coincide with f.

Proposition 4.3.4. *If α_y has slope λ at z, then there exists some $(z', y') \in \mathbf{D}$ such that $f_{y'}$ has slope λ at z'.*

Proof. Recall the definitions of $\alpha'_{y,w}$ and $\alpha''_{y,w}$ from the previous proof and for convenience, let $w = (z', y')$. The slope of $\alpha'_{y,w}$ at z is λ if and only if the slope of $f_{y'-y}$ is λ at $(z' - z)$. Likewise the slope of $\alpha''_{y,w}$ at z is λ if and only if the slope of $f_{y'+y}$ at $z + z'$ is λ.

If the slope of α_y at some z is λ, then there exists some $w_1, w_2 \in W$ (possibly the same) such that the slope of α'_{y,w_1} or α''_{y,w_1} to the right of z is λ and the slope of α'_{y,w_2} or α''_{y,w_2} to the left of z is λ. $\square$

The function α possesses the desirable properties that we exploited in the previous section to characterize non-dominance. Moreover, it behaves nicely at discontinuities and is much easier to deal with computationally than γ. Therefore, we instead operate with α at a very slight expense. One slight nuance is that α may not be maximal anymore as we show in the following example.

Example 4.3.1. We revisit the example from the previous chapter. Let f be defined by

$$f(z) = \begin{cases} 0 & 0 \le z \le 1 \\ 1 & 1 < z \le 4. \end{cases}$$

Omitting the details α is given by

$$\alpha(z) = \begin{cases} 0 & 0 \le z \le 3 \\ 1 & 3 < z \le 4. \end{cases}$$

Note that $f(3) + f(z) = 1 = f(3+z)$ for $0 \le z \le 1$. However as $\alpha(3) < f(3)$, it follows that α is not maximal.

Recall that $\mathbf{E} = \{u \in \mathbf{D} : f(u) + f(v) \le f(u+v),\ \forall v, u+v \in \mathbf{D}\}$. A superadditive approximation $g \le f$ is maximal if $g(u) = f(u)$ for all $u \in \mathbf{E}$. Let

$$\mathbf{E}' = \left\{u \in \mathbf{D} : f(u) + \bar{f}(v) \le f(u+v),\ \forall v, u+v \in \mathbf{D}\right\}.$$

Clearly $\mathbf{E}' \subseteq \mathbf{E}$. We show how α behaves over $\mathbf{E}'$ and where $\mathbf{E}'$ and $\mathbf{E}$ differ.

Proposition 4.3.5. *$\alpha(u) = f(u)$ if and only if $u \in \mathbf{E}'$.*

Proof. If $u \in \mathbf{E}'$, then clearly $\alpha(u) = f(u)$. Otherwise, there exists some v such that $f(u) + \bar{f}(v) > f(u+v)$, hence $\alpha(u) \le f(u+v) - \bar{f}(v) < f(u)$. $\square$

Proposition 4.3.6. *If $u \in \mathbf{E} \setminus \mathbf{E}'$, then $u = w_1 - w_2$ for some $w_1, w_2 \in W$.*

Proof. Suppose that $u \neq w_1 - w_2$ for some $w_1, w_2 \in W$. By applying Proposition 4.2.6 with $g = f$, it follows that $f(u+v) - f(v)$ is always at least

$$\bar{f}(u+w) - \bar{f}(w),$$

for $w \in W$. However, this implies that $u + w \notin W$, and therefore $\bar{f}(u+w) = f(u+w)$. Likewise, $f(u+v) - f(v)$ is at least

$$f(w) - f(w-u),$$

for some $w \in W$. Necessarily $w - u \notin W$, thus $f(w-u) = \bar{f}(w-u)$. This implies that α and γ can differ only when $u = w_1 - w_2$ for $w_1, w_2 \in W$. In particular, if $u \in \mathbf{E} \setminus \mathbf{E}'$, then $\alpha(u) < \gamma(u)$ so $u = w_1 - w_2$. □

Therefore, we see that α and γ can only differ in at most $\mathcal{O}(|W|^2)$ points. As a slight modification, we now say that a superadditive approximation $g \leq f$ is *maximal* if $g(u) = f(u)$ for all $u \in \mathbf{E}'$ with the understanding that only a small number of points need to be tested and adjusted to satisfy maximality in its original sense.

4.3.2 Nested Application of α

We described earlier that γ is typically dominated. In this regard, α is no different. We show that we can strengthen α by updating the function in steps. Let

$$\mathbf{D}^i = \left\{ z \in \mathbf{D} : z \leq u^i \right\},$$

for $i = 1, \ldots, r$, such that $u^1 \leq \cdots \leq u^r$. Further, assume that $\emptyset \subsetneq \mathbf{D}^1 \subsetneq \cdots \subsetneq \mathbf{D}^r = \mathbf{D}$.

Next we slightly modify our definition of α so that it is parametrized by f. We define the function $\alpha(\cdot : f)$, by

$$\alpha(u : f) = \min_{v, u+v \in \mathbf{D}} \left\{ f(u+v) - \bar{f}(v) \right\}. \tag{63}$$

Next we define a collection of functions for $i = 1, \ldots, r$:

$$f^i(u) = \begin{cases} \alpha\left(u : f^{i-1}\right) & u \in \mathbf{D}^i \\ f(u) & u \in \mathbf{D} \setminus \mathbf{D}^i, \end{cases} \tag{64}$$

where $f^0 = f$. Despite its simplicity, this iterative procedure is easily seen to strengthen α.

Observation 4.3.7. $f^r \le f^{r-1} \le \cdots \le f^1 \le f^0 = f$.

Proof. This follows trivially by induction and the validity of $\alpha\left(u : f^i\right)$. ☐

Observation 4.3.8. $\alpha\left(u : f^i\right) \le f(u)$ *for all* $u \in \mathbf{D}$.

Proof. As $f^i \le f$, it follows trivially that $\alpha\left(u : f^i\right) \le f(u)$. ☐

Theorem 4.3.9. *For* $i = 1, \dots, r$, $\alpha\left(u : f^i\right) \ge \alpha\left(u : f^{i-1}\right)$ *for all* $u \in \mathbf{D}$.

Proof. We only need to show the result for $i = 1$. Let $u \in \mathbf{D}^1$. We claim that for $u \in \mathbf{D}^1$,

$$f^1(u) + \bar{f}^1(v) \le f^1(u+v),$$

for all $v, u+v \in \mathbf{D}$. First observe that since $\alpha(\cdot : f)$ is superadditive, it follows from Theorem 4.2.3 that

$$\alpha(u : f) + \bar{\alpha}(v : f) \le \alpha(u+v : f).$$

In particular, if $v \in \mathbf{D}^1$, then because $\alpha(u+v : f) \le f^1(u+v)$, $f^1(u) + f^1(v) \le f^1(u+v)$. Otherwise, $v \in \mathbf{D} \setminus \mathbf{D}^1$, and thus $u+v \in \mathbf{D} \setminus \mathbf{D}^1$. Therefore

$$\alpha(u : f) + \bar{f}(v) \le f(u+v) - \bar{f}(v) + \bar{f}(v) = f(u+v).$$

It follows from Proposition 4.3.5, $\alpha\left(u : f^1\right) = \alpha(u : f)$ for all $u \in \mathbf{D}^1$.

Now for $u \in \mathbf{D} \setminus \mathbf{D}^1$, $f^1(u+v) = f(u+v)$ for all $v \in \mathbf{D}$. Observe that $f^1(v) \le f(v)$ for all $v \in \mathbf{D}$, and therefore

$$\alpha\left(u : f^1\right) = f(u+v) - \bar{f}^1(v) \ge f(u+v) - \bar{f}(v) \ge \alpha(u : f).$$

Thus the approximation $\alpha(\cdot : f^1)$ dominates $\alpha(\cdot : f)$. ☐

Corollary 4.3.10. $\alpha(\cdot : f^i)$ *is maximal for all* $i = 0, \dots, r$.

Proof. For all $u \in \mathbf{E}'$, $f(u) = \alpha(u : f) \le \alpha(u : f^i) \le f(u)$. ☐

Note that we need not restrict ourselves to using α. The same construction works just as well with γ. We conclude this discussion by providing an example demonstrating that this approach can construct stronger approximations than α on its own.

Example 4.3.2. Recall the function from Example 4.3.1. Let $u^1 = 1+z$ for some $0 < z < 2$. Then f^1 is given by

$$f^1(u) = \begin{cases} 0 & 0 \leq u \leq 1+z \\ 1 & 1+z < u \leq 4. \end{cases}$$

Whenever $z \geq 1$, f^1 is superadditive, so $\alpha(\cdot : f^1) = f^1$. Therefore assume that $0 < z < 1$. Observe that for $3 - z < u \leq 4$, $f^1(u) + f^1(v) = 1$ for all v such that $u + v \in \mathbf{D}$, so $\alpha(u : f^1) = 1$. For $1 + z < u \leq 3 - z$, $f(4) - \bar{f}(4 - u) = 0$, thus $\alpha(u : f^1) = 0$. So $\alpha(\cdot : f^1)$ dominates $\alpha(\cdot : f^0)$.

4.4 Computing α

In this section we discuss the task of computing α. Here we describe a restriction on f and g that applies to many known families of lifting functions. We split this section into two parts: first, we describe and simplify this restriction. Next, we discuss how the task of computing α can be achieved in polynomial time with and without our structural assumption on f.

4.4.1 Two-slope Functions

When talking of f, we rely upon properties of f_y such as lower semicontinuity and piecewise linearity. Likewise we can capitalize on the structure of f when f is a two-slope function.

Definition 4.4.1. A piecewise linear function f is a *two-slope function* if there exists some λ and μ such that the slope of f_y is either λ or μ wherever it is defined. If $\lambda = 1$ and $\mu = 0$, then we say that f is a *0-1 function.*

As far as superadditive approximations are concerned, there is essentially no difference between two-slope functions and 0-1 functions.

Theorem 4.4.1. *Without loss of generality we may assume that* $\lambda = 0, \mu = 1$, $w \in \mathbf{Z}^{m+1}$, *and* $f(w), \bar{f}(w) \in \mathbf{Z}$ *for all* $w \in W$.

Proof. Assume for convenience that $\lambda < \mu$. Replace f with the function f^1 defined by $f^1(u) = f(u) - \lambda \cdot z$ where $u = (z, y)$. By construction, for all $y \in \mathbf{D}_1$, the slope of f_y^1 is either $\lambda' = 0$ or $\mu' = \mu - \lambda > 0$.

Next for all $u = (z, y) \in \mathbf{D}$, let $u' = (z/\mu', y)$. Define $f^2(u') = f^1(u)$ and accordingly define $\mathbf{D}' = \mathbf{D}_0' \times \mathbf{D}_1$, with $\mathbf{D}_0' = [0, \mu' \cdot d_0]$. Hence the slope of f_y^2 must either be 0 or 1 for all $y \in \mathbf{D}_1$.

Finally, the breakpoints of f^2 may not be integral. As there are finitely many breakpoints, and all data are rational, there exists an $N \in \mathbf{Z}_+$ such that Nw, $Nf^2(w)$ and $N\bar{f}^2(w)$ are all integral. Thus for $u = (z, y) \in \mathbf{D}'$, we set $u' = (z/N, y)$. Let $\mathbf{D}'' = \mathbf{D}_0'' \times \mathbf{D}_1$, with $\mathbf{D}_0'' = [0, \mu' \cdot Nd_0]$. The function $\tilde{f}$ defined by

$$\tilde{f}(u) = Nf^2(u') = N\left[f\left(\frac{z}{N(\mu-\lambda)}, y\right) - \frac{\lambda z}{N(\mu-\lambda)}\right]$$

satisfies the conditions of the theorem

Now we claim that the superadditive under-approximations of $\tilde{f}$ are in one-to-one correspondence with superadditive approximations of f. First suppose that g is a superadditive under-approximation of f. Then we claim that $\tilde{g}$ defined by

$$\tilde{g}(z, y) = N\left[g\left(\frac{z}{N(\mu-\lambda)}, y\right) - \frac{\lambda z}{N(\mu-\lambda)}\right],$$

is an under-approximation of $\tilde{f}$. Indeed, this is an under-approximation as

$$\begin{aligned}
\tilde{f}(z) - \tilde{g}(z) &= N\left[f\left(\frac{z}{N(\mu-\lambda)}\right) - \frac{\lambda z}{N(\mu-\lambda)}\right] - N\left[g\left(\frac{z}{N(\mu-\lambda)}\right) - \frac{\lambda z}{N(\mu-\lambda)}\right] \\
&= N\left[f\left(\frac{z}{N(\mu-\lambda)}\right) - g\left(\frac{z}{N(\mu-\lambda)}\right)\right] \\
&\geq 0.
\end{aligned}$$

Next we verify that superadditivity is preserved. Clearly,

$$\tilde{g}(z_1, y_1) + \tilde{g}(z_1, y_1) = N\left[g\left(\frac{z_1}{N(\mu-\lambda)}, y_1\right) + g\left(\frac{z_2}{N(\mu-\lambda)}, y_2\right) - \frac{\lambda(z_1+z_2)}{N(\mu-\lambda)}\right],$$

which by the superadditivity of g and non-negativity of N is at most

$$N\left[g\left(\frac{z_1+z_2}{N(\mu-\lambda)}, y_1+y_2\right) - \frac{\lambda(z_1+z_2)}{N(\mu-\lambda)}\right] = \tilde{g}(z_1+z_2, y_1+y_2).$$

Thus we have shown how to produce a superadditive approximation of $\tilde{f}$ from a superadditive approximation of f.

Next we must show that we can produce a superadditive under-approximation g of f from a superadditive under-approximation $\tilde{g}$ of $\tilde{f}$. Essentially, the proof is the same, so we will just show how to invert $\tilde{f}$ to recover f. Thus we have

$$f(z,y) = \frac{1}{N}\tilde{f}\left(N(\mu-\lambda)z, y\right) + \lambda z.$$

Thus we recover g from $\tilde{g}$ by setting

$$g(z,y) = \frac{1}{N}\tilde{g}\left(N(\mu-\lambda)z, y\right) + \lambda z.$$

By repeating the same arguments as above we can show that this indeed produces a valid superadditive approximation of f. □

From Proposition 4.3.4, if f is a 0-1 function, then α will be a 0-1 function. As we will see later, this carries through to the non-dominated approximation that we construct.

It will also be useful to identify the breakpoints of f of g by their behavior. In particular, only certain breakpoints of f and g will be necessary for testing superadditivity and non-dominance.

Definition 4.4.2. We say a breakpoint $w = (z, y) \in W$ is a *slope 0 breakpoint* if, for small $\epsilon > 0$, $f_y(z-\epsilon) = f_y(z)$.

Definition 4.4.3. We say a breakpoint $w = (z, y) \in W$ is a *slope 1 breakpoint* if, for small $\epsilon > 0$, $f_y(z-\epsilon) = f_y(z) - \epsilon$.

Definition 4.4.4. We say a breakpoint $w = (z, y) \in W$ is a *jump* or *discontinuity* if f_y is discontinuous at z.

Note that by definition $(0, y)$ and (d_0, y) are jumps. Breakpoints have played an important role in describing efficiently testable conditions for superadditivity and non-dominance. In the case of 0-1 functions, we can further refine these ideas to help accelerate an implementation.

Proposition 4.4.2. *Suppose that f and g are 0-1 functions. Then for fixed $u \in \mathbf{D}$,*

$$\min_{v, u+v \in \mathbf{D}} \{f(u+v) - \bar{g}(v)\},$$

is minimized when either $u + v$ is a slope 0 breakpoint of f or v is a slope 1 breakpoint or jump of g.

Proof. Either $u+v \in W$ or $v \in V$. The proposition follows as a straightforward application of the definitions. □

By applying this result we can reduce the number of points that we actually need to test for determining superadditivity and non-dominance.

4.4.2 Constructing α in Polynomial Time

Assuming that f is a 0-1 function, we are able to construct α in polynomial time. As in Proposition 4.3.3, we can view α as a minimum of a small number of functions. We must show that the number of breakpoints of this minimum does not blow up.

Theorem 4.4.3. *α has $\mathcal{O}(|W|^2)$ breakpoints.*

Proof. We can define the function α'_w by setting $\alpha'_w(z, y) = \alpha'_{y,w}(z)$. Now observe that α'_w has $\mathcal{O}(|W|)$ breakpoints. Indeed, let $w = (\hat{z}, \hat{y})$, then $\alpha'_{y,w}$ has $\mathcal{O}(|W_{\hat{y}-y}|)$ breakpoints. Summing over all y, and noting that $\sum_y |W_y| = |W|$, the claim follows. We can similarly define α''_w and conclude that it has at $\mathcal{O}(|W|)$ breakpoints.

Now it is entirely possible that there are breakpoints of α that are not breakpoints of α'_w or α''_w for any w, but there cannot be too many of these breakpoints.

Consider α_y. Let f^i, $i = 1 \ldots r$ denote the collection of functions $\alpha'_{y,w}$ and $\alpha''_{y,w}$, and let W^i denote the collection of breakpoints of f^i. Define $T_y = \bigcup_i W^i = \{t_1, \ldots, t_p\}$ with $t_1 < \cdots < t_p$.

Consider the interval (t_j, t_{j+1}). We claim that α_y contains at most one breakpoint in this interval. Observe that none of the functions f^i have a breakpoint in (t_j, t_{j+1}); therefore, the slopes of these functions in this interval is either 0 or 1. Now suppose that there exists some $z^* \in (t_j, t_{j+1})$ such that z^* is a breakpoint of α_y. Because the f^i are continuous in

(t_j, t_{j+1}), α_y must be continuous in (t_j, t_{j+1}), and in particular z^* is a slope 1 breakpoint of α_y. Thus $\alpha_y(z) = \alpha_y(z^*)$ for all $z \in (z^*, t_{j+1}]$.

Putting this together with the previous observation, we conclude that α must have $\mathcal{O}(|W|^2)$ breakpoints. □

As a natural consequence, we can compute α in polynomial time. Therefore, given some function f in this form, it is always possible to efficiently construct a non-trivial superadditive approximation.

Theorem 4.4.4. *α can be constructed in time $\mathcal{O}(|W|^3)$.*

Proof. As before we construct the set T_y, which can be performed in linear time. Then we sort the elements in time $\mathcal{O}\left(|T_y| \log |T_y|\right)$. Observe that $\mathcal{O}(\log |T_y|) = \mathcal{O}(\log |W|)$. Thus the total effort to construct and sort T is

$$\sum_y \mathcal{O}\left(|T_y| \log |W|\right) = \mathcal{O}\left(|W|^2 \log |W|\right).$$

Next we compute α_y for each y.

Let f^i be defined as before, and let $T_y = \{t_1, \ldots, t_p\}$. Now for $j = 1, \ldots, p-1$, compute for each i, $\bar{f}^i(t_j)$ and $f^i(t_{j+1})$. Identify the minimum of each of these quantities, which can be achieved in time $\mathcal{O}(|W|)$. Let $\bar{\alpha}_y(t_j)$ and $\alpha_y(t_{j+1})$ denote these quantities. If $\bar{\alpha}_y(t_j) < \alpha_y(t_{j+1})$, then we must test whether their exists some breakpoint in (t_j, t_{j+1}). To do so, simply compute $z^* = \alpha_y(t_{j+1}) - \bar{\alpha}_y(t_j) + t_j$. If $z^* \in (t_j, t_{j+1})$, then it is a breakpoint of α_y.

Therefore, we can compute α only by considering each of the $\mathcal{O}(|W|^2)$ points belonging to T. Each such iteration takes, $\mathcal{O}(|W|)$ time; thus the statement of the theorem follows. □

Our insistence that α is 0-1 function is not necessary for computing α in polynomial time. Suppose now that we relax this restriction.

We proceed as before, defining T, and considering some interval (t_j, t_{j+1}). In this interval each of the functions f^i are again continuous and have constant slope. The next proposition shows that there can only be $\mathcal{O}(|W|)$ breakpoints in this interval.

Proposition 4.4.5. *Let $f^1, \ldots, f^r : \mathbf{R} \to \mathbf{R}$ be affine functions over (u, v). Then $\alpha = \min\left\{f^i, \ldots, f^r\right\}$ has at most $r - 1$ breakpoints in (u, v).*

Proof. It suffices to show that the slope of α is monotone decreasing. As there are at most $r-1$ distinct slopes, this shows that there can be at most $r-1$ points where the slope changes.

For convenience let f^i have slope λ_i, and assume that $\lambda_1 \geq \cdots \geq \lambda_r$. Suppose that $\alpha(z^*) = f^i(z^*)$, for some $z^* \in (u, v)$. Thus $\alpha(z) \leq f^i(z)$ for all $z \in [z^*, u)$. In particular,

$$f^j(z) = f^j(z^*) + \lambda_j(z - z^*) \geq f^i(z) + \lambda_i(z - z^*) = f^i(z)$$

for all $j > i$. Therefore, the only functions that can possibly intersect f^i in $[z^*, u)$ must have slope strictly less than λ_i (see Figure 14). □

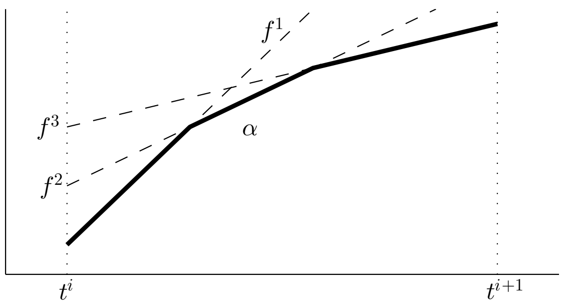

Figure 14: Computing α

As a conservative estimate, there are at most $|W|$ distinct slopes of f, so there can be at most $|W|-1$ breakpoints in (t_j, t_{j+1}). This bound can be tightened if we know the exact number of slopes that f can attain, but in general we have the following theorem:

Theorem 4.4.6. *If f is not a 0-1 function, then α has $\mathcal{O}(|W|^3)$ breakpoints.*

We can similarly modify our algorithm to compute the breakpoints of α_y to consider all pairs of functions to determine where they intersect. We first compute the $\mathcal{O}(|W|^2)$ intersections and sort them. For each intersection, we can verify in constant time (by interpolating) whether it is a breakpoint of α_y. Thus we have the following theorem.

Theorem 4.4.7. *If f is not a 0-1 function, then α can be computed in time $\mathcal{O}(|W|^4 \log |W|)$.*

This analysis is fairly crude and it may be possible to obtain better bounds using more sophisticated approaches. However, this is a first step in demonstrating the tractability of the problem of constructing a superadditive approximation.

