Algorithms and Relaxations for Optimization over Bilinear Covering Sets

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

by

Hamidur Rahman

Roll No. 11I190010

Under the guidance of

Prof. Ashutosh Mahajan

Industrial Engineering and Operations Research

INDIAN INSTITUTE OF TECHNOLOGY BOMBAY

Mumbai - 400 076, India

January 2019

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

Name: **Hamidur Rahman**

Roll No.: **11I190010**

Date: ______________

Abstract

Mixed-Integer Nonlinear Programs (MINLPs) constitute an important class of optimization problems. Such problems are difficult to solve by current state-of-art algorithms and solvers, and are known to be in the complexity class NP-Hard. This difficulty arises mainly because (1) some of the functions in the description of the problem are nonconvex, and (2) some or all the variables are required to assume integer values. The commonly used algorithms for solving MINLP are the Branch-and-Cut Algorithm and its variants.

One crucial step in branch-and-cut algorithms is to find a tractable relaxation of the problem. Solving a relaxation of a minimization problem provides lower bounds on the optimal value. This relaxation, which is usually a convex optimization problem, is then either subdivided (by branching) or further tightened (by adding more inequalities or cuts) to obtain tighter bounds. A starting tighter relaxation to the feasible region of the MINLP gives a lower bound closer to the optimal value of the MINLP and leads to faster algorithms. Finding the tightest convex relaxation for the problem is often quite difficult, and one has to resort to tight convex relaxations of individual constraints separately. In this thesis, we derive the convex hull description of the mixed-integer bilinear covering set and its variants. These nonconvex sets appear in models for trim-loss problem, pattern minimization problem, pre-pack problem etc. We also study the effectiveness of the facet defining inequalities of the convex hull in theoretical and computational sense.

In our first study, we derive the closed form description of the convex hull of mixed-integer bilinear covering set with upper bounds on the integer variables. This convex hull description is determined by considering some orthogonal disjunctive sets defined in a certain way. This description does not introduce any new variables, but consists of exponentially many inequalities. We then present an extended formulation with a few extra variables and much smaller number of constraints. We also derive a linear time separation algorithm for the facet defining inequalities of this convex hull. We study the effectiveness of the new inequalities and the extended formulation computationally using some examples.

Next, we derive the closed form description of the convex hull of mixed-integer bilinear covering set with box constrained integer variables. We describe the extreme points and extreme rays first and then find the facet defining inequalities of the convex hull with their help. We then apply the facet defining inequalities to solve Pattern Minimization Problem in a novel branch-and-cut type of algorithm. Unlike currently available algorithms, this approach does not require column-generation or pricing. We also provide computational experiments to show the effectiveness of our algorithm.

Lastly, we consider the mixed-integer bilinear covering set without any bounds on the variables, and study the facet defining inequalities of its convex hull. In particular, we study these inequalities as split cuts and more generally as disjunctive cuts. Our motivation behind this approach is to find those inequalities which are computationally more useful and easier to obtain. Viewing these facet defining inequalities through the lens of split disjunctions, we see that some of them have split-rank one, and can be obtained easily. We derive the necessary and sufficient condition on the linear objective functions for which the facet defining inequalities with split-rank one are sufficient to give the same optimal value as the convex hull. A particular relaxation of the trim loss problem has this property. When the relaxation is slightly different and the necessary conditions do not hold, our computational experiments showed that these rank-one inequalities still give the same bound as all the facet defining inequalities of the convex hull. We further identify facet defining inequalities that have split-rank more than one, but can be obtained using other disjunctions.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

We study a mathematical set known as the bilinear covering set and its two variants. The set is defined formally as

$$S = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r \right\},$$

and its two variants as

$$S^U = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r, x \leq u \right\}, \text{ and}$$
$$S^B = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r, l \leq x \leq u \right\},$$

where $n \in \mathbb{N}, r > 0, l \in \mathbb{Z}_+^n, u \in \mathbb{N}_+^n$ are given. The inequality constraint $\sum_{i=1}^{n} x_i y_i \geq r$ is known as bilinear covering constraint. These sets appear in modeling optimization problems like the Trim-Loss Problem (or the Cutting Stock Problem) [66], Pattern Minimization Problem [116] and the Pre-pack Problem [69]. Minimizing an objective function over the set S, S^U or S^B is a special case of Mixed-Integer Nonlinear Optimization.

Mixed-Integer Nonlinear Optimization problems, including those with bilinear covering sets are known to be computationally difficult to solve; even small sized problems may take hours or more time on a computer. As we describe later, these problems are not 'convex', which makes these problems difficult. This thesis is devoted to analyzing the structure of the above three sets, finding their convex approximations or relaxations, and deriving properties that make it easier to optimize over them.

We start this introductory chapter by describing optimization problems including the specific problems of our interest: the trim-loss and the pattern minimization problems. We then describe the commonly used terms and concepts like convexity that are used in the thesis. Next we describe algorithms for solving these problems highlighting the difficulties these algorithms face. Finally, we close this chapter with an outline of the rest of the thesis and a summary of our contributions.

1.1 Mixed-Integer Nonlinear Optimization and Related Problems

A mathematical optimization problem is one where we want to find a point or a vector where a given objective function assumes the minimum (or maximum) value amongst all those that satisfy certain given constraints. An optimization problem can be classified depending upon the type of the objective function and constraints it has. One such broad class is that of Mixed-Integer Nonlinear Optimization Problems. We refer to a problem of this class as an MINLP (Mixed-Integer Nonlinear Program).

An MINLP consists of one or more nonlinear constraints and objective function. In addition, some or all of the variables (or unknowns) are required to assume integer values in the solution. Mathematically, an MINLP can be written as

$$\begin{aligned} &\min_{x,y}\ f(x,y) \\ \text{subject to: }\ & g_i(x,y) \le 0,\ \text{ for } i \in I, \\ & x \in \mathbb{Z}_+^{n_1}, y \in \mathbb{R}_+^{n_2}, \end{aligned} \tag{P}$$

where n_1, n_2 are given positive integers and I is a given index set. The functions $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ and $g_i : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ for $i \in I$ are also given. In the MINLP (P), x and y are called decision variables or just variables. We call x integer variables and y continuous variables. The function f is the objective function and $g_i(x,y) \le 0$, for all $i \in I$ are the constraints.

MINLPs arise in a wide variety of real life problems in different sectors in industry, business, sports, medicine, etc. In particular, MINLPs appear in telecommunication network design [16], biological problems like molecular biology [81] and protein folding [93], electricity transmission [89], gas network design [84], industrial problems like cutting stock (or trim loss) problem [66, 110], pre-pack problem [69], supply chain design and inventory management [117], facility layout problem [115], chemical process design [54, 63, 72] and more.

The set of points satisfying all constraints of an optimization problem like MINLP (P)

$$S_P = \left\{(x,y) \in \mathbb{Z}_+^{n_1} \times \mathbb{R}_+^{n_2} : g_i(x,y) \leq 0,\ \text{for } i \in I\right\} \tag{S_P}$$

is known as the feasible region of the MINLP (P). The problem (P) is infeasible when S_P is empty. MINLPs can be viewed as generalization of other classes of optimization problems, and the solution methods for MINLP depend upon the methods for these classes. Some of the important classes are described next.

1.1.1 Nonlinear Optimization

When the set I is empty and $n_1 = 0$, the problem (P) is called a nonlinear unconstrained optimization problem. When $n_1 = 0$, but I is not empty, we get a general nonlinear optimization problem (NLP) of the form

$$\begin{aligned} \min_{y}\ & f(y) \\ \text{s.t. } & g_i(y) \leq 0,\ \text{for } i \in I, \\ & y \in \mathbb{R}_+^n, \end{aligned} \tag{NLP}$$

with feasible region $S_{NLP} = \{y \in \mathbb{R}_+^n : g_i(y) \leq 0,\ \text{for } i \in I\}$. In the context of NLPs, there are two notions of optimal points.

Definition 1.1.1 (Local Minimum). *A point $\bar{y} \in S_{NLP}$ is called a local minimum if there exists an $\epsilon > 0$ such that $f(\bar{y}) \leq f(y)$ for all $y \in S_{NLP} \cap \{y \in \mathbb{R}^{n_2} : ||y - \bar{y}|| < \epsilon\}$.*

Definition 1.1.2 (Global Minimum). *A point $\bar{y} \in S_{NLP}$ is called a global minimum if*

$f(\bar{y}) \leq f(y)$ for all $y \in S_{NLP}$.

Since our goal is to find a globally minimum solution to the problems, we use minimum and global minimum interchangeably in the thesis. Clearly, every global minimum is also a local minimum, but the converse is true only for some special cases.

1.1.2 Convex Nonlinear Optimization

The notion of convexity plays an important role in determining whether any local minimum of a given NLP is also its global minimum. We introduce this notion through some definitions next.

Definition 1.1.3 (Convex Set)**.** *A set $C \subseteq \mathbb{R}^n$ is called a convex set if for any $\lambda \in [0,1], \lambda x^1 + (1-\lambda)x^2 \in C$, for all $x^1, x^2 \in C$.*

Definition 1.1.4 (Convex Combination)**.** *Let $T = \{x^1, x^2, \ldots, x^k\}$ be a finite subset in $\mathbb{R}^n$. The element $\sum_{i=1}^{k} \lambda_k x^k$ is called a convex combination of the elements in T if $\lambda_i \in [0,1], i = 1, \ldots, k$ and $\sum_{i=1}^{k} \lambda_k = 1$.*

Definition 1.1.5 (Convex Hull)**.** *Let T be a subset of $\mathbb{R}^n$ (possibly uncountable). The collection of all possible convex combinations of points in T is known as the convex hull of T, and we denote it as $conv(T)$.*

It can be easily proved that $conv(T)$ is the smallest convex set that contains the set T. In other words, $conv(T)$ is the intersection of all convex sets containing T. Related to the convex hull is the notion of a conic hull.

Definition 1.1.6 (Conic Combination)**.** *Let $T = \{x^1, x^2, \ldots, x^k\}$ be a finite subset in $\mathbb{R}^n$. The element $\sum_{i=1}^{k} \lambda_k x^k$ is called a conic combination of the elements in T if $\lambda_i \geq 0, i = 1, \ldots, k$.*

Definition 1.1.7 (Conic Hull)**.** *Let T be a subset of $\mathbb{R}^n$ (possibly uncountable). The collection of all possible conic combinations of the points in T is known as the conic hull of T, and we denote it as $\mathcal{C}(T)$.*

Definition 1.1.8 (Convex Function)**.** *Let C be a nonempty convex set in $\mathbb{R}^n$. A function $f : C \to \mathbb{R}$ is called a convex function if $f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2)$, for all $x^1, x^2 \in C$ and for any $\lambda \in [0,1]$.*

A function f is called a concave function if $-f$ is a convex function. Only affine functions are both convex and concave. A well known and important result for nonlinear programs is the following.

Theorem 1.1.1 ([99]). *Let $f : C \to \mathbb{R}$ be a convex function where C is a convex set in $\mathbb{R}^n$. Let $\bar{w} \in C$ be a local minimum of f over C, then $\bar{w}$ is also a global minimum of f.*

Proposition 1.1.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $\delta \in \mathbb{R}$. Then the set $S_\delta = \{x \in \mathbb{R}^n : f(x) \leq \delta\}$ is convex.*

The set S_δ is known as δ-Sublevel set of f. Thus, any sublevel set of a convex function (or a superlevel set of a concave function) is a convex set. If for a given NLP $g_i, i \in I$ are all convex functions, the feasible region S_{NLP} is also convex. If such an NLP additionally has a convex objective function f, it is called a convex nonlinear program (convex NLP).

1.1.3 Mixed-Integer Linear Optimization

Another important subclass of MINLPs is Mixed-Integer Linear Optimization. If the functions $f, g_i, i \in I$ with $I = \{1, \dots, m\}$ are all linear (or affine) in the variables, and $n_1 \geq 1$, then this optimization problem is known as a Mixed-Integer Linear Program (MILP). An MILP can be written in the following form.

$$
\begin{aligned}
\min_{x,y} \ & c^T x + d^T y \\
\text{s.t.} \ & Ax + By \leq b, \\
& x \in \mathbb{Z}_+^{n_1}, y \in \mathbb{R}_+^{n_2},
\end{aligned}
\tag{MILP}
$$

where A is an $m \times n_1$ rational matrix, B is an $m \times n_2$ rational matrix, $c \in \mathbb{R}^{n_1}, d \in \mathbb{R}^{n_2}$ and $b \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ are rational vectors. When $n_2 = 0$, i.e., there are no continuous variables in (MILP) then we call it Integer Program (IP). The feasible region of MILP or IP is, in general, not convex.

1.1.4 Linear Optimization

In the problem (MILP), if $n_1 = 0$, i.e., there are no integer constrained variables, then it is called a Linear Program (LP). Mathematically, an LP is of the form

$$\begin{aligned} \min_{y} \ & d^T y \\ \text{s.t.} \ & By \le b, \\ & y \in \mathbb{R}^n_+. \end{aligned} \tag{LP}$$

Since all linear functions are convex, LPs are also a special and an important case of convex NLPs.

1.2 Algorithms for Optimization Problems

An algorithm is a step-by-step recipe or procedure for solving a given problem. The design of algorithms for solving optimization problems is largely dependent on the problem class. MINLPs are known to be theoretically unsolvable [71], that is, there can not be any algorithm which can solve MINLPs in finite time. The result is true even when we restrict all functions in the MINLP (P) to the quadratic form. When the constraints of the MINLP include explicit bounds on all variables, then there are algorithms that can obtain an $\epsilon-$optimal solution, that is, a point which is at most ϵ distance away from the actual optimal solution. Even for such bounded MINLPs, there are no fast or efficient algorithms. Since MINLPs include LPs, MILPs and NLPs as special cases, they are at least as hard as solving these special cases. MILPs, in particular belong to the complexity class NP-Hard [75]. That is, the time for solving them using any known algorithm increases exponentially as the input size increases.

Methods for solving MINLPs are based on those used for solving LPs, MILPs and NLPs. We briefly describe the main concepts and algorithms for solving these problems first. We start by defining a few key terms. Readers are referred to books on convex optimization (e.g. [28, 99]) for more details.

Definition 1.2.1 (Extreme Point)**.** *Let C be a convex set in $\mathbb{R}^n$. A point $x \in C$ is called* *an extreme point of C if it can not be expressed as a convex combination of any other two points in C.*

Definition 1.2.2 (Recession Direction and Recession Cone)**.** *Let C be a convex set in $\mathbb{R}^n$. A vector d is called a recession direction if $x + \alpha d \in C$ for all $x \in C$ and for all $\alpha \geq 0$. The set of all recession directions of C is called the recession cone of C and in this thesis we denote it as $0^+(C)$.*

Definition 1.2.3 (Half Space)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a linear function, i.e., $f(x) = c^T x, x \in \mathbb{R}^n$ for a given $c \in \mathbb{R}^n$. Also let $h \in \mathbb{R}$. Then the set $H = \left\{x \in \mathbb{R}^n_+ : c^T x \leq h\right\}$ is known as a half space.*

Definition 1.2.4 (Polyhedron)**.** *A set Q in $\mathbb{R}^n$ is called a polyhedron (or a polyhedral set) if Q can be expressed as an intersection of finite number of half spaces.*

Any polyhedron Q can be written in the form

$$Q = \left\{x \in \mathbb{R}^n_+ : Ax \leq b\right\},$$

where A is an $m \times n$ real matrix, $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0\}$ and $b \in \mathbb{R}^m$. Thus, Q is an intersection of $m + n$ half spaces (m inequalities in $Ax \leq b$ and n inequalities in $x \geq 0$). A bounded polyhedron is known as a polytope. When a polyhedron in described by specifying the half-spaces, the description (like the form above for Q) is called an H-Description. On the other hand, any polyhedron with at least one extreme point can be uniquely described by the sum of its recession cone and the convex hull of its extreme points. A description which specifies the extreme rays of the recession cone and the extreme points of the polyhedron is known as the V-Description of a polyhedron. Consider the example

$$Q_E = \{(x_1, x_2) \in \mathbb{R}^2_+ : x - 2y \leq 2, -2x + y \leq 2, x + y \geq 1\},$$

for illustration. The set of extreme points $Ext(Q_E)$ is $\{(2, 0), (1, 0), (0, 1), (0, 2)\}$, and the recession cone:

$$0^+(Q_E) = \{(x_1, x_2) \in \mathbb{R}^2_+ : x - 2y \leq 0, -2x + y \leq 0, x + y \geq 0\}.$$

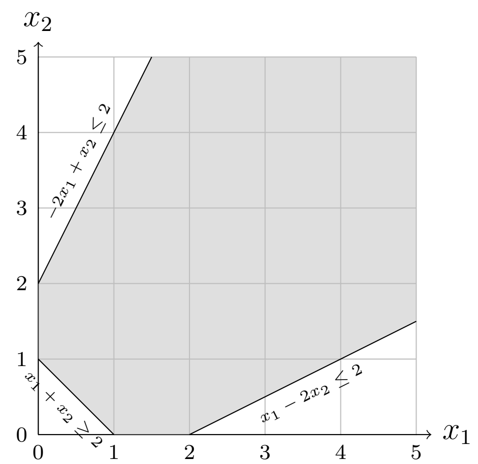

Figure 1.1: The H-Description

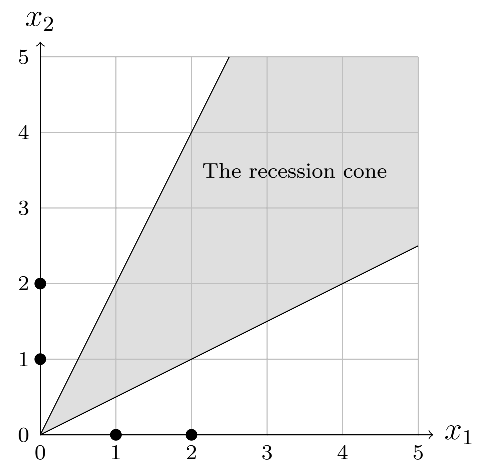

Figure 1.2: The V-Description

Therefore, Q_E can be wrtten as:

$$Q_E = conv\left(Ext(Q_E)\right) + 0^+\left(Q_E\right).$$

Figures 1.1 and 1.2 illustrate it graphically. We now describe algorithms for solving optimization problems, starting with linear optimization.

1.2.1 Linear Optimization

The feasible region of a linear program is a polyhedron. The following property of a linear programs makes it easier to solve than other convex programs.

Theorem 1.2.1 ([22]). *Suppose the problem (LP) has an optimal solution. Then one of the extreme points of the feasible region of (LP) must be an optimal solution.*

Further, every linear program has a finite number of extreme points. Dantzig [42] exploited these properties to develop the simplex method for solving a linear program. Since then the method and theory have been refined and improved in several ways. This method iteratively jumps from one extreme point of the polyhedron to another in search of an optimal solution. The topic of linear programming is rich in theory, and is an important bridge between continuous and integer optimization problems. As we shall see later, many other classes of optimization problems are solved by iteratively solving many linear programs. More details on linear programming and its applications can be found in books like those by Chvátal [34] and Bertsimas and Tsitsiklis [22].

In theory, the simplex algorithm is not efficient, i.e., it does not run in time polynomial in size of the input. For some cleverly designed instances this algorithm takes exponentially many steps. Khachiyan [76] developed the first polynomial time algorithm for linear programs. Though this algorithm does not perform well practically, it has some useful theoretical properties. Karmarkar developed an interior point algorithm [74] that runs in polynomial time and performs well computationally.

Several modern LP solvers, such as CLP [56], CPLEX [3], Gurobi [70] and XPRESS [1] can solve large linear programs fast. They provide implementations of both, the simplex algorithm and the interior point algorithm. They also provide useful routines to re-optimize an LP if some of the inputs change or new constraints and variables are added. Good theoretical properties of a linear programs, fast algorithms and availability of robust solvers make it an useful component of algorithms for other classes of optimization, most notably for integer optimization problems.

1.2.2 Convex Nonlinear Optimization

We need to find only a locally optimal solution of a convex NLP because such a solution is also globally optimal. There are several algorithms available for solving these problems. For more details of such algorithms, see Nesterov and Nemirovskii [92] and Boyd and Vandenberghe [29]. Two celebrated algorithms to solve convex programs are the ellipsoidal method and the interior point method.

The ellipsoidal method was developed by Yudin and Nemirovski [118]. The basic idea behind this method is to generate a sequence of ellipsoids of decreasing volumes containing an optimal solution. At each iteration it divides the current ellipsoid into halves and determines the half that contains an optimal solution and discards the other half. In the next iteration, a new ellipsoid is constructed having smaller volume covering the selected half of the ellipsoid of the previous iteration, and the process is continued. Though this algorithm runs in polynomial time in the size of the input, implementing this algorithm for practical computational purpose is not that easy.

Nesterov and Nemirovskii [92] extended the interior point method of Karmarkar [74] for solving linear programs to more general convex nonlinear programs, and showed that polynomial time convergence can be achieved for convex programs also using a barrier function. An interested reader is referred to the book [92] for details. Other practical methods for solving convex NLPs are based on Sequential Quadratic Programming [51,95] and Augmented Lagrangians [23]. Solvers for NLPs include ALGENCAN [24], FilterSQP [53] etc, IPOPT [111], Knitro [32], and several others.

An important subclass of nonlinear convex optimization is that of Conic Optimization. This class includes second-order conic programs (SOCP), semidefinite programs (SDP), and even linear programs. Readers are referred to [20] and [21] for more details. Solvers like MOSEK [9], SeDuMi [105], CSDP [27] are available for such problems. While these problems are easier to solve as compared to general convex NLPs, the solvers are not as fast and robust as LPs.

1.2.3 Mixed-Integer Linear Optimization

The additional restriction of integrality on certain variables makes MILP more difficult to solve than an LP. In fact, the problem (MILP) is an NP-Hard problem [37], and all known algorithms take time exponential in the size of the input. When the number of variables is fixed, MILPs can be solved in theoretically polynomial time using Lenstra's [79] algorithm.

The most commonly used algorithm to solve an MILP is the Branch-and-Cut Algorithm. A linear program obtained by removing the integer restrictions on the variables is first solved. This LP is a relaxation of the MILP as every feasible point to the MILP is also feasible to the LP. If the LP relaxation is infeasible, then the MILP is also infeasible. If we find an optimal solution to LP in which the integer variables assume integer values, then it is also an optimal solution to the MILP. Otherwise, the LP optimal solution has a fractional (non-integer) value on at least one integer variable. The branch-and-cut algorithm proceeds further in one of the two broad ways:

Branch and Bound

Let $(\bar{x}, \bar{y})$ denote the solution of the LP relaxation. Then one can divide the problem into two or more subproblems in a way that $(\bar{x}, \bar{y})$ is not feasible to any of the subproblems, but each solution to the MILP is feasible in at least one subproblem. This procedure is called branching. Suppose $\bar{x}_i$ is a fractional (non-integer) value for some $i \in \{1, \ldots, n_1\}$. Then one possible branching is to have the additional constraint $x_i \leq \lfloor \bar{x}_i \rfloor$ in one branch and $x_i \geq \lceil \bar{x}_i \rceil$ in the second.

The subproblems created after branching are restrictions of the original MILP, and they are also MILPs. The above procedure can be continued applied again to these subproblems. The process is continued until all the subproblems are explored. For each subproblem, one needs to solve its LP relaxation. If the LP relaxation is infeasible, the subproblem need not to be explored further. Similarly, if the LP relaxation of a subproblem yields a solution that satisfies integer restrictions, it is a feasible solution to the original MILP, and hence provides an upperbound to the optimal value. Such a subproblem is also not explored further as we can not find a better solution by exploring it further. A subproblem whose LP relaxation gives a value equal or higher than the best known upperbound is also discarded as there are no better solutions to be found there. This bounding procedure is often quite useful in eliminating many subproblems.

The above procedure is called Branch-and-Bound. Introduced by Land and Doig [78] for pure integer programming problems, it was later extended by Driebeek [47] and Dakin [41] for mixed-integer linear programs. The progress of this algorithm on a particular instance of MILP can be denoted using a Branch-and-Bound tree. As the name suggests, it is a tree graph where each node represents a subproblem. The root of the tree denotes the original problem. Any children of a node correspond to the subproblems created by branching on the problem associated with the node. Let us consider the following pure integer problem as an example.

$$\begin{gathered} \min\; -8x_1 - 11x_2 - 6x_3 - 4x_4 \\ \text{s.t. } 5x_1 + 7x_2 + 4x_3 + 3x_4 \leq 14, \\ x_i \in \{0, 1\}, i = 1, 2, 3, 4. \end{gathered}$$

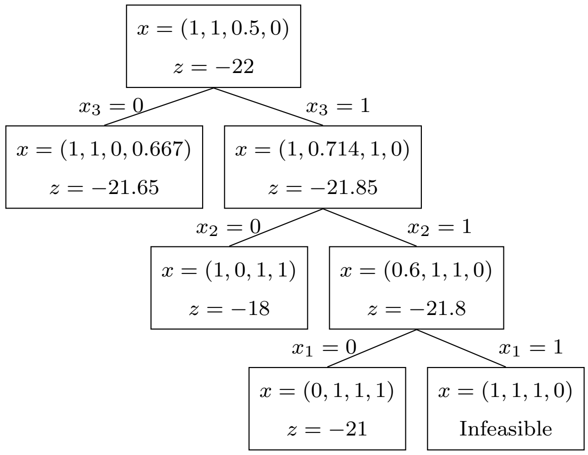

Figure 1.3: A Branch-and-Bound Tree

- The linear relaxation solution is $x = (1, 1, 0.5, 0)$ with a value of -22. The solution is not integral.
- The value $x_3 = 0.5$ is fractional. Choose x_3 to branch. The two new subproblems will have the additional constraints $x_3 = 0$ and $x_3 = 1$ respectively. This process can then be continued over each subproblem.
- Figure 1.3 illustrates the rest of the branch-and-bound algorithm on this instance.

While branch-and-bound still remains the main underlying method in solving MILPs today, it has seen several augmentations, most notably in the form of generating cuts.

Cutting Plane Algorithm

Given a non-integer solution $(\bar{x}, \bar{y})$ of the LP relaxation of a given MILP, if we can add a linear constraint that is violated by $(\bar{x}, \bar{y})$, but satisfied by all feasible points to MILP, then we do not need to branch. Such an inequality that is satisfied by all feasible points of MILP is called a 'valid inequality' or a 'cutting plane'. See Figure 1.4 for an illustration. The LP obtained by adding such a cut is still a relaxation to the MILP, but tighter than the original LP relaxation. Gomory [61] developed a Cutting Plane algorithm to solve integer programs (without continuous variables) that progressively added cuts to an LP. The method stops when the LP is found to be infeasible or when it gives an integer solution. Several classes of cuts have been introduced over the years and many theoretical

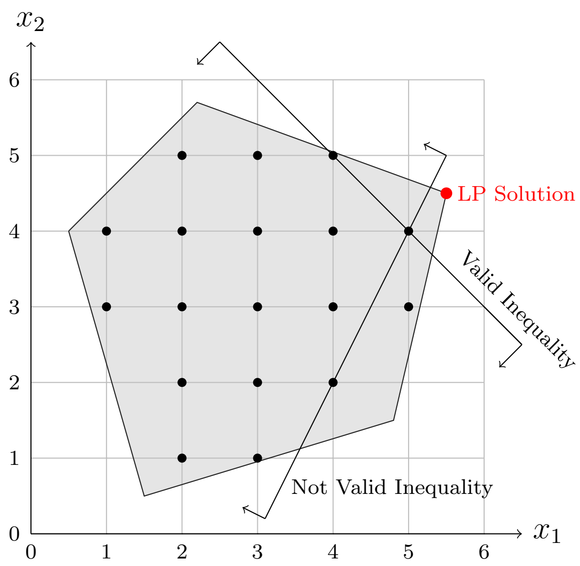

Figure 1.4: Cutting Plane

insights on their properties gained [39]. We will revisit this topic in more detail later in this chapter.

Most modern solvers for MILP combine branch-and-bound with cutting planes to get what is called a Branch-and-Cut algorithm [90, 113]. These solvers also implement advanced presolving techniques for automatically simplifying problems, heuristic methods for finding good upper bounds, and intelligent search routines to speed up the algorithm. CBC [55], CPLEX [3], Gurobi [70], SCIP [114], XPRESS [1] are some of the solvers available for MILP today.

1.2.4 Mixed-Integer Nonlinear Optimization

Solving MINLPs is difficult mainly because of two reasons: (1) some of the functions in the description of (P) are nonconvex (defined in the next section), and (2) some variables are required to assume integer values.

When the objective function f and all the constraint functions $g_i, i \in I$ are convex in MINLP (P), we call it a Convex MINLP. Methods for solving Convex MINLPs are quite similar to those of MILP because, like MILP, one can get a convex relaxation by removing integer restrictions on the variables. The continuous relaxation obtained can be solved relatively easily using algorithms for convex NLPs. Thus a branch-and-bound method based on convex NLPs can be used to solve convex MINLPs. Cutting planes for convex MINLPs can also be derived by first creating LP relaxations of the convex feasible region and then applying the machinery developed for MILP cutting planes. Developing procedures that exploit convexity to generate tighter relaxations that also are practically faster to solve have received considerable attention [48,52,97]. Solvers for convex MINLP include Bonmin [26], FilMINT [4] and MINOTAUR [83].

When some of the functions in the MINLP (P) are not convex, then the problems become even harder to solve. Relaxing integer constraints in such problems creates a nonconvex nonlinear problem which is difficult to solve. In words of Rockafellar, *"In fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."* [100].

In order to use a branch-and-bound based procedure, we require a tractable relaxation to solve. Finding such a relaxation for general nonlinear problems is difficult. Instead, the problem is first reformulated by adding several new variables and constraints so that all nonconvex relations are restricted to atomic functions (like product of two variables, or log of a variable). These constraints are then relaxed individually to obtain a convex relaxation [19, 50, 85, 101, 107].

The solution of the relaxation obtained for a nonconvex MINLP may not be feasible to the continuous relaxation, but it still provides a lower bound. In order to proceed with a branch-and-bound algorithm, one may have to branch on continuous variables in addition to branching on integer constraints. Hence, such a method is also sometimes called Spatial Branch-and-Bound. For more detailed on the existing solution methods for MINLPs, an interested reader may refer to surveys by Belotti et al. [18] and Neumaier [94], and books by Horst and Pardalos [67] and Horst and Tuy [68]. Solvers implementing Branch-and-Cut algorithms for nonconvex MINLP include BARON [108], Couenne [19], LINDO [2] and SCIP [114].

The quality or tightness of the relaxation of a given MINLP determines the quality of the lower bound on the problem and also the number of subproblems that are required to solve it. We next discuss relaxation techniques.

1.3 Relaxation Techniques

A relaxation that is not too large compared to the feasible region is important for developing fast algorithms for MINLP. Ideally, we would like the description of the convex hull of the feasible region of the MINLP. However, that is quite hard, in fact, as hard as finding the optimal solution. A more practical approach is to find a relaxation, which may be weaker but is found relatively easily. We describe some of the main relaxation techniques with some examples in this chapter. We start with some defninitions.