4.5 Constructing Non-dominated Approximations

In this section, we build on the ideas from the previous sections to construct a non-dominated approximation. Let f_y be piecewise non-decreasing and lower semicontinuous for all $y \in \mathbf{D}_1$. In this section, we prove Theorem 4.2.6, i.e. a valid superadditive approximation g of f is non-dominated if and only if $\Delta(u) = 0$ for all $u \in \mathbf{D}$. We show in the next example that this condition is not sufficient to maintain lower semicontinuity.

Example 4.5.1. Let f be defined by

$$f(z) = \begin{cases} z & 0 \leq z \leq 1 \\ 2 & 1 < z \leq 2. \end{cases}$$

The superadditive approximation g defined by

$$g(z) = \begin{cases} 0 & 0 \leq z < 1 \\ 1 & z = 1 \\ 2 & 1 < z \leq 2 \end{cases}$$

is non-dominated, but is not lower semicontinuous.

To preserve lower semicontinuity, we weaken our non-dominance condition at a finite number of points. We modify our definition of $\Delta(u)$:

$$\delta(u) = \min_{v, u+v \in \mathbf{D}} \left\{ f(u+v) - [g(u) + \bar{g}(v)] \right\}, \tag{65}$$

and instead demand that $\delta(u) = 0$ for all $u \neq (0, y)$ for $y \in \mathbf{D}_1$. If g is a 0-1 function, then we can apply Proposition 4.4.2 to conclude that either $u + v$ is a slope 0 breakpoint or v is a slope 1 breakpoint or a jump.

As it turns out, the condition that $\delta(u) > 0$ implies that there is a gap with respect to the superadditivity condition. We will exploit this property as we strengthen g.

Proposition 4.5.1. *If* $\delta(u + v) = 0$*, then*

$$g(u+v) - [g(u) + \bar{g}(v)] \geq \delta(u).$$

Proof. If $v = 0$ then the claim is trivial so assume that $v \neq 0$. By assumption there exists some w such that

$$f(u + v + w) - [g(u + v) + \bar{g}(w)] = 0.$$

In particular,

$$g(u + v) - [g(u) + \bar{g}(v)] = f(u + v + w) - [g(u) + \bar{g}(v) + \bar{g}(w)]$$
$$\geq f(u + v + w) - [g(u) + \bar{g}(v + w)] \geq \delta(u).$$

The first inequality follows by superadditivity and the second by the definition of δ. □

Much like Δ, δ can be computed efficiently by computing a small list of points. Applying the proof of Proposition 4.2.6, $\delta(u)$ can be characterized as follows:

Proposition 4.5.2. *Let*

$$\delta_1(u) = \min_{w \in W}\{f(w) - [g(u) + \bar{g}(w - u)] : w - u \in \mathbf{D}\}$$

$$\delta_2(u) = \min_{v \in V}\{f(u + v) - [g(u) + \bar{g}(v)] : u + v \in \mathbf{D}\}.$$

Then $\delta(u) = \min\{\delta_1(u), \delta_2(u)\}$.

Note that when $\mathbf{D}$ is discrete, δ and Δ coincide. Otherwise, δ and Δ may differ when u can be expressed as $w - v$ for some $w \in W$ and $v \in V$ where both breakpoints are discontinuities.

We will first consider the much simpler discrete case. In this case, a non-dominated superadditive approximation can be constructed in quadratic time. Next we describe the process for constructing a non-dominated approximation when f_y is piecewise linear achieved by using a specially chosen strengthening operation. However, establishing its correctness and finite termination will be considerably more involved.

4.5.1 The Discrete Case

We begin by describing a simpler case for constructing a non-dominated approximation. Let f and g be functions defined over $\mathbf{D}$ with $\mathbf{D}_0 = \{0\}$, i.e. the domain is discrete. Further assume that $g(0) = 0$, $g \leq f$, and g is superadditive.

Observe that the set $\mathbf{D}_1$ is partially ordered under the standard definition of $\leq$ for $\mathbf{R}^m$. For convenience, let $|\mathbf{D}_1| = q$, and let $\mathbf{D}_1 = \{y_1, \ldots, y_q\}$ be ordered such that if $y_i \geq y_j$, then $i \leq j$. For ease of notation, we let $f(0, y) = f_y$, $g(0, y) = g_y$, and $\delta(0, y) = \delta(y)$.

For $i = 1, \ldots, q$, we will identify an appropriate update for g_{y_i} and update g accordingly. First compute

$$\delta(y_i) = \min_{y, y+y_i \in \mathbf{D}_1} \{f_{y+y_i} - [g_{y_i} + g_y]\}.$$

Next compute

$$\delta'(y_i) = \begin{cases} (f_{2y_i} - 2g_{y_i})/2 & 2y_i \in \mathbf{D}_1 \\ +\infty & 2y_i \notin \mathbf{D}_1. \end{cases}$$

Let $\delta = \min\{\delta(y_i), \delta'(y_i)\}$. We construct a function h, by setting

$$h_y = \begin{cases} g_y & y \neq y_i \\ g_{y_i} + \delta & y = y_i. \end{cases}$$

Proposition 4.5.3. *h is valid and superadditive.*

Proof. Validity of h is clear as $\delta \leq f_{y_i} - g_{y_i}$. By Proposition 4.5.1,

$$g_{y_i+y} - [g_{y_i} + g_y] \geq \delta(y_i).$$

for all $y \neq 0$. Therefore if $y \neq y_i$, then $h_{y_i} + h_y \leq h_{y_i+y}$ as $\delta \leq \delta(y_i)$. Now let $\delta''(y_i) = (g_{2y_i} - 2g_{y_i})/2$, and suppose to the contrary that $\delta''(y_i) < \delta$. Then

$$\begin{aligned} 0 &= g_{2y_i} - \left[g_{y_i} + \delta''(y_i) + g_{y_i} + \delta''(y_i)\right] \\ &= f_{2y_i+y'} - \left[g_{y_i} + \delta''(y_i) + g_{y_i} + g_{y'} + \delta''(y_i)\right] \\ &\geq f_{2y_i+y'} - \left[g_{y_i} + g_{y_i+y'} + \delta''(y_i)\right]. \end{aligned}$$

Thus $\delta''(y_i) \geq f_{2y_i+y'} - \left[g_{y_i} + g_{y_i+y'}\right] \geq \delta(y_i)$. Therefore if $\delta'(y_i) > \delta''(y_i)$ then $\delta \leq \delta''(y_i)$. Hence $2h_{y_i} \leq h_{2y_i}$, so h is superadditive. □

Proposition 4.5.4. *There exists some y' such that $h_{y_i} + h_y = f_{y+y'}$.*

Therefore, by iteratively strengthening g, we eventually construct a non-dominated function. Because each strengthening step takes at most $\mathcal{O}(|W|)$ time, we have the following theorem.

Theorem 4.5.5. *If* $\mathbf{D}$ *is discrete, then a non-dominated approximation can be constructed in time* $\mathcal{O}(|W|^2)$.

Despite its apparent limitation, many combinatorial problems can produce lifting functions defined over a discrete domain. Moreover, the surrogate approach to superadditive lifting described in the previous chapter allows us to manage $|W|$. Hence we can use this algorithm to help produce good approximations for more complex problems.

4.5.2 Strengthening Piecewise Linear Functions

We now assume that both f and g are functions such that f_y and g_y are increasing, piecewise linear, and that $g \leq f$ is superadditive. The strengthening operation in this case is by necessity more complex than for discrete functions. Nevertheless, some of the ideas naturally translate. Recalling that $\mathbf{D}_1 = \{y_1, \ldots, y_q\}$ is partially ordered, we again sort the elements such that $y_i \geq y_j$ implies that $i \leq j$. Starting with $i = 1$, we compute

$$u^*_{y_i} = \inf\left\{u' \in \mathbf{D}_0 : \delta(u, y_i) = 0,\ \forall u \in \mathbf{D}_0, u \geq u'\right\}.$$

If $u^*_{y_i} > 0$, then we strengthen g_{y_i} until $u^*_{y_i} = 0$. We increment i once this condition is met. When $u^*_0 = 0$, the resulting function is non-dominated.

The high level idea is simple enough; however, we must identify a strengthening procedure that decreases $u^*_{y_i}$. To make the discussion easier, we will describe the procedure for one-dimensional functions, and note that extending the results to the higher-dimensional setting is straightforward.

Assume that f and g have slopes $\lambda_1 < \cdots < \lambda_k$. To facilitate later discussion about convergence, let g^t, δ^t, u^t, and V^t denote g, δ, u^* and V at iteration t. For now we only consider $t = 0$ and $t = 1$, i.e. the initial iteration and the first update. Further, let $W(u) = \{w \in W : w \geq u\}$ and $V^t(u) = \{v \in V^t : u + v \leq d\}$.

For all $w \in W$ such that $f(w) - [g^0(u^0) + \bar{g}^0(w - u^0)] = \delta^0(u^0)$, let $\mu(w)$ denote the slope of g^0 to the right of $w - u^0$, i.e.

$$\bar{g}^0(w - u^0 + \epsilon) = \bar{g}^0(w - u^0) + \mu(w)\epsilon.$$

Similarly, for all $v \in V^0(u)$ such that $f(u^0 + v) - [g^0(u^0) + \bar{g}^0(v)] = \delta^0(u^0)$, let $\mu(v)$ denote the slope of f to the left of $u^0 + v$. Let μ^0 denote the maximum of these slopes.

We will choose some $\ell^0 < u^0$ and construct g^1 from g^0 by setting

$$g^1(u) = \begin{cases} g^0(u) & u \notin (\ell^0, u^0] \\ g^0(u^0) + \delta^0(u^0) - \mu^0(u^0 - u) & u \in (\ell^0, u^0]. \end{cases}$$

Figure 15 depicts g^0 and g^1.

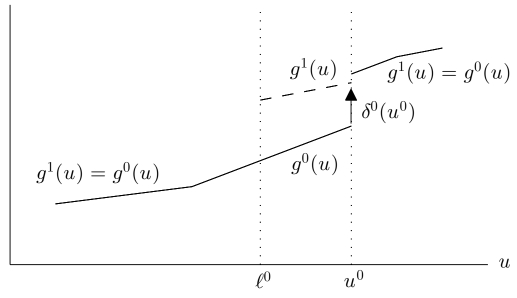

Figure 15: Strengthening of g^0.

First, we must identify conditions on ℓ^0 that ensure that g^1 remains superadditive and valid. Moreover, we must show that g^1 dominates g^0. For convenience, we will let $h^0(u) = g^0(u^0) + \delta^0(u^0) - \mu^0(u^0 - u)$.

The first of these conditions is that $W(u) = W(u^0)$ and $V^0(u) = V^0(u^0)$ for all $u \in (\ell^0, u^0]$. For this, define

$$\ell_1^0 = \max\{w \in W : w < u^0\}.$$

Clearly if $\ell^0 > \ell_1^0$, then $W(u) = W(u^0)$. Similarly, let

$$\ell_2^0 = \max\{u \in \mathbf{D} : u < u^0, d - u \in V^0\}.$$

Again, if $\ell^0 > \ell_2^0$, then $V^0(u) = V^0(u^0)$, satisfying the first condition.

Next, we require that the updated function never "crosses" f. A more formal interpretation of this restriction is that for all $u \in (\ell^0, u^0]$, $g^1(u) + g^1(v) \leq f(u + v)$. Indeed, this condition must be satisfied if g^1 is both valid and superadditive. This introduces three different bounds on ℓ^0.

The first of these bounds is given by

$$\ell_3^0 = \max_{w \in W(u^0)} \sup\{u < u^0 : f(w) - h^0(u) < \bar{g}^0(w-u)\}.$$

Note that ℓ_3^0 can be calculated efficiently in $|V^0|$ and $|W|$ by considering finitely many suprema, each of which can be easily calculated. Examples of ℓ_3^0 are shown in Figure 16 below:

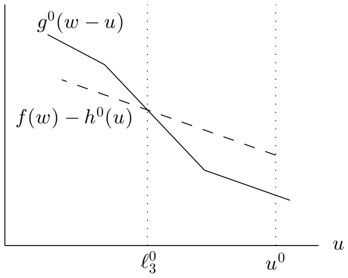

Figure 16: When $f(w)-h^0(u)$ and $g^0(w-u)$ intersect (left) or cross at a breakpoint (right).

The next of these bounds is given by

$$\ell_4^0 = \max_{v \in V^0(u^0)} \sup\{u < u^0 : f(u+v) < h^0(u) + \bar{g}^0(v)\}.$$

Again ℓ_4^0 can be efficiently computed in $|V^0|$ and $|W|$. Possible examples of ℓ_4^0 are likewise depicted in Figure 17.

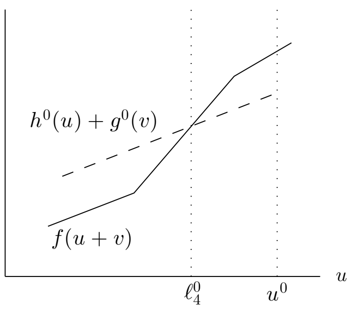

Figure 17: When $h^0(u) + g^0(v)$ and $f(u+v)$ (a) intersect or (b) cross at a breakpoint.

Proposition 4.5.6. *If $\ell^0 \geq \max\{\ell_1^0, \ell_2^0, \ell_3^0, \ell_4^0\}$, then $\delta^0(u) \geq h^0(u) - g^0(u)$ for $u \in (\ell^0, u^0]$.*

Proof. As $\ell^0 \geq \ell_1^0$ and $\ell^0 \geq \ell_3^0$,

$$\begin{aligned} 0 &\leq f(w) - [h^0(u) + \bar{g}^0(w-u)] \\ &= f(w) - [g^0(u) + \bar{g}^0(w-u)] - [h^0(u) - g^0(u)] \end{aligned}$$

for all $w \in W(u)$. Thus $\delta_1^0(u) \geq h^0(u) - g^0(u)$. Applying the same argument for $v \in V^0(u)$ and instead using that $\ell^0 \geq \ell_2^0$ and $\ell^0 \geq \ell_4^0$, it similarly follows that $\delta_2^0(u) \geq h^0(u) - g^0(u)$. By Proposition 4.5.2, the claim immediately follows. □

The last bound relating g^1 and f is given by

$$\ell_5^0 = \sup\{u < u^0 : f(2u) < 2h^0(u)\}.$$

This can be easily computed, and in conjunction with ℓ_3^0 and ℓ_4^0 ensures that the two functions never cross.

By our choice of μ^0, there exists some ℓ_6^0 such that $\delta^0(u) = h^0(u) - g^0(u)$ for all $u \in (\ell_6^0, u^0]$. To compute a valid choice of ℓ_6^0, we compute variants of ℓ_3^0, ℓ_4^0, and ℓ_5^0. For this purpose define sets

$$\widetilde{W}(u) = \left\{w \in W^0(u^0) : f(w) - [g^0(u^0) + \bar{g}^0(w - u^0)] = \delta^0(u^0)\right\},$$

and

$$\widetilde{V}(u) = \left\{v \in V^0(u^0) : f(u^0 + v) - [g^0(u^0) + \bar{g}^0(v)] = \delta^0(u^0)\right\}.$$

By the definition of δ^0 one of these sets is guaranteed to be non-empty.

For all $w \in \widetilde{W}(u)$, let

$$v(w) = \min\{v \in V^0 : v > w - u^0\}.$$

If $\bar{g}^0(u) = \bar{g}^0(w - u^0) + \mu^0\left(u - (w - u^0)\right)$ for all $u \in \left[w - u^0, v(w)\right)$, then set $u(w) = w - v(w)$; otherwise, set $u(w) = +\infty$. Let $\tilde{\ell}_3^0 = \min_{w \in \widetilde{W}(u)} u(w)$.

Analogously, for all $v \in \widetilde{V}(u)$, let

$$w(v) = \max\{w \in W^0 : w < u^0 + v\}.$$

If $f(u) = f(u^0 + v) - \mu^0\left((u^0 + v) - u\right)$ for all $u \in \left(w(v), u^0 + v\right]$, then set $u(v) = (u^0 + v) - w(v)$; otherwise, set $u(v) = +\infty$. Define $\tilde{\ell}_4^0 = \min_{v \in \widetilde{V}(u)} u(v)$.

Finally, if $2u^0 > d$, then let $\tilde{\ell}_5^0 = +\infty$; otherwise, let

$$\tilde{\ell}_5^0 = \min\left\{u < u^0 : f(2u') = 2h^0(u'), \forall u < u' \leq u^0\right\}.$$

Note that this can again be identified by considering the breakpoints of f. This is shown in Figure 18.

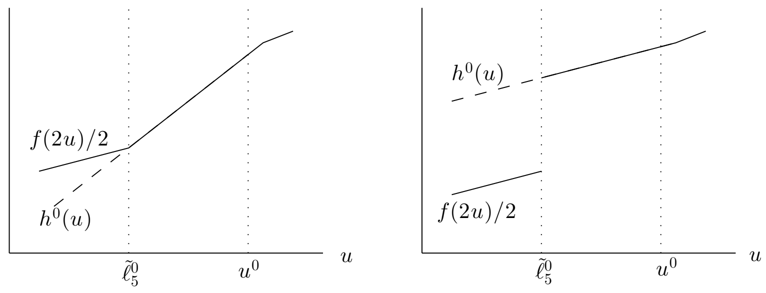

Figure 18: When $2h^0(u)$ and $f(2u)$ no longer intersect.

Define $\ell_6^0 = \min\{\tilde{\ell}_3^0, \tilde{\ell}_4^0, \tilde{\ell}_5^0\}$. By the choice of μ^0 it immediately follows that, $\ell_6^0 < u^0$. Specifically, equality is guaranteed to hold in at least one of these cases, and hence the minimum is guaranteed to be finite.

Proposition 4.5.7. *For all* $u \in (\ell^0, u^0]$, $\delta^0(u) = h^0(u) - g^0(u)$.

Proof. If $\ell_6^0 = \tilde{\ell}_3^0$ or $\ell_6^0 = \tilde{\ell}_4^0$, then the claim is clear. So suppose that $\ell_6^0 = \tilde{\ell}_5^0$. By taking the limit as v decreases to 0 in Proposition 4.5.1, it easily follows that $\bar{g}^0(u^0) - g^0(u^0) \geq \delta^0(u^0)$. For any $u \in (\tilde{\ell}_5^0, u^0]$,

$$f(u^0 + u) = 2h^0\left(\frac{u^0 + u}{2}\right) = h^0(u) + g^0(u^0) + \delta^0(u^0) \leq h^0(u) + \bar{g}^0(u^0).$$

Thus,

$$f(u^0 + u) - [h^0(u) + \bar{g}^0(u^0)] = f(u^0 + u) - [g^0(u) + \bar{g}^0(u^0)] - [h^0(u) - g^0(u)] \leq 0,$$

implying that $\delta^0(u^0) \leq h^0(u) - g^0(u)$. In conjunction, with Proposition 4.5.6 the claim follows. $\square$

One final bound arises from superadditivity,

$$\ell_7^0 = \begin{cases} u^0/2 & h(u^0) > \mu^0 u^0 \\ 0 & h(u^0) \leq \mu^0 u^0. \end{cases}$$

Letting $\ell^0 = \max\{\ell_1^0, \ell_2^0, \ell_3^0, \ell_4^0, \ell_5^0, \ell_6^0, \ell_7^0\}$, we show that the function g^1 dominates g^0, is valid, and is superadditive.

Proposition 4.5.8. *$g^1(u) \geq g^0(u)$ for all $u \in \mathbf{D}$.*

Proof. Observe that $\delta^0(u) = h^0(u) - g^0(u)$ for all $u \in (\ell^0, u^0]$. Suppose to the contrary that $h^0(u) < g^0(u)$ for any such u. Then this implies $\delta(u) < 0$, so there exists some v such that

$$f(u+v) < g^0(u) + \bar{g}^0(v) \leq g^0(u+v) \leq f(u+v).$$

However, this is a contradiction. Thus $h^0(u) \geq g^0(u)$ for all $u \in (\ell^0, u^0]$. □

We now show the main result that g^1 is superadditive and valid.

Theorem 4.5.9. *g^1 is superadditive and valid.*

Proof. That $g^1 \leq f$ follows immediately from $g^0 \leq f$ and $\delta^0(u) \leq f(u) - g^0(u)$. So it remains to show that g^1 is superadditive. Therefore we consider

$$g^1(u+v) - [g^1(u) + g^1(v)] \tag{66}$$

and show that it is non-negative for all $u, v, u+v \in \mathbf{D}$. To do so, we consider several cases.

For the first case, we assume that $u, v \notin (\ell^0, u^0]$. For this case, the non-negativity of (66) follows from the superadditivity of g^0 and $g^1 \geq g^0$.

The next case we consider is when $u \in (\ell^0, u^0]$ and $v, u+v \notin (\ell^0, u^0]$. The non-negativity of (66) is an immediate application of Proposition 4.5.1 and $\delta^0(u) = g^1(u) - g^0(u)$.

So suppose that $u, u+v \in (\ell^0, u^0]$. Within this case there are two possibilities to consider. If $v \in (\ell^0, u^0]$, then $\ell^0 \geq \ell_7^0$ implies that $g^1(v) \leq \mu^0 v$. Therefore, $g^1(u) + g^1(v) \leq g^1(u) + \mu^0 v = g^1(u+v)$.

Alternately, suppose that $v < \ell^0$. By increasing u, we may assume without loss of generality that $u + v = u^0$. It suffices to show that $\bar{g}^1(v) = \bar{g}^0(v) \leq \mu^0 v$. Let v^* satisfy

$$g^0(u^0) + \delta^0(u^0) + g^0(v^*) = f(u^0 + v^*).$$

Next let $w \in (\ell^0, u^0]$, and add $\bar{g}^0(u^0 - w) - \mu^0(u^0 - w)$ to the left-hand side. Suppose to the contrary that $\bar{g}^0(u^0 - w) - \mu^0(u^0 - w) > 0$. Then

$$\begin{aligned} f(u^0 + v^*) &< \left[g^0(u^0) + \delta^0(u^0) - \mu^0(u^0 - w)\right] + \left[g^0(v^*) + \bar{g}^0(u^0 - w)\right] \\ &\leq h^0(w) + \bar{g}^0(v^* + u^0 - w), \end{aligned}$$

where the last inequality follows by substitution and superadditivity. However, we can rewrite $h^0(w) = g^0(w) + \delta^0(w)$, which implies that

$$f(u^0 + v^*) - [g^0(w) + \bar{g}^0(v^* + u^0 - w)] < \delta^0(w).$$

But this contradicts the definition of $\delta^0(w)$. In particular, taking $w = u$, it follows that $\bar{g}^0(v) \leq \mu^0 v$ proving that (66) is non-negative in this case.