Definition 1.3.1 (Underestimator). *Let D be a subset of $\mathbb{R}^n$ and $f, \underline{f} : D \to \mathbb{R}$ be two functions. We say $\underline{f}$ is an underestimator of f over D if $\underline{f}(x) \leq f(x)$ for all $x \in D$.*

Now, consider the problem (P) and define the following related optimization problem.

$$\begin{aligned} z_R = \min_{x,y} \ & \underline{f}(x, y) \\ \text{s.t.} \ & (x, y) \in R. \end{aligned} \tag{RP}$$

Definition 1.3.2 (Relaxation). *The optimization problem (RP) defined above is called a relaxation of the problem (P) if the following two conditions hold.*

1. *$\underline{f}$ is an underestimator of f over the feasible region S_P, and*
2. *S_P is a subset of R.*

From the above definition we can see that the optimization problem (P) can be relaxed in either or both of the following ways: (1) replacing the objective function f with some of its underestimators over S_P, (2) replacing the feasible region S_P with some superset of S_P. From the definition of relaxation, it can be easily seen that z_R is a lower bound on z^*.

Proposition 1.3.1. *Let $(\bar{x}, \bar{y}) \in R$ be an optimal solution for (RP). Then $(\bar{x}, \bar{y})$ is an optimal solution for (P) if $(\bar{x}, \bar{y}) \in S_P$ and $f(\bar{x}, \bar{y}) = \underline{f}(\bar{x}, \bar{y})$. Moreover, we have the optimal value $z^* = z_R = f(\bar{x}, \bar{y})$.*

The relaxation (RP) of the problem (P) is useful when the problem (RP) is easier to solve and the optimal value z_R of (RP) is close to the optimal value z^* of (P). Note that when the underestimating function $\underline{f}$ is convex and the feasible region R of (RP) is a convex set, then the problem (RP) becomes a convex program and can be solved efficiently. This leads to the next definition.

Definition 1.3.3 (Convex Relaxation)**.** *The optimization problem (RP) defined above is called a convex relaxation of the problem (P) if the following two conditions hold.*

1. *$\underline{f}$ is a convex underestimator of f over the feasible region S_P, and*

2. *S_P is a subset of R, and R is the feasible region of a convex NLP.*

For a given nonconvex problem, there are uncountably many convex relaxations. If $\underline{f}$ is a convex underestimator of the function f, then $\underline{f} - \delta$ for any $\delta > 0$ is also a convex underestimator of the objective function. Moreover, there are uncountably many convex supersets of the feasible region. Our goal is to find the highest such underestimator and the smallest such superset.

Definition 1.3.4 (Convex Envelope)**.** *Let $f : D \to \mathbb{R}$ be a function (possibly nonconvex) where $D \subset \mathbb{R}^n$ is a convex and compact set. A convex function $\underline{f}_{CE}$ is called convex envelope of f over D if it satisfies the following conditions.*

1. *The function $\underline{f}_{CE}$ is an underestimator of f over D, and*

2. *If h is another convex underestimator of f over D then $\underline{f}_{CE}(x) \geq h(x)$ for all $x \in D$.*

The above definition implies that $\underline{f}_{CE}$ is the tightest possible convex underestimator of f and $\underline{f}_{CE}$ is unique in a given domain.

Recall that the tightest possible convex relaxation of a given set is called its convex hull. Thus the tightest convex NLP relaxation of a MINLP is in the same space

$$
\begin{aligned}
z_{CR} = \min_{x,y} \ & \underline{f}_{CE}(x, y) \\
\text{s.t.} \ & (x, y) \in conv\,(S_P)\,,
\end{aligned}
\tag{CRP}
$$

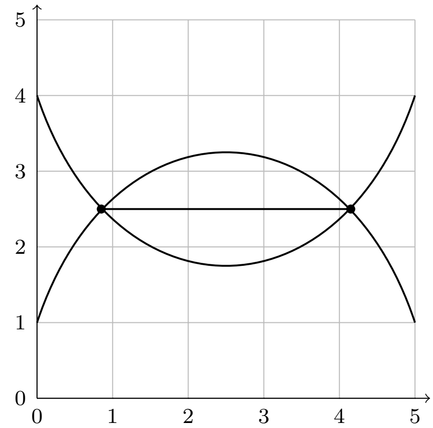

Figure 1.5: The convex hull of the intersection of constraints

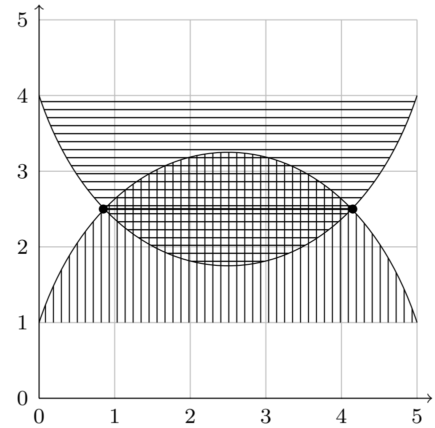

Figure 1.6: Intersection of the convex hulls

where $conv\,(S_P)$ is the convex hull of S_P and $\underline{f}_{CE}$ is the convex envelope of the objective function of (P) over $conv\,(S_P)$. In practice, obtaining convex hull of a nonconvex set is difficult. One way to mitigate this difficulty is to relax each constraint separately, i.e., find the convex relaxation

$$\begin{aligned} &\min_{x,y}\ \underline{f}(x,y) \\ &\text{s.t.}\ (x,y) \in conv\,(G_i)\,,\ i \in I, \end{aligned} \tag{1.1}$$

where $G_i = \{(x,y) \in \mathbb{Z}_+^{n_1} \times \mathbb{R}_+^{n_2} : g_i(x,y) \leq 0\}$ denotes the feasible region of constraint i. The feasible region of problem (1.1) is a convex set which is usually larger than $conv(S_P)$. The difference between the convex hull of the whole problem and the intersection of convex hull of feasible region of individual constraints is illustrated in Figures 1.5 and 1.6. The illustrated problem consists of two equality constraints denoting the curves, and the feasible region is only two discrete points. Taking the intersection of convex hulls of the two constraints yields the hatched area.

Finding convex hull of a single constraint may also be a hard problem. Therefore, simpler convex relaxations are typically derived to generate lower bounds. The methods for creating such relaxations are called convexification methods. Convexification techniques usually depend on the function defining a given set. Since our work is limited to bilinear constraints, we discuss such techniques for quadratic constraints next.

1.3.1 McCormick Relaxation of a Bilinear Constraint

Consider the following bilinear set in $\mathbb{R}^3$

$$B = \{(w, y_1, y_2) : w = y_1 y_2, l_1 \leq y_1 \leq u_1, l_2 \leq y_2 \leq u_2\}. \tag{1.2}$$

The set B is nonconvex. An example is shown in Figure 1.7. The convex hull of this set is described by the McCormick Relaxation [85] consisting of four linear constraints:

$$w \geq l_2 y_1 + l_1 y_2 - l_1 l_2, \tag{1.3}$$

$$w \geq u_2 y_1 + u_1 y_2 - u_1 u_2, \tag{1.4}$$

$$w \leq l_2 y_1 + u_1 y_2 - u_1 l_2, \tag{1.5}$$

$$w \leq l_1 y_2 + u_2 y_1 - l_1 u_2. \tag{1.6}$$

The McCormick Relaxation has some nice properties. It is linear, and hence can be solved using state-of-the-art LP solvers. Further, the relaxation becomes tighter when the difference between lower and upper bounds is reduced. In fact, if one of the variables y_1, y_2 assumes a value at one of the bounds, the McCormick constraints ensure that the original $w = y_1 y_2$, must be satisfied. As we shall see next, these relaxations can be applied to any general quadratic constraints. These properties make McCormick relaxations a popular choice in Branch-and-Cut algorithms for nonconvex MINLPs.

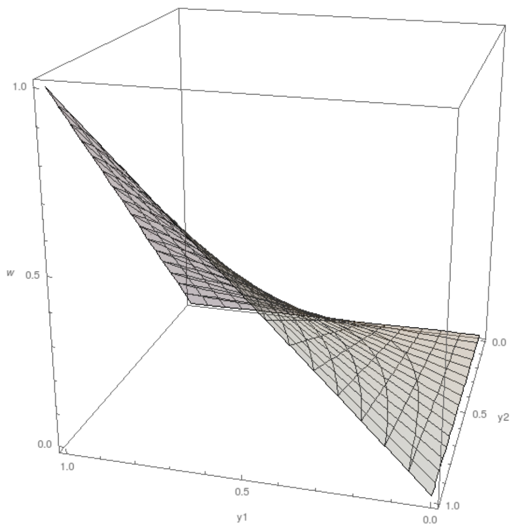

Figure 1.7: $w = y_1 y_2$ in the domain $y_1 \in [0,1], y_2 \in [0,1]$

1.3.2 Relaxations of Quadratically Constrained Sets

McCormick inequalities can be used to obtain a polyhedral relaxation of other bounded quadratically constrained sets as well. For example, let $B = \{(w,y) \in \mathbb{R}\times\mathbb{R} : w = y^2, 0 \leq y \leq 1\}$ be a nonconvex set. One can apply McCormick relaxation to the constraint $w = y^2$ to obtain the following polyhedral relaxation of B

$$
\begin{aligned}
&w \geq 0 \\
&w \geq 2y - 1, \\
&w \leq y, \\
&0 \leq y \leq 1.
\end{aligned}
$$

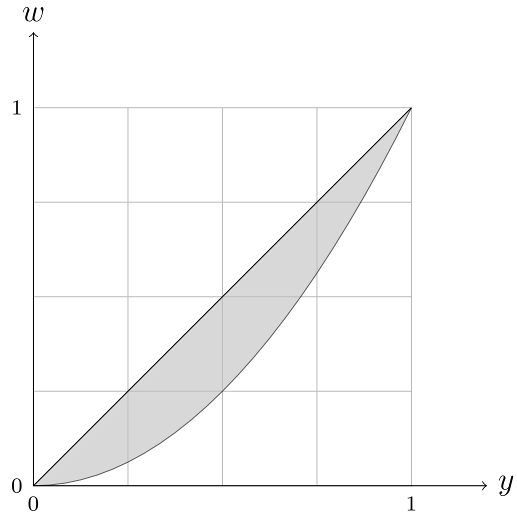

Figure 1.8: Convex hull of the set $\{(w, y) \in \mathbb{R} \times \mathbb{R} : w = y^2, 0 \leq y \leq 1\}$.

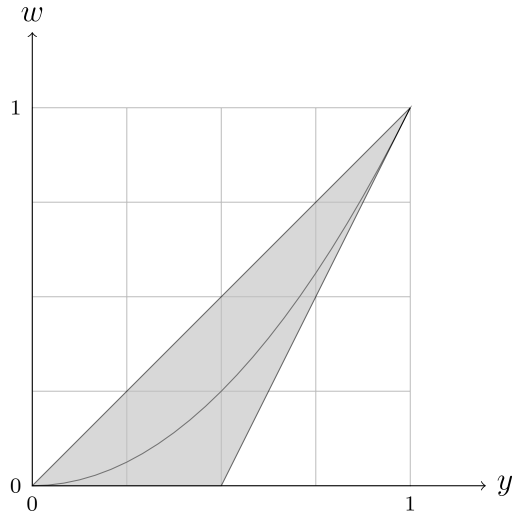

Figure 1.9: McCormick Relaxation of the set $\{(w, y) \in \mathbb{R} \times \mathbb{R} : w = y^2, 0 \leq y \leq 1\}$.

As seen from Figures 1.8 and 1.9, we do not get the convex hull from McCormick relaxation. McCormick relaxation can be applied to any quadratic constraint, for example, consider the general set

$$Q = \left\{ y \in \mathbb{R}^n : \sum_{i=1}^{n} \sum_{j=i}^{n} a_{ij} y_i y_j + \sum_{i=1}^{n} b_i y_i \leq h, l \leq y \leq u \right\}, \tag{1.7}$$

where a, b, h, l, u are given. First introduce new variables w_{ij} to model the relation $w_{ij} = y_i y_j$ and then apply McCormick relaxation on each term. The constraints of the relaxation obtained in this manner are:

$$w_{ij} \geq l_j y_i + l_i y_j - l_i l_j, \quad i = 1, \ldots, n, \;\; j = 1, \ldots, n, \tag{1.8}$$

$$w_{ij} \geq u_j y_i + u_i y_j - u_i u_j, \quad i = 1, \ldots, n, \;\; j = 1, \ldots, n, \tag{1.9}$$

$$w_{ij} \leq l_j y_i + u_i y_j - u_i l_j, \quad i = 1, \ldots, n, \;\; j = 1, \ldots, n, \tag{1.10}$$

$$w_{ij} \leq l_i y_j + u_j y_i - l_i u_i, \quad i = 1, \ldots, n, \;\; j = 1, \ldots, n, \tag{1.11}$$

$$\sum_{i,j} a_{ij} w_{ij} + \sum_{i} b_i y_i \leq h, \tag{1.12}$$

$$l \leq y \leq u, \tag{1.13}$$

Again, the above constraints do not yield the convex hull of Q, in fact the above relaxation is quite weak. Moreover, if bounds on variables are not available, one can not construct this relaxation. Some results on strengths of McCormick relaxations have be obtained by Luedtke et al. [82].

In order to obtain stronger relaxations, several other schemes have been proposed. Semidefinite Relaxation (SDP) [103], Lagrangian Relaxation based on duality [112], Relaxation Based on KKT Optimality Conditions [6,31], Completely Positive Matrix (CPP) based relaxations [30], Doubly Non-negative Matrix (DNN) based relaxations [7, 8] are some such examples. For more description of such relaxation techniques, one may refer to a review by Bao et al. [15].

1.3.3 Valid Inequalities

Relaxations of MILPs and MINLPs can be strengthened or tightened by adding valid inequalities.

Definition 1.3.5 (Valid Inequalities)**.** *Let C be a subset of $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. A linear inequality $c^T x + d^T y \geq b$ is called a valid inequality for C if $c^T \bar{x} + d^T \bar{y} \geq b$ for all $(\bar{x}, \bar{y}) \in C$.*

As discussed above, finding convex hull descriptions of feasible regions of an MILP or MINLP is difficult, and we have to resort to weaker relaxations. By adding inequalities valid for the feasible region of the MINLP, we can get tighter relaxations. Often these inequalities are generated algorithmically to cut off the current solution of the relaxation. Some types of valid inequalities can be generated for all problems of a general class. Examples include Chvátal-Gomory cuts [60] for pure IPs and Gomory Mixed-Integer cuts [62], lift and project cuts [14], and mixed-integer rounding cuts [91]. The above cuts can also be used for polyhedral relaxations of MINLP. Recently, new classes of cuts for quadratic constraints have been introduced and some classes of cuts for MILP, like disjunctive cuts [77, 104] and lift and project cuts [25] have been extended to MINLPs.

Valid inequalities can also be derived from specific types of constraints. If such constraints are present in a problem, these inequalities may be used. Examples include covering inequalities for knapsack constraints [11], comb inequalities for traveling salesperson problem [65] and perspective cuts [57] for convex MINLPs. Our thesis is focussed on deriving tight relaxations specific for the bilinear covering sets by exploiting geometric and algebraic properties of the constraints defining these sets. The valid inequalities we find are therefore structure-specific cuts. While they are not generally applicable to all MINLPs, they are strong and provide tight relaxations.

We also study the cuts for bilinear sets under the lens of general purpose cuts to see whether they can be obtained using these general methods. One of the most general classes of cuts is that of disjunctive cuts which includes split cuts. Many classes of valid inequalities for MILP including all the general classes listed above for MILP are known to be special cases of split cuts [39]. We describe disjunctive and split cuts next.

Let $P \subset \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$ be the feasible set of a convex MINLP of the form

$$\begin{aligned} \min_{x,y}\ & f(x,y) \\ \text{s.t.}\ & g_i(x,y) \leq 0, i = 1, \ldots, m, \\ & x \in \mathbb{Z}^{n_1}, y \in \mathbb{R}^{n_2}, \end{aligned} \tag{MINLP_{CV}}$$

where $f, g_i, i = 1, \ldots, m$ are convex functions. Let P_C be the continuous relaxation of P obtained by removing integer restrictions on the x variables. Assume the convex set P_C is closed.

Definition 1.3.6 (Disjunction [13])**.** *Let $D_k = \left\{(x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : A^k x \leq b^k\right\}$ for $k \in K$, where K is an index set (not necessarily finite). Define $D = \bigcup_{k \in K} D_k$. If $\mathbb{Z}^{n_1} \times \mathbb{R}^{n_2} \subseteq D$, then we call D a disjunction (or a valid disjunction) and each $D_k, k \in K$ is known as an atom of the disjunction D.*

Definition 1.3.7 (Disjunctive Cut)**.** *A linear inequality is called a disjunctive cut for P obtained from the disjunction $D = \cup_{i \in K} D_k$, if it is valid for $P_C \cap D_k$ for all $k \in K$.*

We say that a linear inequality is valid for the disjunction D if it is valid for $P_C \cap D_k$ for all $k \in K$. For some positive integer m, let us define the set notation:

$$\left[G^1 x \leq h_1, \ldots, G^m x \leq h_m\right] = \left\{(x,y) \in \mathbb{R}^{n_1+n_2} : G^1 x \leq h_1, \ldots, G^m x \leq h_m\right\},$$

where $G^i, i = 1, \ldots, m$ are rational matrices of suitable dimension and $h_i, i = 1, \ldots, m$ are rational numbers. Split cuts are a special class of the disjunctive cuts which are obtained from a split disjunction, a special type of disjunction with only two atoms.

Definition 1.3.8 (Split Disjunction)**.** *Given a non-zero integer vector* $\pi \in \mathbb{Z}^{n_1}$ *and an integer* π_0*, the disjunction* $\left[\pi^T x \leq \pi_0\right] \vee \left[\pi^T x \geq \pi_0 + 1\right]$ *is known as Split Disjunction. In a simpler way we write this disjunction as* (π, π_0)*.*

Note that any $(x, y) \in \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$ satisfies either $\pi^T x \leq \pi_0$ or $\pi^T x \geq \pi_0 + 1$. Without loss of generality we can assume $gcd(\pi, \pi_0) = 1$. Let us define the two sets as follows.

$$P_L = P_C \cap \left[\pi^T x \leq \pi_0\right], \text{ and } P_R = P_C \cap \left[\pi^T x \geq \pi_0 + 1\right].$$

Clearly $P \subset P_L \cup P_R$. Therefore, $P \subseteq P_L \cup P_R \subseteq conv\left(P_L \cup P_R\right)$.

Definition 1.3.9 (Split Cut)**.** *An inequality* $c^T x + d^T y \geq b$ *that is valid for both the sets* P_L *and* P_R *(or consequently valid for* $conv\left(P_L \cup P_R\right)$*) is known as a split cut.*

Let us consider a linear inequality $c^T x + d^T y \geq b$. In order to check whether the inequality $c^T x + d^T y \geq b$ is valid for P_R and P_L, one can solve the following two optimization problems

$$\begin{aligned} \zeta_R = \min_{x,y}\ & c^T x + d^T y \\ \text{s.t. } & (x, y) \in P_R \end{aligned} \qquad\qquad \begin{aligned} \zeta_L = \min_{x,y}\ & c^T x + d^T y \\ \text{s.t. } & (x, y) \in P_L. \end{aligned}$$

Clearly the inequality $c^T x + d^T y \geq b$ is valid for $conv\left(P_R \cap P_L\right)$ if and only if $\zeta_R \geq b$ and $\zeta_L \geq b$.

The subset of P_C obtained by adding all possible split cuts to P_C is known as the first split closure or the elementary split closure of P_C. Let us denote it by P_1. Clearly, P_1 is closed since P_C is closed. Similarly applying the split closure procedure to the set P_1 will give the second split closure P_2. Let P_t be the t^{th} split closure. Cook et al. [38] showed that, if P_C is a polyhedral set, then P_t is also polyhedral, for all $t \in \mathbb{N}$. For more results on this can be found in [40, 44, 45].

Definition 1.3.10 (Split Rank)**.** *For a valid inequality* $c^T x + d^T y \geq b$ *for the set* $conv(P)$*, the split rank of the inequality is defined as the smallest integer* t *such that the inequality is valid for* P_t *but not for* P_{t-1}*.*

Before describing the bilinear covering sets, we close this section by defining facet-defining inequalities, i.e., valid inequalities that are the tightest for a given polyhedral set.

Definition 1.3.11 (Affinely Independent Points). *The points $w^1, \ldots, w^n$ are called affinely independent if the $n-1$ points $w^2 - w^1, \ldots, w^n - w^1$ are linearly independent.*

Definition 1.3.12 (Facet Defining Inequality). *A vallid inequality $c^T x + d^T y \geq b$ is called a facet defining inequality for P, if there exist $n(= n_1 + n_2)$ affinely independent points in P that are active at the inequality, i.e., lie on the hyperplane $c^T x + d^T y = b$.*

If all facet defining inequalities of the convex hull of a set are known, then we can optimize over these inequalities to find the minimum value of a linear function over this set. Sometimes the number of facet-defining inequalities for a set may be prohibitively large, as is the case for the bilinear covering sets we consider. Algorithms for selecting only those inequalities that may be useful (or are violated by a given point) are required for such cases.

1.4 Trim-Loss and Pattern Minimization Problems

Trim-loss and pattern minimization problems are encountered in industry where smaller pieces of a certain material need to be cut from large sheets. In a trim-loss (or a cutting-stock) problem, we want to determine the best way to cut large rolls of raw materials into smaller pieces (or finals) using different patterns, so that the demand of finals is met and as few rolls as possible are used. Let $N = \{1, \ldots, n\}$ be the index set that denotes the cutting patterns used, and F be the index set of different sizes of the finals that are to be cut. Let L be the size of each large roll and $l_j, j \in F$ be the lengths of the finals. The demands of the finals, say $d_j, j \in F$ are known. Let x_{ij} be the number of final j cut according in the pattern $i, i \in N, j \in F$, and y_i be the number of rolls cut with cutting

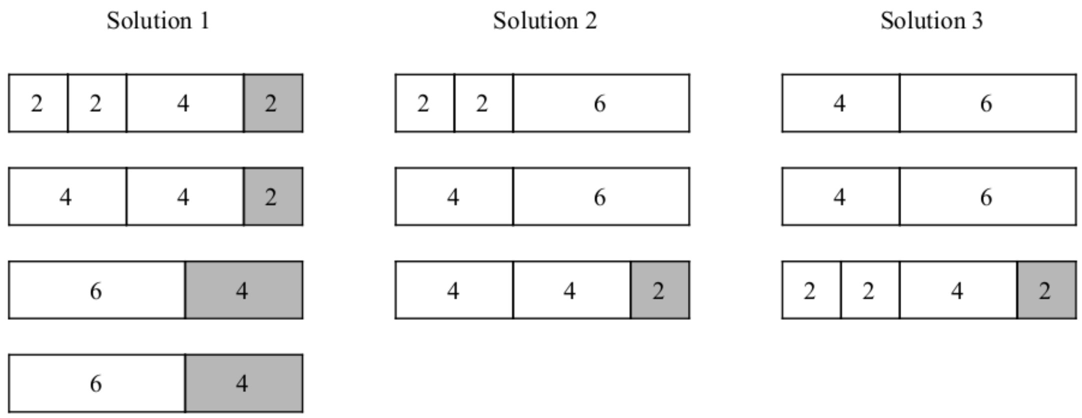

Figure 1.10: Few Possible Patterns

pattern $i, i \in N$. Therefore, we have the following formulation:

$$\begin{aligned} \min \sum_{i=1}^{n} y_i & \\ \text{subject to: } \sum_{i \in N} x_{ij} y_i \geq d_j, \;\; j \in F, & \\ \sum_{j \in F} l_j x_{ij} \leq L, i \in N, & \\ x_{ij} \in \mathbb{Z}_+, y_i \in \mathbb{R}_+, \forall i \in N, j \in F, & \end{aligned} \tag{CS}$$

In the above formulation we considered the variable y to be continuous. When the demands are large, we can let y be continuous without much affecting the optimal value.

To understand it in a better way, consider the following example. Let the final lengths be 2, 4, and 6 with demands 2, 3, and 2 respectively. The length of the large roll is 10. Firgure 1.10 shows few patterns to meet the demand and the scrap amount of the sheet (shaded in gray).

Pattern Minimization Problem (PMP) is an extension of the trim-loss problem and is much more difficult. Many times an optimal solution to a trim-loss problem may require many patterns. Each pattern corresponds to setting up the equipment to make the finals, and hence having too many patterns may not be beneficial. Pattern minimizing problem tries to find as few patterns for cutting as possible while limiting the wastage of raw material.

Let η denote the number of rolls that we are allowed to use for making the finals. This input parameter can be found by first solving a trim-loss problem like the one above. Let $z_i, i \in N$ be a decision variable which takes value 1 if pattern i is used (i.e., $y_i \geq 1$), and 0 otherwise. Then the problem may be formulated as:

$$
\begin{aligned}
\min_{x,y,z} \; & \sum_{i=1}^{n} z_i \\
\text{s.t.} \; & \sum_{i=1}^{n} x_{ij} y_i \geq d_j, j \in F, \\
& \sum_{j \in F} \mu_j x_{ij} \leq z_i L, i \in N, \\
& y_i \leq v z_i, i \in N, \\
& \sum_{i=1}^{n} y_i \leq \eta, \\
& x_{ij}, y_i \in \mathbb{Z}_+, z_i \in \{0,1\}, i \in N, j \in F,
\end{aligned}
\tag{PMP}
$$

In the Figure 1.10, we see that three patterns are required to meet the demands with three finals. If we allow more finals, we can meet the demands but the waste will be higher (Solution 1 in Figure 1.10).

Several studies [66, 69, 110] to solve these problems have been carried out in the past. Recently, Yaodong et al. [116] presented a sequential grouping procedure to generate the patterns and to minimize material and setup cost. Kallrath et al. [73] have developed polylithic techniques including a heuristic and column-generation based approaches to solve cutting stock problems. Theoretical bounds on the number of patterns that may be required in an optimal solution have been obtained only recently [49, 59].

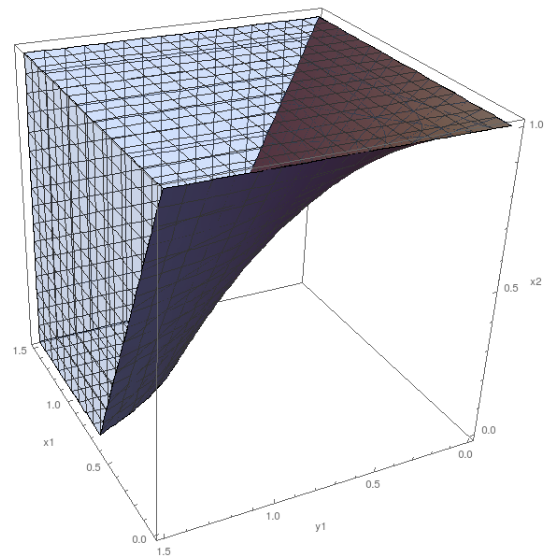

Figure 1.11: The points in positive orthant satisfying $x_1y_1 + x_2y_2 \geq 1$ in the box $x_1 \in [0, 1.5], y_1 \in [0, 1.5]x_2 \in [0, 1]$ with $y_2 = 1$.

1.5 Outline of Thesis and Summary of Contributions

In this thesis, we consider the following three variants of bilinear covering set.

$$
\begin{aligned}
S &= \left\{(x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n x_i y_i \geq r\right\}, r > 0, \\
S^U &= \left\{(x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n x_i y_i \geq r, x \leq u\right\}, r > 0, u \in \mathbb{N}_+^n, \\
S^B &= \left\{(x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n x_i y_i \geq r, l \leq x \leq u\right\}, r > 0, l \in \mathbb{Z}_+, u \in \mathbb{N}_+^n.
\end{aligned}
$$

The continuous relaxation of S obtained by ignoring integer restrictions on x is nonconvex (see Figure 1.11), and hence problems like PMP that include this constraint are nonconvex MINLPs.

Tawarmalani et al. [106], derived the convex hull description of the set S. It consists of countably infinite number of linear inequalities and consequently $conv(S)$ is not a polyhedron. In practical problems, the variables x and y have bounds, either defined explicitly or determined implicitly. The description of $conv(S)$ along with the bounds on the variables x does not give the description of $conv\left(S^U\right)$ or $conv\left(S^B\right)$, instead it gives weaker convex relaxation. We discuss this issue in the Chapter 2 and extend the approach of using orthogonal disjunctions to include the upper bounds on the variables.

In Chapter 3, we find convex hull of the set S^B and use this description to solve the PMP problem in a novel way. Chapter 4 analyses the valid inequalities for $conv(S)$ as split cuts and disjunctive cuts. The contributions of this thesis are enumerated below.

1. We derive the closed form description of the convex hull of the mixed-integer bilinear covering set S^U using orthogonal disjunctive technique. The set $conv\left(S^U\right)$ is a polyhedron. We give both the H-Description (linear inequality or intersection of finite number half spaces) and V-Description (description by vertex and the recession) of $conv\left(S^U\right)$.

2. We also provide an exact extended formulation of $conv\left(S^U\right)$ that consists of much less number of linear inequalities than the description of $conv\left(S^U\right)$.

3. We derive separation algorithms for the facet defining inequalities for both the sets $conv\left(S^U\right)$ and $conv(S)$ which runs in $O(n)$ times in the input size. Computational results show the effectiveness of our approach on real life and artificially generated cutting stock (or trim-loss) instances.

4. Next, we extend our results of $conv\left(S^U\right)$ to derive the description of $conv\left(S^B\right)$. Here also we derive the V-Description and H-Description of $conv\left(S^B\right)$.

5. We describe a new algorithm the inequalities for solving Pattern Minimization Problem and show the effectiveness of our algorithm.

6. We study the facet defining inequalities of $conv(S)$ as split and disjunctive cuts and identify all the split rank-one facet defining inequalities.

7. We also study the gap between the set that is constructed by the facet defining inequalities with split rank-one and $conv(S)$, and then derive the necessary and sufficient condition for the gap to be zero.

8. Finally, we show the effectiveness of the rank-one split facet defining inequalities using some real life and randomly generated cutting stock instances.

1.6 Notation

Unless otherwise mentioned, we use the following notation throughout this thesis. For a given set A, we use $cl(A)$ to denote the closure of $A, conv(A)$ to denote the convex hull of $A, \mathcal{C}(A)$ to denote the conic hull of A and $0^+(A)$ to denote the recession cone of A. $\mathbb{R}^n_+ = [0,\infty)^n = \{x \in \mathbb{R}^n : x \geq 0\}$. We use N to denote the set $\{1, 2, \ldots, n\}$. We use $\mathcal{S}_n$ to denote the set of all $n \times n$ symmetric matrices. For a point $(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$, we write (x, y) in the form $(x_1, y_1, x_2, y_2, \ldots, x_n, y_n)$. We use $\mathcal{L}(i, x_i, y_i)$ to denote the point $(0, 0, \ldots, x_i, y_i, \ldots, 0, 0)$, i.e., $x_j = 0, y_j = 0, \forall j \in N, j \neq i$. The sign $\vee$ means "or" and $\wedge$ means "and". For an integer vector $\mu \in \mathbb{Z}^n$ and an integer μ_0, we use $gcd(\mu, \mu_0) = gcd(\mu_1, \ldots, \mu_n, \mu_0)$ to denote the greatest common divisor of $\mu_1, \ldots, \mu_n$ and μ_0.

Chapter 2

Mixed-Integer Bilinear Covering Set With Upper Bounds on Variables

2.1 Introduction

In this chapter we consider the mixed-integer bilinear covering sets

$$S = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r \right\}, \text{ and}$$
$$S^U = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r, x \leq u \right\}, \text{ where } r > 0 \text{ and } u \in \mathbb{N}^n \text{ are given.}$$

Both S and S^U are nonconvex, even their continuous relaxation is nonconvex for $n \geq 2$. Recall the following two sets of constraints from the cutting stock problem (CS) defined in Chapter 1:

$$\sum_{i \in N} x_{ij} y_i \geq d_j, \ j \in F, \tag{2.1}$$
$$\sum_{j \in F} l_j x_{ij} \leq L, i \in N, \tag{2.2}$$

where $x_{ij} \in \mathbb{Z}_+, y_i \in \mathbb{R}_+$. Bounds on the variables $x_{ij}, i \in N, j \in F$ can be either given explicitly or be implicit from the knapsack constraints (2.2).

Tawarmalani et al. [106] developed a scheme to get a tight convex relaxation using orthogonal disjunctive subsets for a class of sets including S. They applied the scheme to obtain the convex hull description of S consisting of countably infinite number of facet defining inequalities. But these facet defining inequalities of $conv(S)$ along with the bound constraints are not sufficient to describe $conv\left(S^U\right)$. Consider, for example,

$$\begin{aligned} \min\ & -x_1 + 10y_1 - 2x_2 + 12y_2 \\ \text{s.t. }\ & x_1y_1 + x_2y_2 \geq 20, \\ & x_1 \leq 5, x_2 \leq 6, \\ & x_i \geq 0, y_i \geq 0, x_i \in \mathbb{Z}_+, i = 1, 2. \end{aligned} \tag{E}$$

Here, $r = 20, n = 2, u = (5, 6)$. The point $(x_1, y_1, x_2, y_2) = (5, 4, 6, 0)$ is a global optimal solution with optimal value 23. But, if we solve the relaxation defined by the facet defining inequalities of $conv(S)$ (which we describe later), along with the bound constraints on x, we get the solution $\omega = \left(5, 1, 6, \frac{5}{6}\right)$ with objective value 3. As expected, ω is not feasible for S but lies in $conv(S)$ because this point is the mid point of the two points $(10, 2, 0, 0), \left(0, 0, 12, \frac{5}{3}\right) \in S$. Therefore, no facet defining inequality of $conv(S)$ can cut off the point ω from $conv\left(S^U\right)$. We will see later that the inequality $\frac{5y_1}{20} + \frac{6y_2}{20} \geq 1$ is valid for S^U, and it cuts off the point w. In fact we show that this inequality is a facet defining inequality for $conv\left(S^U\right)$.

In this chapter, we derive the closed form description of the convex hull of the mixed-integer bilinear covering set S^U. We note that, the orthogonal disjunctive technique of Tawarmalani et al. [106] is not directly applicable for the set S^U to find $conv\left(S^U\right)$. So, we relax the orthogonal subsets of S^U in such a way that the result is applicable. Our work mainly addresses the following issues of the model of Tawarmalani et al. Their model has infinitely many facet defining inequalities and these inequalities along with the bound constraints gives us a weak relaxation of our set. We show that $conv\left(S^U\right)$ is a polyhedron. We derive both V-Polyhedron (i.e., description by sum of convex hull of the extreme points and its recession cone) and H-Polyhedron (i.e., description by intersection of finite number of half spaces) description of $conv\left(S^U\right)$. We provide fast separation algorithms to find a violated facet defining inequality for both the sets $conv\left(S^U\right)$ and $conv(S)$. We also provide an extended formulation of $conv\left(S^U\right)$. Lastly, we provide some computational results that show the effectiveness of our cuts and the extended formulation.

2.2 Convexification via Orthogonal Disjunction

We start by a general result derived by Tawarmalani et al. [106] for which some more notations are required. We use the same notation as in [106] for convenience. Let $(z, u) \in \mathbb{R}^{\sum_{i=1}^n d_i} \times \mathbb{R}^{\sum_{i=1}^n d'_i}$, where $z_i \in \mathbb{R}^{d_i}$ and $u_i \in \mathbb{R}^{d'_i}$. Moreover, let us define the functions $t^j : \mathbb{R}^{\sum_{i=1}^n d_i} \times \mathbb{R}^{\sum_{i=1}^n d'_i} \to \mathbb{R}$ for $j \in J$, $v^k : \mathbb{R}^{\sum_{i=1}^n d_i} \times \mathbb{R}^{\sum_{i=1}^n d'_i} \to \mathbb{R}$ for $k \in K$ and $w^l : \mathbb{R}^{\sum_{i=1}^n d_i} \times \mathbb{R}^{\sum_{i=1}^n d'_i} \to \mathbb{R}$ for $l \in L$ where J, K and L are some index sets. Let us also define the sets $A\left(t^J, v^K, w^L\right)$ and $C\left(t^J, v^K, w^L\right)$ as below:

$$A\left(t^J, v^K, w^L\right) = \left\{(z, u) : t^j(z, u) \geq 1, \forall j \in J, v^k(z, u) \geq -1, \forall k \in K, w^l(z, u) \geq 0, \forall l \in L\right\},$$
$$C\left(t^J, v^K, w^L\right) = \left\{(z, u) : t^j(z, u) \geq 0, \forall j \in J, v^k(z, u) \geq 0, \forall k \in K, w^l(z, u) \geq 0, \forall l \in L\right\}.$$

To describe the results, we need to additionally define positively-homogeneous functions. The following definition is taken from Rockafellar [99] (1970).

Definition 2.2.1 (Positively Homogeneous Function). *Let $f : \mathbb{R}^n \to [-\infty, \infty]$ be a function. f is said to be a positively homogeneous function if, $f(\lambda x) = \lambda f(x), \forall \lambda > 0$.*

For example, $f(x, y) = \sqrt{xy}$ is positively homogeneous. Also, any linear function is positively homogeneous.

Theorem 2.2.1 (Tawarmalani et al. [106]). *Let $z = (z_1, ..., z_i, ..., z_n) \in \mathbb{R}^{\sum_{i=1}^n d_i}$, where $z_i \in \mathbb{R}^{d_i}$ and $Z \subseteq \mathbb{R}^{\sum_{i=1}^n d_i}$. Let $Z_i \subseteq Z$ for $i \in N = \{1, ..., n\}$. Now let us consider the following assumptions:*

A1: $(z_1, ..., z_i, ..., z_n) \in Z_i \Rightarrow z_j = 0 \ \forall j \in N, j \neq i,$

A2: $conv(Z) = conv\left(\bigcup_{i=1}^n Z_i\right),$

A3: $conv(Z_i) \subseteq proj_z(A_i) \subseteq cl(conv(Z_i)),$ *where*

$$A_i = \left\{\mathcal{L}(i, z_i, u_i) : (z_i, u_i) \in A\left(t_i^{J_i}, v_i^{K_i}, w_i^{L_i}\right)\right\}$$

such that $t_i^{j_i}$, $\forall j_i \in J_i, v_i^{k_i}$, $\forall k_i \in K_i$ *and* $w_i^{l_i}$, $\forall l_i \in L_i$ *are positively-homogeneous functions for all* $i \in N$ *for some index sets* J_i, K_i *and* L_i, *and* $\mathcal{L}(i, z_i, u_i) = (0, ..., 0, z_i, u_i, 0, ..., 0) \in \mathbb{R}^{\sum_{i=1}^n d_i} \times \mathbb{R}^{\sum_{i=1}^n d'_i}$,

A4: *For all* $i = 1, ..., n$, $proj_z(C_i) \subseteq 0^+ (cl\,(conv\,(Z)))$, *where*

$$C_i = \left\{\mathcal{L}(i, z_i, u_i) : (z_i, u_i) \in C\left(t_i^{J_i}, v_i^{K_i}, w_i^{L_i}\right)\right\},$$

Then, $conv(Z) \subseteq proj_z(X) \subseteq cl(conv(Z))$, *where,*

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

Using the above theorem, we can derive the convex hull for those sets which satisfy assumptions A1 - A4. Checking whether A1, A3 and A4 are satisfied by a given set is relatively easy. Verifying A2 might be difficult in practice. To overcome this difficulty, Tawarmalani et al. [106] have used an alternative criterion, called convex extension property which is more general than assumption A2.

Definition 2.2.2 (Convex Extension Property)**.** *Let* Z *be a set in* $\mathbb{R}^n$ *and* $Z_i \subseteq Z, i \in N$. *The convex extension property holds for* Z *if it satisfies the following two properties.*

(i) If $z \in Z_i$, *then* $z_j = 0$ *for all* $j \in N, j \neq i$.

(ii) If $z \in Z$, *then* z *can be expressed as a sum of a convex combination of some points* $\chi_i \in cl(conv(Z_i)), i \in N$ *and a conic combination of rays* $\psi_i \in 0^+(cl(conv(Z_i))), i \in$

N, *i.e.*

$$z = \sum_{i \in N} \lambda_i \chi_i + \sum_{i \in N} \mu_i \psi_i \tag{CE}$$

where $\mu_i \in \mathbb{R}_+, i \in N$ *and* $\lambda_i \in \mathbb{R}_+, i \in N$ *with* $\sum_{i \in N} \lambda_i = 1$.

A collection of sets $Z_i, i \in N$ that satisfy condition (i) in Definition 2.2.2 are known as orthogonal sets. By definition a union of orthogonal sets satisfies the convex extension property. Some other sets that are not defined as union of orthogonal sets, for example, bilinear mixed-integer and pure-integer covering sets without variable bounds also satisfy this property. The convex extension property (CE) is equivalent to the following criterion given in Tawarmalani et al. [106].

$$cl(conv(Z)) = cl\left(conv\left(\bigcup_{i=1}^{n} Z_i\right)\right) \tag{CE-P}$$

Now, if we assume (CE) or (CE-P) instead of the assumption A2 in Theorem 2.2.1, we get $cl\,(proj_z X) = cl\,(conv\,(\bigcup_{i=1}^{n} Z_i)) = cl(conv(Z))$ (Tawarmalani et al. [106]). Since in many cases we only need $cl(conv(Z))$, it is useful to consider (CE) or (CE-P) instead of the assumption A2.

2.3 On The Mixed-Integer Bilinear Covering Set S

We start by revisiting the set $S = \left\{(x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r\right\}, r > 0$, and the facet defining inequalities of its convex hull. Then we derive a property of extreme points of $conv(S)$ that we will later extend to $conv\left(S^U\right)$.

2.3.1 The Convex Hull Description of S

Tawarmalani et al. [106] showed that the set S satisfies the assumptions A1, A3 and A4 of Theorem 2.2.1 and the convex extension property (CE) with respect to the orthogonal disjunctive subsets $S_i, i \in N$, where,

$$S_i = \left\{\mathcal{L}(i, x_i, y_i) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : x_i y_i \geq r\right\}.$$

Therefore, we can apply Theorem 2.2.1 to construct the description of $conv(S)$. For this, first, we have to find the description of $conv\left(S_i\right)$. The continuous relaxation of the set S_i is a convex set and the points of the form $\mathcal{L}\left(i, k, \frac{r}{k}\right), k \in \mathbb{N}$ are the extreme points of $conv\left(S_i\right)$. The convex hull description $conv(S_i)$ can be given as

$$conv\left(S_i\right) = \left\{\mathcal{L}(i, x_i, y_i) : a_k x_i + b_k y_i \geq 1, k \in \mathbb{N}\right\}, \tag{2.3}$$

where $a_k x_i + b_k y_i = 1$ is the line passing through $\left(k, \frac{r}{k}\right)$ and $\left(k-1, \frac{r}{k-1}\right)$ for $k \in \mathbb{N} \setminus \{1\}$ and $a_1 = 1, b_1 = 0$. Hence, we have $a_k = \frac{1}{2k-1}$ and $b_k = \frac{k(k-1)}{r(2k-1)}$ for all $k \in \mathbb{N}$. We note that $conv(S_i)$ has countably infinite number of extreme points and facet defining inequalities. Consequently, $conv(S_i)$ is not a polyhedral set. Note that the recession cone $0^+(conv(S_i))$ of $conv\left(S_i\right)$ is the following set

$$\left\{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : x_j = 0, y_j = 0, j \in N, j \neq i\right\}.$$

All sets $S_i, i \in N$ are identical to each other except for relabeling of indices. Thus, the coefficients a_k and $b_k, k \in \mathbb{N}$ are identical for each $conv\left(S_i\right), i \in N$. Therefore, finding the coefficients $a_k, b_k, k \in \mathbb{N}$ for $conv\left(S_1\right)$ is sufficient to get all the facets of $conv(S)$. The following collection of columns (M) with countably infinite number of rows can be used to generate all the facet defining inequalities of $conv(S)$.

$$\begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ a_2x_1 + b_2y_1 & a_2x_2 + b_2y_2 & a_2x_3 + b_2y_3 & \dots & a_2x_n + b_2y_n \\ a_3x_1 + b_3y_1 & a_3x_2 + b_3y_2 & a_3x_3 + b_3y_3 & \dots & a_3x_n + b_3y_n \\ \dots & \dots & \dots & \dots & \dots \\ a_kx_1 + b_ky_1 & a_kx_2 + b_ky_2 & a_kx_3 + b_ky_3 & \dots & a_kx_n + b_ky_n \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \tag{M}$$

Theorem 2.2.1 states that a facet defining inequality of $conv(S)$ is constructed by adding n terms from (M) taking exactly one term from each column and constraining their sum to be greater than or equal to one. All the facet defining inequalities are constructed this way. It is also clear that $conv(S)$ also has countably infinite number of facet defining inequalities. Since $a_k, b_k \geq 0, \forall k \in \mathbb{N}$, the recession cone $0^+(conv(S))$ of

$conv(S)$ is the entire non-negative orthant $\mathbb{R}^n_+ \times \mathbb{R}^n_+$.

2.3.2 Properties of The Extreme Points of $conv(S)$

Here we derive the description of the extreme points of $conv(S)$ that we use later. We first note that $conv(S)$ is a closed set. This is because, if $(x, y) \notin conv(S)$, there exists a facet defining inequality of $conv(S)$ that strongly separates the point (x, y) from $conv(S)$. Therefore, the point (x, y) can not be a limit point of $conv(S)$, and consequently $conv(S)$ is a closed set. The convex extension property (CE-P) applied to S gives $conv(S) = conv\left(\bigcup_{i\in N} S_i\right)$.

Theorem 2.3.1. *$(\bar{x}, \bar{y})$ is an extreme point of $conv(S)$ if and only if $(\bar{x}, \bar{y})$ is an extreme point of $conv(S_i)$ for some $i \in N$.*

Proof. Let $(\bar{x}, \bar{y})$ be an extreme point of $conv(S)$. If $(\bar{x}, \bar{y})$ belongs to S_i for some $i \in N$, then it has to be an extreme point of $conv(S_i)$ as $S_i \subset S$. On the other hand, if $(\bar{x}, \bar{y})$ does not belong to any $S_i, i \in N$, then by convex extension property (CE-P), $(\bar{x}, \bar{y})$ can be written as a convex combination of points in $S_i, i \in N$ which contradicts the extremality of the point $(\bar{x}, \bar{y})$.

Conversely, let $(\bar{x}, \bar{y})$ be an extreme point of $conv(S_i)$ for some $i \in N$. Then, $\bar{x}_j = 0, \bar{y}_j = 0, \forall j \in N, j \neq i$. For contradiction, let $(\bar{x}, \bar{y})$ be expressed as a convex combination of two distinct points $(\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$ in S. Since $S \subset \mathbb{R}^n_+ \times \mathbb{R}^n_+$, then $\bar{x}^t_j = 0, \bar{y}^t_j = 0, \forall j \in N, j \neq i, t = 1, 2$. This implies that $(\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$ belong to S_i. This is a contradiction to the fact that $(\bar{x}, \bar{y})$ is an extreme point of $conv(S_i)$. Therefore, $(\bar{x}, \bar{y})$ must be an extreme point of $conv(S)$. □

It is clear from Theorem 2.3.1 that any point of the form $\mathcal{L}\left(i, k, \frac{r}{k}\right), k \in \mathbb{N}$ is an extreme point of $conv(S)$ and vice versa, for all $i \in N$.

2.4 On The Mixed-Integer Bilinear Covering Set S^U

In this section we obtain a description of the convex hull of S^U defined in Section 2.1 and show that, unlike $conv(S)$, $conv\left(S^U\right)$ is a polyhedron.

Proposition 2.4.1. *The set $conv\left(S^U\right)$ is a polyhedron.*

Proof. Since there is an upper bound u on the integer variable x, we have finitely many choices for x in S^U. For each $i \in N$, we have $u_i + 1$ different choices for x_i. Since $x = 0$ is not a feasible choice for S^U, the total number of different choices for x is $\prod_{i=1}^{n}(u_i + 1) - 1 = \eta$ (say). Let us denote them by $x^k, k = 1, \ldots, \eta$. Now, define the following polyhedral sets:

$$F_k = \left\{(x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r, x = x^k\right\}, k = 1, \ldots, \eta.$$

Note that the set F_k is constructed from S^U by fixing $x = x^k$, and $S^U = \bigcup_{k=1}^{\eta} F_k$. Further, the recession cone of F_k is the set $\{(0, y) \in \mathbb{R}^n \times \mathbb{R}^n : y \geq 0\}$ for all $k = 1, \ldots, \eta$. Therefore, S^U is a union of finite number of nonempty polyhedra with identical recession cones. So, from Corollary 4.44 in Conforti et al. [36], we have $conv\left(S^U\right)$ is a polyhedron. □

2.4.1 The Extreme Point Description of *conv* $\left(S^U\right)$

Since $conv\left(S^U\right)$ is a polyhedron, it is closed and, therefore, it contains all its extreme points. In this section we give a closed form description of the extreme points of $conv\left(S^U\right)$.

Theorem 2.4.1. *Let $(\bar{x}, \bar{y})$ be an extreme point of $conv\left(S^U\right)$. Then, $\bar{x}_t = p_t, \bar{y}_t = \frac{r}{p_t}$ for some $t \in N$, where $p_t \in \{1, \ldots, u_t\}$, and $\bar{x}_j \in \{0, u_j\}, \bar{y}_j = 0, \forall j \in N, j \neq t$, i.e., $(\bar{x}, \bar{y})$ has the following form,*

$$\left(\bar{x}_1, 0, \bar{x}_2, 0, \ldots, \bar{x}_{t-1}, 0, p_t, \frac{r}{p_t}, \bar{x}_{t+1}, 0, \ldots, \bar{x}_n, 0\right)$$

where $p_t \in \{1, \ldots, u_t\}$ for some $t \in N, \bar{x}_j \in \{0, u_j\}, \forall j \in N, j \neq t$.

Proof. Let $(\bar{x}, \bar{y})$ be an extreme point of $conv\left(S^U\right)$. Then $(\bar{x}, \bar{y}) \in S^U$. Therefore, $(\bar{x}, \bar{y}) \in F_k$ for some $k \in \{1, \ldots, \eta\}$, and is an extreme point of F_k where F_k is defined in the proof of Proposition 2.4.1. Note that in the description of F^k, there are n bound constraints: $y_i \geq 0, i \in N$ and one linear constraint: $\sum_{i \in N} \bar{x}_i y_i \geq r$.

Since $(\bar{x}, \bar{y})$ is an extreme point of F_k, n linear constraints of F_k must be active at

$(\bar{x}, \bar{y})$. One can not choose the n constraints given by $y \geq 0$, otherwise $\sum_{i \in N} \bar{x}_i y_i \geq r$ will be violated. So, the constraint $\sum_{i \in N} \bar{x}_i y_i \geq r$ must be active. Therefore, there exists a $t \in N$ such that $\bar{x}_t \bar{y}_t = r$ and $y_j = 0, j \in N, j \neq t$.

We now show that $\bar{x}_j \in \{0, u_j\}$ for $\forall j \neq t$. Since $j \neq t$, then $\bar{y}_j = 0$. Therefore, if $\bar{x}_j \in (0, u_j), j \in N$, then $(\bar{x}, \bar{y})$ can be written as a convex combination of the two points $(\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$ having the exact same components as $(\bar{x}, \bar{y})$, except for the j^{th} components of the variable x, and $\bar{x}_j^1 = 0, \bar{x}_j^2 = u_j$. Multipliers $1 - \lambda$ and λ respectively provide the convex combination of $(\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$, where $\lambda = \frac{\bar{x}_j}{u_j}$. This is a contradiction to the supposition that $(\bar{x}, \bar{y})$ is an extreme point of $conv\left(S^U\right)$.

Moreover, if $\bar{x}_j \in \{0, u_j\}$ for $j \neq i$, then we can not write $(\bar{x}, \bar{y})$ as a convex combination of two different points in S^U. This is because, if two such points exist, one of the points' j^{th} component of the variable x has to be more than u_j or less than 0, neither of which is allowed. $\square$

Corollary 2.4.1. *$conv\left(S^U\right)$ has $2^{n-1} \sum_{i=1}^n u_i$ extreme points and n extreme rays.*

Proof. We see from the proof of Theorem 2.4.1, for a single choice of $\bar{x}_i \in \{1, \ldots, u_i\}$, we have 2^{n-1} different extreme points, and we have $\sum_{i=1}^n u_i$ distinct such choices. Therefore, the total number of extreme points of $conv\left(S^U\right)$ is $2^{n-1} \sum_{i=1}^n u_i$, which is finite. Consequently, $conv\left(S^U\right)$ is a polyhedral set.

On the other hand, we see that the recession cone $0^+\left(conv\left(S^U\right)\right)$ of $conv\left(S^U\right)$ is the set $\{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : x = 0\}$ which has n extreme rays. $\square$

Note that Theorem 2.4.1 and Corollary 2.4.1 give us the V-Description of $conv\left(S^U\right)$. We now turn our attention to the H-Description of $conv\left(S^U\right)$.

2.4.2 The Convex Hull Description of S^U

We have orthogonal disjunctive subsets of S^U,

$$S_i^U = \left\{\mathcal{L}(i, x_i, y_i) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : x_i y_i \geq r, x_i \leq u_i\right\}, i = 1, \ldots, n.$$

We note that $S_i^U \subset S^U$, and the recession cone of $cl\left(conv\left(S_i^U\right)\right)$ is the set:

$$\left\{(x,y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : x = 0, y_j = 0, \forall j \in N, j \neq i\right\}.$$

We see that the assumption A1 of Theorem 2.2.1 is satisfied by the set S^U with respect to the orthogonal disjunctive subsets S_i^U. The polyhedral description of $conv\left(S_i^U\right)$ is

$$conv\left(S_i^U\right) = \left\{\mathcal{L}(i, x_i, y_i) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : a_k x_i + b_k y_i \geq 1, x_i \leq u_i, \forall k \in K_i\right\},$$

where $K_i = \{1, \ldots, u_i\}$, and as defined earlier, $a_k = \frac{1}{2k-1}, b_k = \frac{k(k-1)}{r(2k-1)}, k \in K_i$. Therefore, assumption A3 of Theorem 2.2.1 is satisfied by the set S^U with respect to its orthogonal subsets $S_i^U, i \in N$.

On the other hand, the assumption A2 and convex extension property are not satisfied by the set S^U with respect to the subsets $S_i^U, i \in N$. An extreme point of $conv\left(S^U\right)$ can have all x components nonzero which does not belong to S_i^U for any $i \in N$. So, if it were in $conv\left(\bigcup_{i \in N} S_i^U\right)$, then it has to be a convex combination of two points in S^U which contradicts the extremality of the point.

In order to find the description of $conv\left(S^U\right)$, we use the following approach. The two inequalities $x_i y_i \geq r$ and $x_i \leq u_i$ in the description of S_i^U together imply $y_i \geq \frac{r}{u_i}$. Let $\frac{r}{u_i} = \bar{u}_i$. Let us now define the following sets:

$$S_i^L = \left\{\mathcal{L}(i, x_i, y_i) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : x_i y_i \geq r, y_i \geq \bar{u}_i\right\}, i = 1, \ldots, n.$$

By adding the lower bound on y_i and ignoring the upper bound on x_i, we have a relaxation of S_i^U. The two sets $conv\left(S_i^U\right)$ and $conv\left(S_i^L\right)$ have exactly the same set of extreme points that are u_i in number. Figure 2.1 and 2.2 illustrate this observation.

We have the description of $conv\left(S_i^L\right)$ as following:

$$conv\left(S_i^L\right) = \left\{\mathcal{L}(i, x_i, y_i) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : a_k x_i + b_k y_i \geq 1, y_i \geq \bar{u}_i, \forall k \in K_i\right\}, \tag{2.4}$$

where $K_i = \{1, \ldots, u_i\}$, and $a_k, b_k, k \in K_i$ are defined earlier. We also note that the

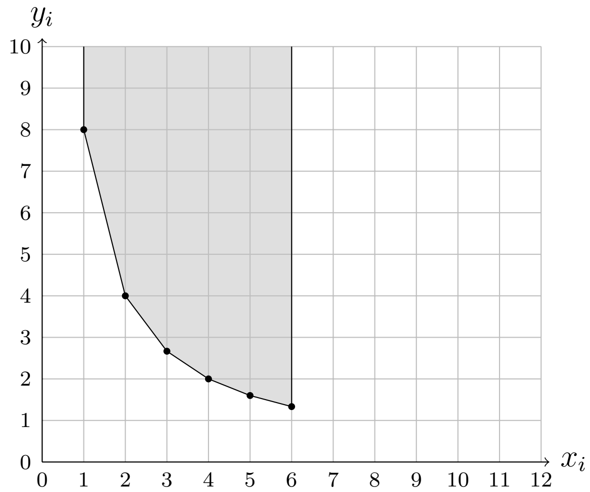

Figure 2.1: $conv\left(S_i^U\right)$ for $r = 8, x_i \leq 6$

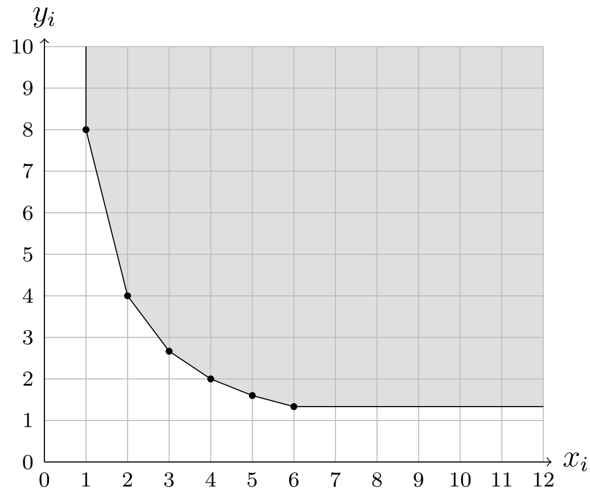

Figure 2.2: $conv\left(S_i^L\right)$ for $r = 8, y_i \geq \frac{8}{6}$

recession cone of $conv\left(S_i^L\right)$ is the set:

$$\left\{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : x_j = 0, y_j = 0, j \in N, j \neq i\right\}.$$

Let us now define a new set

$$S^L = \bigcup_{i=1}^{n} S_i^L.$$

We will later derive the description of $conv\left(S^U\right)$ using $cl\left(conv\left(S^L\right)\right)$. We first observe that, since $S^L = \bigcup_{i=1}^n S_i^L$, we have,

$$cl\left(conv\left(S^L\right)\right) = cl\left(conv\left(\bigcup_{i=1}^{n} S_i^L\right)\right),$$

i.e., the set S^L satisfies the condition (CE-P) with respect to the orthogonal disjunctive subsets $S_i^L, i \in N$.

Proposition 2.4.2. *The set S^L satisfies all the assumptions A1 - A4 of Theorem 2.2.1 with respect to the orthogonal disjunctive subsets $S_i^L, i \in N$.*

Proof. We see that the assumption A1 holds from the definition of S^L. For the assumption A2, we have the convex extension property that is satisfied as noted above. Since we have the polyhedral description of $conv\left(S_i^L\right)$, the assumption A3 is satisfied. Lastly, we see that $0^+\left(cl\left(conv\left(\bigcup_{i=1}^n S_i^L\right)\right)\right)$ is the entire non-negative orthant $\mathbb{R}_+^n \times \mathbb{R}_+^n$, which implies that the assumption A4 is also satisfied. $\square$

We can now apply Theorem 2.2.1 to obtain a description of $cl\left(conv\left(S^L\right)\right)$. We have $conv\left(S_i^L\right) = \left\{\mathcal{L}(i,x_i,y_i) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : a_k x_i + b_k y_i \geq 1, y_i \geq \bar{u}_i, k \in K_i\right\}$, where, as defined earlier $K_i = \{1,\ldots,u_i\}$. Let us write it using a single index set as following:

$$conv\left(S_i^L\right) = \left\{\mathcal{L}(i,x_i,y_i) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : l^{k_i}(x_i,y_i) \geq 1, k_i \in \bar{K}_i\right\},$$

where, $\bar{K}_i = K_i \bigcup \{u_i+1\}, l^{k_i}(x_i,y_i) = a_{k_i}x_i + b_{k_i}y_i$, where $a_{k_i} = \frac{1}{2k_i-1}, b_{k_i} = \frac{k_i(k_i-1)}{r(2k_i-1)}, k_i \in K_i$ and $l^{(u_i+1)_i}(x_i,y_i) = \frac{y_i}{\bar{u}_i}$. Note that the extreme points of $conv\left(S_i^L\right)$ are $\mathcal{L}\left(i,x_i,\frac{r}{x_i}\right), x_i = 1,\ldots,u_i$. Therefore, we have

$$conv\left(S_i^L\right) = conv\left(\left\{\mathcal{L}\left(i,x_i,\frac{r}{x_i}\right) : x_i = 1,\ldots,u_i\right\}\right) + \mathcal{C}\left(\mathcal{L}(i,1,0),\mathcal{L}(i,0,1)\right), \quad (2.5)$$

where $\mathcal{C}\left(\mathcal{L}\left(i,1,0\right),\mathcal{L}(i,0,1)\right)$ is the conic hull of $\left\{\mathcal{L}\left(i,1,0\right),\mathcal{L}(i,0,1)\right\}$. Now applying Theorem 2.2.1 we have,

$$cl\left(conv\left(S^L\right)\right) = \left\{(x,y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n l^{k_i}(x_i,y_i) \geq 1, \forall\ (k_i)_{i=1}^n \in \prod_{i=1}^n \bar{K}_i\right\}. \quad (2.6)$$

The set $cl\left(conv\left(S^L\right)\right)$ is a polyhedral set as it has finite number of facet defining inequalities in its description, and the number of facets is $\prod_{i=1}^n |\bar{K}_i| = \prod_{i=1}^n (u_i+1)$ (which is exponentially large). Also, $0^+\left(cl\left(conv\left(S^L\right)\right)\right)$ is the entire non-negative orthant $\mathbb{R}_+^n \times \mathbb{R}_+^n$. Let us now derive some properties of the set $cl\left(conv\left(S^L\right)\right)$.