The last case we consider is $u, v \in (\ell^0, u^0]$ and $u + v > u^0$. As $u + v > u_0$, there must exist some w such that

$$g^1(u + v) - \left[g^1(u) + g^1(v)\right] = f(u + v + w) - \left[g^1(u) + \bar{g}^0(w) + g^1(v))\right].$$

If $w + v \geq u^0$, then it follows by Proposition 4.5.1 and $g^1(v) = g^0(v) + \delta^0(v)$ that

$$g(v + w) \geq g^1(v) + \bar{g}^0(w).$$

Therefore, because $v + w > u^0$,

$$g^1(u + v) - [g^1(u) + g^1(v)] \geq f(u + v + w) - [g^1(u) + g^1(v + w)] \geq 0.$$

On the other hand, if $v + w \in (\ell^0, u^0]$, then

$$g^1(u + v) - [g^1(u) + g^1(v)] \geq f(u + v + w) - 2g^1\left(\frac{u + v + w}{2}\right) \geq 0,$$

where the final inequality holds because $\ell^0 \geq \ell^0_5$.

Therefore, we have covered all possibilities showing that g^1 is superadditive. □

Before moving on, we revisit Theorem 4.2.7, and fill in the remaining details to its proof. First, we show an analog to Proposition 4.5.1.

Proposition 4.5.10. *Suppose that* $\Delta(u + v) = 0$, *then*

$$g(u + v) - [g(u) + g(v)] \geq \Delta(u). \tag{67}$$

Proof. If $v = 0$ then this assertion is trivial, so assume $v \neq 0$. By Proposition 4.2.6, either there exists some $w \in W$ such that $g(u+v) + g(w-u-v) = f(w)$ or there exists some $v' \in V$ such that $g(u+v) + \bar{g}(v') = \bar{f}(u+v+v')$.

First suppose that $g(u+v) + g(w-u-v) = f(w)$. Then

$$\begin{aligned} g(u+v) - [g(u) + g(v)] &= f(w) - [g(u) + g(v) + g(w-u-v)] \\ &\geq f(w) - [g(u) + g(w-v)] \\ &\geq \Delta_1(u) \geq \Delta(u), \end{aligned}$$

where the first inequality follows because g is superadditive and the second follows from the definition of $\Delta(u)$.

Alternately, suppose there exists some w such that $g(u+v)+\bar{g}(w) = \bar{f}(u+v+w)$. Then

$$\begin{aligned} g(u+v) - [g(u) + g(v)] &= \bar{f}(u+v+w) - [g(u) + g(v) + \bar{g}(w)] \\ &\geq \bar{f}(u+v+w) - [g(u) + g(v+w)] \\ &\geq \Delta(u). \end{aligned}$$

Here, the first inequality follows by superadditivity and lower semicontinuity, and the last follows as $\bar{f} > f$. □

Theorem 4.5.11. *Suppose f_y and g_y are non-decreasing, piecewise linear, and lower semi-continuous. If $g \leq f$ is superadditive, then g is non-dominated if and only if $\Delta(u) = 0$ for all $u \in \mathbf{D}$.*

Proof. Recall that if $u \neq w - v$ for some $w \in W$ and $v \in V$, then $\delta(u) = \Delta(u)$. Let $u^* = (u^*_{y^*}, y^*)$ be maximal with the property that $\Delta(u) = 0$ for all $u > u^*$.

If $\Delta(u^*) = \delta(u^*)$, then the above strengthening procedure can be used to produce an approximation that dominates g. Therefore assume that $\Delta(u^*) > \delta(u^*) \geq 0$. For any non-zero v, $g(u^* + v) - [g(u^*) + g(v)] \geq \Delta(u^*) > 0$. Therefore, by increasing $g(u^*)$ by $\Delta(u^*)/2$, we produce a valid superadditive approximation that dominates g. □

4.5.3 Finite Convergence to a Non-Dominated Approximation

Next we show that the choice of ℓ^0 in the definition of g^1 allows repeated application of the strengthening procedure to produce a non-dominated approximation in finitely many iterations.

Adapting the notation from the previous section, we modify superscripts to denote a given iteration. To show finite termination, we show that $\ell^t = \ell_j^t$ only finitely many times for $j = 1, \ldots, 7$. Throughout, we assume that $u^t > 0$; otherwise, g^t is non-dominated.

A crucial observation to proving termination follows:

Observation 4.5.12. *Suppose that in iteration t, we modify $(\ell^t, u^t]$. Then $u^{t+1} \leq \ell^t$.*

Proof. As $\delta^t(u) = g^{t+1}(u) - g^t(u)$ for all $u \in (\ell^t, u^t]$ the claim immediately follows. □

Therefore, at each iteration we are closer to our goal of a non-dominated approximation, and the algorithm cannot repeatedly modify the same interval.

The first step is to explore how the algorithm behaves when $\ell^t = \ell_3^t$, $\ell^t = \ell_4^t$, or $\ell^t = \ell_5^t$. This will correspondingly enable us to bound how many times we can produce breakpoints of a given form.

The proofs of all these propositions are essentially the same; however, we include them for completeness.

Proposition 4.5.13. *Suppose that $\ell^t = \ell_3^t \neq w_i - v_j$ for any $w_i \in W(u^t)$ and $v_j \in V^t(u^t)$. Then there exists some $w_i \in W(u^t)$ such that $f(w_i) = \bar{g}^{t+1}(\ell^t) + g^{t+1}(w_i - \ell^t)$. Moreover, if $\ell^t \neq \ell_5^t$ then the slope g^{t+1} to the right of $w_i - \ell^t$ for any such w_i is strictly greater than μ^t.*

Proof. For each $w_i \in W(u^t)$, let

$$\ell_3^t(w_i) = \sup \left\{ u < u^0 : f(w_i) < h^t(\ell^t) + \bar{g}^t(w_i - \ell^t) \right\}.$$

Suppose to the contrary that for some $w_i \in W(u^t)$ such that $\ell_3^t(w_i) = \ell^t$,

$$f(w_i) < h^t(\ell^t) + \bar{g}^t(w_i - \ell^t).$$

We claim that this implies that g^t is discontinuous at $w_i - \ell^t$. For all $\epsilon > 0$ such that $\ell^t + \epsilon \leq u^t$,

$$\bar{g}^t(w_i - \ell^t - \epsilon) \leq f(w_i) - h^t(\ell^t + \epsilon).$$

Taking the limit as ϵ decreases to 0,

$$g^t(w_i - \ell^t) \leq f(w_i) - h^t(\ell^t).$$

However, this implies that $\bar{g}^t(w_i - \ell^t) > g^t(w_i - \ell^t)$, and thus $w_i - \ell^t$ is a breakpoint contradicting our assumption that $\ell^t \neq w_i - v_j$.

On the other hand, suppose that $f(w_i) > h^t(\ell^t) + \bar{g}^t(w_i - \ell^t)$. As g^t is continuous at $w_i - \ell^t$, it immediately follows that for small $\epsilon > 0$,

$$f(w_i) > h^t(\ell^t - \epsilon) + \bar{g}^t(w_i - \ell^t + \epsilon).$$

But this contradicts that $\ell_3^t = \ell^t$. Thus $f(w_i) = h^t(\ell^t) + \bar{g}^t(w_i - \ell^t)$.

Furthermore, $\bar{g}^{t+1}(w_i - \ell^t) = \bar{g}^t(w_i - \ell^t) = g^t(w_i - \ell^t)$. Indeed $g^{t+1} \geq g^t$, and if the inequality is strict at $w_i - \ell^t$, then the function g^{t+1} cannot simultaneously be superadditive and valid. However, this contradicts Theorem 4.5.9.

Next, let w_i satisfy $\ell_3^t(w_i) = \ell^t$, and let λ denote the slope of g^t at $w_i - \ell^t$. Assume to the contrary that $\lambda \leq \mu^t$. Then for small $\epsilon > 0$,

$$\begin{aligned} f(w_i) - h^t(\ell^t - \epsilon) &= f(w_i) - h^t(\ell^t) + \mu^t \epsilon \\ &= g^t(w_i - \ell^t) + \mu^t \epsilon \\ &\geq g^t(w_i - \ell^t) + \lambda \epsilon \\ &= g^t(w_i - \ell^t + \epsilon). \end{aligned}$$

However, this contradicts the definition of $\ell_3^t(w_i)$; thus $\lambda > \mu^t$.

Let $\ell_3^t(w_i) = \ell^t$. If $w_i - \ell^t \notin [\ell^t, u^t)$, then

$$g^{t+1}(w_i - \ell^t + \epsilon) = g^t(w_i - \ell^t + \epsilon) > \bar{g}^t(w_i - \ell^t) + \mu^t \epsilon.$$

Instead, assume that $w_i - \ell^t \in [\ell^t, u^t)$. By the assumption that $\ell^t \neq \ell_5^t$, $w_i - \ell^t \neq \ell^t$, and therefore $w_i - \ell^t \in (\ell^t, u^t)$. By Proposition 4.5.8 and the condition that

$$\bar{g}^{t+1}(\ell^t) + g^t(w_i - \ell^t) = f(w_i),$$

$g^{t+1}(w_i - \ell^t) = g^t(w_i - \ell^t)$. However,

$$g^t(w_i - \ell^t + \epsilon) = g^t(w_i - \ell^t) + \lambda^t > g^t(w_i - \ell^t) + \mu^t \epsilon = g^{t+1}(w_i - \ell^t + \epsilon),$$

contradicting Proposition 4.5.8. Therefore, this cannot happen, so $w_i - \ell^t \notin [\ell^t, u^t)$. □

We now address when $\ell^t = \ell_4^t$. The proof of the following proposition uses many of the same ideas as the previous argument, but it is considerably more simple.

Proposition 4.5.14. *Suppose that $\ell^t = \ell_4^t \neq w_i - v_j$ for any $w_i \in W(u^t)$ and $v_j \in V^t(u^t)$. Then there exists some $v_j \in V^{t+1}(u^t)$ such that $f(\ell^t + v_j) = \bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(v_j)$. Moreover, the slope of f at $\ell^t + v_j$ for any such v_j is strictly greater than μ^t.*

Proof. As in the previous proof, define

$$\ell_4^t(v_j) = \sup\{u < u^0 : f(u + v_j) < h^t(u) + \bar{g}^t(v_j)\}.$$

Suppose that $\ell_4^t(v_j) = \ell_4^t$. We claim that $f(\ell^t + v_j) = h^t(u) + \bar{g}^t(v_j)$. Assume to the contrary that

$$f(\ell^t + v_j) < h^t(\ell^t) + \bar{g}^t(v_j).$$

For all $\epsilon > 0$ such that $\ell^t + \epsilon \leq u^t$,

$$f(\ell^t + v_j + \epsilon) \geq h^t(\ell^t + \epsilon) + \bar{g}^t(v_j).$$

Taking the limit as ϵ decreases to 0, $\bar{f}(\ell^t + v_j) \geq h^t(\ell^t) + \bar{g}^t(v_j)$. However, this implies that $f(\ell^t + v_j) < \bar{f}(\ell^t + v_j)$, contradicting that f is continuous at $\ell^t + v_j$.

Otherwise, assume to the contrary that

$$f(\ell^t + v_j) > h^t(\ell^t) + \bar{g}^t(v_j).$$

Then for small $\epsilon > 0$, $f(\ell^t + v_j - \epsilon) > h^t(\ell^t - \epsilon) + \bar{g}^t(v_j)$ by the continuity of f at $\ell^t + v_j$. However, this contradicts that $\ell^t = \ell_4^t(v_j)$.

We now address the slope of f at any such point $\ell^t + v_j$. Again let λ denote this slope. Assume to the contrary that $\lambda \leq \mu^t$. Then for small $\epsilon > 0$

$$\begin{aligned} f(\ell^t + v_j - \epsilon) &= f(\ell^t + v_j) - \lambda\epsilon \\ &= \left[h^t(\ell^t) - \lambda\epsilon\right] + \bar{g}^t(v_j) \\ &\geq \left[h^t(\ell^t) - \mu^t\epsilon\right] + \bar{g}^t(v_j) \\ &= h^t(\ell^t - \epsilon) + \bar{g}^t(v_j). \end{aligned}$$

However, this contradicts that $\ell^t(v_j) = \ell^t$. Thus $\lambda > \mu^t$.

It remains to verify that $v_j \in V^{t+1}(u^t)$. If $v_j \notin V^{t+1}(u^t)$, then $v_j \in (\ell^t, u^t]$. This implies that for small $\epsilon > 0$,

$$\bar{g}^{t+1}(v_j - \epsilon) = \bar{g}^{t+1}(v_j) - \mu^t \epsilon$$

By noting that the slope of f at $\ell^t + v_j$ exceeds μ^t,

$$\bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(v_j - \epsilon) = f(\ell^t + v_j) - \mu^t \epsilon > f(\ell^t + v_j).$$

This contradicts that g is both superadditive and valid, thus v_j must be a breakpoint of g^{t+1}. □

Recalling Figures 16 and 17, we are able to observe this behavior in ℓ_3^t and ℓ_4^t. Whenever ℓ_3^t and ℓ_4^t are defined at an intersection, we see that the slopes g at $w - u$ and f at $u + v$ exceed μ^t.

Without much difficulty, we are able to extend this analysis to ℓ_5^t.

Proposition 4.5.15. *Suppose that $\ell^t = \ell_5^t \neq w_i/2$ for any $w_i \in W(u^t)$. Then $f(2\ell^t) = 2\bar{g}^{t+1}(\ell^t)$. Moreover, the slope of f at $2\ell^t$ is strictly greater than μ^t.*

Proof. Suppose that $f(2\ell^t) > 2h^t(\ell^t)$. Then because f is continuous at $2\ell^t$, $f(2\ell^t - 2\epsilon) > 2h^t(\ell^t - \epsilon)$ for sufficiently small $\epsilon > 0$. But this contradicts the definition of ℓ_5^t.

Otherwise, suppose that $f(2\ell^t) < 2h^t(\ell^t)$. Then observe that for small $\epsilon > 0$,

$$f(2\ell^t + 2\epsilon) \geq 2h^t(\ell^t + \epsilon).$$

Taking the limit as ϵ decreases to 0, it follows that $\bar{f}(2\ell^t) \geq 2h^t(\ell^t) > f(2\ell^t)$. This contradicts our assumption that $2\ell^t \notin W(u^t)$. Therefore $f(2\ell^t) = 2h^t(\ell^t)$.

Next, let λ denote the slope of f at $2\ell^t$. If $\lambda \leq \mu^t$, then

$$\begin{aligned} f(2\ell^t - 2\epsilon) &= f(2\ell^t) - 2\lambda\epsilon \\ &= 2\left(h^t(\ell^t) - \lambda\epsilon\right) \\ &\geq 2(h^t(\ell^t) - \mu^t\epsilon) \\ &= 2h^t(\ell^t - \epsilon). \end{aligned}$$

Again, this contradicts the definition of ℓ_5^t. Therefore, $\mu^t < \lambda$. □

These propositions suggest an important property of the strengthening operation. If $\ell^t \neq w_i - v_j$ and $\ell^t \neq w_i/2$, then the next iteration will preserve continuity and simply increase the slope to the left of ℓ^t. We make this formal in the next two propositions:

Theorem 4.5.16. *Suppose that all of the following conditions hold:*

- $u^{t+1} = \ell^t$
- $\ell^t \neq w_i - v_j$ *for any* $w_i \in W(u^t)$, $v_j \in V^t(u^t)$,
- $\ell^t \neq w_i/2$ *for any* $w_i \in W(u^t)$,
- $\ell^t > \max\{\ell_1^t, \ell_2^t, \ell_6^t, \ell_7^t\}$.

Then $\delta^{t+1}(u^{t+1}) = \bar{g}^{t+1}(u^{t+1}) - g^{t+1}(u^{t+1})$ *and* $\mu^{t+1} > \mu^t$.

Proof. As $\delta^{t+1}(u^{t+1}) = \delta^{t+1}(\ell^t)$,

$$\begin{aligned}\delta^{t+1}(\ell^t) &= \min\{f(\ell^t + v) - [g^{t+1}(\ell^t) + \bar{g}^{t+1}(v)] : v, u + v \in \mathbf{D}\} \\ &= \min\{f(\ell^t + v) - [\bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(v)] : v, u + v \in \mathbf{D}\} \\ &\quad + \left(\bar{g}^{t+1}(\ell^t) - g^{t+1}(\ell^t)\right).\end{aligned}$$

Therefore, to prove the theorem it suffices to show that

$$\min\{f(\ell^t + v) - [\bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(v)] : v, u + v \in \mathbf{D}\} = 0 \tag{68}$$

Again, this is simplified by noting that $\bar{g}^{t+1}(\ell^t)$ is fixed; therefore either $v = w_i - \ell^t$ for some $w_i \in W(\ell^t)$ or $v = v_j$ for some $v_j \in V^{t+1}(\ell^t)$.

Consider the case when $v = w_i - \ell^t$. If $w_i - \ell^t \notin [\ell^t, u^t)$, then $\bar{g}^{t+1}(w_i - \ell^t) = \bar{g}^t(w_i - \ell^t)$; hence, by Proposition 4.5.13

$$f(w_i) - [\bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(w_i - \ell^t)] \geq 0.$$

Otherwise, if $w_i - \ell^t \in [\ell^t, u^t)$, then $w_i/2 \in [\ell^t, u^t)$. Applying Proposition 4.5.15,

$$f(w_i) - [\bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(w_i - \ell^t)] = f(w_i) - 2\bar{g}^{t+1}(w_i/2) \geq 0.$$

Thus, we have established that (68) is non-negative for $v = w_i - \ell^t$.

Next suppose that $v = v_j$ for some $v_j \in V^{t+1}(\ell^t)$. If $v_j \notin [\ell^t, u^t)$, then $\bar{g}^{t+1}(v_j) = \bar{g}^t(v_j)$. Thus by Proposition 4.5.14,

$$f(\ell^t + v_j) - [\bar{g}^{t+1}(\ell^t) + \bar{g}^{t+1}(v_j)] \geq 0.$$

Lastly, if $v_j \in [\ell^t, u^t)$, then $v_j = \ell^t$, and again by Proposition 4.5.15,

$$f(2\ell^t) - 2\bar{g}^{t+1}(\ell^t) \geq 0.$$

Therefore, (68) is non-negative for $v = v_j$.

Further observe that ℓ^t is equal to one of ℓ_3^t, ℓ_4^t, or ℓ_5^t. By our assumptions, this implies that (68) equals 0, implying that $\delta^{t+1}(\ell^t) = \bar{g}^{t+1}(\ell^t) - g^{t+1}(\ell^t)$. Moreover, by applying Propositions 4.5.13, 4.5.14, or 4.5.15 as appropriate, we further conclude that $\mu^{t+1} > \mu^t$. □

As f has k slopes, this theorem implies that after $k-1$ iterations the assumptions of this theorem can no longer hold. We address these assumptions one at a time, and show that each condition can only happen finitely many times.

Proposition 4.5.17. *If $u^{t+1} < \ell^t$, then $u^{t+1} = v_j$ for some $v_j \in V^{t+1}$ or $u^{t+1} = w_i - v_j$ for some $w_i \in W$ and $v_j \in V^{t+1}$.*

Proof. Suppose to the contrary that neither condition holds. Recall by Proposition 4.5.2, that $\delta^{t+1}(u^{t+1})$ is minimized when either $v = w_i - v_j$ for some $w_i \in W(u^{t+1})$ and $v_j \in V^{t+1}(u^{t+1})$ or $v = v_j$ for some $v_j \in V^{t+1}(u^{t+1})$.

Consider first the quantity

$$f(w_i) - \left[g^{t+1}(u^{t+1}) + \bar{g}^{t+1}(w_i - u^{t+1})\right]. \tag{69}$$

Let λ and μ denote the slopes of g^{t+1} at u^{t+1} and $w_i - u^{t+1}$ respectively. By assumption neither of these two points are breakpoints of g^{t+1}, so these slopes do indeed exist. If (69) is equal to 0, then $\lambda = \mu$; as otherwise one of the quantities

$$f(w_i) - \left[g^{t+1}(u^{t+1} \mp \epsilon) + \bar{g}^{t+1}(w_i \pm u^{t+1} + \epsilon)\right]$$

is strictly less than 0 for small $\epsilon > 0$. However, this contradicts the superadditivity and validity of g^{t+1}. But then, u^{t+1} is not minimal, again yielding a contradiction. Thus (69) must be strictly positive.

Likewise consider the quantity

$$f(u^{t+1} + v_j) - \left[g^{t+1}(u^{t+1}) + \bar{g}^{t+1}(v_j)\right]. \tag{70}$$

Applying analogous arguments, it similarly follows that (70) is strictly positive.

Therefore, $\delta^{t+1}(u^{t+1}) > 0$. For sufficiently small $\epsilon > 0$,

$$f(w_i) - \left[g^{t+1}(u^{t+1} + \epsilon) + \bar{g}^{t+1}(w_i - u^{t+1} - \epsilon)\right] > 0,$$

and

$$f(u^{t+1} + v_j + \epsilon) - \left[g^{t+1}(u^{t+1} + \epsilon) + \bar{g}^{t+1}(v_j)\right] > 0.$$

Thus $\delta^{t+1}(u^{t+1} + \epsilon) > 0$. But this contradicts our choice of u^{t+1}. □

We are interested in certain milestones in the execution of the algorithm. Accordingly, define

$$t_i = \min\{t : u^t \leq w_i/2\}.$$

To show finiteness, it suffices to show that the algorithm only adds a finite number of breakpoints between t_i and t_{i-1}. Let n^i denote the number of breakpoints the algorithm introduces between t_i and t_{i-1}. We define four related quantities, each specifying the number of breakpoints of a certain form added between t_i and t_{i-1}:

- n_1^i: breakpoints of the form $w_{i'} - v_{j'}$ with $w_{i'} \in W$ and $v_{j'} \in V^0$
- n_2^i: breakpoints of the form $w_{i'} - v_{j'}$ with $w_{i'} \in W$ and $v_{j'} \notin V^0$,
- n_3^i: breakpoints of the form $w_{i'}/2$,
- n_4^i: breakpoints of the form $\ell^t = \ell_7^t > 0$.

By exploring the conditions of Theorem 4.5.16, we arrive at the following proposition connecting n^i with $n_1^i, \ldots, n_4^i$.

For the remainder of the discussion, we will let $|W| = r$ and $|V^0| = p$

Proposition 4.5.18. $n^i \leq k(n_1^i + n_2^i + n_3^i + n_4^i)$.

Proof. Consider the cases in which the conditions of Theorem 4.5.16 fail to hold. Each of these will coincide with some n_j^i.

If $u^{t+1} < \ell^t$, then either $u^{t+1} = v_j$ for some $v_j \in V^0$ or $u^{t+1} = w_{i'} - v_j$ for some $v_j \in V^{t+1}$. In the first case, no new breakpoint is introduced. In the latter, a breakpoint is added that is either accounted for in n_1^i or n_2^i.

Next suppose that $\ell^t = \ell_1^t$ or $\ell^t = \ell_2^t$. In the first case, $\ell^t = w_i = w_i - 0$, which is accounted for in n_1^i. In the latter $\ell^t = d_0 - v_{j'} = w_r - v_j$ for some $v_j \in V^t$. This is either included in n_1^i or n_2^i.

If $\ell^t = \ell_6^t$ then either $\ell^t = w_{i'} - v_j$ or $\ell^t = w_{i-1}/2$. Again the first case is either included in n_1^i or n_2^i and the latter is counted in n_3^i.

Next observe that when the conditions of Theorem 4.5.16 do hold that the slope of the approximation decreases. Therefore, there can be at most $k-1$ intermediate breakpoints between any two such described points. Including the iteration when ℓ^t is among $n_1^i, \ldots, n_4^i$, the algorithm takes at most k iterations for any such point. $\square$

It remains to bound each of the above quantities. This will correspondingly enable us to bound n^i for all i. For convenience, let $t_{r+1} = 0$ and $n^{r+1} = 0$.

Proposition 4.5.19. *For all i, $n_1^i \leq rp$.*

Proof. As $|W| = r$ and $|V^0| = p$, there are at most rp pairs $w_{i'}$ and v_j. $\square$

Proposition 4.5.20. *For $i = 1, \ldots, r$, $n_2^i \leq r\left(n^{i+1} + \cdots + n^{r+1}\right)$.*

Proof. Suppose that v_j is introduced in iteration t with $t_i \leq t < t_{i-1}$. Then by assumption $w_{i-1}/2 < v_j < w_i/2$. In particular, $w_{i'} - v_j \geq w_i/2$ for all $i' \geq i$. On the other hand $w_{i''} - v_j < w_{i''}/2$ for all $i'' < i$. Therefore, if $w_{i'} - v_j$ is introduced for some $v_j \notin V^0$, then $v_j \in V^{t_i}$. Trivially, there are at most $n^{i+1} + \cdots + n^{r+1}$ such points. For any such v_j there are at most r different $w_{i'}$. $\square$

Proposition 4.5.21. *For $i = 1, \ldots, r$, $n_3^i \leq 1$.*

Proof. By assumption $i' < i$, therefore the proposition trivially holds. $\square$

Proposition 4.5.22. *There exist at most k distinct t such that $\ell^t = \ell^t_7$. Thus $n^i_4 \leq k$.*

Proof. Suppose that $\ell^{\tau_1} = \ell^{\tau_1}_7$ and $\ell^{\tau_2} = \ell^{\tau_2}_7$ and $\tau_1 < \tau_2$. We show that $\mu^{\tau_1} > \mu^{\tau_2}$.

Assume to the contrary that $\mu^{\tau_1} \leq \mu^{\tau_2}$. By assumption

$$g^{\tau_i+1}(u) = \bar{g}^{\tau_i+1}\left(\frac{u^{\tau_i}}{2}\right) + \mu^{\tau_i}\left(u - \frac{u^{\tau_i}}{2}\right).$$

for $u \in (u^{\tau_i}/2, u^{\tau_i}]$ and $i = 1, 2$. By Observation 4.5.12, $u^{\tau_2} \leq u^{\tau_1}/2$. Thus letting $u = 5u^{\tau_1}/8$ and $v = 3u^{\tau_2}/4$, $u + v \in (u^{\tau_1}/2, u^{\tau_1}]$. Therefore

$$\begin{aligned} g^{\tau_2+1}(u) + g^{\tau_2+1}(v) &> g^{\tau_2+1}(u) + \mu^{\tau_2} v \\ &\geq g^{\tau_2+1}(u) + \mu^{\tau_1} v \\ &= g^{\tau_2+1}(u + v). \end{aligned}$$

However, this contradicts that g^{τ_2+1} is superadditive. As there are only k distinct slopes the proposition immediately follows. ☐

By combining these results we conclude that the algorithm produces a non-dominated approximation in finitely many iterations.

Theorem 4.5.23. *t_1 is finite. Thus there exists some t such that g^t is a non-dominated superadditive approximation of f.*

Proof. By applying the Propositions 4.5.19 through 4.5.22

$$n^i \leq k\left(r \sum_{i'=i+1}^{r} n^{i'} + rp + 1 + k\right).$$

Observe that n^r is finite; therefore, by induction it immediately follows that n^i must be finite. Thus n^1 is guaranteed to be finite, and thus $t_1 = \sum_{i=1}^{r} n^i$ is finite. ☐

4.5.4 On the Running Time

The running time yielded by the previous analysis is unfortunately exponential. Before concluding this chapter, we take a moment to explore the proof at a high level. In particular, we address specifically why the bound is exponential and produce (loosely speaking) an example that demonstrates the limitations of this analysis. This, however, does not preclude the possibility that the algorithm runs in polynomial time, but rather shows that any such proof must exploit the problem structure more intimately.

At its heart, the proof relies on a simple counting argument. It counts the maximum number of breakpoints that could possibly be added by the algorithm, but not the actual number of breakpoints that the algorithm introduces.

Revisiting Propositions 4.5.19, 4.5.21, and 4.5.22, $k(n_1^i + n_3^i + n_4^i)$ is polynomial in k, r, and p. Thus, it is plain to see that the exponential blow-up results directly from the term n_2^i. From Proposition 4.5.20,

$$n_2^i \leq r\left(n^{i+1} + \cdots + n^{r+1}\right).$$

This bound originates from a simple observation: if some breakpoint v is introduced by the algorithm, and $w_i - v$ is also introduced by the algorithm, then $v \geq w_i/2$. On the other hand, the analysis does not consider information about the functions at v or $w_i - v$.

Thus to show the limitation of this analysis, we only need to construct W and V^0 in such a way that there are potentially an exponential number of breakpoints. We produce such sets in the following example.

Example 4.5.2. Fix some $n > 0$ and let $d = 2^{2n}$. For $i = 1, \ldots, n$, let $w_i = d - (2^{n-i+1} - 2)$. Set

$$W = \{0, d/2, d/2 + 1, w_1, \ldots, w_n\},$$

and let $V^0 = \{0, d\}$. We are interested in the collection of points of the form $w - v$ that can arise in the interval $[d - 2^n, d]$.

Quite naturally, the first possible point that can arise is $v_1 = w_n - (d/2 + 1) = d/2 - 1$. Next consider $w_{n-1} - v$. There are two possible points, $v_2 = w_{n-1} - d/2 = d/2 - 2$ and $v_3 = w_{n-1} - (d/2 + 1) = d/2 - (4 - 1)$.

Let $T^0 = \{d/2 + 1, d/2\}$ and let

$$T^i = \left\{w_{n-i+1} - v : v \in T^{i'}, i' < i\right\}.$$

We claim inductively that $T^i = \left\{d/2 - (2^{i-1} - 1), \ldots, d/2 - (2^i - 1)\right\}$. Observe that

$$w_{n-i} = 2\left(d/2 - (2^i - 1)\right),$$

and that

$$w_{n-i} - (d/2 + 1) = d/2 - (2^{i+1} - 1).$$

Finally, noting that,

$$\bigcup_{j=0}^{i} T^j = \left\{d/2 + 1, d/2, \ldots, d/2 - (2^i - 1)\right\}.$$

it follows that $T^{i+1} = \left\{d/2 - (2^i - 1), \ldots, d/2 - (2^{i+1} - 1)\right\}$.

Therefore, by induction, there are at least $2^n - 1$ points that can possibly be added.

This example shows that the exponential bound on the running time of the algorithm cannot be improved upon unless more information is incorporated into the analysis. However, our experience with two-slope functions and step functions suggests that at least in these cases, the algorithm tends to perform quite well. Thus, it is entirely conceivable that this algorithm is polynomial or pseudo-polynomial.

4.6 Closing Remarks

In the discrete case, the existence of a polynomial time algorithm is unsurprising as the problem can be easily expressed as a linear program. By moving to a continuous setting, we are no longer afforded this luxury, and it is quite remarkable that a finite algorithm even exists. This heavily depended on our structural assumptions.

It is also possible to consider "decreasing" functions, in which the slopes are all non-positive and the function value decreases at discontinuities. With slight modification, the results presented in this chapter translate fairly naturally to this setting. The strengthening operation used to produce a non-dominated approximation would necessarily be different, but nevertheless is a straightforward extension of its counterpart for increasing functions.

CHAPTER V

APPLICATIONS OF LIFTING IN HIGH DIMENSION

In this chapter, we apply results from Chapters 3 and 4 to high-dimensional lifting problems. By way of example, we demonstrate that the tools we developed in the previous chapters can be used offline to effectively guide the search for non-dominated superadditive approximate lifting functions.

We begin with two traditional mixed integer programs–namely the knapsack and the fixed-charge flow sets–and we introduce a complicating constraint in the form of an additional knapsack constraint. When considering this additional constraint, exact lifting becomes much more challenging; nevertheless, we are able to relax the side-constraints using the techniques of Chapter 3 and derive valid inequalities that are not valid for the single-row systems. Next, applying the algorithm of Chapter 4, we construct non-dominated approximations for specific instances that we then generalize to closed form approximations.

We conclude this chapter by deriving cuts for the stable set polytope obtained by lifting odd-hole inequalities. In the original space of variables, the dimension of the lifting function is typically too large to consider in its entirety. Identifying an appropriate superadditive approximation in this setting may require us to consider a prohibitively large number of points. By reformulating the stable set problem and using the approximation scheme of Chapter 3, we show how to manage the problem dimension. Unlike the problems previously considered in this chapter, however, we may not be able to obtain a closed form description of the lifting function or its superadditive approximation; nevertheless, we show that it is possible to evaluate the lifting function and construct an approximation that yields the deepest cut with respect to separating a current solution.

5.1 *Knapsack Intersections*

We first consider the intersection of multiple knapsacks and show how the approximation scheme of Chapter 3 can be nested. Let

$$X = \left\{ x \in \{0,1\}^n : \begin{array}{ll} a^0 x \le b^0 & \\ a^i x \le b^i, & (i = 1, \dots, m) \end{array} \right\}.$$

Let C satisfy $\sum_{j \in C} a_j^0 = b + \lambda$ for $\lambda > 0$, and $\sum_{j \in C} a_j^i + \Delta^i = b^i$ with $\Delta^i \ge 0$ for $i = 1, \dots, m$. Such a subset C exists, otherwise $a^0 x \le b^0$ is redundant. Further, assume that C is minimal with this property.

We restrict X by fixing $x_j = 0$ for $j \in N \setminus C$ to obtain the restricted system X'. Because $\sum_{j \in C} a_j^i \le b^i$, these knapsack constraints are redundant. Therefore the cover inequality

$$\sum_{j \in C} x_j \le |C| - 1$$

is facet defining for $\text{conv}(X')$. Although redundant, these additional knapsack constraints alter the lifting function:

$$\begin{aligned} f\left(z^0, z^1, \dots, z^m\right) = \min \quad & (|C| - 1) - \sum_{j \in C} x_j \\ \text{s.t.} \quad & \sum_{j \in C} a_j^0 \le b^0 - z^0 \\ & \sum_{j \in C} a_j^i \le b^i - z^i, \quad i = 1, \dots, m \\ & x_j \in \{0,1\}, \quad j \in C. \end{aligned} \tag{71}$$

We cannot in general hope to have a nice closed form expression of f. Our goal is to replace the remaining knapsack constraints with a single cardinality constraint:

$$\begin{array}{lcl} \sum_{j \in C} a_j^i \le b^i - z^i, \quad i = 1, \dots, m & & \sum_{j \in C} x_j \le |C| - \Phi\left(z^1, \dots, z^m\right) \\ & \mapsto & \\ x_j \in \{0,1\}, \quad j \in C & & x_j \in \{0,1\}, \quad j \in C. \end{array}$$

The task of computing Φ exactly is in general no easier than computing f; hence we can construct a collection of nested relaxations for computing Φ. Let

$$\begin{aligned}\Phi^i\left(z^i,\ldots,z^m\right) = \min \quad & |C| - \sum_{j\in C} x_j \\ \text{s.t.} \quad & \sum_{j\in C} a_j^k x_j \leq b^k - z^k, \quad (k = i,\ldots,m) \\ & x_j \in \{0,1\}, \quad j \in C\end{aligned}$$

for $i = 1,\ldots,m$; thus $\Phi = \Phi^1$. Similarly, define Ψ^i,

$$\begin{aligned}\Psi^i\left(z^i,z^{i+1},\ldots,z^m\right) = \min \quad & |C| - \sum_{j\in C} x_j \\ \text{s.t.} \quad & \sum_{j\in C} a_j^i x_j \leq b^i - z^i \\ & \sum_{j\in C} x_j \leq |C| - \Psi^{i+1}\left(z^{i+1},\ldots,z^m\right) \\ & x_j \in \{0,1\}, \quad j \in C.\end{aligned} \tag{72}$$

and for convenience let $\Psi^{m+1} = 0$. If we let $\Psi = \Psi^1$, then we have the following proposition:

Proposition 5.1.1. *For $i = 1,\ldots,m$, $\Psi^i \leq \Phi^i$. In particular $\Psi \leq \Phi$.*

Proof. The claim trivially holds for Ψ^m, and the general result follows by inductively applying Proposition 3.3.2. □

It will also be convenient to have a succinct representation of the cardinality constraint restricted to individual knapsacks. Hence, we introduce $\widehat{\Phi}^i$:

$$\begin{aligned}\widehat{\Phi}^i\left(z^i\right) = \min \quad & |C| - \sum_{j\in C} x_j \\ \text{s.t.} \quad & \sum_{j\in C} a_j^i x_j \leq b^i - z^i, \\ & x_j \in \{0,1\}, \quad j \in C\end{aligned} \tag{73}$$

Proposition 5.1.2. $\Psi\left(z^1,\ldots,z^m\right) = \max\left(\widehat{\Phi}^1(z^1),\ldots,\widehat{\Phi}^m(z^m)\right)$.

Proof. We proceed by induction on i. Trivially if $i = m$ the statement of the proposition holds. So consider some arbitrary $i < m$. We claim that

$$\Psi^i = \max\left(\widehat{\Phi}^i(z^1),\Psi^{i+1}(z^{i+1},\ldots,z^m)\right).$$

Indeed if $\widehat{\Phi}^i(z^i) < \Psi^{i+1}(z^{i+1}, \ldots, z^m)$, then taking an optimal solution x^i of $\widehat{\Phi}^i$, we must set $x_j = 0$ for precisely $\Psi^{i+1}(z^{i+1}, \ldots, z^m) - \widehat{\Phi}^i(z^i)$ additional elements to satisfy the cardinality constraint. Otherwise, if $\widehat{\Phi}^i(z^i) \geq \Psi^{i+1}(z^{i+1}, \ldots, z^m)$, then this solution satisfies the cardinality constraint. Applying the inductive hypothesis, the proposition holds. □

Any superadditive approximation of Ψ can in turn be used to obtain a superadditive approximation of (71). We show how such an approximation of Ψ can be attained by composing functions.

Let $\Gamma^{m+1} = 0$, and define Γ^i for $i = 1, \ldots, m$ by

$$\begin{aligned} \Gamma^i\left(z^i, y\right) = \min \quad & |C| - \sum_{j \in C} x_j \\ \text{s.t.} \quad & \sum_{j \in C} a^i_j x_j \leq b^i - z^i \\ & \sum_{j \in C} x_j \leq |C| - y \\ & x_j \in \{0, 1\} \qquad\qquad j \in C. \end{aligned}$$

As in the previous proposition, $\Gamma^i(z^i, y) = \max(\widehat{\Phi}^i(z^i), y)$. Now let $\gamma^i \leq \Gamma^i$ be increasing, superadditive, and satisfy the condition that $\gamma^i(z^i, y) \in \{0, \ldots, |C|\}$ for $i = 1, \ldots, m$.

Proposition 5.1.3. *Let $\psi^{m+1} = \gamma^{m+1}$ and define ψ^i by*

$$\psi^i\left(z^i, z^{i+1}, \ldots, z^m\right) = \gamma^i\left(z^i, \psi^{i+1}(z^{i+1}, \ldots, z^m)\right)$$

for $i = 1, \ldots, m$. Then $\psi^i \leq \Psi^i$ is superadditive and non-decreasing. In particular $\phi = \psi^1$ is a superadditive, non-decreasing, under-approximation of Φ.

Proof. We proceed by induction on i. By assumption $\psi^{m+1} \leq \Psi^{m+1}$ is superadditive and non-decreasing. So assume that $\psi^{i+1} \leq \Psi^{i+1}$ is superadditive and non-decreasing for $i < m$.

The composition of non-decreasing functions is itself non-decreasing, therefore ψ^i is non-decreasing. Observe that $\psi^{i+1} \leq \Psi^{i+1}$ and γ^i is superadditive and non-decreasing. Thus by Theorem 3.3.3, $\psi^i \leq \Psi^i$ is superadditive. □

Once we have computed ϕ, we compute

$$
\begin{aligned}
\hat{g}\left(z^0, y\right) = \min \quad & (|C| - 1) - \sum_{j \in C} x_j \\
\text{s.t.} \quad & \sum_{j \in C} a_j^0 x_j \leq b^0 - z^0 \\
& \sum_{j \in C} x_j \leq |C| - y \\
& x_j \in \{0, 1\}, \quad j \in C.
\end{aligned}
\tag{74}
$$

If we let $\widehat{\Phi}^0(z^0)$ denote the exact lifting function without the cardinality constraint, then $\hat{g}(z^0, y) = \max\left(\widehat{\Phi}^0(z^0), y - 1\right)$.

Note that the maximum of superadditive functions is often not superadditive; however, any one of these superadditive functions is trivially a superadditive under-approximation of the maximum. It is easy to construct an approximation ϕ that performs at least as well as this simple approximation; however, we can typically do better with our nesting scheme.

We now explore the structure of $\widehat{\Phi}^i$, γ^i, and $\hat{g}$ in greater detail and give a concrete example of each of these functions in the computation of an approximate superadditive lifting function. As a matter of notation, when we refer to $\widehat{\Phi}^i$, we assume that $i > 0$.

The evaluation of $\widehat{\Phi}^i$ is no different than the standard knapsack cover inequality. Let $j_1, \ldots, j_{|C|}$ satisfy

$$a^i_{j_1} \geq \cdots \geq a^i_{j_{|C|}},$$

and define $A^i_k = A^i_{k-1} + a^i_{j_k}$ with $A^i_0 = 0$.

Proposition 5.1.4.

$$
\widehat{\Phi}^i(z^i) = \begin{cases} 0 & 0 \leq z^i \leq \Delta^i \\ k & \Delta^i + A^i_{k-1} < z^i \leq \Delta^i + A^i_k, \quad (j = 1, \ldots, |C|) \end{cases}
$$

Proof. If $x_{j_k} = 0$ and $x_{j_{k-1}} = 1$ in a solution then there is an equivalent solution with $x_{j_k} = 1$ and $x_{j_{k-1}} = 0$. Therefore we set $x_{j_k} = 0$ in increasing order of k. □

If $\Delta^i < a^i_{j_1}$, then $\widehat{\Phi}^i$ is not superadditive. Let k^i be defined

$$k^i = \min_{k \geq 1} \left\{ k : \frac{\Delta^i + A^i_{k-1}}{k} \geq a^i_{j_k} \right\}, \tag{75}$$

and let $\bar{a}^i$ denote this quantity. We define $\widehat{\phi}^i$ as follows:

$$\widehat{\phi}^i(z^i) = \begin{cases} 0 & z = 0 \\ k & k\bar{a}^i < z^i \leq (k+1)\bar{a}^i, \quad (k = 0, \ldots, k^i - 1), \\ \widehat{\Phi}^i(z^i) & z^i > k^i \bar{a}^i = \Delta^i + A^i_{k^i-1}. \end{cases} \tag{76}$$

Proposition 5.1.5. *$\widehat{\phi}^i \leq \widehat{\Phi}^i$, and $\widehat{\phi}^i$ is superadditive.*

Proof. It will be easier to consider $\widehat{\Phi}^i$ and $\widehat{\phi}^i$ as members of a family of step functions. Let $v_1, v_2, \ldots, v_p > 0$. Let $V_0 = 0$ and $V_j = V_{j-1} + v_j$. Define the function

$$\chi_V(z) = \begin{cases} 0 & z = 0 \\ j & V_j < z \leq V_{j+1}, \quad (j = 0, \ldots, p-1). \end{cases}$$

Analogously define $\bar{V}_j$ for coefficients $\bar{v}_j$ and a function $\chi_{\bar{V}}$. Clearly $\chi_V \geq \chi_{\bar{V}}$ if and only if $\bar{V}_j \geq V_j$ for all j.

Therefore to show $\widehat{\phi}^i \leq \widehat{\Phi}^i$, it suffices to show that $k\bar{a}^i \geq \Delta^i + A^i_{k-1}$ for $k < k^i$. Let $\bar{a}^i_k = \left(\Delta^i + A^i_{k-1}\right)/k$. If $k < k^i$, then $\bar{a}^i_k < a^i_{j_k}$. By definition,

$$\bar{a}^i_{k+1} = \bar{a}^i_k \cdot \frac{k}{k+1} + a^i_{j_k} \cdot \frac{1}{k+1} > \bar{a}^i_k.$$

Now observe that $k \cdot \bar{a}^i_k = \Delta^i + A^i_{k-1}$; hence $k \cdot \bar{a}^i \geq \Delta^i + A^i_{k-1}$ for all $k \leq k^i$. Thus $\widehat{\phi}^i \leq \widehat{\Phi}^i$.

Next we show that $\widehat{\phi}^i$ is superadditive. We show that if $v_1 \geq v_2 \geq \cdots \geq v_p$, then χ_V is superadditive. Indeed consider

$$\chi_V(u+v) - [\chi_V(u) + \chi_V(v)].$$

Let $u = V_s + \epsilon_1$ and $v = V_t + \epsilon_2$, where $0 < \epsilon_1 \leq v_{s+1}$ and $0 < \epsilon_2 \leq v_{t+1}$; thus $\chi_V(u) + \chi_V(v) = s + t$. Next, observe that $V_s + V_t \geq V_{s+t}$; therefore $\chi_V(u+v) \geq s + t$. Noting that $\bar{a}^i \geq a^i_{j_k} \geq \cdots \geq a^i_{j_{|C|}}$, it follows that $\widehat{\phi}^i$ is superadditive. □

We can now use this superadditive approximation to seed our algorithm to construct a non-dominated superadditive approximation of Γ^i. Before moving to a more general setting, we work out an explicit example.