Proposition 2.4.3. *The set $cl\left(conv\left(S^L\right)\right)$ is a polyhedral relaxation of S^U.*

Proof. Since $S^L = \bigcup_{i=1}^n S_i^L$, from (2.5) using Lemma 4.41 in [36] we have

$$cl\left(conv\left(S^L\right)\right) = conv\left(\bigcup_{i\in N}\left\{\mathcal{L}\left(i,x_i\frac{r}{x_i}\right) : x_i = 1,\ldots,u_i\right\}\right) + \mathbb{R}_+^n \times \mathbb{R}_+^n \quad (2.7)$$

Since $0^+\left(conv\left(S^U\right)\right)$ is a subset of $\mathbb{R}_+^n \times \mathbb{R}_+^n = 0^+\left(cl\left(conv\left(S^L\right)\right)\right)$, it is sufficient to show that all the extreme points of $conv\left(S^U\right)$ belong to $cl\left(conv\left(S^L\right)\right)$. Let $(\bar{x},\bar{y})$ be an extreme point of *conv* $\left(S^U\right)$, then from Theorem 2.4.1 we have

$$(\bar{x}, \bar{y}) = \mathcal{L}\left(i, \bar{x}_i, \frac{r}{\bar{x}_i}\right) + (\bar{x}_1, 0, \ldots, \bar{x}_{i-1}, 0, 0, 0, \bar{x}_{i+1}, 0, \ldots, \bar{x}_n, 0)$$

for some $i \in N$. This clearly shows that $(\bar{x}, \bar{y}) \in cl\left(conv\left(S^L\right)\right)$. $\square$

Theorem 2.4.2. *$(\bar{x}, \bar{y})$ is an extreme point of $cl\left(conv\left(S^L\right)\right)$ if and only if $(\bar{x}, \bar{y})$ is an extreme point of $conv\left(S_i^L\right)$ for some $i \in N$.*

Proof. The statement follows from (2.7). $\square$

Corollary 2.4.2. *$(\bar{x}, \bar{y})$ is an extreme point of $cl\left(conv\left(S^L\right)\right)$ if and only if $(\bar{x}, \bar{y})$ is an extreme point of $conv\left(S_i^U\right)$ for some $i \in N$.*

Proof. Since *conv* $\left(S_i^U\right)$ and *conv* $\left(S_i^L\right)$ have exactly same set of extreme points, the result follows from Theorem 2.4.2. $\square$

Here we observe that $cl\left(conv\left(S^L\right)\right)$ is a polyhedral relaxation of S^U such that each extreme point of $cl\left(conv\left(S^L\right)\right)$ lies in S^U. Now we prove our main result.

Theorem 2.4.3. *Let $\bar{S} = \{(x, y) \in cl\left(conv\left(S^L\right)\right) : x \leq u\}$. Then, $conv\left(S^U\right) = \bar{S}$.*

Proof. Minkowski Resolution Theorem (Theorem 4.15 in [22]) states that any polyhedral set having at least one extreme point can be described by its extreme points and recession cone. The polyhedral sets *conv* $\left(S^U\right)$ and $\bar{S}$ have the same recession cone $\{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : x = 0\}$.

The constraint $x_i \leq u_i$ passes through only one extreme point $\mathcal{L}\left(i, u_i, \frac{r}{u_i}\right)$ of $cl\left(conv\left(S^L\right)\right)$ and does not cut off any of its extreme points. Therefore, adding this constraint to $cl\left(conv\left(S^L\right)\right)$ only creates new extreme points of the form

$$\left(w_1, 0, w_2, 0, \ldots, w_{i-1}, 0, p_i, \frac{r}{p_i}, w_{i+1}, 0, \ldots, w_n, 0\right),$$

where $w_j \in \{0, u_j\}, j \in N, j \neq i, p_i \in \{1, \ldots, u_i\}, i \in N$. From Theorem 2.4.1, we see that such points lie in S^U, in fact, they are extreme points of *conv* $\left(S^U\right)$. Again, since $conv\left(S^U\right) \subseteq \bar{S}$, we have $\bar{S} = conv\left(S^U\right)$. $\square$

2.4.3 Facet Defining Inequalities of $conv\left(S^U\right)$

We now focus our attention on the new inequalities that are generated by our procedure and their effectiveness. We have seen from Theorem 2.4.3 that each facet defining inequality of $conv\left(S^U\right)$ is either a bound constraint $x_i \leq u_i$ for some $i \in N$ or a facet defining inequality of $cl\left(conv\left(S^L\right)\right)$ of the following form:

$$\sum_{i=1}^{n} l^{k_i}(x_i, y_i) \geq 1, (k_i)_{i=1}^{n} \in \prod_{i=1}^{n} \bar{K}_i \tag{F$_{\text{SL}}$}$$

where, $\bar{K}_i = K_i \bigcup \{u_i + 1\}, K_i = \{1, \ldots, u_i\}, l^{k_i}(x_i, y_i) = a_{k_i} x_i + b_{k_i} y_i, a_{k_i} = \frac{1}{2k_i - 1}, b_{k_i} = \frac{k_i(k_i-1)}{r(2k_i-1)}, k_i \in K_i$ and $l^{(u_i+1)_i}(x_i, y_i) = \frac{y_i}{\bar{u}_i}, \bar{u}_i = \frac{r}{u_i}$. The inequality $\sum_{i=1}^{n} l^{k_i}(x_i, y_i) \geq 1$ is identical to one of the facet defining inequalities of $conv(S)$ if $(k_i)_{i=1}^{n} \in \prod_{i=1}^{n} K_i$. Now let $Q \subseteq N$ be a non-empty index set such that $k_i = (u_i + 1)_i$ for all $i \in Q$. Then, the inequalities of the form

$$\sum_{i \in Q} \frac{y_i}{\bar{u}_i} + \sum_{i \in N \setminus Q} l^{k_i}(x_i, y_i) \geq 1, (k_i)_{i=1}^{n} \in \prod_{i=1}^{n} \bar{K}_i \tag{NF}$$

are generated by applying our approach and they are not valid for $conv(S)$.

2.4.4 An Extended Formulation of $conv\left(S^U\right)$

We saw that the description of $conv\left(S^U\right)$ consists of exponentially many facet defining inequalities. Let us consider the following set:

$$S^E = \left\{ (x, y, w) \in \mathbb{R}_+^{n+n+n} : \sum_{i \in N} w_i \geq 1, w_i \leq l^{k_i}(x_i, y_i), k_i \in \bar{K}_i, i \in N, x \leq u \right\}.$$

Proposition 2.4.4. *The set S^E is an extended formulation of conv $\left(S^U\right)$.*

Proof. If $(x, y, w) \in S^E$, then clearly $(x, y) \in conv\left(S^U\right)$. Now let $(x, y) \in conv\left(S^U\right)$ and define $w_i = \min_{k_i}\{l^{k_i}(x_i, y_i), k_i \in \bar{K}_i\}, i \in N$. Since the point (x, y) is feasible for the set $conv\left(S^U\right), \sum_{i \in N} \min_{k_i}\{l^{k_i}(x_i, y_i), k_i \in \bar{K}_i\} \geq 1$, and consequently $\sum_{i \in N} w_i \geq 1$. Thus S^E is an extended formulation of $conv\left(S^U\right)$. □

Even though the description of S^E consists of far fewer number of constraints than

$conv\left(S^U\right)$, it has $\sum_{i\in N} u_i + n + 1$ linear inequalities in addition to the the bound constraints, which is pseudopolynomial in the input size because of its dependency on u.

This extended formulation can be solved as a linear program to optimize a linear function over $conv\left(S^U\right)$. When the components of u are small (as in some cutting stock problems), this linear program can be solved fast.

2.5 The Separation Problem

We now describe a linear time separation algorithm to separate a given point $(\bar{x},\bar{y})$ from $conv\left(S^U\right)$. Let $(\bar{x},\bar{y})$ be a point in $\mathbb{R}^n \times \mathbb{R}^n$. If $\bar{x} \not\leq u$, then a bound constraint is sufficient to separate $(\bar{x},\bar{y})$. We thus consider the separation problem for the facet defining inequalities of $cl\left(conv\left(S^L\right)\right)$.

The facet defining inequalities of $cl\left(conv\left(S^L\right)\right)$ given by ($\mathrm{F_{SL}}$) can be listed in a different way for easier understanding. Consider the following collection of columns.

$$\begin{bmatrix} l^{1_1}(x_1,y_1) & l^{1_2}(x_2,y_2) & l^{1_3}(x_3,y_3) & \dots & l^{1_n}(x_n,y_n) \\ l^{2_1}(x_1,y_1) & l^{2_2}(x_2,y_2) & l^{2_3}(x_3,y_3) & \dots & l^{2_n}(x_n,y_n) \\ \dots & \dots & \dots & \dots & \dots \\ l^{(u_1+1)_1}(x_1,y_1) & l^{(u_2+1)_2}(x_2,y_2) & l^{(u_3+1)_3}(x_3,y_3) & \dots & l^{(u_n+1)_n}(x_n,y_n) \end{bmatrix} \tag{$\mathrm{M_U}$}$$

Note that ($\mathrm{M_U}$) may have a different number of elements in each column depending upon u, and thus it is not a matrix. The facet defining inequalities of $cl\left(conv\left(S^L\right)\right)$ can be constructed by adding n terms from ($\mathrm{M_U}$), taking exactly one term from each column and constraining the sum to be at least one.

Let us revisit the example (E) in Section 2.1. As discussed in Section 2.1, the point $\left(5,1,6,\frac{5}{6}\right)$ lies in $conv(S)$. But we see that this point is violated by the inequality $\frac{5y_1}{20} + \frac{6y_2}{20} \geq 1$ which is of the form (NF). Note that the inequalities $\frac{5y_1}{20} \geq 1$ and $\frac{6y_2}{20} \geq 1$ are valid for S_1 and S_2, respectively and combining them in the way described above we obtain the inequality $\frac{5y_1}{20} + \frac{6y_2}{20} \geq 1$. Adding this inequality to $conv(S)$, we get the optimal solution $(5,4,6,0)$ with optimal value 23.

If $(\bar{x},\bar{y}) \notin conv\left(S^U\right)$, then it must be violated by at least one inequality of the form ($\mathrm{F_{SL}}$). In order to find such a violated inequality, we have to find one term from each column of ($\mathrm{M_U}$) so that the sum of these is less than 1.

2.5.1 Efficient Separation for $conv\left(S^U\right)$

In order to separate a point $(\bar{x}, \bar{y})$ from $conv\left(S^U\right)$, we find a minimum element from each column of ($\mathrm{M_U}$) at $(\bar{x}, \bar{y})$ and add them. Clearly, if the the sum is greater than or equal to 1, the point $(\bar{x}, \bar{y})$ is feasible to $conv\left(S^U\right)$. Otherwise, adding the corresponding terms from each column and setting it to greater than or equal to 1, will give us a violated facet defining inequality.

Column i of ($\mathrm{M_U}$) has $(u_i + 1)$ terms, $i \in N$. To solve the separation problem, we need to find the minimum value at $(\bar{x}, \bar{y})$ from each column. This step takes $O(u_i)$ time which is pseudo-polynomial in the size of input. We now present a linear time algorithm for the separation problem.

Proposition 2.5.1. *There exists an efficient separation of the facet defining inequalities of* $conv\left(S^U\right)$.

Proof. Since the bound constraints can be checked easily, let $(\bar{x}, \bar{y}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ such that $\bar{x} \leq u$ be a given point. For each column of ($\mathrm{M_U}$), we want to find the term that gives the minimum evaluation at the point $(\bar{x}, \bar{y})$. Let

$$\xi_i = \min\left\{\frac{\bar{x}_i}{2w-1} + \frac{\bar{y}_i w(w-1)}{r(2w-1)}, \frac{\bar{y}_i}{\bar{u}_i}, w = 1, \ldots, u_i\right\}, \text{ where } \bar{u}_i = \frac{r}{u_i}.$$

Note that $\xi_i \geq 0$. To find ξ_i, we consider the following cases:

Case 1: If $\bar{y}_i = 0$, then clearly $\xi_i = 0$ at the last term, i.e., at $\frac{y_i}{\bar{u}_i}$ since $\frac{\bar{y}_i}{\bar{u}_i} = 0$.

Case 2: If $\bar{x}_i = 0$, then again $\xi_i = 0$ at $w = 1$.

Case 3: $\bar{x}_i > 0$ and $\bar{y}_i > 0$. Let us consider the following function:

$$f(w) = \frac{\bar{x}_i}{2w-1} + \frac{\bar{y}_i w(w-1)}{r(2w-1)}, w \geq 1.$$

Our goal is to find a positive integer q that minimizes $f(w)$ among all the integers in

$[1, u_i]$. The function f is continuously differentiable in the domain $w \geq 1$ with

$$f'(w) = -\frac{2\bar{x}_i}{(2w-1)^2} + \frac{\bar{y}_i}{r} \cdot \frac{2w^2 - 2w + 1}{(2w-1)^2} \text{ and } f''(w) = \frac{2(4\bar{x}_i r - \bar{y}_i)}{r(2w-1)^3}.$$

We have the following two subcases:

Case 3.1: When $4\bar{x}_i r - \bar{y}_i > 0$, the function f is strictly convex and has unique minimizer, say $\bar{w}_i$. Now $f'(\bar{w}_i) = 0$ occurs at

$$\bar{w}_i = \frac{1}{2} + \frac{\sqrt{\frac{4\bar{x}_i r}{\bar{y}_i} - 1}}{2}. \tag{2.8}$$

When $\bar{w}_i \leq 1$, the integer minimizer of f is $q = 1$. When $u_i > \bar{w}_i > 1$, $q = \lceil \bar{w}_i \rceil$ or $\lfloor \bar{w}_i \rfloor$ whichever gives a lower $f(q)$ is the required q. Finally, $q = u_i$ when $\bar{w}_i > u_i$.

Case 3.2: When $4\bar{x}_i r - \bar{y}_i \leq 0$, the function f is concave for $w \geq 1$. Therefore, the minimum value will be attained at a boundary point, i.e., either at 1 or at u_i. Moreover, we see that

$$\begin{aligned} f'(w) &= -\frac{2\bar{x}_i}{(2w-1)^2} + \frac{\bar{y}_i}{r} \cdot \frac{2w^2 - 2w + 1}{(2w-1)^2} \\ &= \frac{2\bar{y}_i w(w-1) + \bar{y}_i - 2\bar{x}_i r}{r(2w-1)^2} > 0. \end{aligned}$$

Thus, f is strictly increasing function, and is minimized at $q = 1$. Now one more comparison is required to find the value of ξ_i. If $\frac{\bar{x}_i}{2q-1} + \frac{\bar{y}_i q(q-1)}{r(2q-1)} \leq \frac{\bar{y}_i}{\bar{u}_i}$ then $\xi_i = \frac{\bar{x}_i}{2q-1} + \frac{\bar{y}_i q(q-1)}{r(2q-1)}$, else $\xi_i = \frac{\bar{y}_i}{\bar{u}_i}$.

The term corresponding to any column of ($\mathrm{M_U}$) can be computed in $O(1)$ time, and since there are n columns, a violated inequality can be found in $O(n)$ time.

If $\sum_{i=1}^{n} \xi_i \geq 1$, the point $(\bar{x}, \bar{y})$ is feasible to *conv* $\left(S^U\right)$. In Algorithm 1 we provide the separation algorithm in pseudocode. □

Corollary 2.5.1. *The optimization problem having a linear objective function over the set conv* $\left(S^U\right)$ *can be solved in time polynomial in size of the input.*

Proof. Since there is a polynomial time separation algorithm of the facet defining inequalities of *conv* $\left(S^U\right)$, the optimization of a linear function over *conv* $\left(S^U\right)$ can also be done

Algorithm 1 Separation of the facet defining inequalities of $conv\left(S^U\right)$

Input : A point $(\bar{x}, \bar{y}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+, x \leq u$

Output : Decide whether $(\bar{x}, \bar{y}) \in conv\left(S^U\right)$, and if not then provide a facet defining inequality of $conv\left(S^U\right)$ that cuts off $(\bar{x}, \bar{y})$

for $i = 1, \dots, n$ **do**
    if $\bar{y}_i = 0$ **then**
        $\hat{w}_i = u_i + 1$
    else if $\bar{x}_i = 0$ **then**
        $\hat{w}_i = 1$
    else
        $q = 0$
        if $4\bar{x}_i r > \bar{y}_i$ **then**
            if $\frac{1}{2} + \frac{\sqrt{\frac{4\bar{x}_i r}{\bar{y}_i} - 1}}{2} > 1$ **then**
                if $\frac{1}{2} + \frac{\sqrt{\frac{4\bar{x}_i r}{\bar{y}_i} - 1}}{2} < u_i$ **then**
                    $p = \left\lfloor \frac{1}{2} + \frac{\sqrt{\frac{4\bar{x}_i r}{\bar{y}_i} - 1}}{2} \right\rfloor$
                    if $\frac{\bar{x}_i}{2p-1} + \frac{\bar{y}_i p(p-1)}{r(2p-1)} \leq \frac{\bar{x}_i}{2(p+1)-1} + \frac{\bar{y}_i p(p+1)}{r(2(p+1)-1)}$ **then**
                        $q = p$
                    else
                        $q = p + 1$
                    end if
                else
                    $q = u_i$
                end if
            else
                $q = 1$
            end if
        else
            $q = 1$
        end if
        if $\frac{\bar{x}_i}{2q-1} + \frac{\bar{y}_i q(q-1)}{r(2q-1)} \leq \frac{\bar{y}_i}{\bar{u}_i}$ **then**
            $\hat{w}_i = q$
        else
            $\hat{w}_i = u_i + 1$
        end if
    end if
end for
$R = \sum_{i \in N : \hat{w}_i \leq u_i} \frac{\bar{x}_i}{2\hat{w}_i - 1} + \frac{\bar{y}_i \hat{w}_i(\hat{w}_i - 1)}{r(2\hat{w}_i - 1)} + \sum_{i \in N : \hat{w}_i = u_i + 1} \frac{\bar{y}_i}{\bar{u}_i}$
if $R \geq 1$ **then**
    The point $(\bar{x}, \bar{y})$ is feasible to $conv\left(S^U\right)$.
else
    The inequality $\sum_{i \in N : \hat{w}_i \leq u_i} \frac{x_i}{2\hat{w}_i - 1} + \frac{y_i \hat{w}_i(\hat{w}_i - 1)}{r(2\hat{w}_i - 1)} + \sum_{i \in N : \hat{w}_i = u_i + 1} \frac{y_i}{\bar{u}_i} \geq 1$ separates $(\bar{x}, \bar{y})$.
end if

in polynomial time (Grötschel et al. [64]). We present an algorithm in Appendix B. $\square$

2.5.2 Efficient Separation for $conv(S)$

The separation problem in the case of $conv(S)$ can also be solved in similar way with some modification. We use this algorithm to compare the effectiveness of our new cuts derived for $conv\left(S^U\right)$ in computational experiments.

Proposition 2.5.2. *There exists an efficient separation of the facet defining inequalities of $conv(S)$.*

Proof. Given a point $(\bar{x}_i, \bar{y}_i) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$, let

$$\xi_i = \min\left\{\frac{\bar{x}_i}{2w-1} + \frac{\bar{y}_i w(w-1)}{r(2w-1)}, w \in \mathbb{N}\right\}.$$

Note that $\xi \geq 0$ for $w \geq 1$. Our goal is to find a positive integer that minimizes $f(w)$. We consider the following cases.

Case 1: When $\bar{x}_i = 0$, then clearly $\xi_i = 0$ at $\hat{w}_i = 1$.

Case 2: When $\bar{y}_i = 0, \bar{x}_i \neq 0$, then $\inf\left\{\frac{\bar{x}_i}{2k-1} + \frac{\bar{y}_i k(k-1)}{r(2k-1)}, k \in \mathbb{N}\right\} = 0$, since $\frac{\bar{x}_i}{2k-1} \to 0$ as $k \to \infty$. Therefore ξ_i can be taken as 0 in this case.

Case 3: When $\bar{x}_i > 0, \bar{y}_i > 0$, the same logic used for $conv\left(S^U\right)$ can be deployed. Let $\hat{w}_i$ be the desired integer value. Then

$$\hat{w}_i = \begin{cases} 1, \text{ when } 4\bar{x}_i r - \bar{y}_i > 0 \text{ and } \bar{w}_i \leq 1, \\ \lceil \bar{w}_i \rceil, \text{ when } 4\bar{x}_i r - \bar{y}_i > 0, \bar{w}_i > 1 \text{ and } f\left(\lceil \bar{w}_i \rceil\right) \leq f\left(\lfloor \bar{w}_i \rfloor\right), \\ \lfloor \bar{w}_i \rfloor, \text{ when } 4\bar{x}_i r - \bar{y}_i > 0, \bar{w} > 1 \text{ and } f\left(\lceil \bar{w}_i \rceil\right) \geq f\left(\lfloor \bar{w}_i \rfloor\right), \\ 1, \text{ when } 4\bar{x}_i r - \bar{y}_i \leq 0, \end{cases}$$

where $\bar{w}_i$ is defined by (2.8). If $\sum_{i=1}^n \xi_i \geq 1$, the point $(\bar{x}, \bar{y})$ is feasible to $conv(S)$. Otherwise, it is infeasible, and we have to find a violated facet defining inequality. We know the required value of $\hat{w}_i$ for Case 1 and 3. Let $t \in \mathbb{N}$ such that the following holds,

$$\sum_{i=1}^n \xi_i + \sum_{i \in N: \bar{x}_i > 0, \bar{y}_i = 0} \frac{\bar{x}_i}{2t-1} < 1. \tag{2.9}$$

Such a t can always be found by the Archimedean property. A simple calculation shows that any integer greater than $\left\lfloor \frac{1-\xi+v}{2(1-\xi)} \right\rfloor$ where $\xi = \sum_{i=1}^{n} \xi_i$, $v = \sum_{i \in N: \bar{x}_i > 0, \bar{y}_i = 0} \bar{x}_i$ is sufficient. Therefore, the following inequality is violated by the point $(\bar{x}, \bar{y})$:

$$\sum_{i \in N: \bar{x}_i = 0} x_i + \sum_{i \in N: \bar{x}_i > 0, \bar{y}_i > 0} \left[\frac{x_i}{2\hat{w}_i - 1} + \frac{y_i \hat{w}_i (\hat{w}_i - 1)}{r(2\hat{w}_i - 1)} \right] + \sum_{i \in N: \bar{x}_i > 0, \bar{y}_i = 0} \left[\frac{x_i}{2t - 1} + \frac{y_i t (t-1)}{r(2t-1)} \right] \geq 1,$$

where $t \in \mathbb{N}$ such that $t \geq \left\lfloor \frac{1-\xi+v}{2(1-\xi)} \right\rfloor + 1$. In Algorithm 2, we provide the pseudocode of the separation algorithm. $\square$

Note that for any positive integer $t \geq \left\lfloor \frac{1-\xi+v}{2(1-\xi)} \right\rfloor + 1$, we get a violated inequality. From (2.9) we see that as t increases, the violation also increases and equals $1 - \sum_{i=1}^{n} \xi_i$ in the limiting case. Following this argument, one may conclude that the best inequality is the one with t arbitrarily large. However, this conclusion may not be correct because our measure of violation is not normalized properly. Ideally we should find an inequality farthest from the given point. Such a measure can be considered in future studies.

Corollary 2.5.2. *The optimization problem having a linear objective function over the set S (or equivalently over $conv(S)$) can be solved in polynomial time.*

We present a polynomial time algorithm to optimize a linear function over $conv(S)$ in Appendix A.

2.6 Computational Results

We now study the effectiveness of the cuts obtained for S^U by doing computational experiments on cutting stock instances. Recall from Chapter 1 the formulation of cutting stock

Algorithm 2 Separation of the facet defining inequalities of $conv(S)$

Input : A point $(\bar{x}, \bar{y}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$
Output : Decide whether $(\bar{x}, \bar{y}) \in conv(S)$, and if not then provide a facet defining inequality that cuts off $(\bar{x}, \bar{y})$
for $i = 1, \dots, n$ **do**
  if $\bar{x}_i = 0$ **then**
    $\hat{w}_i = 1, \xi_i = 0$
  else if $\bar{x}_i \bar{y}_i > 0$ **then**
    if $4\bar{x}_i r > \bar{y}_i$ **then**
      if $\frac{1}{2} + \frac{\sqrt{\frac{4\bar{x}_i r}{\bar{y}_i} - 1}}{2} > 1$ **then**
        $p = \left\lfloor \frac{1}{2} + \frac{\sqrt{\frac{4\bar{x}_i r}{\bar{y}_i} - 1}}{2} \right\rfloor$
        if $\frac{\bar{x}_i}{2p-1} + \frac{\bar{y}_i p(p-1)}{r(2p-1)} \leq \frac{\bar{x}_i}{2(p+1)-1} + \frac{\bar{y}_i p(p+1)}{r(2(p+1)-1)}$ **then**
          $\hat{w}_i = p$
        else
          $\hat{w}_i = p + 1$
        end if
      else
        $\hat{w}_i = 1$
      end if
    else
      $\hat{w}_i = 1$
    end if
    $\xi_i = \frac{\bar{x}_i}{2\hat{w}_i - 1} + \frac{\bar{y}_i \hat{w}_i(\hat{w}_i - 1)}{r(2\hat{w}_i - 1)}$
  else
    $\xi_i = 0$
  end if
end for
$\xi = \sum_{i \in N} \xi_i$
if $\xi \geq 1$ **then**
  The point $(\bar{x}, \bar{y})$ is feasible to $conv(S)$.
else
  $v = \sum_{i \in N : \bar{y}_i = 0} \bar{x}_i$
  $t = \left\lfloor \frac{1 - \xi + v}{2(1 - \xi)} \right\rfloor + \gamma$, where γ can be taken as any positive integer.
  for $i = 1, \dots, n$ **do**
    if $\bar{x}_i > 0$ and $\bar{y}_i = 0$ **then**
      $\hat{w}_i = t$
    end if
  end for
  The inequality $\sum_{i=1}^{n} \frac{x_i}{2\hat{w}_i - 1} + \frac{y_i \hat{w}_i(\hat{w}_i - 1)}{r(2\hat{w}_i - 1)} \geq 1$ cuts off the point $(\bar{x}, \bar{y})$.
end if

problem (CS). In the following formulation we consider the variable y to be continuous.

$$
\begin{aligned}
\min \sum_{i=1}^{n} & y_i \\
& \sum_{i \in N} x_{ij} y_i \geq d_j, \ \ j \in F, \\
& \sum_{j \in F} l_j x_{ij} \leq L, i \in N, \\
& x_{ij} \in \mathbb{Z}_+, y_i \in \mathbb{R}_+, \forall i \in N, j \in F,
\end{aligned}
$$

where the notation is the same as that in Chapter 1. These instances have stocks of one length L from which n different sizes of finals are to be cut. So, there are n mixed-integer bilinear covering constraints modeling demand satisfaction. The upper bounds $x_{ij} \leq \left\lfloor \frac{L}{l_j} \right\rfloor = \nu_j(\text{say}), \forall i \in N, j \in F$ of the integral variables are implicit from the knapsack constraints present in the formulation. Here, our objective is to minimize the total number of stocks that are used. Since not more than n finals are usually seen in solutions to (CS), we assume $|N| = |F| = n$.

We have selected for our experiments ten instances used in Umetani et al. [109] taken from applications in a chemical fiber company in Japan (Fiber-xx-xxxx), six instances generated by CUTGEN (Gau and Wascher [58]) (CutGen-xx-xx) and five randomly generated instances (Rand-xx). These random instances were generated by fixing L to 1030 and selecting specifc problem size n (denoted as 'xx' in the name). The final lengths l_j were generated randomly between 75 and 600, and d_j between 300 and 5000.

We perform three sets of experiments. In all three we have used PuLP (Mitchell et al. [86]) version 1.6.2 (installed in Python 2.7.12) to model the linear programs and CBC (Forrest et al. [55]) solver to solve them. The system we used to run our code has Linux (Ubuntu 16.04) operating system with 4x Intel(R) Core(TM) i5-3570 CPU@3.40 GHz processor and 8 GB of RAM. All experiments were carried out on a single core.

In our first study we compare the bounds generated by our cuts for $conv\left(S^U\right)$ to those by Tawarmalani et al. [106] for $conv(S)$. In both the cases we consider the facet defining inequalities of each mixed-integer bilinear covering constraint. Adding facet defining inequalities for each mixed-integer bilinear covering constraint together gives a polyhedral relaxation for the actual problem. For each instance, in both the cases, we start our iterations with the facet defining inequalities $\sum_{i=1}^{n} x_{ij} \geq 1$, for all $j \in F$, the bound constraints and the knapsack inequalities, i.e., we start our iterations by solving the following linear program.

$$
\begin{aligned}
\min \ & \sum_{i=1}^{n} y_i \\
\text{s.t.} \ & \sum_{i=1}^{n} x_{ij} \geq 1, \forall j \in F, \\
& 0 \leq x_{ij} \leq \nu_j, \forall i \in N, j \in F, \\
& \sum_{j \in F} l_j x_{ij} \leq L, i \in N, \\
& y \geq 0.
\end{aligned}
\tag{LP-I}
$$

Then, we add violated inequalities (if any) obtained from our separation procedures and resolve (LP-I). This process is continued until we can not find any more violated inequalities, or the number of LPs solved exceeds a predefined limit of 800, or the total time used exceeds two hours. If we can not find any more violated inequalities, then the solution of the current LP lies in the convex hull of the set S^U associated with each of the bilinear constraints. This solution may not be feasible to the original problem (CS).

We run the above experiment in two different settings using facet defining inequalities derived (i) for *conv* $\left(S^U\right)$ and (ii) for *conv*(S). We consider the sets S^U and S by looking at each bilinear covering constraint separately and add one most violated cut for each such constraint using Algorithm 1 and Algorithm 2 respectively. So, we add at most n cuts in every iteration (LP solve) which are not deleted in further iterations. This means, at iteration k, we solve an LP relaxation of the instance with at most $k|F|$ number of linear inequalities in addition to those in (LP-I).

Table 2.1 compares the effects of new cuts for S^U to the cuts derived for S. We observe that cuts for *conv* $\left(S^U\right)$ improve the lower bounds with fewer cuts and in lesser time as compared to cuts for *conv*(S). In the Figures 2.3, 2.4 and 2.5, we present iteration-wise bound comparisons for three instances Fiber-15-5180, CutGen-01-25 and Rand16 respectively. We see that the cuts for *conv* $\left(S^U\right)$ improve bounds faster than those for

$conv(S)$.