Example 5.1.1. In this example, we consider only a single knapsack. Suppose that $C = \{1, \ldots, 5\}$ and consider the knapsack

$$8x_1 + 6x_2 + 5x_3 + 5x_4 + 4x_5 \leq 32.$$

Therefore, $\Delta^i = 4$. By Propositions 5.1.4 and 5.1.5, we can easily compute $\widehat{\Phi}^i$ and its superadditive under-approximation $\widehat{\phi}^i$ given in Table 5 below:

Table 5: Superadditive approximation $\widehat{\phi}^i$ of $\widehat{\Phi}^i$

$\widehat{\Phi}^i(z)$	z	$\widehat{\phi}^i(z)$	z
0	$0 \leq z \leq 4$	0	$0 \leq z \leq 6$
1	$4 < z \leq 12$	1	$6 < z \leq 12$
2	$12 < z \leq 18$	2	$12 < z \leq 18$
3	$18 < z \leq 23$	3	$18 < z \leq 23$
4	$23 < z \leq 28$	4	$23 < z \leq 28$
5	$28 < z \leq 32$	5	$28 < z \leq 32$

Recall that $\Gamma^i(z, y) = \max\left(\widehat{\Phi}^i(z), y\right)$. Thus to construct an initial superadditive approximation, we set $\gamma_0^i(z, y) = \widehat{\phi}^i(z)$ for all y. By applying the algorithm from Chapter 4, we strengthen γ_0^i to obtain the non-dominated approximation γ^i of Γ^i represented in Table 6.

Table 6: Non-dominated approximation γ^i

$\gamma^i(z,y)$	$y=0,1,2$	$y=3$	$y=4$	$y=5$
0	$0 \le z \le 6$	$0 \le z \le 5$	$0 \le z \le 5$	$0 \le z \le 4$
1	$6 < z \le 12$	$5 < z \le 11$	$5 < z \le 10$	$4 < z \le 9$
2	$12 < z \le 18$	$11 < z \le 17$	$10 < z \le 16$	$9 < z \le 14$
3	$18 < z \le 23$	$17 < z \le 23$	$16 < z \le 22$	$14 < z \le 20$
4	$23 < z \le 28$	$23 < z \le 28$	$22 < z \le 28$	$20 < z \le 26$
5	$28 < z \le 32$	$28 < z \le 32$	$28 < z \le 32$	$26 < z \le 32$

When $y = 0$, $y = 1$, and $y = 2$ the values of $\gamma^i(z, y)$ all coincide. In the above table, the left-hand side indicates the value of $\gamma^i(z, y)$ and the right-hand side indicates which values of z attain this value for a given y. This function is represented graphically in Figure 19 by considering each slice of the function associated with a fixed value of y.

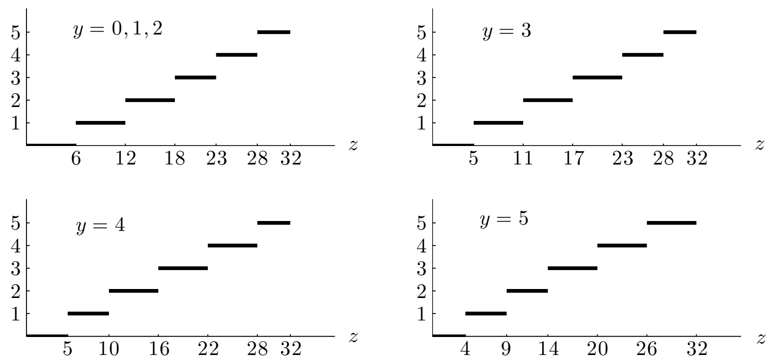

Figure 19: Graphical depiction of $\gamma^i(z, y)$ from Example 5.1.1

Despite its small size, this numerical example is quite telling about what form a non-dominated superadditive approximation might take. Consider the interval lengths of $\widehat{\phi}^i$ in the example: 6, 6, 6, 5, 5, 4. These interval lengths reappear in γ^i, for every fixed y, but sometimes their order changes. In the case of $y = 3$, the interval lengths appear in the order 5, 6, 6, 6, 5, 4. When $y = 5$, the interval lengths appear in the reverse order: namely 4, 5, 5, 6, 6, 6.

This suggests a more general behavior; specifically, for each y the first $y+1$ intervals appear in the reverse order. To express this more succinctly, we return to the abstraction χ_V used in the proof of Proposition 5.1.5, and express $\widehat{\phi}^i$ in terms of its interval lengths: $v_1 \geq v_2 \geq \cdots \geq v_p$.

We propose the following approximation of Γ^i:

$$\gamma^i(z,y) = \begin{cases} 0 & z = 0 \\ y - \bar{\chi}_V(V_{y+1} - z) & 0 < z \leq V_{y+1} \\ \chi_V(z) & z > V_{y+1}, \end{cases} \tag{77}$$

where $\bar{\chi}_V(u) = \limsup_{z \to u} \chi_V(u)$. We show in the next two theorems that γ^i is valid, superadditive, and non-dominated.

Theorem 5.1.6. *$\gamma^i \leq \Gamma^i$ and γ^i is superadditive.*

Proof. Observe that $\gamma^i(z,y) \leq y \leq \Gamma^i(z,y)$ for all $z \in [0, V_{y+1}]$. Otherwise, by the validity of $\widehat{\phi}^i$, whenever $z > V_{y+1}$, it again follows that $\widehat{\phi}^i(z) \leq \widehat{\Phi}^i(z) \leq \Gamma^i(z,y)$. Thus $\gamma^i \leq \Gamma^i$ establishing validity.

To prove superadditivity, we now must show that

$$\gamma^i(z_1,y_1) + \gamma^i(z_2,y_2) \leq \gamma^i(z_1+z_2, y_1+y_2) \tag{78}$$

whenever (z_1,y_1), (z_2,y_2), and (z_1+z_2, y_1+y_2) are all in the domain of γ^i. In the same spirit as Theorem 4.2.3, it suffices to consider whenever (z_1,y_1) and (z_2,y_2) are breakpoints.

We break the proof of superadditivity into cases. First suppose that $z_1 \leq V_{y_1+1}$ and $z_2 \leq V_{y_2+1}$. In this case, express z_1 and z_2 as

$$z_1 = \sum_{j=0}^{s} v_{y_1+1-j} \qquad z_2 = \sum_{j=0}^{t} v_{y_2+1-j},$$

for some $s \leq y_1$ and $t \leq y_2$. Thus $\gamma^i(z_1,y_1) = s$ and $\gamma^i(z_2,y_2) = t$. Within this case, there are two possibilities to consider: either $s+t < y_1+y_2$ or $s+t = y_1+y_2$.

In the first case, we may assume without loss of generality that $s < y_1$. We claim that

$$\sum_{j=0}^{s+t+1} v_{y_1+y_2+1-j} \leq z_1 + z_2. \tag{79}$$

Splitting the sum in (79), we obtain

$$\sum_{j=0}^{s+t+1} v_{y_1+y_2+1-j} = \sum_{j=0}^{s} v_{y_1+y_2+1-j} + \sum_{j=0}^{t} v_{y_2+(y_1-(s+1))+1-j}.$$

From the sorting of elements in V,

$$\sum_{j=0}^{s} v_{y_1+y_2+1-j} \leq \sum_{j=0}^{s} v_{y_1+1-j} = z_1.$$

Furthermore observe that $y_1 - (s+1) \geq 0$; therefore,

$$\sum_{j=0}^{t} v_{y_2+(y_1-(s+1))+1-j} \leq \sum_{j=0}^{t} v_{y_2+1-j} = z_2,$$

proving (79). Thus $\gamma^i(z_1 + z_2, y_1 + y_2) \geq s + t + 1$ with equality holding if and only if (79) holds at equality. In particular, we must also have that

$$\gamma^i(z_1 + \epsilon_1, y_2) + \gamma^i(z_2 + \epsilon_2, y_2) = s + t + 2 \leq \gamma^i(z_1 + z_2 + \epsilon_1 + \epsilon_2, y_1 + y_2),$$

for any infinitesimally small $\epsilon_1, \epsilon_2 > 0$.

Next suppose that $s + t = y_1 + y_2$. Applying the same arguments, we can conclude that

$$V_{y_1+y_2} = \sum_{j=0}^{s+t} v_{y_1+y_2+1-j} \leq z_1 + z_2 - v_{y_2+1}.$$

Observe that $y_1 + y_2 \geq y_2$; therefore, the next interval has length $v_{y_1+y_2+2} \leq v_{y_2+1}$. In particular, this implies that $\gamma^i(z_1+z_2, y_1+y_2) \geq s+t+1$ and $\gamma^i(z_1+z_2+\epsilon, y_1+y_2) \geq s+t+2$ for any $\epsilon > 0$.

Thus we have resolved when $z_1 \leq V_{y_1+1}$ and $z_2 \leq V_{y_2+1}$. So suppose that

$$z_1 = V_s \qquad z_2 = \sum_{j=0}^{t} v_{y_2+1-j}$$

for $s > y_1 + 1$, and $t \leq y_2 + 1$. Either $s + t < y_1 + y_2$ or $s + t \geq y_1 + y_2$.

If $s + t < y_1 + y_2$, then

$$\sum_{j=0}^{s+t+1} v_{y_1+y_2+1-j} \leq z_1 + z_2.$$

This claim follows as in the previous case by splitting the sum:

$$\sum_{j=0}^{s+t+1} v_{y_1+y_2+1-j} = \sum_{j=0}^{t} v_{y_1+y_2+1-j} + \sum_{j=0}^{s} v_{y_1+y_2-(t+1)+1-j}.$$

As $y_1 + y_2 - (t+1) \geq (s+t+1) - (t+1) = s$, it follows that

$$\sum_{j=0}^{s} v_{y_1+y_2-(t+1)+1-j} \leq \sum_{j=0}^{s} v_{s+1-j} = V_s.$$

Therefore (78) follows as before.

Next suppose that $s+t \geq y_1+y_2$. In this case, we show that $V_{s+t+1} \leq z_1+z_2$. Rewriting V_{s+t+1}, we obtain

$$V_{s+t+1} = V_s + \sum_{j=0}^{t} v_{s+t+1-j} \leq V_s + \sum_{j=0}^{t} v_{y_1+y_2+1-j} \leq V_s + \sum_{j=0}^{t} v_{y_2+1-j},$$

where the first inequality follows from the relation $s+t \geq y_1+y_2$ and the last inequality follows from the ordering imposed on of the elements of V. Again (78) follows as in the previous cases.

The last case we consider is when both $z_1 = V_s$ and $z_2 = V_t$ for $s \geq y_1+1$ and $t \geq y_2+1$. In this case, (78) immediately follows from the superadditivity of $\widehat{\phi}^i$. $\square$

Next, we show that γ^i is non-dominated in the weaker sense of Chapter 4.

Theorem 5.1.7. *γ^i is non-dominated.*

Proof. First we show that for $(0, y_1)$ there exists some (z_2, y_2) such that

$$\gamma^i(0, y_1) + \gamma^i(z_2, y_2) = \Gamma^i(z_2, y_1 + y_2).$$

As $\gamma^i(b^i, y) = |C| = \Gamma^i(b^i, y)$ for all y, $(z_2, y_2) = (b^i, 0)$ satisfies this requirement.

Next for each (z_1, y_1) in the domain of γ^i with $z_1 > 0$, we show that there exists some (z_2, y_2) such that

$$\gamma^i(z_1, y_1) + \bar{\gamma}^i(z_2, y_2) = \Gamma^i(z_1 + z_2, y_1 + y_2).$$

There are two cases to consider: either $y_1 \leq k^i - 1$ or $y_1 \geq k^i$ with k^i defined as in (75).

Consider first when $y_1 \leq k^i - 1$, and observe that

$$\gamma^i(k^i \bar{a}^i, y_1) = k^i - 1 = \Gamma^i\left(k^i \bar{a}^i, y_1\right).$$

Suppose now that $z_1 = k\bar{a}^i + \epsilon$ for some $k < k^i$ and $0 < \epsilon \leq \bar{a}^i$. Then let

$$z_2 = (k^i - k - 1)\bar{a}^i + (\bar{a}^i - \epsilon).$$

By construction $z_1 + z_2 = k^i \bar{a}^i$, and therefore

$$\gamma^i(z_1, y_1) + \bar{\gamma}^i(z_2, 0) = k + (k^i - k - 1) = k^i - 1 = \Gamma^i(z_1 + z_2, y_1 + 0).$$

On the other hand, if $z_1 > k^i \bar{a}^i$, then $\gamma^i(z_1, y_1) = \Gamma^i(z_1, y_1)$.

Therefore, it only remains to consider when $y_1 \geq k^i$. In this case,

$$\gamma^i \left(\Delta^i + A^i_{y_1}, y_1\right) = y_1 = \Gamma^i \left(\Delta^i + A^i_{y_1}, y_1\right).$$

If $z_1 < \Delta^i + A^i_{y_1}$, then we set $z_2 = \Delta^i + A^i_{y_1} - z_1$. Thus

$$\gamma^i(z_1, y_1) + \bar{\gamma}^i(z_2, 0) = y_1 - \bar{\gamma}^i(z_2, 0) + \bar{\gamma}^i(z_2, 0) = y_1 = \Gamma^i(z_1 + z_2, y_1 + 0).$$

Similarly if $z_1 > \Delta^i + A^i_{y_1}$, then $\gamma^i(z_1, y_1) = \Gamma^i(z_1, y_1)$, concluding the proof that γ^i is non-dominated. □

We now consider $\hat{g}$. The structure of $\hat{g}$ is quite similar to Γ^i, so many of the above results translate with only slight modification. Again, we consider a numerical example to guide our study of the properties of $\hat{g}$.

Example 5.1.2. Consider a knapsack and a minimal cover, $C = \{1, \ldots, 5\}$. The restricted system is given by

$$9x_1 + 8x_2 + 6x_3 + 5x_4 + 4x_5 \leq 29.$$

Therefore, this cover exceeds the knapsack capacity by $\lambda = 3$. Thus we can compute the lifting function for the minimal cover inequality:

$$\widehat{\Phi}^0(z) = \min \left\{ 4 - (x_1 + x_2 + x_3 + x_4 + x_5) : \begin{array}{l} 9x_1 + 8x_2 + 6x_3 + 5x_4 + 4x_5 \leq 29 - z \\ x_1, x_2, x_3, x_4, x_5 \in \{0, 1\} \end{array} \right\}.$$

The closed form description in Table 7 of $\widehat{\Phi}^0$ and its superadditive approximation $\widehat{\phi}^0$ can be obtained by appropriately modifying Propositions 5.1.4 and 5.1.5.

Table 7: $\widehat{\Phi}^0$ and its approximation $\widehat{\phi}^0$

$\widehat{\Phi}^0$	z	$\widehat{\phi}^0(z)$	z
0	$0 \le z \le 6$	0	$0 \le z \le 7$
1	$6 < z \le 14$	1	$7 < z \le 14$
2	$14 < z \le 20$	2	$14 < z \le 20$
3	$20 < z \le 25$	3	$20 < z \le 25$
4	$25 < z \le 29$	4	$25 < z \le 29$

Note that $\hat{g}(z, y) = \max\left(\widehat{\Phi}^0(z), y - 1\right)$. Thus to construct an initial superadditive approximation, we set $\hat{h}(z, y) = \widehat{\phi}^0(z)$ for all y. By applying the algorithm from Chapter 4, we again strengthen $\hat{h}_0$ to obtain the non-dominated approximation $\hat{h}$ of $\hat{g}$ in Table 8.

Table 8: Approximation of $\hat{g}$ for knapsack intersections

$\hat{h}(z, y)$	$y = 0, 1, 2$	$y = 3$	$y = 4$	$y = 5$
0	$0 \le z \le 7$	$0 \le z \le 6$	$0 \le z \le 5$	$0 \le z \le 4$
1	$7 < z \le 14$	$6 < z \le 13$	$5 < z \le 11$	$4 < z \le 9$
2	$14 < z \le 20$	$13 < z \le 20$	$11 < z \le 18$	$9 < z \le 15$
3	$20 < z \le 25$	$20 < z \le 25$	$18 < z \le 25$	$15 < z \le 22$
4	$25 < z \le 29$	$25 < z \le 29$	$25 < z \le 29$	$22 < z \le 29$

When $y = 0$, $y = 1$, and $y = 2$, the values of $\hat{h}$ all coincide. Note that this function can also be depicted graphically as a collection of one-dimensional functions for each fixed value of y.

As in Example 5.1.1, the order of the interval lengths also reverses for higher values of y. This motivates a similar construction for a non-dominated superadditive approximation of $\hat{g}$. Representing $\widehat{\phi}^0$ by χ_V, with $v_1 \ge v_2 \ge \cdots \ge v_p$, we propose the following approximation of $\hat{g}$:

$$\hat{h}(z, y) = \begin{cases} 0 & z = 0 \\ (y-1)^+ - \bar{\chi}_V(V_y - z) & 0 < z \le V_y \\ \chi_V(z) & V_y < z. \end{cases}$$

The proof that $\hat{h}$ is valid, superadditive, and non-dominated is no different from γ^i, and thus is omitted.

Theorem 5.1.8. *The function $\hat{h}$ defined is a non-dominated valid superadditive approximation of $\hat{g}$.*

We conclude with one final example showing how all these ideas are combined to obtain lifting coefficients that dominate those obtained by the individual knapsack constraints.

Example 5.1.3. Consider the following set obtained by identifying some cover C and setting $x_j = 0$ for all $j \in N \setminus C$:

$$X = \left\{ x \in \{0,1\}^5 : \begin{array}{r} 9x_1 + 8x_2 + 6x_3 + 5x_4 + 4x_5 \le 29 \\ 6x_1 + 5x_2 + 8x_3 + 4x_4 + 5x_5 \le 32 \\ 8x_1 + 4x_2 + 6x_3 + 5x_4 + 5x_5 \le 32 \end{array} \right\}.$$

Note that these are the coefficients from Examples 5.1.1 and 5.1.2, so we may use the evaluations of ϕ^i, γ^i, and $\hat{h}$ from each of these examples.

Suppose now that we want to reintroduce x_6, and

$$\left(a_6^1, a_6^2, a_6^3\right) = (12, 23, 24).$$

By Example 5.1.1, we have that $\widehat{\phi}^3(24) = 4$. Next we determine $\gamma^2(23, 4) = 4$. Finally from Example 5.1.2, we determine $\hat{h}(12, 4) = 2$. Thus we obtain the inequality

$$x_1 + x_2 + x_3 + x_4 + x_5 + 2x_6 \le 4.$$

In particular, this is stronger than the inequality we would obtain by ignoring the side-constraints.

5.2 *Knapsack-Constrained Flow Covers*

We consider the fixed charge flow set with only outflow arcs and additionally impose a knapsack side-constraint on the binary variables. The constraint set is given by

$$X = \left\{ (x,y) \in \mathbf{R}^n \times \mathbf{Z}^n : \begin{array}{ll} \sum_{j \in N} x_j \le b & \\ \sum_{j \in N} a_j y_j \le d & \\ 0 \le x_j \le u_j y_j, & \forall j \in N \\ y_j \in \{0,1\}, & \forall j \in N \end{array} \right\}, \tag{80}$$

with $0 < a_j \le d$ for all $j \in N$.

Let C be a flow cover, i.e. $\sum_{j \in C} u_j = b + \lambda$ for $\lambda > 0$, such that $\sum_{j \in C} a_j \le d$. If no such set C exists, then $\sum_{j \in N} x_j \le b$ is redundant. Additionally, assume that C is minimal with this property.

Let $C = \{1, \ldots, r\}$, with $u_1 \ge \cdots \ge u_r > \lambda$, and set $(x_j, y_j) = (0,0)$ for $j \in N \setminus C$. The flow cover inequality is

$$\sum_{j \in C} x_j + \sum_{j \in C} (u_j - \lambda)(1 - y_j) \le b, \tag{81}$$

and the restricted system is

$$X' = \left\{ (x,y) \in \mathbf{R}^{|C|} \times \mathbf{Z}^{|C|} : \begin{array}{ll} \sum_{j \in C} x_j \le b & \\ \sum_{j \in C} a_j y_j \le d & \\ 0 \le x_j \le u_j y_j, & \forall j \in C \\ y_j \in \{0,1\}, & \forall j \in C \end{array} \right\}. \tag{82}$$

Because $\sum_{j \in C} a_j \le d$, the knapsack constraint is redundant in this system. Therefore, the flow cover inequality (81) is facet defining for $\text{conv}(X')$. However, when we perform lifting, the knapsack constraint influences the lifting function:

$$
\begin{aligned}
f(z_1, z_2) = \min \quad & b - \left[\sum_{j \in C} x_j + \sum_{j \in C} (u_j - \lambda)(1 - y_j) \right] \\
\text{s.t.} \quad & \sum_{j \in C} x_j \le b - z_1 \\
& \sum_{j \in C} a_j y_j \le d - z_2 \\
& 0 \le x_j \le u_j y_j, && \forall j \in C \\
& y_j \in \{0, 1\}, && \forall j \in C.
\end{aligned}
\tag{83}
$$

Observe that if z_2 is made sufficiently large, the minimum will increase as we are forced to set $y_j = 0$. It may happen that there exists some i and j such that $u_i < u_j$, but $a_i > a_j$. Therefore as we increase z_2, we may disable arcs in a different order as when we increase z_1. This introduces a challenge in explicitly computing f. To overcome this difficulty, we apply Proposition 3.3.2 with (μ, μ_0) given by

$$\sum_{j \in C} y_j \le |C|.$$

Letting Φ denote the lifting function

$$\Phi(z_2) = \min \left\{ |C| - \sum_{j \in C} y_j : \begin{array}{c} \sum_{j \in C} a_j y_j \le d - z_2 \\ y_j \in \{0, 1\}, \quad j \in C \end{array} \right\},$$

we can (76) and Proposition 5.1.5 to construct an appropriate approximation ϕ.