We also study the time taken to solve the extended formulations (both LP and MILP) of Section 2.4.4. While the LP defines the set $conv\left(S^U\right)$, the MILP is an even tighter relaxation of cutting stock problem. Recall that the extended formulation has $2n^2 + n$ variables, that means n^2 more variables than the original formulation. Table 2.2 lists the bounds and time taken to solve the two relaxations. We set computational time limit to two hours. We write "7200*" for the instances where this time limit is reached, and for such instances we report the relative gap $\frac{ub-lb}{lb}$ of the MILP. The lb of the MILP is a lower bound for the optimal value of (CS). We also compute an upper bound to optimal solution of (CS) obtained by fixing the variable x to the MILP solution in (CS) and solving a linear program in y only. This bound is reported in the last column ("UB") of Table 2.2.

We see that the extended formulation LP takes much less time compared to the cutting plane algorithm using cuts for $conv\left(S^U\right)$, and even the MILP is often faster than the cut based iterative LP approach. This observation suggests that extended formulation is quite good for these instances when the implied bounds ν on x are small. The extended MILP for randomly generated instances seems to be unusually difficult for the solver. The bounds given by the LP and MILP of the extended formulation are the same for all instances except for Rand10. We do not have an explanation of this phenomenon currently.

Lastly, we consider an exact MILP formulation of (CS). Let $w_{ijh} = 1$ if $x_{ij} = h$ and $w_{ijh} = 0$ otherwise for $i \in N, j \in F$ and $h \in \{0, 1, \ldots, \nu_j\}$. Replacing the terms $w_{ijh}y_i$ with z_{ijh} and using the linear inequalities to model the products $z_{ijh}y_i$ we have the

Table 2.1: Comparison of iterations taken to optimize over the convex hull and the lower bounds obtained. (Here "Iter" means number of LP iterations, and "LB" means Lower Bound obtained after termination, "Cuts" column indicates the number of cuts added, Time is in seconds). A * mark indicates time or iteration limit is reached

Instance	n	Using inequalities for $conv\left(S^U\right)$				Using inequalities for $conv(S)$			
		Iter	Cuts	LB	Time	Iter	Cuts	LB	Time
Fiber10-5180	10	156	1212	27.00	18.23	226	1917	6.88	55.09
Fiber10-9080	10	215	1663	15.00	23.86	223	2045	3.85	46.15
Fiber11-5180	11	118	939	26.00	9.22	288	2673	6.10	89.22
Fiber11-9080	11	463	3151	14.44	137.83	335	2946	3.40	162.3
Fiber14-5180	14	147	1343	22.00	17.27	473	5417	3.34	547.81
Fiber14-9080	14	136	1522	11.00	19.22	476	6211	1.90	658.27
Fiber15-5180	15	335	2350	28.80	98.96	560	7219	3.74	1412.46
Fiber15-9080	15	623	2861	16.00	317.27	800*	8890	2.09	1881.71
Fiber16-5180	16	800*	2393	27.20	282.51	756	9763	5.17	3086.56
Fiber16-9080	16	223	2566	15.11	63.9	723	10330	2.93	2692.61
CutGen01-01	10	211	1359	2.43	30.79	252	2244	1.24	95.33
CutGen01-02	10	235	1888	2.57	43.86	270	2443	0.97	103.32
CutGen01-25	10	180	1238	3.40	21.95	246	2157	0.99	77.18
CutGen01-100	10	194	1276	3.80	18.82	244	2131	1.25	52.82
CutGen02-40	10	186	1292	26.00	22.48	262	2272	10.41	92.60
CutGen02-60	10	186	1322	33.80	24.05	275	2480	10.10	76.32
Rand10	10	64	557	1520.50	3.06	185	1601	697.22	28.86
Rand15	15	114	1272	2122.00	18.15	792	8485	576.69	3786.64
Rand16	16	800*	2293	2724.00	257.04	800*	8780	686.27	4611.77
Rand20	20	173	2698	2250.00	66.91	725	13131	631.02	7200*
Rand25	25	800*	9835	1437.00	3382.22	686	17175	517.15	7200*

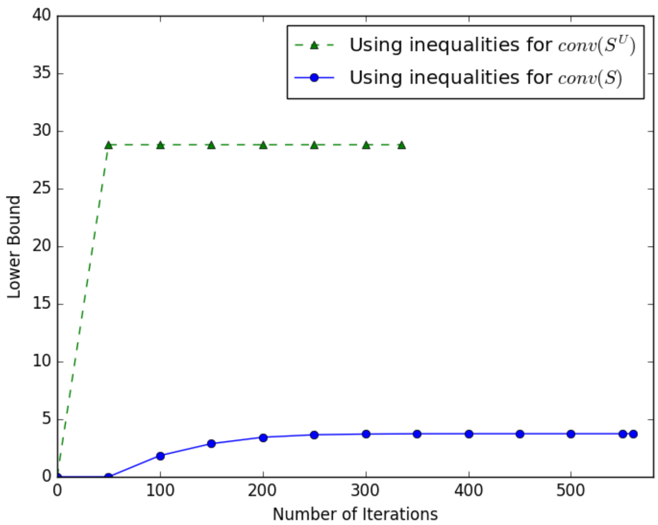

Figure 2.3: Bound comparisons for Fiber-15-5180

Table 2.2: Comparison for Extended Formulation. ("Const." column contains total number of constraints in the extended formulation)

Instance	Const.	Extended LP		Extended MILP			UB
		LB	Time	LB(Rel. Gap)	Time	Nodes	
Fiber10-5180	610	27.00	0.17	27.00	1.64	41	135.00
Fiber10-9080	1010	15.00	0.25	15.00	2.99	61	68.06
Fiber11-5180	792	26.00	0.22	26.00	6.21	91	75.28
Fiber11-9080	1331	14.44	0.37	14.44	3.55	4	45.20
Fiber14-5180	1274	22.00	0.34	22.00	4.32	176	66.00
Fiber14-9080	2114	11.00	0.59	11.00	1.68	1	40.86
Fiber15-5180	1470	28.80	0.57	28.80	5.80	91	88.00
Fiber15-9080	2430	16.00	0.86	16.00	8.54	91	35.02
Fiber16-5180	1648	27.20	0.69	27.20	9.09	615	136.00
Fiber16-9080	2800	15.11	0.74	15.11	7.38	171	82.50
CutGen01-01	1740	2.43	0.68	2.43	18.78	635	13.25
CutGen01-02	1300	2.57	0.55	2.57	9.76	208	10.77
CutGen01-25	1550	3.40	0.44	3.40	6.19	91	17.00
CutGen01-100	1190	3.80	0.32	3.80	7.62	146	19.00
CutGen02-40	2170	26.00	0.67	26.00	13.56	116	149.00
CutGen02-60	1780	33.80	0.53	33.80	6.58	80	117.36
Rand10	320	1520.50	0.11	1557.87(1.46)	7200*	7562839	7407.00
Rand15	720	2122.00	0.27	2122.00(1.89)	7200*	3818886	9688.00
Rand16	832	2724.00	0.30	2724.00(1.58)	7200*	2392670	12899.00
Rand20	1440	2250.00	0.53	2250.00(2.59)	7200*	1758597	13990.00
Rand25	3750	1437.00	1.67	1437.00(3.67)	7200*	681074	15041.25

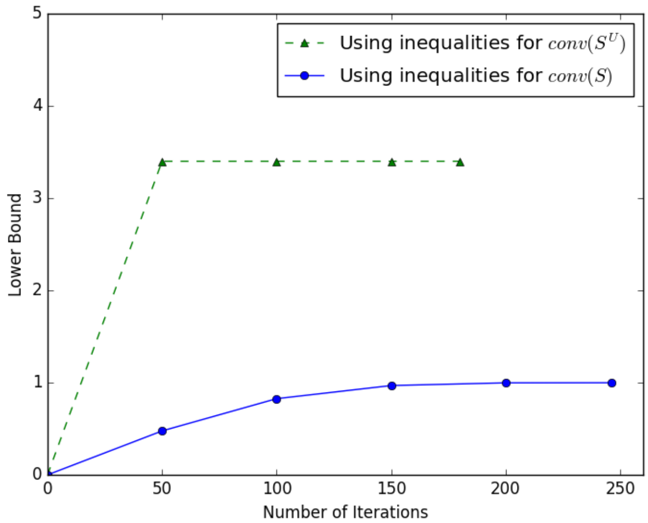

Figure 2.4: Bound comparisons for CutGen-01-25

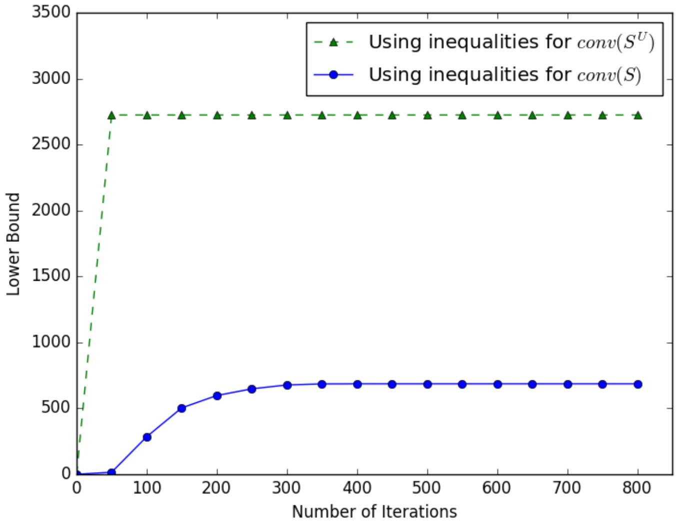

Figure 2.5: Bound comparisons for Rand16

following formulation:

$$
\begin{aligned}
&\min \sum_{i=1}^{n} y_i \\
&\quad \sum_{i \in N} \sum_{h=0}^{\nu_j} h z_{ijh} \geq d_j, \ \ j \in F, \\
&\quad \sum_{j \in F} l_j x_{ij} \leq L, i \in N, \\
&\quad \sum_{h=0}^{\nu_j} w_{ijh} = 1, i \in N, j \in F, \\
&\quad \sum_{h=0}^{\nu_j} h w_{ijh} = x_{ij}, i \in N, j \in F, \\
&\quad z_{ijh} \geq y_i + B w_{ijh}, i \in N, j \in F, h \in \{0, 1, \ldots, \nu_j\}, \\
&\quad z_{ijh} \leq y_i, i \in N, j \in F, h \in \{0, 1, \ldots, \nu_j\}, \\
&\quad z_{ijh} \leq B w_{ijh}, i \in N, j \in F, h \in \{0, 1, \ldots, \nu_j\}, \\
&\quad w_{ijh} \in \{0, 1\}, z_{ijh} \in \mathbb{R}_+, x_{ij} \in \mathbb{Z}_+, y_i \in \mathbb{R}_+, \forall i \in N, j \in F, \in \{0, 1, \ldots, \nu_j\}.
\end{aligned}
\tag{R$_{\mathrm{CS}}$}
$$

The formulation ($\mathrm{R_{CS}}$) is an exact reformulation of (CS) because w_{ijh} are binary. Here B is an upper bound for the variables y. For our experiment, we used the UB value reported in Table 2.2 for B.

Table 2.3: Bounds generated by Binary MILP (R_{CS}) after two hours of computational time.

Instance	LB	UB	Rel. Gap	Nodes	LB ($conv(S^U)$)	LB ($conv(S)$)
Fiber10-5180	9.00	69.80	6.76	754875	27.00	6.88
Fiber10-9080	3.00	39.19	12.06	232134	15.00	3.85
Fiber11-5180	8.67	67.15	6.75	651330	26.00	6.10
Fiber11-9080	2.89	39.79	12.77	216931	14.44	3.40
Fiber14-5180	11.00	48.83	3.44	311882	22.00	3.34
Fiber14-9080	3.14	32.44	9.32	91340	11.00	1.90
Fiber15-5180	9.60	60.73	5.33	220926	28.80	3.74
Fiber15-9080	3.20	34.51	9.78	51267	16.00	2.09
Fiber16-5180	9.50	85.97	8.05	347264	27.20	5.17
Fiber16-9080	3.39	66.67	18.65	78161	15.11	2.93
CutGen01-01	0.62	13.56	20.90	215080	2.43	1.24
CutGen01-02	0.64	10.62	15.52	138215	2.57	0.97
CutGen01-25	1.13	10.69	8.43	119516	3.40	0.99
CutGen01-100	1.27	13.40	9.58	210161	3.80	1.25
CutGen02-40	8.67	117.97	12.61	114179	26.00	10.41
CutGen02-60	11.27	114.66	9.18	136551	33.80	10.10
Rand10	1013.67	7101.50	6.01	1501923	1520.50	697.22
Rand15	2122.00	8834.67	3.16	554079	2122.00	576.69
Rand16	2724.00	10876.00	2.99	1310800	2724.00	686.27
Rand20	2250.00	13608.83	5.05	183486	2250.00	631.02
Rand25	958.00	14200.16	13.82	103304	1437.00	517.15

In Table 2.3, we list both lower and upper bounds to the objective value of (CS) by solving the MILP reformulation (R_{CS}). We also report the lower bounds obtained from cuts for $conv(S^U)$ and $conv(S)$ from the earlier tables for comparison. The time limit was again set to two hours. We observe that the solver reached the time limit for all instances. Further the lower bound at the root relaxation was the same as the lower bound after two hours for all instances. From the table we see that the lower bounds generated by solving (R_{CS}) are smaller than the lower bounds generated by the facets of $conv\left(S^U\right)$ in all instances except Rand15, Rand16 and Rand20 for which they are equal. On the other hand, the bounds from $conv(S)$ are sometimes weaker than that from the MILP (R_{CS}).

2.7 Conclusion

When bounds on integer variables in a bilinear covering set are finite, we are able to obtain the polyhedral description of the convex hull. Even though one can not directly apply the orthogonal disjunctive procedure here, we are still able to compute the convex hull by first creating a suitable relaxation and then applying the procedure. It would be interesting to see if similar procedures can be applied to other restrictions of the set as well. Our examples and experiments show that the new facet defining inequalities of $conv\left(S^U\right)$ improve the bounds as compared to the case when bounds are not considered. The extended formulation for many cutting stock problem can be solved fast, even the MILP can be solved fast. Also the procedure of finding facet defining inequalities to separate a given point from the convex hull is fast.

Our results for the set S^U can be applied in a straight-forward manner to the following set also:

$$S^\delta = \left\{(x,y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n \delta_i x_i y_i \geq r, x \leq u\right\},$$

where $r > 0, u \in \mathbb{N}$ and $\delta_i > 0$ for all $i \in N$. Using our analysis, we can show that the facet defining inequalities of $conv\left(S^\delta\right)$ also can be separated in $O(n)$ time. Some questions related to our results are still open and can be considered as future work. These are discussed in Chapter 5.

Chapter 3

Convex Hull of S^B and an Algorithm to Solve PMP

3.1 Introduction

In this chapter, we consider the following variant of mixed-integer bilinear covering set with box constraint on the integer variables.

$$S^B = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r, l \leq x \leq u \right\},$$

where $r > 0, l \in \mathbb{Z}_+^n, u \in \mathbb{N}^n$ are given. While solving these problems using Branch-and-Cut framework, we deal with nonzero lower bounds, and this motivated us to extend our work of previous chapter to derive tighter linear relaxation. Here, we present the V-Description and H-Description of the set $conv\left(S^B\right)$.

Unlike, the description of $conv\left(S^U\right)$, here we do not use the orthogonal disjunctve approach. Extreme points and extreme rays of $conv\left(S^B\right)$ are first identified. Using these points and rays, we obtain a description of all facet defining inequalities. The number of these inequalities grows exponentially fast with n. An extended formulation of $conv\left(S^B\right)$ which has a structure similar to that of $conv\left(S^U\right)$ is derived. The advantage of a smaller formulation is that, it can be deployed in a Branch-and-Cut algorithm based on linear programming for solving the Pattern Minimization Problem (PMP) [116]. Unlike the existing approach of [110], the new algorithm does not require any column generation or decomposition, and hence is easier to implement.

We implemented the algorithm by overriding branching and cutting-plane functions of Branch-and-Cut solver and performed tests on some real life instances. This algorithm obtains tighter bounds and solves the problem faster than the global optimization solver SCIP [114].

3.2 The Convex Hull Description of S^B

First we derive the description of *conv* $\left(S^B\right)$ as V-Description and then we construct the H-Description with the help of the V-Description.

3.2.1 The V-Description

Proposition 3.2.1. *The set conv* $\left(S^B\right)$ *is a polyhedral set.*

Proof. The proof is similar to that of Proposition 2.4.1 in Chapter 2. □

Proposition 3.2.2. *Let* $(\bar{x}, \bar{y})$ *be an extreme point of conv* $\left(S^B\right)$*. Then* $\bar{x}_t = p_t, \bar{y}_t = \frac{r}{p_t}$ *for some* $t \in N$*, where* $p_t \in \{l_t, \ldots, u_t\} \setminus \{0\}$*, and* $\bar{x}_j \in \{l_j, u_j\}$ *for all* $j \in N, j \neq t$*, i.e.,* $(\bar{x}, \bar{y})$ *is of the following form:*

$$\left(\bar{x}_1, 0, \bar{x}_2, 0, \ldots, \bar{x}_{t-1}, 0, p_t, \frac{r}{p_t}, \bar{x}_{t+1}, 0, \ldots, \bar{x}_n, 0\right)$$

where $p_t \in \{l_t, \ldots, u_t\} \setminus \{0\}$ *for some* $t \in N, \bar{x}_j \in \{l_j, u_j\}, \forall j \in N, j \neq t$*.*

Proof. Since $(\bar{x}, \bar{y})$ is an extreme point of *conv* $\left(S^B\right)$, by the same arguments as in the proof of Theorem 2.4.1 in Chapter 2, there exists $t \in N$ such that $\bar{x}_t = p_t, \bar{y}_t = \frac{r}{p_t}, p_t \in \{l_t, \ldots, u_t\} \setminus \{0\}$ and $\bar{y}_j = 0$ for all $j \in N, j \neq t$, i.e., only one y component of $\bar{y}$ is nonzero. Now we have to show that $\bar{x}_j \in \{l_j, u_j\}$ for all $j \in N, j \neq t$.

If possible, let $\bar{x}_j \in (l_j, u_j)$ for some $j \in N, j \neq t$. Now consider two points $(\bar{x}^1, \bar{y}^1)$ and $(\bar{x}^2, \bar{y}^2)$ in S^B having exactly same components as $(\bar{x}, \bar{y})$ except the j^{th} components of x variable, and $\bar{x}^1_j = l_j$ and $\bar{x}^2_j = u_j$. Therefore, we have

$$(\bar{x}, \bar{y}) = \lambda(\bar{x}^1, \bar{y}^1) + (1 - \lambda)(\bar{x}^2, \bar{y}^2)$$

where $\lambda = \frac{u_j - \bar{x}_j}{u_j - l_j}$. This contradicts the extremality of $(\bar{x}, \bar{y})$.

Moreover, if $\bar{x}_j \in \{l_j, u_j\}$ for $j \neq t$, then we can not write $(\bar{x}, \bar{y})$ as a convex combination of two different points in S^U. This is because, if two such points exist, one of the points' j^{th} component of the variable x has to be more than u_j or less than l_j, neither of which is allowed. □

Note that the recession cone of $conv\left(S^B\right)$ is $\{(x,y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : x = 0\}$. Therefore, form Proposition 3.2.2, we have the V-Description of $conv\left(S^B\right)$. Now we give the H-Description of $conv\left(S^B\right)$ following.

3.2.2 The H-Description

We construct the description of $conv\left(S^B\right)$ without using orthogonal disjunctive procedure. Consider the following collection of columns:

$$\begin{bmatrix} \theta^{(l_1+1)_1}(x_1,y_1) & \theta^{(l_2+1)_2}(x_2,y_2) & \dots & \theta^{(l_n+1)_n}(x_n,y_n) \\ \theta^{(l_1+2)_1}(x_1,y_1) & \theta^{(l_2+2)_2}(x_2,y_2) & \dots & \theta^{(l_n+2)_n}(x_n,y_n) \\ \dots & \dots & \dots & \dots \\ \theta^{(u_1+1)_1}(x_1,y_1) & \theta^{(u_2+1)_2}(x_2,y_2) & \dots & \theta^{(u_n+1)_n}(x_n,y_n) \end{bmatrix} \tag{$\mathrm{M_B}$}$$

where,

$$\theta^{k_i}(x_i,y_i) = \begin{cases} \frac{x_i - l_i}{2k_i - 1 - l_i} + \frac{y_i k_i (k_i - 1)}{r(2k_i - 1 - l_i)}, k_i \in \{l_i + 1, \dots, u_i\}, \\ \frac{y_i u_i}{r}, k_i = u_i + 1, \end{cases} \quad i \in N. \tag{3.1}$$

Note that different columns of ($\mathrm{M_B}$) have different number of elements. Let us define the index set $T_i = \{l_i + 1, \dots, u_i + 1\}, i \in N$. Now define the following set:

$$S^{BR} = \left\{ (x,y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : \sum_{i=1}^{n} \theta^{k_i}(x_i,y_i) \geq 1, (k_i)_{i=1}^n \in \prod_{i=1}^{n} T_i \right\}.$$

The inequalities of S^{BR} are constructed adding n terms taking exactly one term from each column of ($\mathrm{M_B}$) and constraining the sum to greater than or equal to one. Clearly, S^{BR} is a polyhedral set. Define the set $\bar{S}^B = \left\{(x,y) \in S^{BR} : l \leq x \leq u\right\}$. We show that $\bar{S}^B = conv\left(S^B\right)$. First we need two small results for the proof.

Proposition 3.2.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be function of t defined as $f(t) = \frac{t-h}{2k-1-h} + \frac{k(k-1)}{r(2k-1-h)}\frac{r}{t}, t > 0$ where $h, k \in \mathbb{Z}_+, u \in \mathbb{N}$ are given parameters such that $h \leq u$ and $k \geq h+1$. Then the minimum value of f for $t \in \{h, \ldots, u\} \setminus \{0\}$ is 1.*

Proof. Consider t as a continuous variable. Note that for $t > 0, f$ is continuously differentiable. Therefore, we have

$$f'(t) = \frac{1}{2k-1-h} - \frac{k(k-1)}{2k-1-h}\frac{1}{t^2}.$$

Now, $f'(t^*) = 0 \implies t^* = \sqrt{k(k-1)}$. Note that $f''(t) \geq 0$, and therefore, t^* minimizes f. Since, t is restricted to be integer, f attains the minimum value at $\lfloor t^* \rfloor = k-1$ or $\lceil t^* \rceil = k$ and the minimum value is 1. □

Proposition 3.2.4. *Let $(\hat{x}, \hat{y})$ be a solution to the following system of equations,*

$$\frac{x-h}{2k_1-1-h} + \frac{k_1(k_1-1)}{r(2k_1-1-h)}y = \alpha \tag{3.2}$$

$$\frac{x-h}{2k_2-1-h} + \frac{k_2(k_2-1)}{r(2k_2-1-h)}y = \alpha \tag{3.3}$$

where $h, k_1, k_2 \in \mathbb{Z}_+$, $k_1 \geq h+1$, $k_2 \geq k_1+2$ and $0 < \alpha \leq 1$ are given parameters. Then,

$$\frac{\hat{x}-h}{2(k_1+1)-1-h} + \frac{(k_1+1)((k_1+1)-1)}{r(2(k_1+1)-1-h)}\hat{y} < \alpha. \tag{3.4}$$

Proof. The solution to the system of equations (3.2) and (3.3) is,

$$\hat{x} = \alpha\left[\frac{2k_1k_2}{k_1+k_2-1} - 1 - h\right] + h \tag{3.5}$$

$$\hat{y} = \frac{2r\alpha}{k_1+k_2-1} \tag{3.6}$$

Then we have,

$$\frac{\hat{x}-h}{2(k_1+1)-1-h} + \frac{(k_1+1)((k_1+1)-1)}{r(2(k_1+1)-1-h)}\hat{y} = \alpha\left[1 + \frac{2(k_1-k_2+1)}{(2k_1+1-h)(k_1+k_2-1)}\right].$$

Since $k_2 \geq k_1+2$, equation (3.4) holds. □

Theorem 3.2.1. *The set $conv\left(S^B\right)$ is a subset of $\bar{S}^B$.*

Proof. It is sufficient to show that the inequalities of S^{BR} are valid for $conv\left(S^B\right)$. In fact we show that the inequalities of S^{BR} are facet defining inequalities for $conv\left(S^B\right)$. Since, it can be generalized for general n, we prove our result for $n = 2$. The inequalities of S^{BR} for $n = 2$ are of the form $\sum_{i=1}^{2} \theta^{k_i}(x_i, y_i) \geq 1$ for all $(k_i)_{i=1}^{2} \in \prod_{i=1}^{2} T_i$. We consider the following cases:

CASE 1: $l_1+1 \leq k_1 \leq u_1$ and $l_2+1 \leq k_2 \leq u_2$. The corresponding inequality is following:

$$\frac{x_1 - l_1}{2k_1 - 1 - l_1} + \frac{y_1 k_1(k_1 - 1)}{r(2k_1 - 1 - l_1)} + \frac{x_2 - l_2}{2k_2 - 1 - l_2} + \frac{y_2 k_2(k_2 - 1)}{r(2k_2 - 1 - l_2)} \geq 1. \tag{3.7}$$

From Proposition 3.2.2, an extreme point $(\bar{x}, \bar{y})$ of $conv\left(S^B\right)$ when $n = 2$ is of the form either $\left(t, \frac{r}{t}, \bar{x}_2, 0\right)$ where $t \in \{l_1, \ldots, u_1\} \setminus \{0\}$ and $\bar{x}_2 \in \{l_2, u_2\}$, or $\left(\bar{x}_1, 0, t, \frac{r}{t}\right)$ where $t \in \{l_2, \ldots, u_2\} \setminus \{0\}$ and $\bar{x}_1 \in \{l_1, u_1\}$. So, using Proposition 3.2.3, we can see that at $(\bar{x}, \bar{y})$, either $\frac{\bar{x}_1 - l_1}{2k_1 - 1 - l_1} + \frac{\bar{y}_1 k_1(k_1-1)}{r(2k_1-1-l_1)} \geq 1$ or $\frac{\bar{x}_2 - h_2}{2k_2 - 1 - l_2} + \frac{\bar{y}_2 k_2(k_2-1)}{r(2k_2-1-l_2)} \geq 1$. Since, $(\bar{x}, \bar{y})$ is arbitrary, the inequality (3.7) is valid for $conv\left(S^B\right)$.

The points $\left(k_1, \frac{r}{k_1}, l_2, 0\right)$, $\left(k_1 - 1, \frac{r}{k_1 - 1}, l_2, 0\right)$, $\left(l_1, 0, k_2, \frac{r}{k_2}\right)$ and $\left(l_1, 0, k_2 - 1, \frac{r}{k_2 - 1}\right)$ lie in S^B and are affinely independent and the inequality (3.7) is active at these points. Therefore, the inequality (3.7) is a facet defining inequality of $conv\left(S^B\right)$.

CASE 2: $l_1 + 1 \leq k_1 \leq u_1$ and $k_2 = u_2 + 1$. The corresponding inequality is following:

$$\frac{x_1 - l_1}{2k_1 - 1 - l_1} + \frac{y_1 k_1(k_1 - 1)}{r(2k_1 - 1 - l_1)} + \frac{y_2 u_2}{r} \geq 1. \tag{3.8}$$

From Proposition 3.2.2, at any extreme point $(\bar{x}, \bar{y})$ of $conv\left(S^B\right)$, if $\bar{y}_2 > 0$, then $\bar{y}_2 \geq \frac{r}{u_2}$. Again, if $\bar{y}_2 = 0$, then by Proposition 3.2.3, we have $\frac{\bar{x}_1 - l_1}{2k_1 - 1 - l_1} + \frac{\bar{y}_1 k_1(k_1-1)}{r(2k_1-1-l_1)} \geq 1$. Therefore, in either case, the inequality (3.8) is valid for $conv\left(S^B\right)$.

Now, the points $\left(k_1, \frac{r}{k_1}, h_2, 0\right)$, $\left(k_1, \frac{r}{k_1}, u_2, 0\right)$, $\left(k_1 - 1, \frac{r}{k_1 - 1}, l_2, 0\right)$ and $\left(l_1, 0, u_2, \frac{r}{u_2}\right)$ lie in S^B and are affinely independent and the inequality (3.8) is active at these points, and consequently, the inequality (3.8) is a facet defining for $conv\left(S^B\right)$.

CASE 3: $k_1 = u_1 + 1$ and $l_2 + 1 \leq k_2 \leq u_2$. This case is similar as CASE 2.

CASE 4: $k_1 = u_1 + 1$ and $k_2 = u_2 + 1$. The corresponding inequality is following:

$$\frac{y_1 u_1}{r} + \frac{y_2 u_2}{r} \geq 1. \tag{3.9}$$

By similar logic as Case 2, we can show that the inequality (3.9) is valid for $conv\left(S^B\right)$. Again, the points $\left(u_1, \frac{r}{u_1}, h_2, 0\right), \left(u_1, \frac{r}{u_1}, u_2, 0\right), \left(l_1, 0, u_2, \frac{r}{u_2}\right)$ and $\left(u_1, 0, u_2, \frac{r}{u_2}\right)$ lie in S^B and are affinely independent. Therefore, the inequality (3.9) is facet defining for $conv\left(S^B\right)$.

Therefore, all the inequalities of S^{BR} are valid for $conv\left(S^B\right)$, and consequently we have $conv\left(S^B\right) \subseteq \bar{S}^B$. □

Theorem 3.2.2 ([98]). *The set $\bar{S}^B$ is a subset of $conv\left(S^B\right)$.*

Proof. It is sufficient to show that each extreme point of $\bar{S}^B$ lies in S^B, because $\bar{S}^B$ and $conv\left(S^B\right)$ have the same recession cone. To create an extreme point of $\bar{S}^B$, four facet defining inequalities of S^B should intersect at a point and the intersection point should be feasible to $\bar{S}^B$. We consider the following cases.

Case 1: Let the extreme point is constructed by the intersection of four simple facets. Then we must have its y components be zero. Such points can not be feasible for $\bar{S}^B$ as it will violate the inequality $\frac{y_1 u_1}{r} + \frac{y_2 u_2}{r} \geq 1$. So, intersection of simple facets can not give any extreme point of $\bar{S}^B$.

Case 2: Let at least one of the four intersecting facets at an extreme point of $\bar{S}^B$ be a non-simple one, say $\theta^{k_1}(x_1, y_1) + \theta^{k_2}(x_2, y_2) \geq 1$. We show that at the extreme point, either $\theta^{k_1}(x_1, y_1) = 0$ or $\theta^{k_1}(x_1, y_1) = 1$.

Consider the case when $\theta^{k_1}(x_1, y_1) = \alpha$ and $\theta^{k_2}(x_2, y_2) = 1 - \alpha$ where $0 < \alpha < 1$. To find the exact values of x_1, y_1 and α, two more equations are needed. Let $\theta^{k_3}(x_1, y_1) + \theta^{k_4}(x_2, y_2) = 1$ be the second intersecting facet. If $\theta^{k_3}(x_1, y_1) < \alpha$, then $\theta^{k_3}(x_1, y_1) + \theta^{k_2}(x_2, y_2) < 1$. If $\theta^{k_3}(x_1, y_1) > \alpha$, then $\theta^{k_1}(x_1, y_1) + \theta^{k_4}(x_2, y_2) < 1$. i.e., in either case the intersecting point is not feasible to $\bar{S}^B$. Therefore, $\theta^{k_3}(x_1, y_1) = \alpha$ and $\theta^{k_4}(x_2, y_2) = 1 - \alpha$. Also, from Proposition 3.2.3, we know that if $k_3 \geq k_1 + 2$, $\theta^{k_1+1}(x_1, y_1) < \alpha$ and, consequently, $\theta^{k_1}(x_1, y_1) + \theta^{k_2}(x_2, y_2) < 1$. Therefore, $k_3 = k_1 + 1$. Similarly, one can argue that $k_4 = k_2 + 1$. The extreme point defined by system of equations $\theta^{k_1}(x_1, y_1) = \alpha$ and $\theta^{k_1+1}(x_1, y_1) = \alpha$ is $x_1 = \alpha(k_1 - l_1) + l_1$ and $y_1 = \frac{r}{k_1}\alpha$. The third interesting facet has to be either $x_1 = l_1, x_1 = u_1$ or $y_1 = 0$. Therefore, α has to be either 0 or 1, i.e., at an extreme point of $\bar{S}^B$ either $\theta^{(k_1)_1}(x_1, y_1) = 0$ or $\theta^{k_1}(x_1, y_1) = 1$. Now we show that, when $\theta^{k_1}(x_1, y_1) = 0$ or 1, the extreme point that is created lies in S^B.

If $\theta^{k_1}(x_1, y_1) = 0$, then we have $x_1 = l_1$ or u_1 (depending on the value of k_1) and $y_1 = 0$. Further, $\theta^{k_2}(x_2, y_2) = 1$ at the extreme point. So, we need one more inequality to fix the values of x_2 and y_2 at the extreme point.

Let the intersecting facet be $x_2 = l_2(\neq 0)$. If $k_2 \geq l_2 + 2$ then it is easy to verify that $\theta^{k_2}(x_2, y_2) < 1$, and therefore the extreme point violates the inequality $\theta^{k_1}(x_1, y_1) + \theta^{k_2}(x_2, y_2) \geq 1$ as $\theta^{k_1}(x_1, y_1) = 0$, a contradiction. So, we must have $k_2 = l_2 + 1$, and consequently the extreme point in this case is $\left(l_1(\text{or } u_1), 0, l_2, \frac{r}{l_2}\right)$ which lies in S^B. Similarly we can show that if the intersecting facet is $x_2 = u_2$, then it will create the extreme point $\left(l_1(\text{or } u_1), 0, u_2, \frac{r}{u_2}\right)$ of $\bar{S}^B$ which again lies in S^B.

Let the intersecting facet be of the form $\theta^{k_5}(x_1, y_1) + \theta^{k_6}(x_2, y_2) = 1$. Since, $\theta^{k_1}(x_1, y_1) = 0$, we ave $\theta^{k_5}(x_1, y_1) = 0$. This is because, if $\theta^{k_1}(x_1, y_1) > 0$ at the intersecting point, we have $\theta^{k_6}(x_2, y_2) < 1$ and consequently $\theta^{k_1}(x_1, y_1) + \theta^{k_6}(x_2, y_2) < 1$, i.e., the intersecting point is not feasible to $\bar{S}^B$. Without loss of generality, assume that $k_6 \geq k_5$. If $k_6 \geq k_2 + 2$. By Proposition 3.2.4 we know that $\theta^{k_2+1}(x_2, y_2) < 1$. But by the definition of $\bar{S}^B$, we know that $\theta^{k_1}(x_1, y_1) + \theta^{k_2+1}(x_2, y_2) \geq 1$, a contradiction. Therefore, $k_6 = k_2 + 1$ and the extreme point of $\bar{S}^B$ is $x_2 = k_2$ and $y_2 = \frac{r}{k_2}$. Clearly this extreme point lies in S^B.

Since each extreme point of $\bar{S}^B$ lies in S^B and the recession cones of set $\bar{S}^B$ and $conv\left(S^B\right)$ are same, we have $\bar{S}^B \subseteq conv\left(S^B\right)$. □

Corollary 3.2.1. $\bar{S}^B = conv\left(S^B\right)$.

Proof. The proof directly follows from Theorem 3.2.1 and 3.2.2. □

3.2.3 An Extended Formulation of $conv\left(S^B\right)$

Like the description of $conv\left(S^U\right)$, (i.e., $conv\left(S^B\right)$ when $l = 0$) in the previous chapter, here also we give an extended formulation of $conv\left(S^B\right)$ in a straightforward way introducing n new variables, say $w_i, i \in N$. Consider the following set:

$$S^{BE} = \left\{(x, y, w) \in \mathbb{R}_+^{n+n+n} : w_i \leq \theta^{k_i}(x_i, y_i), k_i \in T_i, i \in N, \sum_{i \in N} w_i \geq 1, l \leq x \leq u\right\},$$

where $\theta^{k_i}(x_i, y_i)$ is defined in (3.1) and $T_i = \{l_i + 1, \ldots, u_i + 1\}, i \in N$ which is also defined earlier. The same way as in Proposition 2.4.4, we can show easily that S^{BE}

is an exact extended formulation of $conv\left(S^B\right)$. Note that, whereas the description of $conv\left(S^B\right)$ consists of exponentially many linear constraints, the description of S^{BE} consists of $\sum_{i \in N}(u_i + 1 - l_i) + 1$ number of linear constraints other than the bound constraints which is pseudopolynomial in the input size because of its dependency on the parameters l and u. So, when the value of u is small (as in the case of Pattern Minimization Problem), we can solve the problem efficiently using the extended formulation.

3.3 Solving PMP Using Inequalities for $conv\left(S^B\right)$

In this section we present a Branch-and-Cut algorithmic framework to solve Pattern Minimization Problem (PMP) using the extended formulation of $conv\left(S^B\right)$.

3.3.1 The Mathematical Model

Recall the following mathematical formulations of Pattern Minimization Problem (PMP) and the associated Cutting Stock (or trim-loss) problem (CS) defined on Chapter 1.

$$
\begin{aligned}
\min_{x,y,z} & \sum_{i=1}^{n} z_i \\
\text{s.t.} \ & \sum_{i=1}^{n} x_{ij} y_i \geq d_j, j \in F, \\
& \sum_{j \in F} \mu_j x_{ij} \leq z_i L, i \in N, \\
& y_i \leq v z_i, i \in N, \\
& \sum_{i=1}^{n} y_i \leq \eta, \\
& x_{ij}, y_i \in \mathbb{Z}_+, z_i \in \{0,1\}, i \in N, j \in F,
\end{aligned}
$$

and

$$
\begin{aligned}
\min \sum_{i=1}^{n} & y_i \\
\sum_{i\in N} & x_{ij} y_i \geq d_j,\ \ j \in F, \\
\sum_{j\in F} & \mu_j x_{ij} \leq L, i \in N, \\
& x_{ij} \in \mathbb{Z}_+, y_i \in \mathbb{R}_+, \forall i \in N, j \in F.
\end{aligned}
$$

We determine the value of η solving the above associated Cutting Stock problem (CS). Since in this chapter we are using "l" to denote lower bound on the varibale x, to avoid confusion, in this chapter we use $\mu_j, j \in F$ to denote the length of the finals unlike Chapter 1.

3.3.2 McCormick Relaxation

Recall from Chapter 1 that, for a bilinear term $w = xy$, we have the following McCormick inequalities:

$$
\begin{aligned}
w &\geq u_y x + u_x y - u_x u_y, & w &\geq l_y x + l_x y - l_x l_y, \\
w &\leq l_y x + u_x y - u_x l_y, & w &\leq u_y x + l_x y - l_x u_y,
\end{aligned}
$$

where $x \in [l_x, u_x], y \in [l_y, u_y]$. Now, the McCormick relaxation of (PMP) can be given as:

$$
\begin{aligned}
\min_{x,y,z,w} \quad & \sum_{i=1}^{n} z_i \\
\text{s.t.} \quad & \sum_{i=1}^{n} w_{ij} \geq d_j, j \in F \\
& w_{ij} \geq v x_{ij} + u_j y_i - u_j v, i \in N, j \in F, \\
& w_{ij} \leq v x_{ij}, i \in N, j \in F, \\
& w_{ij} \leq u_j y_i, i \in N, j \in F, \\
& \sum_{j \in F} \mu_j x_{ij} \leq z_i L, i \in N, \\
& y_i \leq v z_i, i \in N, \\
& \sum_{i=1}^{n} y_i \leq \eta, \\
& w_{ij}, x_{ij}, y_i \in \mathbb{R}_+, z_i \in [0,1], i \in N, j \in F.
\end{aligned}
\tag{PMP$_{\text{MC}}$}
$$

Proposition 3.3.1 ([98]). *The optimal value of the problem (PMP$_{MC}$) is given by*

$$\max \left\{ \sum_{j \in F} \frac{\mu_j d_j}{vL}, \frac{1}{v} \frac{d_j}{u_j} : j \in F \right\}.$$

3.3.3 A Branch-and-Cut Algorithm

We solve the PMP using a Branch-and-Cut framework. We start with solving the MILP version of the extended formulation of the convex hull of each bilinear constraint along with the other linear inequalities that are there in the PMP. This extended formulation is a relaxation of the original PMP. If the solution satisfies all the demand constraints, then the solution is an optimal solution to the PMP. Otherwise we branch on the integer variables in the following way.

Let $(\bar{x}, \bar{y}, \bar{z}, \bar{w})$ be an integer optimal solution to the extended MILP for which the demand constraint for the j^{th} item is not satisfied, i.e., an integer infeasible solution to the problem PMP. Then we choose a branching variable x_{ij} and we branch $x_{ij} \leq \bar{x}_{ij}$ and $x_{ij} \geq \bar{x}_{ij}+1$ when $\bar{x}_{ij} < u_j$, and we branch $x_{ij} \leq \bar{x}_{ij}-1$ and $x_{ij} = u_j$ when $\bar{x}_{ij} = u_j$. Then for the children nodes, we add the facet defining inequalities for the updated bounds on the variable x_{ij}. For the node, $x_{ij} \leq \bar{x}_{ij}$ we need to add only one new inequality and for the node $x_{ij} \geq \bar{x}_{ij} + 1$, we need to add $u_j - \bar{x}_{ij} - 1$ (as all the terms of the corresponding columns of ($\mathrm{M_B}$) get updated) new inequalities and then we repeat the procedure. We select branching variables lexicographically for the item for which the demand is not satisfied.

3.3.4 Computational Results

For our experiments, we used a system with Intel(R) Xenon(R) CPU E5-2670 v2 @ 2.50GHz processor with 128 GB of RAM and Linux Debian 8.1 operating system. We did our experiments using CPLEX Python API (Python 3 with CPLEX 12.8). We compare our algorithm with the global optimization solver "Solving Constraint Integer Programs (SCIP)" [114]. All experiments are carried out in a single core.

We did our experiments with twenty instances. Among them, ten instances used in Umetani et al. [109] taken from applications in a chemical fiber company in Japan (Fiber-xx-xxxx) and other ten instances generated by CUTGEN (Gau and Wascher [58]) (CutGen-xx-xx). We determined the values of η for each instance solving the binary reformulation of the accociated cutting stock problem (CS) for two hours using CBC [55] solver.

Instance	n	η	MC	LB-R	LB	UB	Time	Nodes	Solved
Fiber10-5180	10	70	0.50	0.43	2.25	NFS	3600*	417533	No
Fiber10-9080	10	40	0.28	0.42	2	NFS	3600*	517499	No
Fiber11-5180	11	68	0.50	0.38	2.21	NFS	3600*	366448	No
Fiber11-9080	11	40	0.29	0.36	2	NFS	3600*	520649	No
Fiber14-5180	14	61	0.70	0.32	3	3	9.95	3942	Yes
Fiber14-9080	14	33	0.40	0.00	2	2	0.15	1	Yes
Fiber15-5180	15	61	0.38	0.47	2	NFS	3600*	187569	No
Fiber15-9080	15	35	0.22	0.46	1.97	NFS	3600*	334860	No
Fiber16-5180	16	86	0.60	0.39	2	NFS	3600*	139885	No
Fiber16-9080	16	67	0.34	0.28	3	3	935.08	1632145	Yes
CutGen01-01	10	14	0.73	0.26	2	2	4.97	1551	Yes
CutGen01-02	10	11	0.54	0.25	2	2	1.85	364	Yes
CutGen01-25	10	11	0.53	0.33	3	3	60.59	150870	Yes
CutGen01-100	10	14	0.62	0.31	3	3	20.44	38971	Yes
CutGen02-16	10	92	0.49	0.30	2	2	10.52	4514	Yes
CutGen02-32	10	109	0.68	0.28	3	3	57.59	136379	Yes
CutGen02-40	10	118	0.68	0.00	2	2	0.41	2	Yes
CutGen02-50	10	119	0.63	0.29	2	NFS	3600*	423614	No
CutGen02-60	10	115	0.60	0.34	2	2	1.84	258	Yes
CutGen02-64	10	136	0.44	0.31	3	3	67.22	148478	Yes

Table 3.1: Computational results using SCIP. Here 'LB-R' means the lower bound at the root node.

Instance	LB	UB	Nodes	Cuts	Int. Inf.	Time	Solved
Fiber10-5180	4	4	857080	2882875	45433	1208.72	Yes
Fiber10-9080	3	3	83731	423064	2119	127.62	Yes
Fiber11-5180	4	7	2691000	9378412	115135	3600*	No
Fiber11-9080	3	3	17957	103607	742	32.73	Yes
Fiber14-5180	3	3	123	455	0	0.90	Yes
Fiber14-9080	2	2	262	1664	1	1.28	Yes
Fiber15-5180	4	5	1589539	6309819	747	3600*	No
Fiber15-9080	3	3	63470	331390	317	167.83	Yes
Fiber16-5180	4	11	1662411	6833527	3485	3600*	No
Fiber16-9080	3	3	43	310	0	0.90	Yes
CutGen01-01	2	2	368	2645	7	0.97	Yes
CutGen01-02	2	2	365	2391	8	1.02	Yes
CutGen01-25	3	3	6656	51120	106	14.03	Yes
CutGen01-100	3	3	384	2244	0	0.94	Yes
CutGen02-16	2	2	4765	52472	793	14.04	Yes
CutGen02-32	3	3	6827	66369	1249	16.70	Yes
CutGen02-40	2	2	147	1622	1	0.84	Yes
CutGen02-50	3	3	1051	7301	12	2.36	Yes
CutGen02-60	2	2	326	2739	4	1.29	Yes
CutGen02-64	3	3	1187	5874	27	2.06	Yes

Table 3.2: Computational results using our algorithm. 'Int. Inf.' means the number of infeasible integer solution found

We observe form the Tables 3.1 and 3.2 that our algorithm hits the time limit for only three instances, whereas SCIP hits the time limit for eight instances, i.e., Our algorithm solves thirteen out of twenty instances, whereas SCIP solve twelve instances. For those instances, our algorithm hits the time limit, so do for SCIP, and for all such instances our algorithm gives better lower bounds. We also see that SCIP solved only three instances faster than our algorithm, and for rest of the instances, our algorithm is faster. We also observe that McCormick bound is weak for all the instances.

3.4 Conclusion

In this article we derived closed form description of $conv\left(S^B\right)$ without using orthogonal disjunctive procedure, but used the concepts of the derivation of $conv\left(S^U\right)$. Like $conv\left(S^U\right)$, we also provide an extended formulation of $conv\left(S^B\right)$. We used this extended formulation to do our experiments and provided a Branch-and-Cut algorithm to solve PMP. Our proof-of-concept implementation of the algorithm showed that it performs better as compared to an off-the-shelf general purpose global optimization solver. Implementing a faster version of the algorithm and testing on larger instances of PMP can be taken up in the future.

Chapter 4

Facet Defining Inequalities of $conv(S)$ as Disjunctive Cuts

4.1 Introduction

In this chapter, we study the facet defining inequalities of the convex hull of the mixed-integer bilinear covering set

$$S = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r \right\},$$

where $r > 0$. As discussed earlier, the set S is a nonconvex set, even the continuous relaxation R of S defined as

$$R = \left\{ (x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r \right\},$$

is nonconvex for $n \geq 2$. Tawarmalani et al. [106] obtained the convex hull of R with the help of their orthogonal disjunctive procedure. This convex hull, which we call $\hat{R}$, can be described using only one constraint:

$$\hat{R} = conv(R) = \left\{ (x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} \sqrt{\frac{x_i y_i}{r}} \geq 1 \right\}.$$

As discussed earlier, Tawarmalani et al. [106] obtained the description of $conv(S)$ with the help of the orthogonal disjunctive procedure. However, unlike $\hat{R}$, the description of $conv(S)$ consists of countably infinite number of facet defining inequalities.

Recall the definitions of Disjunctions, Split-Cut, Split-Rank etc. for valid inequalities in Section 1.3.3 in Chapter 1. We study the facet defining inequalities of $conv(S)$ using the framework of split disjunctions when applied to $\hat{R}$ in an attempt to find those which might be computationally more useful and easy to obtain. Viewing these facet defining inequalities through the lens of split disjunctions, we see that some of them have split-rank one, and can be obtained easily. Further, when minimizing the objective function of trim-loss problems, these rank-one inequalities give the same bound as the $conv(S)$. Some other facet defining inequalities are seen to have split-rank more than one, but can be obtained using other disjunctions. None of the remaining facet-defining inequalities can be obtained by applying any disjunctive procedure on $\hat{R}$, in fact each of them cuts off an integer point from $\hat{R}$. These inequalities can not be derived from $\hat{R}$, but disjunctions can still be applied on the nonconvex set R to derive them. We observe that all facet generating disjunctions have a similar form.

The problem of finding facet defining inequalities of a general nonconvex mixed-integer set is difficult, and there are no general algorithms for finding all facets of such sets. One has to exploit specific structures and properties of the given set in order to find facets, like is done in the orthogonal disjunctive procedure. A more common approach in these methods (see for example, [17, 80, 102]) is to first find a suitable disjunction, and then obtain an inequality that is valid for each subset of the disjunction.

The principle of obtaining disjunctive inequalities [12, 13, 43, 96], in particular split inequalities, has been quite useful for the case of integer linear optimization. Gomory Mixed Integer inequalities [60, 62], Mixed-Integer Rounding (MIR) [90, 91] inequalities, lift-and-project inequalities [14] and several others are all known to be special cases of split cuts [39]. While some of them are equivalent theoretically, they still provide their own computational advantages and insights.

The general approach of obtaining disjunctive inequalities for integer linear optimization has been extended to the class convex-MINLP consisting of MINLPs whose continuous relaxation is convex. Several studies on theoretical aspects of split inequalities [5, 10, 40, 87, 88] and on using them for solving convex MINLPs [25, 33, 77, 104] have been performed recently. Given the relatively well established foundations of convex integer sets, it is tempting to exploit it for nonconvex MINLPs as well. This is the main motivation for our work.

In general, determining the split-rank for a given linear inequality belongs to the hard class of problems, even for MILP. Determining the bounds on the split-rank is relatively easy. Split-rank of an inequality can be finite or infinite [35]. A valid inequality having finite split-rank indicates that the inequality can be obtained by recursively applying the split cuts finite number of times.

4.2 Few properties of the sets $\hat{R}$ and R

It is easy to see that the set $\hat{R}$ is a closed convex set in the positive orthant. In this section we analyze few more properties of the sets $\hat{R}$ and R that are necessary for our further discussion. Some similar results can be found in [46].

Proposition 4.2.1. *Suppose the following optimization problem has an optimal solution.*

$$\begin{aligned} \min\ & c\sum_{i=1}^{n} x_i + d\sum_{i=1}^{n} y_i \\ s.t.\ & \sum_{i=1}^{n} x_i \le (\ge) k, \\ & \sum_{i=1}^{n} \sqrt{\frac{x_i y_i}{r}} \ge 1. \end{aligned} \tag{P1}$$

Then there exists an optimal solution (x^*, y^*) to (P1) such that only one pair of its component is non zero, i.e., there exists $t \in N$ such that $x_i^* = 0, y_i^* = 0$ for all $i \in N \setminus \{t\}$.

Proof. Since the proof can be easily generalized for any positive integer n, we prove our result for $n = 2$ only. Let $(\bar{x}, \bar{y}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2$ be an optimal solution to the optimization problem. Therefore, we have

$$\sum_{i=1}^{n} \bar{x}_i \le (\ge) k, \text{ and} \tag{4.1}$$

$$\sqrt{\bar{x}_1 \bar{y}_1} + \sqrt{\bar{x}_2 \bar{y}_2} \ge \sqrt{r}. \tag{4.2}$$

The objective value at this point is $c\sum_{i=1}^{2}\bar{x}_i + d\sum_{i=1}^{2}\bar{y}_i$. Now consider the point (x^*, y^*) such that

$$x_1^* = \bar{x}_1 + \bar{x}_2, x_2^* = 0,$$
$$y_1^* = \bar{y}_1 + \bar{y}_2, y_2^* = 0.$$

At this point the objective value is same as the optimal value. Now it is sufficient to show that (x^*, y^*) is feasible for (P1). Clearly, from (4.1) we see that, the point (x^*, y^*) satisfies the first constraint. Now we have

$$\begin{aligned}
x_1^* y_1^* &= (\bar{x}_1 + \bar{x}_2)(\bar{y}_1 + \bar{y}_2) \\
&= \left(\sqrt{\bar{x}_1\bar{y}_1} + \sqrt{\bar{x}_2\bar{y}_2}\right)^2 + \left(\sqrt{\bar{x}_1\bar{y}_2} - \sqrt{\bar{x}_2\bar{y}_1}\right)^2 \\
&\geq r, \text{ using (4.2) and } \left(\sqrt{\bar{x}_1\bar{y}_2} - \sqrt{\bar{x}_2\bar{y}_1}\right)^2 \geq 0 \\
\Rightarrow \sqrt{x_1^* y_1^*} &\geq \sqrt{r}
\end{aligned}$$

This implies that $(x^*, y^*) \in \hat{R}$, and thus feasible for (P1). □

Proposition 4.2.2. *Consider the following optimization problem,*

$$\begin{aligned}
z^* = \min\ & a_k x_1 + b_k y_1 \\
s.t.\ & \sqrt{x_1 y_1} \geq \sqrt{r}, \\
& x_1, y_1 \geq 0.
\end{aligned}$$

where $a_k = \frac{1}{2k-1}, b_k = \frac{k(k-1)}{r(2k-1)}, k \in \mathbb{N} \setminus \{1\}$. *Then the unique optimal solution of the above problem is* $\left(\sqrt{k(k-1)}, \frac{r}{\sqrt{k(k-1)}}\right)$, *and* $\frac{1}{2} < z^* < 1$.

Proof. Note that the above problem is a convex problem as the curve $y_1 = \frac{r}{x_1}$ is strictly convex. Therefore, by elementary KKT condition, we have the unique optimal solution is $\left(\sqrt{k(k-1)}, \frac{r}{\sqrt{k(k-1)}}\right)$.

Therefore, the optimal value is $\frac{\sqrt{k(k-1)}}{2k-1} + \frac{k(k-1)}{r(2k-1)}\frac{r}{\sqrt{k(k-1)}} = \frac{2\sqrt{k(k-1)}}{2k-1} = \sqrt{\frac{4k^2-4k}{4k^2-4k+1}} < 1$.

Now the following function

$$f(v) = \frac{4v^2 - 4v}{4v^2 - 4v + 1}, v \geq 2$$

is continuously differentiable in the domain, and $f'(v) = \frac{8v-4}{(4v^2-4v+1)^2} > 0$ for $v \geq 2$. Therefore, f and hence $\sqrt{f}$ is strictly increasing for $v \geq 2$. Now, $\sqrt{f(2)} = \sqrt{\frac{8}{9}} > \frac{1}{2}$. Thus the result follows. $\square$

Corollary 4.2.1. *Let us consider the problem* $z^* = \min_{(x,y)\in\hat{R}} \sum_{i=1}^{n}(a_{k_i}x_i + b_{k_i}y_i)$ *where* $n \in \mathbb{N}, a_{k_i} = \frac{1}{2k_i-1}, b_{k_i} = \frac{k_i(k_i-1)}{r(2k_i-1)}, k_i \in \mathbb{N}\setminus\{1\}$. *Then* $z^* < 1$.

Proof. Note that the point $\left(0, 0, \cdots, \sqrt{k_i(k_i-1)}, \frac{r}{\sqrt{k_i(k_i-1)}}, \cdots, 0, 0\right)$ is feasible for $\hat{R}$. The objective value at this point is $\frac{\sqrt{k_i(k_i-1)}}{2k_i-1} + \frac{k_i(k_i-1)}{r(2k_i-1)}\frac{r}{\sqrt{k_i(k_i-1)}}$ which is less than one from the proof of Proposition 4.2.2. This implies $z^* < 1$. $\square$

Proposition 4.2.3. *Let $n \geq 2$. Consider the set R along with some additional linear constraints on the variables x. Let us call it R_X. Let $(\bar{x}, \bar{y})$ be an extreme point of the set conv (R_X). Then there exists $t \in N$ such that $\bar{x}_t\bar{y}_t = r, y_i = 0, \forall i \in N, i \neq t$, i.e., only one pair of $(\bar{x}_i, \bar{y}_i), i = 1, \ldots, n$ can have both the non zero value.*

Proof. Let $(\bar{x}, \bar{y})$ be an extreme point of *conv* (R_X), then the point $(\bar{x}, \bar{y})$ lies on the surface $\sum_{i=1}^{n} x_iy_i = r$. If possible, let there exist two pairs of components of $(\bar{x}, \bar{y})$ that are strictly greater than zero. Without loss of generality let $(\bar{x}_1, \bar{y}_1)$ and $(\bar{x}_2, \bar{y}_2)$ have all their components greater than zero. Also let $\bar{x}_1\bar{y}_1 + \bar{x}_2\bar{y}_2 = \alpha$. Without loss of generality let us assume $\bar{x}_1\bar{y}_1 \geq \frac{\alpha}{2}$. Now we have $(\bar{x}, \bar{y}) = \frac{1}{2}(\bar{x}, \bar{y})^1 + \frac{1}{2}(\bar{x}, \bar{y})^2$, where

$$(\bar{x}, \bar{y})^1 = \left(\bar{x}_1, \frac{\alpha}{\bar{x}_1}, \bar{x}_2, 0, \bar{x}_3, \bar{y}_3, \ldots, \bar{x}_n, \bar{y}_n\right) \text{ and}$$
$$(\bar{x}, \bar{y})^2 = \left(\bar{x}_1, 2\bar{y}_1 - \frac{\alpha}{\bar{x}_1}, \bar{x}_2, 2\bar{y}_2, \bar{x}_3, \bar{y}_3, \ldots, \bar{x}_n, \bar{y}_n\right).$$

Clearly $(\bar{x}, \bar{y})^1, (\bar{x}, \bar{y})^2 \in \mathbb{R}_+^n \times \mathbb{R}_+^n$. Since the x components of the points $(\bar{x}, \bar{y}), (\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$ are same. Therefore, the points $(\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$ satisfy the additional linear constraints on x that are present in R_X. Now we have,

$$\bar{x}_1 \frac{\alpha}{\bar{x}_1} + \bar{x}_2 0 + \bar{x}_3 \bar{y}_3 + \ldots + \bar{x}_n \bar{y}_n = \alpha + \bar{x}_3 \bar{y}_3 + \ldots + \bar{x}_n \bar{y}_n$$
$$= \bar{x}_1 \bar{y}_1 + \bar{x}_2 \bar{y}_2 + \bar{x}_3 \bar{y}_3 + \ldots + \bar{x}_n \bar{y}_n \geq r.$$

Again,

$$\begin{aligned} &\bar{x}_1 \left(2\bar{y}_1 - \frac{\alpha}{\bar{x}_1} \right) + \bar{x}_2 2\bar{y}_2 + \bar{x}_3 \bar{y}_3 + \ldots + \bar{x}_n \bar{y}_n \\ &= 2(\bar{x}_1 \bar{y}_1 + \bar{x}_2 \bar{y}_2) - \alpha + \bar{x}_3 \bar{y}_3 + \ldots + \bar{x}_n \bar{y}_n \\ &= \bar{x}_1 \bar{y}_1 + \bar{x}_2 \bar{y}_2 + \bar{x}_3 \bar{y}_3 + \ldots + \bar{x}_n \bar{y}_n \geq r. \end{aligned}$$

Thus, $(\bar{x}, \bar{y})^1$ and $(\bar{x}, \bar{y})^2$ lie in R_X. This shows that $(\bar{x}, \bar{y})$ can not be an extreme point of $conv\,(R_X)$. Therefore, our assumption must be wrong which proves that $\bar{x}_i \bar{y}_i = 0$ for all $i \in N, i \neq t$. We still have to show that $\bar{y}_i = 0$ for all $i \in N, i \neq t$.

Now let $\bar{x}_t \bar{y}_t = r$. If possible, let there exist $j \in N, j \neq t$ such that $\bar{y}_j > 0$. Therefore, using the above arguments, $\bar{x}_j = 0$. Let $\epsilon > 0$ be such that $\bar{y}_j - \epsilon > 0$. Then $(\bar{x}, \bar{y})$ lies in the middle of two points $(\bar{x}, \bar{y})^3$ and $(\bar{x}, \bar{y})^4$ such that $(\bar{x}, \bar{y})^3$ and $(\bar{x}, \bar{y})^4$ have the same components as $(\bar{x}, \bar{y})$ except the j^{th} component of the variable y and $\bar{y}_j^3 = \bar{y}_j - \epsilon$ and $\bar{y}_j^4 = \bar{y}_j + \epsilon$. Since $(\bar{x}, \bar{y})^3, (\bar{x}, \bar{y})^4 \in S$, this contradicts the extremality of $(\bar{x}, \bar{y})$. □

4.3 The facet defining inequalities of $conv(S)$

Recall from Chapter 2, the facet defining inequalities of $conv(S)$ can be constructed using the following collection of columns:

$$\begin{bmatrix} x_1 & x_2 & x_3 & \ldots & x_n \\ a_2 x_1 + b_2 y_1 & a_2 x_2 + b_2 y_2 & a_2 x_3 + b_2 y_3 & \ldots & a_2 x_n + b_2 y_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_k x_1 + b_k y_1 & a_k x_2 + b_k y_2 & a_k x_3 + b_k y_3 & \ldots & a_k x_n + b_k y_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix} \qquad (M)$$

where $a_j = \frac{1}{2j-1}$ and $b_j = \frac{j(j-1)}{r(2j-1)}, j \in \mathbb{N}$ as defined earlier in Chapter 2. Note that the coefficients $a_j, b_j, j \in \mathbb{N}$ are independent of $i \in N$.

To construct the facet defining inequalities for $conv(S)$, we add exactly n terms from (M) taking one term from each column and constraining their sum to greater than or equal to one [106]. Conversely, each facet defining inequality of $conv(S)$ is constructed in this way.