Similar to (74) for the knapsack intersection, we thus we consider the lifting function

$$
\begin{aligned}
\hat{g}(z, v) = \min \quad & b - \left[\sum_{j \in C} x_j + \sum_{j \in C} (u_j - \lambda)(1 - y_j) \right] \\
\text{s.t.} \quad & \sum_{j \in C} x_j \le b - z \\
& \sum_{j \in C} y_j \le |C| - v \\
& 0 \le x_j \le u_j y_j, && \forall j \in C \\
& y_j \in \{0, 1\}, && \forall j \in C.
\end{aligned}
\tag{84}
$$

The cardinality constraint is considerably more manageable than the original knapsack constraint, and we are able to explicitly compute $\hat{g}$. Letting $U_0 = 0$ and $U_k = U_{k-1} + u_k$ for $k > 0$, we obtain the following description of $\hat{g}$.

Theorem 5.2.1. *If* $v = 0$ *then*

$$\hat{g}(z, v) = \begin{cases} k\lambda & U_k \leq z \leq U_{k+1} - \lambda, \quad (k = 0, \ldots, r-1) \\ z - U_k + k\lambda & U_k - \lambda \leq z \leq U_k, \quad (k = 1, \ldots, r-1) \end{cases}$$

If $0 < v \leq |C|$, *then*

$$\hat{g}(z, v) = \begin{cases} (v-1)\lambda & 0 \leq z \leq U_v - \lambda \\ \hat{g}(z, 0) & z \geq U_v - \lambda. \end{cases}$$

Proof. If $v = 0$, then the cardinality-constrained lifting function is no different from its unconstrained counterpart (38). Therefore assume that $v \geq 1$. Clearly if $x_j > 0$ in a solution to (84) then we can assume $x_j > u_j - \lambda$. Next let $j^* = \min\{j : x_j > 0\}$. Without loss of generality, we may assume that $x_j = u_j$ for all $j > j^*$ such that $x_j > 0$. Suppose that $x_{j'} = 0$ for some $j' > j^*$. Then set $(x_{j^*}, y_{j^*}) = (0, 0)$ and $(x_{j'}, y_{j'}) = \big(\min(x_{j^*}, u_{j'}), 1\big)$. The change in the objective function is

$$\big[x_{j^*} + \big(u_{j'} - \lambda\big)\big] - \big[\min(x_{j^*}, u_{j'}) + (u_{j^*} - \lambda)\big] \leq 0.$$

So we can construct an optimal solution as follows: first set $(x_j, y_j) = 0$ for $j = 1, \ldots, v$. Set $(x_j, y_j) = (u_j, 1)$ for all $j > v$. Increasing z, the solution remains optimal until the flow constraint becomes tight. At this point decrease x_{v+1} until $x_{v+1} = (u_v - \lambda)$, and then set $(x_{v+1}, y_{v+1}) = (0, 0)$. We repeat this procedure for the remaining arcs until $z = b$. □

This function is superadditive only in the trivial case when $r = 1$. If $r \geq 2$, then $\hat{g}(0, 2) = \hat{g}(u_1, 2) = \lambda$. But $\hat{g}(u_1, 0) = \lambda$. Therefore

$$\hat{g}(0, 2) + \hat{g}(u_1, 2) = 2\lambda > \hat{g}(u_1, 2).$$

There are a number of superadditive approximations that we can apply. The most trivial approximations is to simply set $\hat{h}(z, v) = \hat{g}(z, 0)$.

Using this as a starting approximation, we use the algorithm of Chapter 4 to construct a numerical example that we can then generalize.

Example 5.2.1. Consider a cardinality-constrained fixed-charge flow set given by the following system of inequalities:

$$x_1 + x_2 + x_3 + x_4 \leq 26$$

$$y_1 + y_2 + y_3 + y_4 \leq 4$$

$$0 \leq x_j \leq u_j y_j$$

with $u = (10, 7, 6, 6)$. From Theorem 5.2.1, we obtain the evaluation of $\hat{g}(z, 0)$ given in Table 9:

Table 9: Evaluation of $\hat{g}(z, 0)$ for cardinality-constrained flow-cover

$\hat{g}(z,0)$	z	$\hat{g}(z,0)$	z
0	$0 \leq z \leq 7$	6	$17 \leq z \leq 20$
$z - 7$	$7 \leq z \leq 10$	$z - 14$	$20 \leq z \leq 23$
3	$10 \leq z \leq 14$	9	$23 \leq z \leq 26$
$z - 11$	$14 \leq z \leq 17$		

Likewise, Theorem 5.2.1 allows us to compute $\hat{g}(z, v)$ for $v > 0$. Beginning with the initial approximation $\hat{h}_0(z, v) = \hat{g}(z, 0)$, we apply the algorithm of Chapter 4 to obtain a non-dominated valid superadditive approximation $\hat{h}$. This is given in Tables 10 through 12, where each table represents the function for a given value of v.

When $v = 0$ and $v = 1$, the final approximation satisfies $\hat{h}(z, v) = \hat{g}(z, 0)$. For the remain values of v, $\hat{h}$ is as follows:

Table 10: Non-dominated $\hat{h}(z,v)$ for cardinality-constrained flow cover with $v = 2$

$\hat{h}(z,2)$	z	$\hat{h}(z,2)$	z
0	$0 \le z \le 4$	6	$17 \le z \le 20$
$z - 4$	$4 \le z \le 7$	$z - 14$	$20 \le z \le 23$
3	$7 \le z \le 14$	9	$23 \le z \le 26$
$z - 11$	$14 \le z \le 17$		

Table 11: Non-dominated $\hat{h}(z,v)$ for cardinality-constrained flow cover with $v = 3$

$\hat{h}(z,3)$	z	$\hat{h}(z,3)$	z
0	$0 \le z \le 3$	6	$13 \le z \le 20$
$z - 3$	$3 \le z \le 6$	$z - 14$	$20 \le z \le 23$
3	$6 \le z \le 10$	9	$23 \le z \le 26$
$z - 7$	$10 \le z \le 13$		

Table 12: Non-dominated $\hat{h}(z,v)$ for cardinality-constrained flow cover with $v = 4$

$\hat{h}(z,4)$	z	$\hat{h}(z,4)$	z
0	$0 \le z \le 3$	6	$12 \le z \le 16$
$z - 3$	$3 \le z \le 6$	$z - 10$	$16 \le z \le 19$
3	$6 \le z \le 9$	9	$19 \le z \le 26$
$z - 6$	$9 \le z \le 12$		

As in the case of the of the knapsack intersection, $\hat{h}$ is depicted in Figure 20 by graphing its individual slices corresponding with each value of v.

This example is quite telling about how we might generally construct a non-dominated approximation of $\hat{g}$. Again, we note the lengths of the slope 0 intervals and the order in

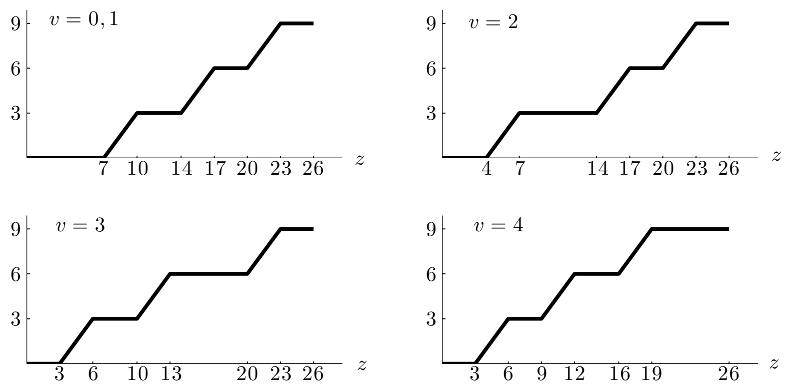

Figure 20: The lifting function $\hat{h}$ from Example 5.2.1

which they appear for each value of v.

For this example when $v = 0$ and $v = 1$ the lengths appear in the order 7, 4, 3, 3; when $v = 2$, the order becomes 4, 7, 3, 3; for $v = 3$, the order again changes to 3, 4, 7, 3; and finally when $v = 4$, the order changes to 3, 3, 4, 7. Once again, we see that the order of these intervals reverses.

Motivated by this example, we propose the following approximation of $\hat{g}$:

$$\hat{h}(z, v) = \begin{cases} \hat{g}(z, 0) & v = 0 \\ (v-1)\lambda - \hat{g}\left((U_v - \lambda) - z\right) & 0 \le z \le U_v - \lambda, \ \ (v > 0) \\ \hat{g}(z, 0) & z \ge U_v - \lambda, \ \ (v > 0). \end{cases} \tag{85}$$

As we did for (77), we show that $\hat{h}$ is a non-dominated valid superadditive approximation.

Theorem 5.2.2. *The function $\hat{h}$ is valid and superadditive.*

Proof. By Theorem 5.2.1, we only need to show that $\hat{h}(z, v) \le \hat{g}(z, v)$ whenever $v > 0$ and $z \le U_v - \lambda$. In this case

$$(v-1)\lambda - \hat{g}\left((U_v - \lambda) - z\right) \le (v-1)\lambda = \hat{g}(z, v).$$

Therefore $\hat{h} \le \hat{g}$, proving validity.

To prove superadditivity, we must show that

$$\hat{h}(z_1, v_1) + \hat{h}(z_2, v_2) \le \hat{h}(z_1 + z_2, v_1 + v_2) \tag{86}$$

for all $0 \le z_1, z_2 \le z_1 + z_2 \le b$ and $0 \le v_1, v_2 \le v_1 + v_2 \le r$. Note that $\hat{h}(z_1, v_1)$ is a 0-1 function with no jumps. By Proposition 4.4.2, we only need to test certain points. Namely, we may assume that (z_1, v_1) is a slope 1 breakpoint (i.e. the slope changes from 1 to 0 at (z_1, v_1)) and that either (z_2, v_2) is a slope 1 breakpoint or $z_1 + z_2 = b$.

Suppose first that $\hat{h}(z_1, v_1) = \hat{g}(z_1, 0)$ and $\hat{h}(z_2, v_2) = \hat{g}(z_2, 0)$. If $z_1 = z_2 = 0$, then (86) clearly holds by the superadditivity of $\hat{g}$. Therefore, assume without loss of generality that $v_2 > 0$. Thus $z_1 = U_s$ for some $s \ge v_1$ and $z_2 \ge U_{v_2} - \lambda$. In particular,

$$U_s + (U_{v_2} - \lambda) \ge U_{s+v_2} - \lambda \ge U_{v_1+v_2} - \lambda.$$

Thus $\hat{h}(z_1 + z_2, v_1 + v_2) = \hat{g}(z_1 + z_2, 0)$ and (86) again holds by the superadditivity of $\hat{g}$.

So assume now that $z_1 \le U_{v_1} - \lambda$ and $v_1 > 0$. As we assumed that z_1 was a slope 1 breakpoint, we may write $z_1 = \sum_{j=0}^{s} u_{v_1-j}$ for some $s < v_1 - 1$. Now suppose that $z_2 = d - z_1$; thus

$$z_2 = (U_r - \lambda) - \sum_{j=0}^{s} u_{v_1-j} \le (U_r - \lambda) - \sum_{j=0}^{s} u_{r-j} = U_{r-(s+1)} - \lambda.$$

Therefore if $z_2 \ge U_{v_2} - \lambda$, then $\hat{h}(z_2, v_2) = \hat{g}(z_2, 0) \le (r - s - 2)\lambda$. In particular,

$$\hat{h}(z_1, v_1) + \hat{h}(z_2, v_2) \le (s + 1)\lambda + (r - s - 2)\lambda = (r - 1)\lambda.$$

On the other hand if $z_2 < U_{v_2} - \lambda$, then $v_2 < r - (s + 1)$. Indeed if $v_2 \ge r - (s + 1)$ then

$$v_1 + v_2 \ge v_1 + r - (s + 1) > r.$$

Thus $\hat{h}(z_2, v_2) \le (r - s - 2)\lambda$, and (86) again follows.

So we may assume now that z_2 is a slope 1 breakpoint. This leaves two possibilities: either $z_2 \le U_{v_2} - \lambda$ or $z_2 > U_{v_2} - \lambda$. In the first case, write $z_2 = \sum_{j=0}^{t} u_{v_2-j}$ with $t < v_2 - 1$. Therefore, $s + t + 1 < v_1 + v_2 - 1$ and

$$\begin{aligned} \sum_{j=0}^{s+t+1} u_{v_1+v_2-j} &= \sum_{j=0}^{s} u_{v_1+v_2-j} + \sum_{j=s+1}^{s+t+1} u_{v_1+v_2-j} \\ &= \sum_{j=0}^{s} u_{v_1+v_2-j} + \sum_{j=0}^{t} u_{v_2+(v_1-(s+1))-j} \\ &\le z_1 + z_2. \end{aligned}$$

Thus $\hat{h}(z_1 + z_2, v_1 + v_2) \geq (s + t + 2)\lambda = (s + 1)\lambda + (t + 1)\lambda = \hat{h}(z_1, v_1) + \hat{h}(z_2, v_2)$ showing that (86) holds.

Otherwise, if $z_2 \geq U_{v_2} - \lambda$, then let $z_2 = U_t$ with $t \geq v_2$. If $v_1 + v_2 > s + t$, then

$$\begin{aligned}\sum_{j=0}^{s+t} u_{v_1+v_2-j} &= \sum_{j=0}^{s} u_{v_1+v_2-j} + \sum_{j=s+1}^{s+t} u_{v_1+v_2-j} \\ &= \sum_{j=0}^{s} u_{v_1+v_2-j} + \sum_{j=1}^{t} u_{v_1+v_2-s-j} \\ &\leq \sum_{j=0}^{s} u_{v_1-j} + \sum_{j=1}^{t} u_{t+1-j} \\ &= z_1 + z_2.\end{aligned}$$

In particular, $\hat{h}(z_1 + z_2, v_1 + v_2) \geq (s + t + 1)\lambda = \hat{h}(z_1, v_1) + \hat{h}(z_2, v_2)$. If $v_1 + v_2 \leq s + t$, then

$$U_{s+t+1} = \sum_{j=0}^{s+t} u_{s+t+1-j} = \sum_{j=0}^{s} u_{s+t+1-j} + \sum_{j=1}^{t} u_{t+1-j} \leq z_1 + z_2.$$

Thus $\hat{h}(z_1 + z_2, v_1 + v_2) \geq (s + t + 1)\lambda$ and (86) follows as before.

The only case left to consider now is when $z_1 \geq U_{v_1} - \lambda$ and $z_2 = d - z_2$. Thus we express $z_1 = U_s$ for some $s \geq v_1$, and let

$$z_2 = (U_r - U_s) - \lambda = \sum_{j=0}^{r-s-1} u_{r-j} - \lambda.$$

If $z_2 \geq U_{v_2} - \lambda$, then $z_2 \leq U_{r-s} - \lambda$, so $\hat{h}(z_2, v_2) \leq (r - s - 1)\lambda$, implying that

$$\hat{h}(z_1, v_1) + \hat{h}(z_2, v_2) \leq s\lambda + (r - s - 1)\lambda = (r - 1)\lambda.$$

Thus (86) holds. Otherwise, if $z_2 \leq U_{v_2} - \lambda$, then

$$\sum_{j=0}^{r-s-1} u_{r-j} - \lambda \leq \sum_{j=0}^{r-s-1} u_{v_2-j} - \lambda.$$

Therefore, $\hat{h}(z_2, v_2) \leq (r - s - 1)\lambda$, and (86) again holds. □

Now that we have established that $\hat{h}$ is valid and superadditive, we show that it is non-dominated again using Theorem 4.2.7.

Theorem 5.2.3. *$\hat{h}$ is non-dominated.*

Proof. For each point (z_1, v_1), we show there exists a second point (z_2, v_2) such that

$$\hat{h}(z_1, v_1) + \hat{h}(z_2, v_2) = \hat{g}(z_1 + z_2, v_1 + v_2).$$

If $v_1 = 0$ or $z_1 \geq U_{v_1} - \lambda$, then $\hat{h}(z_1, v_1) = \hat{g}(z_1, v_1)$. On the other hand, if $z_1 \leq U_{v_1} - \lambda$, then setting $(z_2, v_2) = ((U_{v_1} - \lambda) - z_1, 0)$ clearly suffices. Thus $\hat{h}$ is non-dominated. □

5.3 Stable Set

We now describe how we can apply superadditive lifting to the odd-hole inequalities of the stable set polytope. Before we can do this, we must manage the dimension of the lifting function. We divide this section into three parts: in the first we describe the problem setting and address the lifting function; in the second, we explore how we can specialize our approximation to obtain a deep cut; and in the third, we compare the performance of lifted and non-lifted inequalities.

5.3.1 The Stable Set Polytope and Lifting

Let $G = (V, E)$ be a simple undirected graph. A *stable set* of G is a set $S \subseteq V$ such that for all distinct $u, v \in S$, $(u, v) \notin E$. The complement of a stable set, a *clique*, is a set $S \subseteq V$ such that for all distinct $u, v \in S$, $(u, v) \in E$. For a vertex $v \in V$, its *neighborhood*, $N(v)$, is defined by $N(v) = \{u \in V : (u, v) \in E\}$. We will also be interested in the neighborhood of v belonging to some $U \subseteq V$; thus we define $N_U(v) = N(v) \cap U$.

Given a set $S \subseteq V$, we can encode S by its incidence vector. Thus the set of all stable sets is defined by

$$X = \left\{ x \in \{0,1\}^{|V|} : x_u + x_v \leq 1, \quad \forall (u, v) \in E \right\}. \tag{87}$$

The *stable set polytope* is defined by $P = \text{conv}(X)$. Observe that if we delete an edge (u, v) from G, then this is equivalent to deleting its corresponding constraint $x_u + x_v \leq 1$ from the description of X.

We show that by relaxation and reformulation we can derive valid inequalities for the stable set polytope from a much lower-dimensional set. Fix $U \subseteq V$ and some $S \subseteq U \subseteq V$, and let $V_S = \{v \in V : S = N_U(v)\} \neq \emptyset$. Let $G' = (V', E')$ be the graph defined by setting

$V' = V$ and

$$E' = E \setminus \{(u,v) \in E : v \in V_S, u \in V \setminus (V_S \cup S)\}.$$

Conceptually, G' is obtained by removing the edges leaving V_S that do not end in S. Let X' denote the relaxed set and $P' = \mathrm{conv}(X')$. Clearly $P \subseteq P'$, so valid inequalities for P' are valid for P.

Let $G'' = (V'', E'')$ be obtained by identifying the vertices of V_S with a single vertex v^*. Therefore,

$$V'' = \left(V' \setminus V_S\right) \cup \{v^*\},$$
$$E'' = \left(E' \setminus \left\{(u,v) \in E' : \{u,v\} \cap V_S \neq \emptyset\right\}\right) \cup \{(u,v^*) : u \in S\}.$$

Similarly, we define X'' and P''. The operations transforming G into G'' are shown in Figure 21. We now show that facets of P'' readily translate to facets of P'.

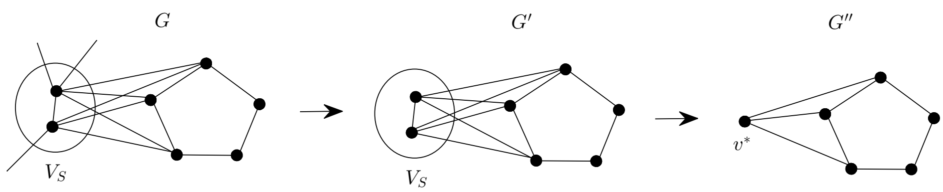

Figure 21: Relaxation and reformulation of G

Theorem 5.3.1. *Let $\pi x \leq \pi_0$ be a non-trivial facet of P'' and $K \subseteq V_S$ be a maximal clique. Define the inequality $\pi' x \leq \pi'_0$ as follows: $\pi'_v = \pi_v$ for $v \in V' \setminus V_S$, $\pi'_0 = \pi_0$, and*

$$\pi'_v = \begin{cases} \pi_{v^*} & v \in K \\ 0 & v \in V_S \setminus K. \end{cases}$$

Then $\pi' x \leq \pi'_0$ defines a facet of P'.

Proof. First we argue validity. At most one vertex from K can be contained in any stable set of G'. Let x' denote the incidence vector of a stable set of G', and let x'' be obtained by setting $x''_{v^*} = \sum_{v \in K} x'_v$ and $x''_v = x'_v$ for $v \notin V_S$. Since the set of neighbors of K not belonging to V_S is precisely S; thus x'' defines a stable set of G''. In particular, if $\pi' x' > \pi'_0$, then $\pi x'' > \pi_0$.

Next we show that this inequality is facet-defining. Because the singletons and the empty set are all stable sets, P' and P'' are full-dimensional, and the non-negativity constraints are facet-defining.

Let $n' = |V'|$ and $n'' = |V''| = n' - |V_S| + 1$. Any facet-defining inequality of P' (respectively P'') is satisfied by n' (respectively n'') affinely independent points. Let $y^1, \ldots, y^{n''} \in X''$ satisfy $\pi y = \pi_0$, and fix some $\hat{v} \in K$. Construct n'' affinely independent points $x^1, \ldots, x^{n''}$ as follows:

$$x_v^i = \begin{cases} y_v^i & v \in V' \setminus V_S \\ y_{v^*}^i & v = \hat{v} \\ 0 & v \in V_S \setminus \hat{v}. \end{cases}$$

By construction $x^i \in X'$ and $\pi' x^i = \pi'_0$.

Now we must construct the remaining $|V_S| - 1$ points. Because $\pi x \leq \pi_0$ is non-trivial, there exists some y^i such that $y_{v^*}^i = 1$. For $v \in K \setminus \hat{v}$, set $x^v = x^i - e_{\hat{v}} + e_v$. Otherwise, for $v \in V_S \setminus K$, there exists some $v' \in K$ such that $(v, v') \notin E'$ by the maximality of K. Therefore, we set $x^v = x^{v'} + e_v$.

It is trivial to verify that $x^v \in X'$ and $\pi' x^v = \pi'_0$. Further, these points are necessarily affinely independent, so we have constructed $n'' + |V_S| - 1 = n'$ tight affinely independent points, proving that $\pi' x \leq \pi'_0$ is facet-defining. □

This theorem gives us a tool for performing sequence-independent lifting on the *odd-hole inequalities*. An *odd-hole* is simply a minimal odd-length cycle: i.e. it contains no shorter length cycles. For $k \geq 1$ integral, let C_{2k+1} be an odd-hole on $2k+1$ vertices. Numbering the vertices from 1 to $2k+1$, the odd-hole inequality is defined by

$$\sum_{j=1}^{2k+1} x_j \leq k. \tag{88}$$

Whenever $G = C_{2k+1}$ this inequality is facet defining for P; however, even when $G \neq C_{2k+1}$, the odd-hole inequality can be used to derive strong valid inequalities for P through lifting.

Given a graph G, let C be an odd-hole. Partition the vertices of $V \setminus C$ into sets $V_1, \ldots, V_t$ such that $S_i = N_C(v) = N_C(u)$ for all $u, v \in V_i$ and $S_i \neq S_j$ for $i \neq j$. Let G' be obtained by deleting all edges (u, v) such that $u \in V_i$ and $v \in V_j$ for $i \neq j$, and construct G'' by identifying V_i with a single vertex v_i^* for $i = 1, \ldots, t$.

By repeated application of Theorem 5.3.1, we can construct valid inequalities for P' (and therefore P) from P''.

We proceed by lifting the odd-hole inequalities. We first construct a restricted set X^0 by setting $x_{v_i^*} = 0$ for $i = 1, \ldots, t$:

$$X^0 = \left\{ x \in \{0,1\}^{2k+1} : \begin{array}{l} x_i + x_{i+1} \leq 1, \quad i = 1, \ldots, 2k \\ x_1 + x_{2k+1} \leq 1 \end{array} \right\}. \tag{89}$$

The odd-hole inequality (88) is facet-defining for $P^0 = \mathrm{conv}(X^0)$, and its associated lifting function is given by

$$\begin{array}{rll} f(z_1, \ldots, z_t) = \min & k - \displaystyle\sum_{j=1}^{2k+1} x_j & \\ \text{s.t.} & x_i \leq 1 - z_j, & i \in S_j,\ j = 1, \ldots, t \\ & x_i + x_{i+1} \leq 1, & i = 1, \ldots, 2k \\ & x_1 + x_{2k+1} \leq 1, & \\ & x_i \in \{0, 1\}, & i = 1, \ldots, 2k+1. \end{array} \tag{90}$$

It is worth noting that $t \leq \min\left\{2^{2k+1}, 2^{n-(2k+1)}\right\}$. Although this may still be quite large, it is often a substantial improvement over lifting in the original variable space.

In general, f is not superadditive as we show in the next example.

Example 5.3.1. Let $C = \{1, 2, 3, 4, 5\}$. Let $S_1 = \{1, 2, 3\}$ and $S_2 = \{2, 3, 4\}$. In this case, it is easy to see that $f(1, 0) = 1$ and $f(0, 1) = 1$. Similarly $x_5 = 1$ is still feasible for $f(1, 1)$. Therefore,

$$f(0, 1) + f(1, 0) = 2 > 1 = f(1, 1),$$

violating superadditivity.

For $i = 1, \ldots, t$, define the function

$$\begin{aligned} \Phi^j(z_j) = \min \quad & k - \sum_{j=1}^{2k+1} x_j \\ \text{s.t.} \quad & x_i \le 1 - z_j, && i \in S_j \\ & x_i + x_{i+1} \le 1, && i = 1, \ldots, 2k \\ & x_1 + x_{2k+1} \le 1, \\ & x_i \in \{0, 1\}, && i = 1, \ldots, 2k+1. \end{aligned}$$

We can easily calculate Φ^j. When $z_j = 0$, clearly $\Phi^j(z_j) = 0$. Otherwise, we consider the graph $C \setminus S_j$, which is a union of isolated vertices and disjoint paths. Letting k_j denote the maximum cardinality stable set on this graph, $\Phi^j(1) = k_j$.

Similarly, define Ψ^j by

$$\begin{aligned} \Psi^j(z_j, v) = \min \quad & k - \sum_{j=1}^{2k+1} x_j \\ \text{s.t.} \quad & x_i \le 1 - z_j, && i \in S_j \\ & \sum_{j=1}^{2k+1} x_j \le k - v \\ & x_i + x_{i+1} \le 1, && i = 1, \ldots, 2k \\ & x_1 + x_{2k+1} \le 1, \\ & x_i \in \{0, 1\}, && i = 1, \ldots, 2k+1. \end{aligned}$$

It is straightforward to see that $\Psi^j(z_j, v) = \max\left(\Phi^j(z_j), v\right)$. In particular, $\Psi^j(0, v) = v$ and $\Psi^j(1, v) = \max(k_j, v)$. Clearly Ψ^j is not superadditive whenever $k_j > 0$. However, there exists a family of non-dominated superadditive approximations for Ψ^j. One such example is given in Figure 22.

Theorem 5.3.2. *Let r_j be some integer satisfying $0 \le r_j \le k_j$. Define the function $\psi^j(z_j, v)$ by $\psi^j(0, v) = (v - r_j)^+$ and*

$$\psi^j(1, v) = \begin{cases} r_j + v & 0 \le v \le k_j - r_j \\ k_j & k_j - r_j \le v \le k_j \\ v & k_j \le v \le k. \end{cases} \tag{91}$$

Then ψ^j is a non-dominated superadditive valid approximation of Ψ^j.

Proof. Validity is clear. Superadditivity follows from case analysis. First consider $\psi^j(0,u)+\psi^j(0,v)$. If $u \leq r_j$, $\psi^j(0,u) = 0$. As ψ^j is increasing,

$$\psi^j(0,u+v) \geq \psi^j(0,v) = \psi^j(0,v) + \psi^j(0,u).$$

Therefore assume that $u, v > r_j$; hence

$$\psi^j(0,u) + \psi^j(0,v) = u + v - 2r_j \leq u + v - r_j = \psi^j(0,u+v).$$

So now we consider $\psi^j(1,u) + \psi^j(0,v)$. Since ψ^j is increasing we can again assume that $v > r_j$. Noting that $\psi^j(1,u) \leq r_j + u$,

$$\psi^j(1,u) + \psi^j(0,v) \leq (r_j + u) + (v - r_j) = u + v \leq \psi^j(1,u+v).$$

proving superadditivity.

Next we argue non-dominance. Observe that for $\psi^j(1,u) = \Psi^j(1,u)$ for $u \geq k_j - r_j$ and cannot be increased. For $0 \leq v \leq r_j$, $\psi^j(1,k_j - r_j) + \psi^j(0,v) = k_j = \psi^j(1,u + v) = \Psi^j(1,u+v)$. Therefore we cannot increase $\psi^j(0,v)$. Similarly, for $r_j \leq v \leq k - r_j$, $\psi^j(1,k_j - r_j) + \psi^j(0,v) = k_j + v - r_j = \psi^j(1,k_j + v - r_j) = \Psi^j(1,k_j + v - r_j)$, so we cannot increase $\psi^j(0,v)$. Lastly, for $0 \leq u \leq k_j - r_j$, we set $v = k - u$, and $\psi^j(1,u) + \psi^j(0,v) = k = \psi^j(1,k) = \Psi^j(1,k)$, and we cannot increase $\psi^j(1,u)$ or $\psi^j(0,v-u)$.

We have exhaustively shown that we cannot increase ψ^j anywhere without having to decrease ψ^j at another point. Hence we cannot increase ψ^j and preserve validity and superadditivity. ☐

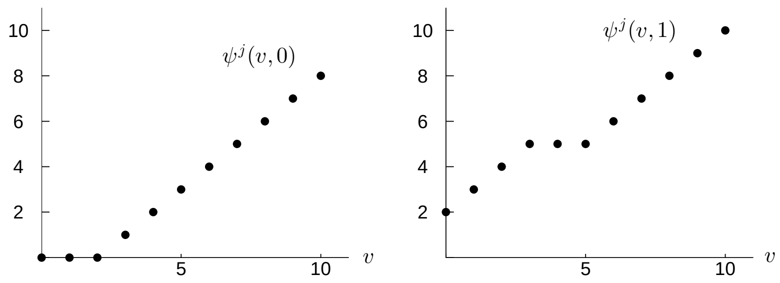

Figure 22: The function ψ^j for the odd-hole inequalities

Noting that ψ^j is superadditive, the next theorem follows as an immediate consequence of Theorem 3.3.3:

Theorem 5.3.3. *Let $\phi^t = \Phi^t$, and define ϕ^i by $\phi^i(z_i, z_{i+1}, \ldots, z_t) = \psi\left(z_i, \phi(z_{i+1}, \ldots, z_t)\right)$. Setting $\phi = \phi^1$,*

$$\phi\left(z_1, \ldots, z_t\right) \leq f\left(z_1, \ldots, z_t\right),$$

and is superadditive.

Using ϕ we are able to determine valid lifting coefficients for $x_{v_i^*}$, by simply calculating ϕ at the corresponding point. By repeatedly applying Theorem 5.3.1, we can obtain valid inequalities for P' and hence P.

5.3.2 Obtaining Deep Cuts

Although it is always possible to perform superadditive lifting using this nesting scheme, we may wish to find an approximation that reflects our goal of separating a fractional solution. A non-dominated approximation is a reasonable starting point, but it is generally possible that many such approximations exist and their performance will vary greatly.

For the lifting problem (90), we only considered sets $S \subseteq C$ such that there exists some vertex v such that $N_C(v) = S$. Let $\mathcal{S} = \{S_1, \ldots, S_t\}$ denote the set of all such subsets of C. Let $J = \{1, \ldots, t\}$, and let $I = \{j \in J : f(e_i) > 0\}$ and $\bar{I} = J \setminus I$. We show that it is possible to further reduce the dimension of the lifting function by excluding vertices v_i^* with $i \in \bar{I}$.

Rewrite $f(z)$ as $f\left(z_I, z_{\bar{I}}\right)$, where z_I and $z_{\bar{I}}$ denote the components of z restricted to I and $\bar{I}$ respectively. Consider the restriction, $f_{\mathrm{R}}\left(z_I\right) = f\left(z_I, 0\right)$. In the next proposition we show that a superadditive approximation of f_{R} always obtains lifting coefficients at least as strong as a superadditive approximation of f.

Proposition 5.3.4. *Let g be a valid superadditive approximation of f. Then there exists some superadditive approximation h_{R} of f_{R} such that the function h defined by $h\left(z_I, z_{\bar{I}}\right) = h_{\mathrm{R}}\left(z_I\right)$ is valid, superadditive, and produces lifting coefficients at least strong as those obtained from g.*

Proof. Let g be a superadditive approximation of f. Then the function g_{R} defined by $g_{\mathrm{R}}(z_I) = g(z_I, 0)$ is a valid superadditive approximation of f_{R}. Now let $g_{\mathrm{R}} \leq h_{\mathrm{R}} \leq f_{\mathrm{R}}$ for some superadditive h_{R}. In the worst case, $h_{\mathrm{R}} = g_{\mathrm{R}}$.

As f is increasing, $f_{\mathrm{R}}(z_I) \leq f(z_I, z_{\bar{I}})$. Thus the function h defined by $h(z_I, z_{\bar{I}}) = h_{\mathrm{R}}(z_I)$ must be a valid approximation of f. Moreover, we claim that h is superadditive. Indeed, this easily follows from the superadditivity of h_{R}, as

$$\begin{aligned} h(u_I, u_{\bar{I}}) + h(v_I, v_{\bar{I}}) &= h_{\mathrm{R}}(u_I) + h_{\mathrm{R}}(u_I) \\ &\leq h_{\mathrm{R}}(u_I + v_I) = h(u_I + v_I, u_{\bar{I}} + v_{\bar{I}}). \end{aligned}$$

Thus it remains to consider the lifting coefficients. For $i \in I$, $h(e_i, 0) \geq g(e_i, 0)$. Likewise for $i \in \bar{I}$, $h(0, e_i) = g(0, e_i) = 0$. In particular, the lifting coefficients obtained from h are at least as strong as those obtained from g. □

Therefore, we are able to further reduce our problem dimension, so we can assume without loss of generality that $f(e_i) > 0$ for all $i \in I$. Next we show that for any non-dominated approximation, g of f, $g(e_i) \geq 0$.

Proposition 5.3.5. *If g is non-dominated, then $g \geq 0$.*

Proof. Assume to the contrary that $g(z_1) < 0$. As g is non-dominated, there exists some z_2 such that

$$g(z_1) + g(z_2) = f(z_1 + z_2).$$

However, this implies that $g(z_2) = f(z_1 + z_2) - g(z_1) > f(z_1 + z_2) > f(z_2)$. But this contradicts that g is valid. □

Suppose now that we have some point $\bar{x} \geq 0$ and an odd-hole C that induces some partition $V_{S_1}, \ldots, V_{S_t}$ of $V \setminus C$, and identify cliques $K_1, \ldots, K_t$ with $K_i \subseteq V_{S_i}$. Therefore, if $|C| = 2k + 1$, the lifted odd-hole inequality takes the form

$$\sum_{i=1}^{t} \left(g(e_i) \sum_{v \in K_i} x_v \right) + \sum_{v \in C} x_v \leq k. \tag{92}$$

Our goal is to maximize the left-hand side of (92) when evaluated at $\bar{x}$. There are two different parts of this inequality that we can control: the cliques K_i and the function g.

Proposition 5.3.6. *Without loss of generality the g maximizing the left-hand side of (92) is non-dominated.*

Proof. Suppose that g is not non-dominated. Then by definition there exists a function h such that $g \leq h \leq f$ and $h(z) > g(z)$ for some z. In particular, as $\bar{x} \geq 0$,

$$\sum_{i=1}^{t} \left(h(e_i) \sum_{v \in K_i} \bar{x}_v \right) + \sum_{v \in C} \bar{x}_v \geq \sum_{i=1}^{t} \left(g(e_i) \sum_{v \in K_i} \bar{x}_v \right) + \sum_{v \in C} \bar{x}_v.$$

Thus we may assume that g is non-dominated. □

This immediately yields the following proposition:

Proposition 5.3.7. *Without loss of generality, any choice of K_i maximizing the left-hand side of (92), maximizes $\sum_{v \in K_i} \bar{x}_v$.*

Proof. As g is non-dominated, $g(e_i) \geq 0$. Thus, if there exists some clique K_i' with

$$\sum_{v \in K_i'} \bar{x}_v > \sum_{v \in K_i} \bar{x}_v,$$

then we can replace K_i with K_i' to obtain a larger value for the left-hand side of (92). □

Therefore, the first step is to identify maximum weight cliques K_i. When V_{S_i} is small (for example $|V_{S_i}| \leq 20$) then K_i can be identified exactly by enumeration; however, if V_{S_i} is large, we may instead opt to use a heuristic, like the greedy algorithm, to quickly identify some clique. These cliques produce weights

$$w_i = \sum_{j \in K_i} \bar{x}_i,$$

that we can then use to identify the best superadditive approximation. Indeed, this can be achieved by solving the following linear program:

$$\begin{aligned} \max \quad & \sum_{i=1}^{t} w_i \cdot g(e_i) & \\ \text{s.t.} \quad & g(z^1) + g(z^2) \leq g(z^1 + z^2) \qquad & z^1, z^2, z^1 + z^2 \in \{0,1\}^t \\ & 0 \leq g(z) \leq f(z) & z \in \{0,1\}^t . \end{aligned} \tag{93}$$

We are not interested in the value of the lifting function at $g(z)$ whenever z is not one of the unit vectors. So the number of variables in (93) greatly exceeds what is needed. In the next proposition, we show a much smaller linear program that achieves the same maximum.

Proposition 5.3.8. *The linear program*

$$\begin{aligned} \max \quad & \sum_{i=1}^{t} w_i \cdot g(e_i) \\ s.t. \quad & \sum_{i=1}^{t} z_i \cdot g(e_i) \leq f(z) \qquad z \in \{0,1\}^t \\ & 0 \leq g(e_i) \leq f(e_i) \qquad i = 1, \ldots, t \\ & g(0) = 0 \end{aligned} \tag{94}$$

achieves the same maximum as (93).

Proof. First consider an optimum solution $\bar{g}$ to (93). We claim that $(\bar{g}(e_1), \ldots, \bar{g}(e_t))$ is feasible to (94). Indeed, by taking a linear combination of the constraints of (93), it follows that

$$\sum_{i=1}^{t} z_i \cdot \bar{g}(e_i) \leq \bar{g}(z) \leq f(z).$$

On the other hand, consider a solution $\hat{g}$ to (94). Then if we set

$$\bar{g}(z) = \sum_{i=1}^{t} z_i \cdot \hat{g}(e_i),$$

we satisfy $\bar{g}(z^1) + \bar{g}(z^2) = \bar{g}(z^1 + z^2)$ for all $z^1, z^2, z^1 + z^2 \in \{0,1\}^t$, and $\bar{g}(z) \leq f(z)$ for all $z \in \{0,1\}^t$. Therefore, this solution is feasible to (93). □

As one final remark, (94) has exponentially many constraints in t. The techniques of Chapter 3 give us a tool to manage the size of t. It is quite easy to extend these results of this section to accommodate this slight modification.

5.3.3 Lifted Versus Non-Lifted Inequalities

Now that we have a method to identify the lifted odd-hole inequalities obtaining the deepest cut coefficients, we demonstrate that they outperform their non-lifted counterparts.

We consider the odd-hole inequalities on a standard model of random graphs $\mathcal{G}_{n,p}$. Let $\mathcal{G}_{n,p} = (V, E)$, where $V = \{1, \ldots, n\}$ and $(u, v) \in E$ with probability p for all unordered pairs

(u, v). We further assume that $\mathcal{G}_{n,p}$ is connected by discarding graphs that are disconnected. This is more of a simplifying assumption than a necessity, as we could test the connected components separately.

The problem we consider is the *maximum cardinality stable set* given by

$$\begin{aligned} z = \max \quad & \sum_{v \in V} x_v \\ \text{s.t.} \quad & x_u + x_v \leq 1 \qquad (u, v) \in E \\ & x_v \in \{0, 1\} \qquad v \in V. \end{aligned} \tag{95}$$

We relax the integrality requirement and instead optimize over the LP relaxation. If the optimum is attained at a fractional point $\bar{x}$, we then attempt to find a violated odd-hole inequality by using Dijkstra's algorithm on a specially constructed bipartite graph (see [57]).

The first test we perform is without lifting. If a violated odd-hole is found, the odd-hole inequality is added and the LP relaxation is resolved. This process is repeated until no violated odd-hole inequality is found. Lifted inequalities are tested similarly; however, the added inequality is strengthened by solving (94) and using the corresponding cut coefficients. Again we stop when no violated odd-hole inequality is found. Note that this terminating condition does not preclude the possibility that there exist violated lifted inequalities, but identifying such violated inequalities seems far more challenging than separating over the non-lifted odd-hole inequalities.

We test a small collection of instances with $n = 50, 100, 200$ and $p = 1/8, 1/4, 3/8$. For these instances, the sizes of V_{S_i} and t are sufficiently small that enumeration is not prohibitively expensive. For each combination of n and p we tested ten instances. As the sampled graphs are independent and identically distributed, the performance tends to be quite similar for fixed n and p. Therefore, we report the average value of the LP solution upon termination, the average number of cuts added, and the average time in seconds per cut. Represent these quantities by N, z, and τ for the non-lifted cuts and N_ℓ, z_ℓ, and τ_ℓ for the lifted cuts. The results of the test are shown in Table 13.

The most apparent behavior from this table is that of z: namely it tends toward $n/3$. This suggests the following proposition:

Table 13: Performance of (non-)lifted odd-hole inequalities

n	p	z	N	τ	z_ℓ	N_ℓ	τ_ℓ
50	1/8	17.8	83.9	0.04	17.4	83.9	0.04
50	1/4	16.7	70.4	0.05	13.8	100.5	0.05
50	3/8	16.7	65.4	0.07	11.5	101.0	0.07
100	1/8	33.4	158.5	0.21	30.9	216.8	0.21
100	1/4	33.3	131.8	0.31	24.6	220.3	0.30
100	3/8	33.3	129.4	0.39	19.1	166.2	0.39
200	1/8	66.7	277.0	1.40	54.3	577.3	1.43
200	1/4	66.7	261.6	1.75	43.9	397.6	1.74
200	3/8	66.7	251.4	2.43	33.0	250.0	2.37

Proposition 5.3.9. *$z \geq n/3$ for all n.*

Proof. Consider the solution $\hat{x}$ obtained by setting $x_v = 1/3$ for all $v \in V$. Clearly, this solution satisfies any edge constraint; thus it remains to consider the odd-hole constraints. Let $C = \{1, \ldots, 2k+1\}$ be some odd-hole. Then the left-hand side of the corresponding odd-hole constraint is given by

$$\sum_{j=1}^{2k+1} \frac{1}{3} = \frac{2k+1}{3}.$$

If $k \geq 1$, then $k \geq (2k+1)/3$. Therefore the odd-hole constraint must be satisfied. □

Thus, without lifting, the odd-hole inequalities are quite limited in their performance. But it is readily observed that the approximate lifting we proposed attains solutions beyond this bound; moreover, their performance tends to improve as the graph contains more edges. This trend is apparent in Figures 23 through 25, which plot the objective as a function of the number of cuts added for various parameter settings.

For these examples, lifting does not seem to significantly impact the time it takes to generate an individual cut; nevertheless, the impact on the final objective is quite noticeable. The overall running times increase on many of the problems as we apply lifting, but this is more closely related to the number cuts that we add. Considering the objective as a function of the number of cuts added, it is evident that the lifted inequalities are the clear winners as they are consistently capable of producing better bounds using fewer cuts.

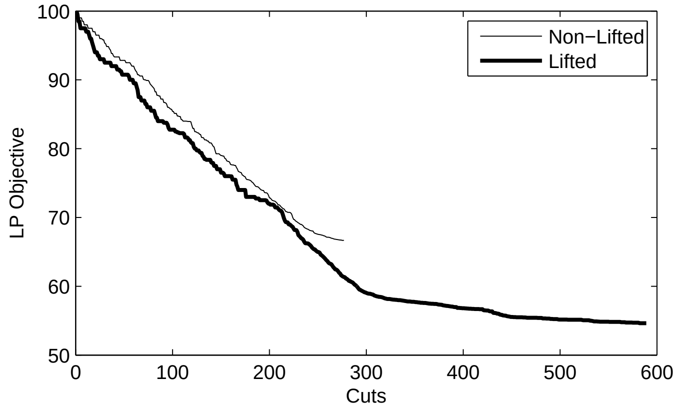

Figure 23: Non-lifted versus lifted odd-hole inequalities with $n = 200$ and $p = 1/8$

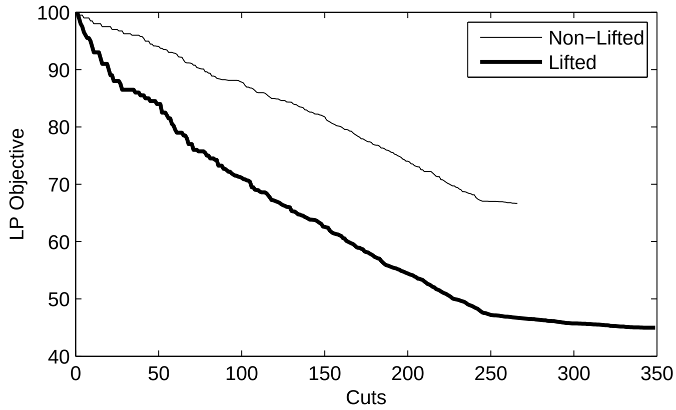

Figure 24: Non-lifted versus lifted odd-hole inequalities with $n = 200$ and $p = 1/4$

5.4 Closing Remarks

In this chapter, we put together ideas appearing throughout this thesis to derive cuts for higher-dimensional problems. We highlighted the potential of our algorithm to explicitly derive non-dominated approximations for concrete examples that we could then generalize. Using the approach of Chapter 3, these general approximate lifting functions could then be applied to considerably more challenging problems that were otherwise inaccessible to lifting.