Note that, in general the facet defining inequalities of $conv(S)$ can be written in the following form:

$$\sum_{i \in J_1}(a_{j_1}x_i + b_{j_1}y_i) + \sum_{i \in J_2}(a_{j_2}x_i + b_{j_2}y_i) + \ldots + \sum_{i \in J_p}(a_{j_p}x_i + b_{j_p}y_i) \geq 1, \tag{I$_\mathrm{G}$}$$

where $j_1, j_2, \ldots, j_p$ are different row numbers of (M) for some $p \in \mathbb{N}$. Without loss of generality we can assume $j_1 < j_2 < \ldots < j_p$. The index sets $J_1, J_2, \ldots, J_p$ define a partition on the set N. Note that, for any such partition, we get one facet defining inequality of $conv(S)$ and vice versa.

4.4 Split-rank of the facet defining inequalities of $conv(S)$

In this section, we derive the ranks of the facet defining inequalities of $conv(S)$. We first analyze the simpler cases for $n = 1$, and then we generalize it for any positive integer n.

4.4.1 When $n = 1$

For $n = 1$, we have the set $S = \{(x_1, y_1) \in \mathbb{Z}_+ \times \mathbb{R}_+ : x_1 y_1 \geq r\}$. In this case the convex hull of S can be written as,

$$conv(S) = \{(x_1, y_1) \in \mathbb{R}_+ \times \mathbb{R}_+ : x_1 \geq 1, a_j x_1 + b_j y_1 \geq 1, \forall j \in \mathbb{N} \setminus \{1\}\}$$

where, $a_j x_i + b_j y_i = 1$ is the straight line joining the two points $(x_i, y_i) = \left(j-1, \frac{r}{j-1}\right)$ and $\left(j, \frac{r}{j}\right), \forall j \in \mathbb{N} \setminus \{1\}$. Moreover, in this case we have $\hat{R} = R$.

Lemma 4.4.1. *Let $n = 1$. Consider a point (u, v) on the boundary of $\hat{R}(= R)$, i.e., $uv = r$. Then the point (u, v) is cut off by the facet defining inequality $a_j x_1 + b_j y_1 = \frac{x_1}{2j-1} + \frac{y_1 j(j-1)}{r(2j-1)} \geq 1$ of conv(S) if and only if $u \in (j-1, j)$. In other words, the optimal value of the optimization problem $\min_{(x_1,y_1)\in R(=\hat{R})} a_j x_1 + b_j y_1$ is less than one if and only if $u \in (j-1, j)$.*

Proof. For $j = 1$, we have the facet defining inequality $x_1 \geq 1$ and therefore, the proof is straightforward. For $j \geq 2$, since the facet defining inequality $\frac{x_1}{2j-1} + \frac{y_1 j(j-1)}{r(2j-1)} \geq 1$ is constructed by joining the points $\left(j-1, \frac{r}{j-1}\right)$ and $\left(j, \frac{r}{j}\right)$, and since the curve $y_1 = \frac{r}{x_1}$ is strictly convex in the positive orthant, the result follows. $\square$

In our further discussion, we consider the following convex mixed-integer relaxation $\hat{S}$ of S obtained by adding integer constraints to $\hat{R}$.

$$\hat{S} = \left\{ (x, y) \in \mathbb{Z}^n_+ \times \mathbb{R}^n_+ : \sum_{i=1}^{n} \sqrt{\frac{x_i y_i}{r}} \geq 1 \right\}.$$

We study the facet defining inequalities of *conv*(S) as split cuts for $\hat{S}$ and determine their split-ranks.

Theorem 4.4.1. *For $n = 1$, every facet defining inequality of conv(S) is a rank-one split inequality for $\hat{S}$.*

Proof. Consider the facet defining inequality $x_1 \geq 1$ of *conv*(S). The point $\left(\frac{1}{2}, 2r\right) \in \hat{R}$ which is cut off by this facet defining inequality of *conv*(S), and therefore it can not have split-rank zero. Now, the inequality $x_1 \geq 1$ is valid for both the sets $\hat{R} \cap [x_1 \leq 0] (= \phi)$ and $\hat{R} \cap [x_1 \geq 1]$, i.e., the inequality $x \geq 1$ is valid for the disjunction $[x_1 \leq 0] \vee [x_1 \geq 1]$. Therefore, the split-rank of this inequality is one.

Now consider a facet defining inequality $a_j x_1 + b_j y_1 = \frac{x_1}{2j-1} + \frac{y_1}{r(2j-1)} \geq 1, j \in \mathbb{N}, j \neq 1$ of *conv*(S). Since $\frac{2j-1}{2} \in (j-1, j)$, by Lemma 4.4.1, the point $\left(\frac{2j-1}{2}, \frac{2r}{2j-1}\right) \in \hat{R}$ is cut off by this inequality. Therefore, it has split-rank at least one. Since the facet $a_j x_1 + b_j y_1 = 1$ is constructed by joining the two points $\left(j-1, \frac{r}{j-1}\right)$ and $\left(j, \frac{r}{j}\right)$ and the curve $y_1 = \frac{r}{x_1}$ is concave, the inequality $a_j x_1 + b_j y_1 \geq 1$ is valid for both the sets $\hat{R} \cap [x_1 \leq j-1]$ and $\hat{R} \cap [x_1 \geq j]$, and consequently it is valid for the disjucntion $[x_1 \leq j-1] \vee [x_1 \geq j]$, and its split rank is 1. The following figure illustrates this geometrically. $\square$

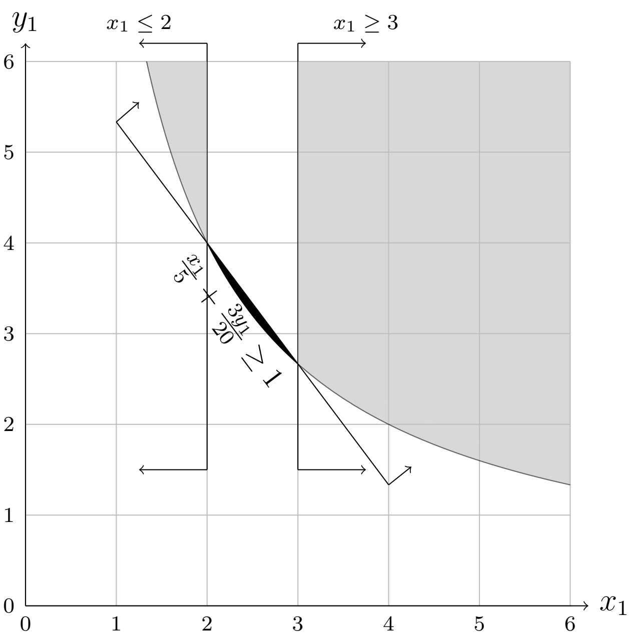

Figure 4.1: The split disjunction $[x \leq 2] \vee [x_1 \geq 3]$ and the split cut $\frac{x_1}{5} + \frac{3y_1}{20} \geq 1, r = 8$

4.4.2 Split-ranks for higher dimension

In this section we discuss about the split ranks for general positive integer n.

Proposition 4.4.1. *Any facet defining inequality of conv(S) that is constructed using exactly one row of (M), i.e., of the form $a_j \sum_{i=1}^n x_i + b_j \sum_{i=1}^n y_i \geq 1$ for any $j \in \mathbb{N}$ is a rank-one split inequality for $\hat{S}$.*

Proof. The point $(x, y) = \left(\frac{2j-1}{2}, \frac{2r}{2j-1}, 0, 0, \ldots, 0, 0\right)$ lies in $\hat{R}$ but violates the given inequality for any $j \in \mathbb{N}$. Therefore, this inequality has split-rank at least one. Consider the disjunction $[\sum_{i=1}^n x_i \leq j-1] \vee [\sum_{i=1}^n x_i \geq j]$ and the following two optimization problems.

$$\begin{aligned} \min_{(x,y)\in\hat{R}} \ & a_j \sum_{i=1}^n x_i + b_j \sum_{i=1}^n y_i \\ \text{s.t.} \ & \sum_{i=1}^n x_i \geq j, \end{aligned} \qquad \begin{aligned} \min_{(x,y)\in\hat{R}} \ & a_j \sum_{i=1}^n x_i + b_j \sum_{i=1}^n y_i \\ \text{s.t.} \ & \sum_{i=1}^n x_i \leq j-1. \end{aligned}$$

Consider the first optimization problem. From Proposition 4.2.1, there exists an optimal solution say $(\bar{x}, \bar{y})$ and an index $t \in N$ such that $x_i = 0, y_i = 0$ for all $i \in N \setminus \{t\}$. Because of symmetry, we assume $t = 1$. Therefore, the problem reduces to the following optimization problem

$$\begin{aligned}\min_{(x_1,y_1)\in\hat{R}} \ & a_j x_1 + b_j y_1 \\ \text{s.t. } \ & x_1 \geq j.\end{aligned}$$

From Lemma 4.4.1 the optimal value of this optimization problem is at least one. In fact the optimal value is exactly one as $(j, \frac{r}{j})$ is a feasible point with objective value one. Thus the inequality $a_j \sum_{i=1}^n x_i + b_j \sum_{i=1}^n y_i \geq 1$ is valid for $\hat{R} \cap [\sum_{i=1}^n x_i \geq j]$.

Similarly we can show that the inequality $a_j \sum_{i=1}^n x_i + b_j \sum_{i=1}^n y_i \geq 1$ is valid for $\hat{R} \cap [\sum_{i=1}^n x_i \leq j-1]$. Consequently it is rank-one split inequality for $\hat{S}$. □

The following results give us lower bound on the split-ranks for rest of the facet defining inequalities.

Theorem 4.4.2. *Consider a facet defining inequality of conv(S) that is constructed using two or more rows of (M). For any such inequality, there does not exist any split disjunction $(\pi, \pi_0) \in \mathbb{Z}^{n+1}, \pi \neq 0$ of $\hat{S}$ for which it is valid.*

Proof. Without loss of generality we assume that two different rows of (M) are used for the variables with the first two indices, i.e., we consider the facet defining inequality

$$a_j x_1 + b_j y_1 + a_k x_2 + b_k y_2 + \sum_{i=3}^{n} (a_{p_i} x_i + b_{p_i} y_i) \geq 1 \tag{$\mathrm{I_S}$}$$

of *conv*(S) where $j \neq k$. Since the set $\hat{R}$ lies entirely in the positive orthant, it is sufficient to consider $\pi_0 \geq 0$ and those π that have at least one positive component. Consider the two optimization problems:

$$\begin{aligned}\min_{(x,y)\in\hat{R}} \ & \left[a_j x_1 + b_j y_1 + a_k x_2 + b_k y_2 + \sum_{i=3}^{n} (a_{k_i} x_i + b_{k_i} y_i)\right] \\ \text{s.t. } \ & \pi^T x \leq \pi_0,\end{aligned} \tag{$Q_{\leq}$}$$

and

$$\min_{(x,y)\in\hat{R}} \left[a_j x_1 + b_j y_1 + a_k x_2 + b_k y_2 + \sum_{i=3}^{n} (a_{k_i} x_i + b_{k_i} y_i)\right]$$
$$\text{s.t. } \pi^T x \geq \pi_0 + 1. \tag{$Q_\geq$}$$

We show that at least one of the above problems has optimal value strictly less than one. We consider the following cases.

Case A: When $\pi_1 \leq 0$, the point $\left(\frac{2j-1}{2}, \frac{2r}{2j-1}, 0, 0, \ldots, 0, 0\right)$ is feasible for ($Q_\leq$) with objective value $a_j \frac{2j-1}{2} + b_j \frac{2r}{2j-1}$. Since $\frac{2j-1}{2} \in (j-1, j)$, by Lemma 4.4.1, the objective value is strictly less than one, and so is the optimal value of ($Q_\leq$).

Case B: When $\pi_2 \leq 0$, we can similarly show that ($Q_\leq$) has optimal value less than one.

Therefore, the inequality ($\mathrm{I_S}$) is not valid for ($Q_\leq$) for both the above cases.

Case C: Now the remaining case is when π_1 and π_2 are both positive integers. Suppose that one of the following relations holds true.

$$\pi_1\sqrt{j(j-1)} \leq \pi_0, \tag{4.3}$$
$$\pi_2\sqrt{k(k-1)} \leq \pi_0, \tag{4.4}$$
$$\pi_1\sqrt{j(j-1)} \geq \pi_0 + 1, \tag{4.5}$$
$$\pi_2\sqrt{k(k-1)} \geq \pi_0 + 1. \tag{4.6}$$

If (4.3) is true then clearly the point $\left(\sqrt{j(j-1)}, \frac{r}{\sqrt{j(j-1)}}, 0, 0, \ldots, 0, 0\right)$ is feasible for ($Q_\leq$) and by Proposition 4.2.2 and its corollary, the objective value at this point is strictly less than one. Again, if (4.4) is true then by the same arguments, ($Q_\leq$) has optimal value less than one. Similarly, using the same arguments, the optimal value of ($Q_\geq$) is less than one when (4.5) or (4.6) hold.

Finally, suppose none of the above four relations hold. Therefore we have,

$$\pi_0 < \pi_1\sqrt{j(j-1)} < \pi_0 + 1, \text{ and}$$
$$\pi_0 < \pi_2\sqrt{k(k-1)} < \pi_0 + 1.$$

Therefore both the values of π_1 and π_2 can not be one from Proposition C.0.1 in

Appendix C. Now by Proposition C.0.2 in Appendix C we have,

$$j - 1 < \frac{\pi_0 + 1}{\pi_1} \leq j \tag{4.7}$$

$$k - 1 < \frac{\pi_0 + 1}{\pi_2} \leq k \tag{4.8}$$

$$j - 1 \leq \frac{\pi_0}{\pi_1} < j \tag{4.9}$$

$$k - 1 \leq \frac{\pi_0}{\pi_2} < k \tag{4.10}$$

Since, $\pi_1, \pi_2 \in \mathbb{N}$, not both equal to one, at least one of the values of $\frac{\pi_0+1}{\pi_1}, \frac{\pi_0+1}{\pi_2}, \frac{\pi_0}{\pi_1}$ and $\frac{\pi_0}{\pi_2}$ must be non-integral.

If $\frac{\pi_0}{\pi_1}$ is non-integral, then we have $j - 1 < \frac{\pi_0}{\pi_1} < j$ form (4.9). Now, the point $\mathcal{L}\left(1, \frac{\pi_0}{\pi_1}, \frac{r\pi_1}{\pi_0}\right) = \left(\frac{\pi_0}{\pi_1}, \frac{r\pi_1}{\pi_0}, 0, 0, \ldots, 0, 0\right)$ is feasible for the optimization problem ($Q_{\leq}$). Since at this point exactly one pair of components is positive, Lemma 4.4.1 is applicable, and the objective value at this point is strictly less than one. Similarly, the point $\mathcal{L}\left(2, \frac{\pi_0}{\pi_2}, \frac{r\pi_2}{\pi_0}\right) = \left(0, 0, \frac{\pi_0}{\pi_2}, \frac{r\pi_2}{\pi_0}, 0, 0, \ldots, 0, 0\right)$ is feasible for the optimization problem ($Q_{\leq}$) with objective value strictly less than one if $\frac{\pi_0}{\pi_2}$ is non-integral.

Similarly, if $\frac{\pi_0+1}{\pi_1}$ $\left(\text{or } \frac{\pi_0+1}{\pi_2}\right)$ is non-integral, from (4.7) (or (4.8)) and Lemma 4.4.1, we can say that the optimization problem ($Q_{\geq}$) has optimal value strictly less than one.

Therefore, for all the possible cases there does not exist any split disjunction (π, π_0) for which the facet defining inequality ($\mathrm{I_S}$) is valid. □

Corollary 4.4.1. *Any facet defining inequality of conv(S) that is constructed using two or more number of rows of (M) is not a rank-one split cut for $\hat{S}$.*

Proof. Since a facet defining inequality for *conv*(S) can not be expressed as a linear combination of any other valid inequalities for *conv*(S), it follows from Theorem 4.4.2 that its split-rank is at least two. □

4.5 Disjunctions for the facet defining inequalities

In the proof of Theorem 4.4.2, we showed that there does not exist any split disjunction for a class of inequalities that are valid. In this section we show that there exist facet defining inequalities of *conv*(S) that can be derived using some other more general disjunctions on the set $\hat{S}$. Furthermore, no other facet defining inequality, besides the above two types can be derived by disjunctive procedure on $\hat{S}$. Then we derive a closed convex relaxation from which any given facet defining inequality of $conv(S)$ can be derived by the disjunctive procedure.

Proposition 4.5.1. *A facet defining inequality of $conv(S)$ that is constructed using two rows of (M), one of which is the first row, is a disjunctive cut for $\hat{S}$.*

Proof. Such a facet defining inequalities of $conv(S)$ is of the following form:

$$\sum_{i\in J} x_i + \sum_{i\in K} (a_k x_i + b_k y_i) \geq 1 \tag{I$_\mathrm{D}$}$$

for some $k \in \mathbb{N}, k \neq 1$, where $J \cup K = N, J \cap K = \phi, J \neq \phi$ and $K \neq \phi$. We show that ($\mathrm{I_D}$) is valid for the disjunction $\left[\sum_{i\in J} x_i \geq 1\right] \vee \left[\sum_{i\in J} x_i \leq 0, \sum_{i\in K} x_i \geq k\right] \vee \left[\sum_{i\in J} x_i \leq 0, \sum_{i\in K} x_i \leq k-1\right]$ applied to $\hat{S}$.

Clearly the disjunction is valid. Like Proposition 4.4.1, we consider each atom separately. Consider the optimization problem:

$$\begin{aligned} \min_{(x,y)\in\hat{R}} \ & \sum_{i\in J} x_i + \sum_{i\in K} (a_k x_i + b_k y_i) \\ \text{s.t.} \ & \sum_{i\in J} x_i \geq 1. \end{aligned}$$

Since $\sum_{i\in J} x_i \geq 1$ is a constraint, the optimal value has to be at least one. Therefore, the inequality ($\mathrm{I_D}$) is valid for $\hat{R} \cap \left[\sum_{i\in J} x_i \geq 1\right]$. Now consider the following optimization problem:

$$\begin{aligned} \min_{(x,y)\in\hat{R}} \ & \sum_{i\in J} x_i + \sum_{i\in K} (a_k x_i + b_k y_i) \\ \text{s.t.} \ & \sum_{i\in J} x_i \leq 0, \\ & \sum_{i\in K} x_i \geq k. \end{aligned}$$

Clearly $x_i = 0$ for all $i \in J$. Using the same logic as in the proof of Proposition 4.4.1 (treating it as $n = |K|$) it is clear that for any $i \in K, \mathcal{L}(i, k, \frac{r}{k})$ is an optimal solution with optimal value one, and consequently ($\mathrm{I_D}$) is valid for $\hat{R} \cap \left[\sum_{i \in J} x_i \leq 0, \sum_{i \in K} x_i \geq k\right]$.

Finally using the proof of Proposition 4.4.1 again, we can show that ($\mathrm{I_D}$) is valid for $\hat{R} \cap \left[\sum_{i \in J} x_i \leq 0, \sum_{i \in K} x_i \leq k-1\right]$. □

The following result shows that there exists facet defining inequality of $conv(S)$ that can not be derived by any disjunctive procedure on $\hat{S}$. In fact we show that many of the facet defining inequalities of $conv(S)$ are not valid for $\hat{S}$.

Proposition 4.5.2. *Let (I) be a facet defining inequality of $conv(S)$ constructed form a set of rows Γ from (M). If there exist two distinct $j, k \in \Gamma$ and $j \neq 1, k \neq 1$, then (I) is not valid for $\hat{S}$.*

Proof. We prove the result for the case when the inequality is constructed taking j^{th} row for first column and k^{th} row for the second column such that $j, k \geq 2, j \neq k$. Without loss of generality we assume $j < k$. The proof for the general case is similar. Now we have the inequality:

$$a_j x_1 + b_j y_1 + a_k x_2 + b_k y_2 + \sum_{i=3}^{n} (a_{p_i} x_i + b_{p_i} y_i) \geq 1, \tag{I}$$

where $a_j = \frac{1}{2j-1}, b_j = \frac{j(j-1)}{r(2j-1)}, a_k = \frac{1}{2k-1}, b_k = \frac{k(k-1)}{r(2k-1)}, a_{p_i} = \frac{1}{2p_i-1}, b_{p_i} = \frac{p_i(p_i-1)}{r(2p_i-1)}$. Consider the following point:

$$(\bar{x}, \bar{y}) = \left(j-1, \frac{r(j-1)b_k^2}{\left((j-1)b_k + b_j^2\right)^2}, 1, \frac{r b_j^2}{\left((j-1)b_k + b_j^2\right)^2}, 0, 0, \ldots, 0, 0 \right)$$

Clearly the point $(\bar{x}, \bar{y})$ lies in $\hat{S}$. Therefore it can not be cut off by applying any disjunctive procedure on $\hat{S}$. We will be done if we can show that the inequality (I) cuts off $(\bar{x}, \bar{y})$, i.e., if

$$\frac{j-1}{2j-1} + \frac{r b_j b_k^2 (j-1)}{\left((j-1)b_k + b_j^2\right)^2} + \frac{1}{2k-1} + \frac{r b_k b_j^2}{\left((j-1)b_k + b_j^2\right)^2} < 1.$$

$$
\begin{aligned}
\text{LHS} &= \frac{j-1}{2j-1} + \frac{rb_k b_j\left((j-1)b_k + b_j^2\right)}{\left((j-1)b_k + b_j^2\right)^2} + \frac{1}{2k-1} \\
&= \frac{j-1}{2j-1} + \frac{r}{\frac{j-1}{b_j} + \frac{1}{b_k}} + \frac{1}{2k-1} \\
&= \frac{j-1}{2j-1} + \frac{1}{\frac{2j-1}{j} + \frac{2k-1}{k(k-1)}} + \frac{1}{2k-1} \\
&= 1 - \frac{1}{2-\frac{1}{j}} + \frac{1}{2-\frac{1}{j}+\frac{2k-1}{k(k-1)}} + \frac{1}{2k-1} \\
&= 1 - \frac{\frac{2k-1}{k(k-1)}}{\left(2-\frac{1}{j}\right)\left(2-\frac{1}{j}+\frac{2k-1}{k(k-1)}\right)} + \frac{1}{2k-1} \\
&< 1 + \frac{1}{2k-1} - \frac{\frac{2k-1}{k(k-1)}}{\left(2-\frac{1}{k}\right)\left(2-\frac{1}{k}+\frac{2k-1}{k(k-1)}\right)}, \text{ since } j<k \\
&= 1.
\end{aligned}
$$

□

Thus, applying disjunctive inequalities (or any other valid inequalities) of $\hat{S}$ is not sufficient to obtain all the facet defining inequalities of $conv(S)$. In order to obtain the facet defining inequalities are disjunctive inequalities, we use the following approach. Recall the from ($\mathrm{I_G}$) of the facet defining inequalities of $conv(S)$.

Theorem 4.5.1. *The facet defining inequality (I_G) of $conv(S)$ is valid for the following disjunction on the nonconvex set R.*
$\left[\sum_{i\in J_1} x_i \geq j_1\right] \vee \left[\sum_{i\in J_1} x_i \leq j_1 - 1, \sum_{i\in J_2} x_i \geq j_2\right] \vee \cdots \vee$
$\left[\sum_{i\in J_1} x_i \leq j_1 - 1, \sum_{i\in J_p} x_i \geq j_p\right] \vee$
$\left[\sum_{i\in J_1} x_i \leq j_1 - 1, \sum_{i\in J_2} x_i \leq j_2 - 1, \ldots, \sum_{i\in J_p} x_i \leq j_p - 1\right].$

Proof. Clearly the disjunction in the statement of the theorem is valid. We prove our result for $n = 2$. It can be easily generalized to any $n \geq 2$. For $n = 2$, the inequality ($\mathrm{I_G}$) can be given as:

$$a_j x_1 + b_j y_1 + a_k x_2 + b_k y_2 \geq 1 \tag{4.11}$$

where $j, k \in \mathbb{N}, j < k$ (assuming $j_1 = j$ and $j_2 = k$). We have to show that the inequality is valid for the following disjunction:

$$[x_1 \geq j] \vee [x_1 \leq j-1, x_2 \geq k] \vee [x_1 \leq j-1, x_2 \leq k-1]\,.$$

Case 1: Suppose $j = 1$. Therefore $a_j = 1$ and $b_j = 0$. Consider the global optimization problem:

$$\begin{aligned} \min_{x,y}\ & x_1 + a_k x_2 + b_k y_2 \\ \text{s.t. }\ & x_1 y_1 + x_2 y_2 \geq r, \\ & x_1 \geq 1, x_i, y_i \geq 0, i = 1, 2. \end{aligned}$$

Since we have $x_1 \geq 1$, $(1, r, 0, 0)$ is an optimal solution with optimal value 1. Next consider the global optimization problem:

$$\begin{aligned} \min_{x,y}\ & x_1 + a_k x_2 + b_k y_2 \\ \text{s.t. }\ & x_1 y_1 + x_2 y_2 \geq r, \\ & x_1 \leq 0, \\ & x_2 \geq k, \\ & x_i, y_i \geq 0, i = 1, 2. \end{aligned}$$

Since $x_1 = 0$, the problem reduces the $n = 1$ case. It is clear to see that $(0, 0, k, \frac{r}{k})$ is an optimal solution with optimal value 1. In an exactly similar way we can show that the optimal value is 1 for $R \cap [x_1 \leq 0, x_2 \leq k-1]$ also. Thus the inequality (4.11) is valid for all the three atoms.

Case 2: When $j \geq 2$. We consider the global optimization problem:

$$\begin{aligned} \min_{x,y}\ & a_j x_1 + b_j y_1 + a_k x_2 + b_k y_2 \\ \text{s.t. }\ & x_1 y_1 + x_2 y_2 \geq r, \\ & x_1 \geq j, \\ & x_i, y_i \geq 0, i = 1, 2. \end{aligned}$$

The point $\left(j, \frac{r}{j}, 0, 0\right)$ is feasible with objective value one. Since the objective function is linear, it is equivalent to optimize over the convex hull of the feasible region of the above problem. Let $(\bar{x}, \bar{y})$ be an an extreme point optimal solution. Therefore either $\bar{x}_1\bar{y}_1 = r$ or $\bar{x}_2\bar{y}_2 = r$ by Proposition 4.2.3. Suppose $\bar{x}_1\bar{y}_1 = r$. Then $\bar{x}_1$ can not be more than j because the value of $a_j\bar{x}_1 + b_j\bar{y}_1$ will be strictly greater than one (Lemma 4.4.1).

If $\bar{x}_2\bar{y}_2 = r$, then by Theorem 4.2.3 we have $y_1 = 0$. At this point the objective value is $\frac{t}{2t-1} + a_k\bar{x}_2 + b_k\bar{y}_2$ for some $t \geq j$. Since we are minimizing the objective function, from Proposition 4.2.2 and its corollary, the minimum value of $a_kx_2 + b_ky_2$ subject to the given constraints will be more than $\frac{1}{2}$ as $k \geq 2$. Also $\frac{t}{2t-1} > \frac{1}{2}$, since $t \geq j \geq 2$. Therefore, the objective value is more than one. Therefore, the optimal value of the above optimization problem is one and the inequality (4.11) is valid for $R \cap [x_1 \geq j]$.

We can show similarly that the inequality is valid for $R \cap [x_2 \geq k]$ and consequently for its subsets. Since $[x_1 \leq j-1, x_2 \geq k]$ is a subset of $[x_2 \geq k]$, the inequality is valid for the set $R \cap [x_1 \leq j-1, x_2 \geq k]$.

Finally consider the global optimization problem:

$$
\begin{aligned}
\min_{x,y}\ & a_jx_1 + b_jy_1 + a_kx_2 + b_ky_2 \\
\text{s.t. } & x_1y_1 + x_2y_2 \geq r, \\
& x_1 \leq j-1, \\
& x_2 \leq k-1, \\
& x_i, y_i \geq 0, i = 1, 2.
\end{aligned}
$$

Let $(\bar{x}, \bar{y})$ be an extreme point optimal solution of the convex hull of the feasible region. Therefore, either $\bar{x}_1\bar{y}_1 = r$ or $\bar{x}_2\bar{y}_2 = r$ (by Proposition 4.2.3). Suppose $\bar{x}_1\bar{y}_1 = r$. If $\bar{x}_1 < j-1$, then by Lemma 4.4.1, the value of $a_j\bar{x}_1 + b_j\bar{y}_1$ is strictly greater than one, and therefore, the point $\left(j-1, \frac{r}{j-1}, 0, 0\right)$ gives the least objective value with objective value one. If $\bar{x}_2\bar{y}_2 = r$, then by the same logic the point $\left(0, 0, k-1, \frac{r}{k-1}\right)$ gives the least objective value with objective value one. Therefore, the optimal solution of the above optimization problem is one. Thus the inequality (4.11) is valid for all the nonconvex atoms. $\square$

Let us now define the set $S_{j_1 j_2 \dots j_p}$ to be the closure of the convex hull of the unions of the atoms of Theorem 4.5.1.

$$S_{j_1 j_2 \dots j_p} = cl\left(conv\left(\bigcup_{q=1}^{p+1} S_{J_q}\right)\right),$$

where $S_{J_q}, q = 1, \dots, p+1$ is defined below:

$$S_{J_q} = \begin{cases} S_C \cap \left[\sum_{i \in J_t} x_i \geq j_t\right], q = t, \\ S_C \cap \left[\sum_{i \in J_1} x_i \geq j_q, \sum_{i \in J_t} x_i \leq j_t - 1\right], q = 1, \dots, p, q \neq t, \\ S_C \cap \left[\sum_{i \in J_1} x_i \leq j_1 - 1, \sum_{i \in J_2} x_i \leq j_2 - 1, \dots, \sum_{i \in J_p} x_i \leq j_p - 1\right], q = p+1, \end{cases}$$

where $t \in \{1, \dots, p\}$ such that $j_t < j_q$ for all $q \in \{1, \dots, p\} \setminus \{t\}$. Now we have the following result.

Corollary 4.5.1. *The facet defining inequality (I_G) of $conv(S)$ can be constructed using disjunctive procedure on the closed convex set $S_{j_1 j_2 \dots j_p}$.*

Corollary 4.5.2. *Consider the set S_{CC} defined below:*

$$S_{CC} = \bigcap_{\substack{j_1,\dots,j_p \in \mathbb{N}, for\ some\ p \in N, \\ \forall J_q \subseteq N, \cup_{q=1}^{p} J_q = N, \\ For\ q \neq s, J_q \cap J_s = \phi, \forall q,s=1,\dots,p.}} S_{j_1 j_2 \dots j_p}.$$

Then $S_{CC} = conv(S)$.

Proof. The set S_{CC} is constructed intersecting $S_{j_1 j_2 \dots j_p}$ over all possible partitions of the index set N. Since each facet defining inequality of $conv(S)$ is associated with some partition on the index set N and vice versa, any facet defining inequality of $conv(S)$ is valid for the set S_{CC}.