Our study of lifted odd-hole inequalities exposed the potential for superadditive lifting to be embedded into a cutting plane framework even without a priori knowledge of the lifting functions involved. By using superadditivity we were able to simultaneously ensure validity

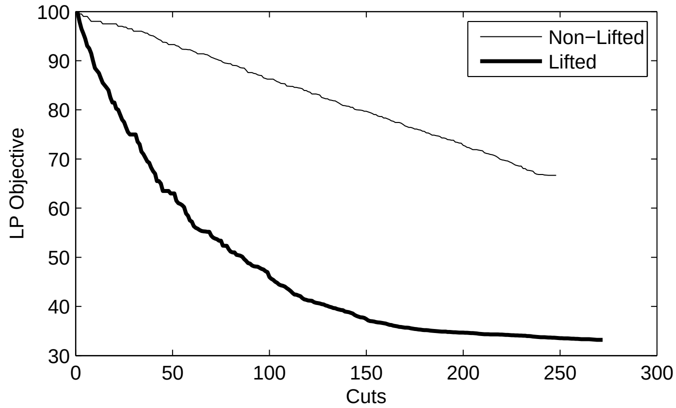

Figure 25: Non-lifted versus lifted odd-hole inequalities with $n = 200$ and $p = 3/8$

and target the deepest cut. This idea might readily be extended to other combinatorial problems. The difficulty of describing valid lifting coefficients in the general and mixed integer case translate to this setting. In this regard, it is an interesting open problem how and when one might be able to extend this idea to these cases.

As we demonstrated, lifting in higher dimensions enables us to strengthen known families of cuts; however, it also offers the potential to obtain entirely new cuts. The tools we have developed can help reveal previously unexplored applications for this powerful technique.

CHAPTER VI

FUTURE WORK

In this thesis, we investigated cutting planes in mixed integer programming. Generally speaking, our main contributions are to the theoretical understanding of the Gomory mixed integer cut and to techniques for deriving multi-row cuts through lifting.

Our study of the mixed integer cut in Chapter 2 approached the problem by considering it as a facet of the master cyclic group polyhedron. We showed that its adjacent facets are precisely tilted knapsack facets that arise from one of two different knapsack polytopes. Given the practical success of the mixed integer cut, this result indicates that the tilted knapsack facets may play an important role in the solution of mixed integer programs. Therefore, one worthy pursuit is the identification of new classes of knapsack facets.

We were also able to generalize our results under automorphic and homomorphic mappings by extending two of Gomory's original results to include adjacency. Under automorphic mapping our result translates in its entirety; however, we are unfortunately left with an incomplete understanding of the lifted mixed integer cut obtained through homomorphisms. The same shooting experiments that frequently hit the mixed integer cut often hit its lifted counterparts; thus, these facets may be quite important. Another avenue of research would be the classification of adjacent facets to the lifted mixed integer cut that do not arise by lifting the tilted knapsack facets.

In Chapter 3, we considered the task of lifting an already lifted flow cover inequality. The lifted simple generalized flow cover inequality is obtained similarly, but reintroduces a different set of variables before performing the second lifting step. This example demonstrated the challenge in characterizing the second lifting function let alone even approximating it. In particular, the standard practice of using the exact lifting function to establish the validity of an approximation becomes very cumbersome. By exploiting the fact that the lifting function arises through the solution of a mixed integer program, we are able to reformulate and relax the lifting problem to prove that the superadditive approximation we proposed was indeed valid.

We take this idea further and give an entire framework for obtaining approximate lifting functions for high-dimensional problems by relaxing sets of complicating constraints with lifting functions. Despite its simplicity, this idea has not yet been explored; moreover, it gives us a tool for incorporating side constraints that would otherwise be discarded into the lifting function. By simplifying the lifting problem, we are able to open entirely new problems to the approach and obtain multi-row cuts.

There are a number of open questions that naturally follow from this work. For the fixed-charge flow set we studied, one very interesting question is whether a non-dominated approximation can actually be constructed and how it would compare computationally against the highly successful lifted simple generalized flow cover inequalities. With regard to the approximation of high-dimensional lifting functions, another worthwhile endeavor is determining conditions under which the non-dominance of the final superadditive approximation is guaranteed.

Next in Chapter 4 we address the computational aspects of constructing non-dominated superadditive approximations. We restricted our attention to one-dimensional lifting functions and simple variants of one-dimensional lifting functions restricted to the positive orthant. Our work revealed that the desirable characteristics of validity, superadditivity, and non-dominance can all be efficiently tested. These results also simplify proofs of superadditivity and non-dominance by reducing the tests to an easily certifiable condition over a small collection of points in the domain. However, our main result in this chapter is the existence of a finite algorithm that constructs non-dominated approximations.

By far the biggest limitation in this work is the assumption we make about the problem domain. Along this line, there are two natural extensions. The first is the development of a necessary and sufficient condition to characterize non-dominance when the domain can include negatives. Even for a one-dimensional discrete domain, the presence of negatives greatly complicates matters. Such work might also lead to an algorithmic construction of non-dominated approximations in this slightly more general setting. Additionally, we limited the continuous part of the function to be one-dimensional. This essentially limits the algorithmic treatment of superadditive approximations to lifting functions with one continuous argument. There is no reason we need to restrict ourselves to this setting. The structure of the lifting function is well understood in high dimensions; thus, it may be possible to use this structure to extend our results to more general lifting functions.

Lastly, note that we were only able to prove the finite termination of our algorithm. The bound we obtained was exponential in the number of breakpoints, but our experience working examples did not seem to agree with this bound. This suggests two possibilities: either the algorithm is polynomial or pseudo-polynomial and a different analysis would reveal this, or the algorithm requires exponential time and we have not yet considered the right example. Conclusively resolving this question one way or the other would give a much more complete picture of this work.

We concluded in Chapter 5 by applying lifting to several high-dimensional problems. Using the algorithm from Chapter 4, we were able to guide our intuition to identify lifting functions that generate cut coefficients that are at least as good and possibly better than could otherwise be obtained from the single-row systems. In this chapter, we only considered knapsack covers and flow covers, but there are many classical problems for which side constraints might be used to further strengthen cuts.

Finally, we considered the odd-hole inequalities for the stable set polytope. The more traditional approach to superadditive lifting relies on an a priori description of the lifting function. However, taking this approach directly for the odd-hole inequalities would completely blow up the size of the lifting function. We are able to overcome this challenge by computing the lifting function as we need it. We can then use a linear program to identify a superadditive approximation that gives the deepest separation with respect to a current fractional solution.

We rely specifically on the variables we reintroduce being binary; therefore our approach seems well suited to combinatorial problems. There is a wide array of such problems where superadditive lifting has not actively been explored as an option for obtaining good cuts. It may be possible to further generalize this technique to general integer and mixed integer problems; however, the choice of lifting coefficients producing the deepest separation becomes far less transparent. Progress on this problem will reveal even more possibilities for lifting, and might be a large step in the search for effective multi-row cuts.

REFERENCES

[1] Applegate, D., Bixby, R., Chvatal, V., and Cook, W., *The traveling salesman problem: a computational study.* Princeton University Press, 2007.

[2] Aráoz, J., *Polyhedral neopolarities.* PhD thesis, University of Waterloo, 1973.

[3] Aráoz, J., Evans, L., Gomory, R., and Johnson, E., "Cyclic group and knapsack facets," *Mathematical Programming*, vol. 96, no. 2, pp. 377–408, 2003.

[4] Atamtürk, A., "Sequence independent lifting for mixed-integer programming," *Operations Research*, vol. 52, no. 3, pp. 487–490, 2004.

[5] Balas, E., "Facets of the knapsack polytope," *Mathematical Programming*, vol. 8, no. 1, pp. 146–164, 1975.

[6] Balas, E., Ceria, S., and Cornuéjols, G., "A lift-and-project cutting plane algorithm for mixed 0–1 programs," *Mathematical programming*, vol. 58, no. 1, pp. 295–324, 1993.

[7] Balas, E., Ceria, S., Cornuéjols, G., and Natraj, N., "Gomory cuts revisited," *Operations Research Letters*, vol. 19, no. 1, pp. 1–9, 1996.

[8] Balas, E. and Zemel, E., "Facets of the knapsack polytope from minimal covers," *SIAM Journal on Applied Mathematics*, vol. 34, no. 1, pp. 119–148, 1978.

[9] Barany, I., Roy, T., and Wolsey, L., "Uncapacitated lot-sizing: The convex hull of solutions," *Mathematical Programming at Oberwolfach II*, pp. 32–43, 1984.

[10] Barnhart, C., Johnson, E., Nemhauser, G., Savelsbergh, M., and Vance, P., "Branch-and-price: Column generation for solving huge integer programs," *Operations Research*, vol. 46, no. 3, pp. 316–329, 1998.

[11] Basu, A., Conforti, M., Cornuéjols, G., and Zambelli, G., "Maximal lattice-free convex sets in linear subspaces," *Mathematics of Operations Research*, vol. 35, no. 3, pp. 704–720, 2010.

[12] Basu, A., Conforti, M., Cornuéjols, G., and Zambelli, G., "Minimal inequalities for an infinite relaxation of integer programs," *SIAM Journal on Discrete Mathematics*, vol. 24, no. 1, pp. 158–168, 2010.

[13] Bertsimas, D. and Tsitsiklis, J. N., *Introduction to linear optimization.* Athena Scientific series in optimization and neural computation, Belmont, Mass.: Athena Scientific, 1997.

[14] Blair, C. and Jeroslow, R., "The value function of a mixed integer program: I," *Discrete Mathematics*, vol. 19, no. 2, pp. 121–138, 1977.

[15] Borozan, V. and Cornuéjols, G., "Minimal valid inequalities for integer constraints," *Mathematics of Operations Research*, vol. 34, no. 3, pp. 538–546, 2009.

[16] Chvátal, V., "Edmonds polytopes and weakly hamiltonian graphs," *Mathematical Programming*, vol. 5, no. 1, pp. 29–40, 1973.

[17] Conforti, M., Di Summa, M., and Wolsey, L., "The mixing set with flows," *SIAM Journal on Discrete Mathematics*, vol. 21, no. 2, pp. 396–407, 2007.

[18] Conforti, M., Di Summa, M., and Wolsey, L., "The mixing set with divisible capacities," *Integer Programming and Combinatorial Optimization*, pp. 435–449, 2008.

[19] Conforti, M., Wolsey, L., and Zambelli, G., "Projecting an extended formulation for mixed-integer covers on bipartite graphs," *Mathematics of operations research*, vol. 35, no. 3, pp. 603–623, 2010.

[20] Cook, W., Cunningham, W., Pulleyblank, W., and Schrijver, A., *Combinatorial optimization*. Wiley-Interscience series in discrete mathematics and optimization, New York: Wiley, 1998.

[21] Cook, W., Kannan, R., and Schrijver, A., "Chvátal closures for mixed integer programming problems," *Mathematical Programming*, vol. 47, no. 1, pp. 155–174, 1990.

[22] Crowder, H., Johnson, E., and Padberg, M., "Solving large-scale zero-one linear programming problems," *Operations Research*, vol. 31, no. 5, pp. 803–834, 1983.

[23] Dantzig, G., "Maximization of a linear function of variables subject to linear inequalities," *New York*, 1951.

[24] Dantzig, G., Fulkerson, R., and Johnson, S., "Solution of a large-scale traveling-salesman problem," *Journal of the operations research society of America*, pp. 393–410, 1954.

[25] de Farias Jr, I. and Nemhauser, G., "A polyhedral study of the cardinality constrained knapsack problem," *Mathematical programming*, vol. 96, no. 3, pp. 439–467, 2003.

[26] Dummit, D. and Foote, R., *Abstract algebra*, vol. 43. Prentice Hall New Jersey, 1999.

[27] Edmonds, J., "Maximum matching and a polyhedron with 0, 1-vertices," *Journal of Research of the National Bureau of Standards B*, vol. 69, pp. 125–130, 1965.

[28] Edmonds, J. and Johnson, E., "Matching: A well-solved class of integer linear programs," in *in: Combinatorial structures and their applications (Gordon and Breach*, Citeseer, 1970.

[29] Ford, L. and Fulkerson, D., "Maximal flow through a network," *Canadian Journal of Mathematics*, vol. 8, no. 3, pp. 399–404, 1956.

[30] Ford, L. and Fulkerson, D., "A suggested computation for maximal multicommodity network flows," *Management Science*, vol. 5, no. 1, pp. 97–101, 1958.

[31] Gilmore, P. and Gomory, R., "A linear programming approach to the cutting-stock problem," *Operations research*, vol. 9, no. 6, pp. 849–859, 1961.

[32] Gilmore, P. and Gomory, R., "A linear programming approach to the cutting stock problem–part ii," *Operations research*, vol. 11, no. 6, pp. 863–888, 1963.

[33] Gomory, R., "An algorithm for the mixed integer problem," tech. rep., DTIC Document, 1960.

[34] Gomory, R., "Outline of an algorithm for integer solutions to linear programs," *Bulletin of the American Mathematical Society*, vol. 64, no. 5, pp. 275–278, 1958.

[35] Gomory, R., "Some polyhedra related to combinatorial problems," *Linear Algebra and its Applications*, vol. 2, no. 4, pp. 451–558, 1969.

[36] Gomory, R. and Johnson, E., "Some continuous functions related to corner polyhedra," *Mathematical Programming*, vol. 3, no. 1, pp. 23–85, 1972.

[37] Gomory, R. and Johnson, E., "Some continuous functions related to corner polyhedra, ii," *Mathematical Programming*, vol. 3, no. 1, pp. 359–389, 1972.

[38] Gomory, R., Johnson, E., and Evans, L., "Corner polyhedra and their connection with cutting planes," *Mathematical Programming*, vol. 96, no. 2, pp. 321–339, 2003.

[39] Grötschel, M. and Padberg, M., "On the symmetric travelling salesman problem ii: lifting theorems and facets," *Mathematical Programming*, vol. 16, no. 1, pp. 281–302, 1979.

[40] Gu, Z., *Lifted Cover Inequalities for 0-1 and Mixed 0-1 Programs.* PhD thesis, Georgia Institute of Technology, 1995.

[41] Gu, Z., Nemhauser, G., and Savelsbergh, M., "Lifted flow cover inequalities for mixed 0-1 integer programs," *Mathematical Programming*, vol. 85, no. 3, pp. 439–467, 1999.

[42] Gu, Z., Nemhauser, G., and Savelsbergh, M., "Sequence independent lifting in mixed integer programming," *Journal of Combinatorial Optimization*, vol. 4, no. 1, pp. 109–129, 2000.

[43] Günlük, O. and Pochet, Y., "Mixing mixed-integer inequalities," *Mathematical Programming*, vol. 90, no. 3, pp. 429–457, 2001.

[44] Johnson, E., "On the group problem for mixed integer programming," *Approaches to Integer Programming*, pp. 137–179, 1974.

[45] Karmarkar, N., "A new polynomial-time algorithm for linear programming," in *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pp. 302–311, ACM, 1984.

[46] Karp, R., "Reducibility among combinatorial problems," *Complexity of Computer Computations*, 1972.

[47] Khachiian, L., "Polynomial algorithms in linear programming," *Doklady Akademii Nauk SSSR*, vol. 224, pp. 1093–1096, 1979.

[48] Khachiian, L., "Polynomial algorithms in linear programming," *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, vol. 20, pp. 51–68, 1980.

[49] Küçükyavuz, S., "On mixing sets arising in chance-constrained programming," *Mathematical programming*, vol. 132, no. 1, pp. 31–56, 2012.

[50] Loparic, M., Marchand, H., and Wolsey, L., "Dynamic knapsack sets and capacitated lot-sizing," *Mathematical Programming*, vol. 95, no. 1, pp. 53–69, 2003.

[51] Lovász, L. and Schrijver, A., "Cones of matrices and set-functions and 0-1 optimization," *SIAM Journal on Optimization*, vol. 1, no. 2, pp. 166–190, 1991.

[52] Luedtke, J., Ahmed, S., and Nemhauser, G., "An integer programming approach for linear programs with probabilistic constraints," *Integer Programming and Combinatorial Optimization*, pp. 410–423, 2007.

[53] Marchand, H. and Wolsey, L., "The 0-1 knapsack problem with a single continuous variable," *Mathematical Programming*, vol. 85, no. 1, pp. 15–33, 1999.

[54] Margot, F., "Pruning by isomorphism in branch-and-cut," *Mathematical Programming*, vol. 94, no. 1, pp. 71–90, 2002.

[55] Margot, F., "Exploiting orbits in symmetric ilp," *Mathematical Programming*, vol. 98, no. 1, pp. 3–21, 2003.

[56] Narisetty, A., Richard, J., and Nemhauser, G., "Lifted tableaux inequalities for 0–1 mixed-integer programs: A computational study," *INFORMS Journal on Computing*, vol. 23, no. 3, pp. 416–424, 2011.

[57] Nemhauser, G. and Sigismondi, G., "A strong cutting plane/branch-and-bound algorithm for node packing," *Journal of the Operational Research Society*, pp. 443–457, 1992.

[58] Nemhauser, G. and Vance, P., "Lifted cover facets of the 0–1 knapsack polytope with gub constraints," *Operations Research Letters*, vol. 16, no. 5, pp. 255–263, 1994.

[59] Nemhauser, G. and Wolsey, L., *Integer and combinatorial optimization.* Wiley-Interscience series in discrete mathematics and optimization, New York: Wiley, 1988.

[60] Nemhauser, G. and Wolsey, L., "A recursive procedure to generate all cuts for 0–1 mixed integer programs," *Mathematical Programming*, vol. 46, no. 1, pp. 379–390, 1990.

[61] Ostrowski, J., Linderoth, J., Rossi, F., and Smriglio, S., "Constraint orbital branching," in *Integer Programming and Combinatorial Optimization* (Lodi, A., Panconesi, A., and Rinaldi, G., eds.), vol. 5035 of *Lecture Notes in Computer Science*, pp. 225–239, Springer Berlin / Heidelberg, 2008.

[62] Ostrowski, J., Linderoth, J., Rossi, F., and Smriglio, S., "Orbital branching," *Mathematical Programming*, vol. 126, no. 1, pp. 147–178, 2011.

[63] Padberg, M., "On the facial structure of set packing polyhedra," *Mathematical programming*, vol. 5, no. 1, pp. 199–215, 1973.

[64] PADBERG, M., VAN ROY, T., and WOLSEY, L., "Valid linear inequalities for fixed charge problems," *Operations Research*, vol. 33, no. 4, pp. 842–861, 1985.

[65] POCHET, Y. and WOLSEY, L., "Polyhedra for lot-sizing with wagnerwhitin costs," *Mathematical Programming*, vol. 67, no. 1, pp. 297–323, 1994.

[66] POCHET, Y. and WOLSEY, L. A., *Production planning by mixed integer programming.* Springer series in operations research and financial engineering, New York ; Berlin: Springer, 2006.

[67] RICHARD, J., DE FARIAS JR, I., and NEMHAUSER, G., "Lifted inequalities for 0-1 mixed integer programming: Basic theory and algorithms," *Mathematical programming*, vol. 98, no. 1, pp. 89–113, 2003.

[68] RICHARD, J., DE FARIAS JR, I., and NEMHAUSER, G., "Lifted inequalities for 0-1 mixed integer programming: Superlinear lifting," *Mathematical programming*, vol. 98, no. 1, pp. 115–143, 2003.

[69] SCHRIJVER, A., *Theory of linear and integer programming.* Wiley-Interscience series in discrete mathematics and optimization, Chichester ; New York: Wiley, 1986.

[70] SHIM, S. and JOHNSON, E., "Cyclic group blocking polyhedra," *Mathematical Programming*, pp. 1–35, 2012.

[71] STIGLER, G., "The cost of subsistence," *Journal of Farm Economics*, vol. 27, no. 2, pp. 303–314, 1945.

[72] VAN ROY, T. and WOLSEY, L., "Valid inequalities for mixed 0–1 programs," *Discrete Applied Mathematics*, vol. 14, no. 2, pp. 199–213, 1986.

[73] VAN VYVE, M., "The continuous mixing polyhedron," *Mathematics of Operations Research*, vol. 30, no. 2, pp. 441–452, 2005.

[74] WAGNER, H. and WHITIN, T., "Dynamic version of the economic lot size model," *Management Science*, vol. 5, no. 1, pp. 89–96, 1958.

[75] WOLSEY, L., "Technical notefacets and strong valid inequalities for integer programs," *Operations Research*, vol. 24, no. 2, pp. 367–372, 1976.

[76] WOLSEY, L., "Valid inequalities and superadditivity for 0–1 integer programs," *Mathematics of Operations Research*, vol. 2, no. 1, pp. 66–77, 1977.

[77] WOLSEY, L., "Valid inequalities for 0–1 knapsacks and mips with generalised upper bound constraints," *Discrete Applied Mathematics*, vol. 29, no. 2, pp. 251–261, 1990.

[78] ZANETTE, A., FISCHETTI, M., and BALAS, E., "Can pure cutting plane algorithms work?," *Integer Programming and Combinatorial Optimization*, pp. 416–434, 2008.

[79] ZEMEL, E., "Lifting the facets of zero–one polytopes," *Mathematical Programming*, vol. 15, no. 1, pp. 268–277, 1978.

[80] ZENG, B. and RICHARD, J., "A framework to derive multidimensional superadditive lifting functions and its applications," *Integer Programming and Combinatorial Optimization*, pp. 210–224, 2007.

[81] Zeng, B. and Richard, J., "A polyhedral study on 0–1 knapsack problems with disjoint cardinality constraints: strong valid inequalities by sequence-independent lifting," *Discrete Optimization*, vol. 8, no. 2, pp. 259–276, 2011.