Since the set $S_{j_1 j_2 \dots j_p}$ is a convex relaxation of S, the set S_{CC} also a convex relaxation of the set S. Again, since any facet defining inequality of $conv(S)$ is valid for the set S_{CC}, we have $S_{CC} = conv(S)$. □

4.6 The gap between rank one facet defining inequalities of $\hat{S}$ and $conv(S)$

Let S^1 be the set of points that satisfy all the facet defining inequalities of $conv(S)$ that have split-rank one for $\hat{S}$. Therefore, S^1 can be given as below:

$$S^1 = \left\{ (x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : \frac{1}{2k-1} \sum_{i=1}^{n} x_i + \frac{k(k-1)}{r(2k-1)} \sum_{i=1}^{n} y_i \geq 1, k \in \mathbb{N} \right\}.$$

In this section we study the "gap" between the set S^1 and $conv(S)$. Here, by the "gap" we mean the difference between the optimal objective values of a linear objective function $c^T x + d^T y$ over S^1 and $conv(S)$. Since both the sets S^1 and $conv(S)$ are unbounded, we can not compare them in terms of their volumes. Note that S^1 may not be the first split closure of $\hat{S}$. Let

$$Z = \min_{(x,y) \in conv(S)} c^T x + d^T y, \text{ and}$$
$$Z^1 = \min_{(x,y) \in S^1} c^T x + d^T y.$$

We derive some conditions for which the gap between S^1 and $conv(S)$ is zero. We also give an example with an arbitrarily large gap.

Proposition 4.6.1. *Consider the optimization problem* $\min_{(x,y) \in S^1} c^T x + d^T y$. *Let* $\lambda, \mu \in N$ *such that* $c_\lambda \leq c_i$ *for all* $i \in N$ *and* $d_\mu \leq d_i$ *for all* $i \in N$. *Then this optimization problem has the same optimal value as the optimization problem* $\min_{(x_\lambda, y_\mu) \in Q} c_\lambda x_\lambda + d_\mu y_\mu$, *where*

$$Q = \left\{ (x_\lambda, y_\mu) \in \mathbb{R}_+ \times \mathbb{R}_+ : \frac{x_\lambda}{2k-1} + \frac{k(k-1) y_\mu}{r(2k-1)} \geq 1, k \in \mathbb{N} \right\}.$$

Proof. We see that if $(x_\lambda, y_\mu) \in Q$ then $\mathcal{L}(1, x_\lambda, y_\mu) = (x_\lambda, y_\mu, 0, 0, \ldots, 0, 0) \in S^1$. Again, if $(x, y) \in S^1$ then we have $(\sum_{i=1}^n x_i, \sum_{i=1}^n y_i) \in Q$. Therefore, the two sets S^1 and Q are feasibility wise equivalent in the sense that if one has a feasible solution, we can construct a feasible solution to the other with the same objective value and vice versa. Note that the set Q is the convex hull of the two dimensional mixed-integer bilinear covering set $\{(x_\lambda, y_\mu) \in \mathbb{Z}_+ \times \mathbb{R}_+ : x_\lambda y_\mu \geq r\}$. We consider the following cases.

Case 1: One of the values of c_λ and d_μ is negative. Then clearly both the optimization problems are unbounded.

Case 2: When $c_\lambda = 0$. Then for both the optimization problems, the optimal value of the objective function is c_λ if $d_\mu = 0$, and infimum is zero if $d_\mu > 0$ (from Appendix A).

Case 3: When $c_\lambda > 0$ and $d_\mu = 0$, then for both the optimization problems the optimal value is c_λ (from Appendix A).

Case 4: $c_\lambda > 0$ and $d_\mu > 0$. We see that the constraints in the descriptions of S^1 are symmetric about the indices of the variables x and y, i.e., interchanging the variables x_i with x_j (or y_i with y_j) for any $i, j \in N$ is not going to affect the feasibility of S^1. Again since we are minimizing $c^T x + d^T y$ over S^1 which lies in the positive orthant, we can always chose an optimal solution $(\bar{x}, \bar{y})$ such that $\bar{x}_\lambda > 0, \bar{x}_i = 0$ for all in $N, i \neq \lambda$ and $\bar{y}_\mu > 0, \bar{y}_i = 0$ for all in $N, i \neq \mu$. Now it is clear that $(\bar{x}_\lambda, \bar{y}_\mu)$ is an optimal solution to the problem $\min_{(x_\lambda, y_\mu) \in Q} c_\lambda x_\lambda + d_\mu y_\mu$. □

Note that $Z^1 = \min_{(x_\lambda, y_\mu) \in Q} c_\lambda x_\lambda + d_\mu y_\mu$ from above proposition. Z is also unbounded when $\min_{i \in N}\{c_i, d_i\} < 0$, just like Z^1. Therefore, we consider $c \geq 0, d \geq 0$. In Appendix A, the algorithm to derive the values of Z and Z^1 are described with closed form solutions for both and they are as follows. It is also described in Appendix A that there exist $q \in N$ such that $Z = \min_{\mathcal{L}(q, x_q, y_q) \in S_q} c_q x_q + d_q y_q$, where S_q is an orthogonal disjunctive subset of S which is defined in the earlier section of this article.

$$Z^1 = \begin{cases} \min\left\{ c_\lambda \left\lfloor \sqrt{\frac{rd_\mu}{c_\lambda}} \right\rfloor + \frac{rd_\mu}{\left\lfloor \sqrt{\frac{rd_\mu}{c_\lambda}} \right\rfloor}, c_\lambda \left\lceil \sqrt{\frac{rd_\mu}{c_\lambda}} \right\rceil + \frac{rd_\mu}{\left\lceil \sqrt{\frac{rd_\mu}{c_\lambda}} \right\rceil} \right\} \text{when } c_\lambda > 0, d_\mu > 0, \\ c_\lambda, \text{when } c_\lambda \geq 0, d_\mu = 0, \\ 0 \quad \text{(actually infimum value), when } c_\lambda = 0, d_\mu > 0. \end{cases}$$

$Z = \min_{i \in N} Z_i,$ where

$$Z_i = \begin{cases} \min\left\{ c_i \left\lfloor \sqrt{\frac{rd_i}{c_i}} \right\rfloor + \frac{rd_i}{\left\lfloor \sqrt{\frac{rd_i}{c_i}} \right\rfloor}, c_i \left\lceil \sqrt{\frac{rd_i}{c_i}} \right\rceil + \frac{rd_i}{\left\lceil \sqrt{\frac{rd_i}{c_i}} \right\rceil} \right\} \text{when } c_i > 0, d_i > 0, \\ c_i, \text{when } c_i \geq 0, d_i = 0, \\ 0 \quad \text{(actually infimum value), when } c_i = 0, d_i > 0. \end{cases}$$

We assume $\frac{1}{\left\lfloor \sqrt{\frac{rd_i}{c_i}} \right\rfloor}$ and $\frac{1}{\left\lfloor \sqrt{\frac{rd_\mu}{c_\lambda}} \right\rfloor}$ are infinite if $\left\lfloor \sqrt{\frac{rd_i}{c_i}} \right\rfloor$ and $\left\lfloor \sqrt{\frac{rd_\mu}{c_\lambda}} \right\rfloor$ are zeros respectively with $c_i > 0, d_i > 0, c_\lambda > 0$ and $d_\mu > 0$.

4.6.1 When the gap is zero

It can be seen clearly that $Z^1 \leq Z$ as S^1 is a relaxation of $conv(S)$. The following result gives us the condition for $Z^1 = Z$.

Proposition 4.6.2. *Let Λ, Δ be two subsets of N such that $c_\lambda = c_i$ for all $i \in \Lambda$ and $d_\mu = d_i$ for all $i \in \Delta$. Then $Z^1 = Z$ if and only if $\Lambda \cap \Delta$ is non empty.*

Proof. Let $p \in \Lambda \cap \Delta$. Then by Proposition 4.6.1, there exists an optimal solution $(\bar{x}, \bar{y})$ of the form $\mathcal{L}(p, \bar{x}_p, \bar{y}_p)$ to the problem $\min_{(x,y) \in S^1} c^T x + d^T y$. Therefore, $(\bar{x}_p, \bar{y}_p)$ satisfies

$$\frac{\bar{x}_p}{2k-1} + \frac{k(k-1)\bar{y}_p}{r(2k-1)} \geq 1, k \in \mathbb{N}.$$

If we show $(\bar{x}, \bar{y}) \in conv(S)$, we will have $Z = Z^1$. We know that

$$conv\,(S_p) = \left\{ \mathcal{L}(p, x_p, y_p) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : \frac{x_p}{2k-1} + \frac{k(k-1)y_p}{r(2k-1)} \geq 1, k \in \mathbb{N} \right\}$$

where, $S_p = \left\{ \mathcal{L}(p, x_p, y_p) \in \mathbb{Z}^n_+ \times \mathbb{R}^n_+ : x_p y_p \geq r \right\}$ is an orthogonal disjunctive subset of

S. This implies that $\mathcal{L}(p, \bar{x}_p, \bar{y}_p) \in conv\,(S_p)$. Again, since $conv(S)$ is closed and satisfies the convex extension property, i.e., $conv(S) = conv\,(\bigcup_{i=1}^{n} S_i)$ [106], where $S_i, i \in N$ are orthogonal disjunctive subsets of S, the point $\mathcal{L}(p, \bar{x}_p, \bar{y}_p) \in conv(S)$. Therefore, we have a point in $conv(S)$ with objective value Z^1, consequently $Z^1 = Z$.

Conversely, let $\Lambda \cap \Delta$ be empty. We know that there exists $q \in N$ such that $Z_{CV} = \min_{\mathcal{L}(q,x_q,y_q) \in S_q} c_q x_q + d_q y_q = \min_{\mathcal{L}(q,x_q,y_q) \in conv(S_q)} c_q x_q + d_q y_q$ (from Appendix A). Since $\Lambda \cap \Delta = \phi$, $c_\lambda \leq c_q, d_\mu \leq d_q$ with $c_\lambda < c_q$ or $d_\mu < d_q$. Again, since the two sets $conv\,(S_q)$ and Q (defined in Proposition 4.6.1) are feasibility wise equivalent, we have $Z^1 < Z$. $\square$

4.6.2 When the gap is arbitrary large : An example

Let $n = 2, r = 16$ and consider the objective function $x_1 + \eta^2 y_1 + \eta^2 x_2 + y_2$ where η is a positive integer.

We see that $c_\lambda = c_1 = 1$ and $d_\mu = d_2 = 1$. Therefore, $Z^1 = 8$ with optimal solution $(4, 0, 0, 4)$ which is not feasible for $conv(S)$. Clearly, if we increase the value of η, the value of Z^1 is not going to change.

On the other hand, we see that $Z = \eta Z^1$. Since for any value of $\eta \geq 1$, the value of Z^1 is constant, the value of Z increases by a factor of η with Z^1, therefore, the gap between the values of Z and Z^1 can be arbitrary large.

4.7 Rank-one facets and the cutting stock problem

We now study the gap between the bounds that are obtained using the inequalities of S^1 and the inequalities for $conv(S)$ doing computational experiments on some cutting stock problem instances, and as defined in Chapter 1, the mathematical formulation (CS) is following:

$$\min \sum_{i=1}^{n} y_i$$
$$\sum_{i\in N} x_{ij}y_i \geq d_j,\ j \in F,$$
$$\sum_{j\in F} l_j x_{ij} \leq L, i \in N,$$
$$x_{ij} \in \mathbb{Z}_+, y_i \in \mathbb{R}_+, \forall i \in N, j \in F,$$

where all the notations have the same meaning as in Chapter 1. The knapsack constraints implies $x_{ij} \leq \left\lfloor \frac{L}{l_j} \right\rfloor, \forall i \in N, j \in F$

Proposition 4.7.1. *Consider the problem (CS) without the knapsack constraints. Then the lower bound obtained by considering all the facet defining inequalities of each bilinear constraint is equal to the lower bound obtained by considering only the rank one facet defining inequalities for each bilinear constraint.*

Proof. Let $(\bar{x}, \bar{y})$ be an optimal solution when we consider only the rank one facet defining inequalities of each bilinear constraint. Therefore, we have

$$\sum_{i\in N} \frac{\bar{x}_{ij}}{2k-1} + \sum_{i\in N} \frac{\bar{y}_i k(k-1)}{d_j(2k-1)} \geq 1,\ \text{for all } j \in F, \tag{4.12}$$

and the optimal value at this point is $\sum_{i\in N} \bar{y}_i$. If we can show that there exists a point satisfying all the facet defining inequalities of each bilinear constraint with objective value $\sum_{i\in N} \bar{y}_i$, we will be done. Now consider the point (x^*, y^*) defined below.

$$x^*_{ij} = \begin{cases} \sum_{i\in N} \bar{x}_{ij}, & \text{if } i = 1 \\ 0, & \text{if } i \neq 1, i \in N \end{cases} \quad \text{for } j \in F$$
$$y^*_i = \begin{cases} \sum_{i\in N} \bar{y}_i, & \text{if } i = 1 \\ 0, & \text{if } i \neq 1, i \in N \end{cases}$$

Therefore, using the relation (4.12) and the construction procedure of the facet defining inequalities using collection of column (M), it can be seen that the point (x^*, y^*)

satisfies all the facet defining inequalities of each bilinear constraint, and the objective value at this point is $\sum_{i\in N} \bar{y}_i$. $\square$

To check whether the result holds with the knapsack constraints also we performed a computational experiment on the same instances that are used in the computational section of Chapter 2.

For each instance, in either case we start the iterations with the same (LP-I) defined in the computational section in Chapter 2, i.e., with the facet defining inequalities $\sum_{i\in N} x_{ij} \geq 1, j \in F$, the bound constraints $x_{ij} \leq \left\lfloor \frac{L}{l_j} \right\rfloor, \forall i \in N, j \in F$ and the knapsack constraints, i.e., the following LP.

$$\begin{aligned}
\min & \sum_{i=1}^{n} y_i \\
\text{s.t.} & \sum_{i=1}^{n} x_{ij} \geq 1, \forall j \in F, \\
& 0 \leq x_{ij} \leq \left\lfloor \frac{L}{l_j} \right\rfloor, \forall i \in N, j \in F, \\
& \sum_{j\in F} l_j x_{ij} \leq L, i \in N, \\
& y \geq 0.
\end{aligned}$$

We use Algorithm 2 in Chapter 2 to separate the facet defining inequalities for $conv(S)$ and the inequalities of S^1. Therefore, we add at most $|F|$ cuts in each iteration for both the cases (i.e., at the k^{th} iteration we solve an LP with $k|F|$ number of linear inequalities in addition to those in the starting LP). We stop when either of two conditions hold: (a) we can not find any more violated inequalities, or (b) both the time limit of 3600 seconds and the number of LPs solved exceeds a limit of 800 LPs (we write "800*" or "3600*" whichever hits later for such cases in Table 4.1).

We have used Python based PuLP [86] to model the problem and CLP [56] to solve the linear programs. The experiment was performed on a 4x Intel(R) Core(TM) i5-3570 CPU@3.40 GHz processor, 8 GB of RAM and Linux (Ubuntu 16.04) operating system. We have used single core to do the experiments. The results are compiled in Table 4.1. The results show that the optimization over S^1 gives the same bound as that over the

Instances	n	Using inequalities for $conv(S)$				Using inequalities for S^1 only			
		Iter	Cuts	LB	Time	Iter	Cuts	LB	Time
Fiber-10-5180	10	226	1917	6.88	55.09	5	37	6.88	0.14
Fiber-10-9080	10	223	2045	3.85	46.15	6	42	3.85	0.16
Fiber-11-5180	11	288	2673	6.10	89.22	4	35	6.10	0.13
Fiber-11-9080	11	335	2944	3.40	162.30	5	45	3.40	0.17
Fiber-14-5180	14	473	5417	3.34	547.81	5	57	3.34	0.25
Fiber-14-9080	14	476	6211	1.90	658.27	5	57	1.90	0.26
Fiber-15-5180	15	560	7219	3.74	1412.46	6	65	3.74	0.32
Fiber-15-9080	15	1092	11592	2.09	3600*	6	72	2.11	0.35
Fiber-16-5180	16	756	9763	5.17	3086.56	5	63	5.17	0.31
Fiber-16-9080	16	723	10330	2.93	2692.61	6	78	2.93	0.38
CutGen-01-01	10	252	2244	1.24	95.33	5	43	1.24	0.16
CutGen-01-02	10	270	2443	0.97	103.32	5	40	0.97	0.14
CutGen-01-25	10	246	2157	0.99	77.18	5	39	0.99	0.16
CutGen-01-100	10	244	2131	1.25	52.82	5	41	1.25	0.15
CutGen-02-40	10	262	2272	10.41	92.6	4	33	10.41	0.14
CutGen-02-60	10	275	2480	10.10	76.32	5	40	10.10	0.16
Rand-10	10	185	1601	697.22	28.86	4	33	697.22	0.14
Rand-15	15	440	5296	650.29	900.01	5	55	650.29	0.27
Rand-16	16	800*	9935	753.17	4543.9	5	52	753.17	0.26
Rand-20	20	800*	14105	624.92	12061.05	5	71	624.92	0.43
Rand-25	25	800*	19516	541.80	18040.75	6	94	544.99	0.75

Table 4.1: Comparison of iterations and time taken to optimize using the inequalities for S^1 only and the convex hull.

convex hull in much fewer iterations and in much less time for all input problems.

4.8 Conclusion

All facet defining inequalities of $conv(S)$ can be viewed as disjunctive cuts derived from disjunctions specified in the discussion above. Some of them have split-rank one for a convex mixed-integer relaxation of S. These cuts are sufficient to find the optimal value over $conv(S)$ for certain objective functions like those in trimloss problems. Finding strong valid inequalities for convex hull of the feasible region of trimloss problems is still open and can be taken up in the future.

Chapter 5

Concluding Remarks and Future Work

The convex hulls of sets considered in Chapter 2 and Chapter 3 turned out to be 'easy', i.e., one can separate over them in polynomial time. It means there is a possible scope of including more constraints of the original problem in to these sets so that we can find even tighter relaxations of these problems. The knapsack constraint of (CS) is one such candidate. Similarly, finding the convex hull of multiple bilinear constraints together can be taken up as future work.

The cut generated for S^U by our criterion of 'maximum violation' without any normalization may not be the cut that improves the lower bound the most, or the cut that is farthest from the infeasible point. Consider the following example:

$$\begin{aligned} \min\ & x_1 + y_1 + x_2 + y_2 \\ \text{s.t.}\ & x_1y_1 + x_2y_2 \geq 20, \\ & x_i \leq 10, \quad i = 1, 2, \\ & x_i \in \mathbb{Z}_+, y_i \geq 0, \quad i = 1, 2. \end{aligned}$$

The point $(x_1, y_1, x_2, y_2) = (5, 4, 0, 0)$ is a global minimizer with optimal value 9. At the first iteration, the LP solution is $(1, 0, 0, 0)$ with objective value 1. The best cut generated by Algorithm 1 to cut this point off is $\frac{y_1}{2} + \frac{y_2}{2} \geq 1$. After adding this inequality, the solution is $(1, 2, 0, 0)$ with objective value 3. But, if we instead add the facet defining inequality $\frac{x_1}{5} + \frac{3y_1}{50} + \frac{x_2}{5} + \frac{3y_2}{50} \geq 1$, we get a better solution $(5, 0, 0, 0)$ with objective value 5. Also, the distance of the latter from the point $(1, 0, 0, 0)$ is nearly 2.7 as compared to 1.41 for the former. Finding the cut that improves the bound the most or that is farthest from the given point is a problem that can be explored.

In Chapter 3, we studied the convex hull structure of mixed-integer bilinear covering set with box constrained integer variables. We implemented the extended formulation in a branch-and-bound algorithm to solve Pattern Minimization Problem. Immediate future work is to speed up our algorithm to find solution faster. For this, exploring different branching strategies on the variables and branching points will be explored.

Lastly, in Chapter 4, we studied the facet defining inequalities of $conv(S)$ as disjunctive cuts. We derived the exact split rank for a class of inequalities and provided lower bound on the same for the others. We derived the disjunctions from which they can be constructed.

We saw that for cutting stock instance, without the knapsack constraints rank one cuts are sufficient to consider. We observed that for all instances of cutting stock with the knapsack constraints also, merely considering the rank one cuts was sufficient to reach the same bound as the convex hull. A systematic study of this extended set may be carried out in the future.

Appendix A

Optimization Over S

Minimizing a linear function $c^T x + d^T y$ over S is equivalent to minimizing it over $conv(S)$. For the following cases, we can solve the optimization problem by inspection.

CASE 1 : If one of the components of c or d is negative, then the problem is unbounded.

CASE 2 : Suppose, $c \geq 0, d \geq 0$ and one of the component of the vector c is zero, say $c_t = 0$. If $d_t = 0, \min_{(x,y) \in S} c^T x + d^T y = 0$ and if $d_t > 0, \inf_{(x,y) \in S} \; c^T x + d^T y = 0$. This is because, in either case we can choose y_t arbitrary small such that $x_t y_t = r$ and all other components are zero, i.e., $\mathcal{L}(t, x_t, y_t)$ is an optimal solution in either case.

CASE 3 : Suppose, $c \geq 0, d = 0$. Let $c_t \leq c_j, \forall j \in N$. Then $\mathcal{L}(t, 1, r)$ is an optimal solution with optimal value c_t.

Now, the only remaining case is, when $c > 0, d \geq 0, d \neq 0$, which we consider in the following proposition.

Proposition A.0.1. *Let us consider the orthogonal disjunctive subset S_i of the set S. Then we can solve the optimization problem $\min_{(x,y) \in S_i} c_i x_i + d_i y_i$ in polynomial time.*

Proof. From the definition, each $(x, y) \in S_i$ is of the form $\mathcal{L}(i, x_i, y_i)$. If $c_i \geq 0, d_i = 0$ for some $i \in N$, then $\mathcal{L}(i, 1, r)$ is an optimal solution with optimal value c_i.

Now, we only have to consider $c_i > 0, d_i > 0$. Let $\mathcal{L}(i, x_i^*, y_i^*)$ be an extreme point optimal solution. Clearly, this point should lie on the surface $x_i y_i = r$. We note that the continuous relaxation of the set S_i is a strictly convex set. Therefore, the optimal solution $\mathcal{L}(i, \bar{x}_i, \bar{y}_i)$ (say) over the continuous relaxation is unique, and we have,

$$\bar{x}_i = \sqrt{\frac{rd_i}{c_i}}, \text{ and } \bar{y}_i = \frac{r}{\bar{x}_i}.$$

If $\sqrt{\frac{rd_i}{c_i}}$ is an integer, then $\mathcal{L}(i, \bar{x}_i, \bar{y}_i)$ is an optimal solution. If not, then from the geometry, it is clear that at the optimal solution either $x_i^* = \left\lceil \sqrt{\frac{rd_i}{c_i}} \right\rceil$ or $x_i^* = \left\lfloor \sqrt{\frac{rd_i}{c_i}} \right\rfloor$ whichever minimizes the objective function and is feasible.

So, to determine an optimal solution, we just have to check the signs of the objective coefficient and compute the value of $\sqrt{\frac{rd_i}{c_i}}$. This can be done in constant time. □

Now we consider the set S. Since the objective function is linear, it is equivalent to minimize over $conv(S)$. There must be an extreme point optimal solution, provided an optimal solution exists. Suppose optimal solution exists. Then by Theorem 2.3.1, there must be an optimal solution that is an extreme point of $conv(S_i)$ for some $i \in N$, because each extreme point of $conv(S)$ is an extreme point of $conv(S_i)$ for some $i \in N$. So, if we solve the n problems $\min_{\mathcal{L}(i,x_i,y_i) \in S_i} c_i x_i + d_i y_i$ for $i \in N$ and pick the minimum of the n objective values, we will get the optimal value and corresponding optimal solution. We have seen earlier that each subproblem takes constant time to solve. So, we can solve this problem in linear time. The algorithm is presented in Algorithm 3

Algorithm 3 Algorithm to solve $\min_{(x,y)\in S} c^T x + d^T y$

if One of the components of the vectors c or d is negative **then**

    The problem is unbounded.

else if $c \geq 0, d \geq 0$ and one of the component of the vector c is zero **then**

    $inf_{(x,y)\in S}\ c^T x + d^T y = 0.$

else if $c \geq 0, d = 0$. Let $c_t \leq c_j, \forall j \in N$ **then**

    $\mathcal{L}(t, 1, r)$ is an optimal solution with optimal value c_t.

else (i.e., when $c > 0, d \geq 0$)

    for $i = 1, \ldots, n$ **do**

        Solve the problem $\min_{\mathcal{L}(i,x_i,y_i)\in S_i} c_i x_i + d_i y_i$.

        Let $\mathcal{L}^i(i, x_i, y_i)$ be an optimal solution with optimal value v_i.

    end for

    Find the minimum of $v_i, i \in N$. Let $t \in N$ such that $v_t \leq v_i$ for all $i \in N$.

    Then $\mathcal{L}^t(t, x_t, y_t)$ is an optimal solution with optimal value v_t.

end if

Appendix B

Optimization Over S^U

Now we consider the following problem:

$$\min_{(x,y)\in S^U} c^T x + d^T y \tag{P$_{\text{SU}}$}$$

This problem is equivalent to minimizing the objective function over *conv* $\left(S^U\right)$. Also, there must be an extreme point optimal solution, provided an optimal solution exists. We know the description of *conv* $\left(S^U\right)$ in terms of the facet defining inequalities, and the number of facet defining inequalities is exponential in the input size. The extreme point descriptions of *conv* $\left(S^U\right)$ is also known. In the following discussion we will see that the problem is efficiently solvable and we will also present the algorithm.

Though there are few similarities between the solution strategy over S and over S^U, in general the solution strategy for the case of S^U is quite different. Here also by inspection we can solve the problem for the following cases.

Case 1 : When $d_t < 0$ for some $t \in N$, the problem is unbounded.

Case 2 : When $c \leq 0, d = 0$, then $\left(u_1, \frac{r}{u_1}, u_2, 0, \ldots, u_n, 0\right)$ is an extreme point optimal solution.

Now the remaining case is $d \geq 0$. We note that $\sum_{i\in N: c_i \leq 0} c_i u_i$ is a lower bound on the objective value. To solve this problem, we will only consider the extreme points and compare their corresponding objective values to find the optimal solution. We first partition the set of extreme points of *conv* $\left(S^U\right)$ and optimize over those partitions. Let us define the following set for each $i \in N$.

$$E_i = \left\{ (x,y) \in \mathbb{R}^{2n}_{+} : x_i \in \{1,\ldots,u_i\}, y_i = \frac{r}{x_i}, x_j \in \{0,u_j\}, y_j = 0, \forall j \in N, j \neq i \right\}.$$

From the discussion in Section 2.4.1, all the points in E_i are extreme points of $conv\left(S^U\right)$ and $E = \bigcup_{i\in N} E_i$ gives the complete set of extreme points of $conv\left(S^U\right)$, i.e., the sets $E_i, i \in N$ defines a partition of the set E. If we minimize $c^T x + d^T y$ over each set $E_i, i \in N$ and compare their values, we will get the optimal solution. Now our goal is to solve the following problem.

$$\zeta_i = \min_{(x,y)\in E_i} c^T x + d^T y \tag{$\mathrm{P^i{}_{SU}}$}$$

Note that only the i^{th} component of the variable y of each point in E_i is non-zero and rest of all are zero. Therefore, the objective function of the above problem ($\mathrm{P^i{}_{SU}}$) reduces to $c_i x_i + d_i y_i + \sum_{j\in N, j\neq i} c_j x_j$. Also, we see that for any point $(x,y) \in E_i$, the choices of the components $x_j \in \{0,u_j\}, j \in N, j \neq i$ are independent of the choice of $x_i \in \{1,\ldots,u_i\}$. Let $(\bar{x},\bar{y})^i \in E_i$ be an optimal solution of ($\mathrm{P^i{}_{SU}}$). Then, $\bar{y}^i_i = \frac{r}{\bar{x}^i_i}, \bar{y}^i_j = 0, \forall j \in N, j \neq i$. Let us consider the following choices of x components of $(\bar{x},\bar{y})^i$.

$$\begin{aligned} &\bar{x}^i_i \in \{1,\ldots,u_i\} \text{ such that } (\bar{x}^i_i, \bar{y}^i_i) \text{ minimzes } c_i x_i + d_i y_i, \\ &\bar{x}^i_j = \begin{cases} 0, \text{ if } c_j > 0, \\ u_j, \text{ if } c_j \leq 0, \end{cases} \qquad \forall j \in N, j \neq i. \end{aligned}$$

It can be seen clearly that such above choice of the components of $(\bar{x},\bar{y})^i$ minimizes the objective function. Now to find the value of $\bar{x}^i_i \in \{1,\ldots,u_i\}$, we consider the following cases.

Case 1 : When $c_i \leq 0$, then $\bar{x}^i_i = u_i$. This is because, since $c_i \leq 0$, the maximum value of x_i in the domain will minimize $c_i x_i$. Moreover, for this choice of $\bar{x}^i_i, \bar{y}^i_i = \frac{r}{u_i}$ is also minimum, and since $d_i \geq 0$, $\left(u_i, \frac{r}{u_i}\right)$ minimizes $c_i x_i + d_i y_i$.

Case 2 : If $c_i > 0$ and $d_i = 0, \bar{x}^i_i = 1, \bar{y}^i_i = r$ as $\bar{x}^i_i \geq 1$.

Case 3 : The remaining case is $c_i > 0, d_i > 0$. Since the points $\mathcal{L}\left(i, p_i, \frac{r}{p_i}\right), p_i \in$

$\{1, \ldots, u_i\}$ are the extreme points of *conv* $\left(S_i^U\right)$, it is equivalent to minimize $c_i x_i + d_i y_i$ over *conv* $\left(S_i^U\right)$. To solve this we will use the same analysis as in the proof of Proposition A.0.1 with slight modification as there is an upper bound u_i on the variable x_i. So, in this case we have the following choice of $\bar{x}_i^i$ and consequently $\bar{y}_i^i = \frac{r}{\bar{x}_i^i}$.

$$\bar{x}_i^i = \begin{cases} \sqrt{\frac{rd_i}{c_i}}, \text{ if } \sqrt{\frac{rd_i}{c_i}} \in \{1, \ldots, u_i\}, \\ 1, \text{ if } \sqrt{\frac{rd_i}{c_i}} < 1, \\ \left\lceil \sqrt{\frac{rd_i}{c_i}} \right\rceil \text{ or } \left\lfloor \sqrt{\frac{rd_i}{c_i}} \right\rfloor, \text{ whichever minimizes } c_i x_i + d_i \frac{r}{x_i}, \\ \qquad\qquad\qquad \text{if } 1 < \sqrt{\frac{rd_i}{c_i}} < u_i \text{ and } \sqrt{\frac{rd_i}{c_i}} \notin \mathbb{Z}_+, \\ u_i, \text{ if } \sqrt{\frac{rd_i}{c_i}} > u_i. \end{cases}$$

So, from the above analysis, we can solve the problem ($\mathrm{P^i}_{\mathrm{SU}}$) in linear time, as we just have to check the signs of $n-1$ entries and have to check the value of $\sqrt{\frac{rd_i}{c_i}}$, whenever it exists and if not then the signs of c_i and d_i. Now, we have the following algorithms to solve the problem (P_{SU}).

Algorithm 4 Algorithm to solve ($\mathrm{P^i}_{\mathrm{SU}}$) when $d \geq 0$

Let $(\bar{x}, \bar{y})^i$ be an optimal solution to ($\mathrm{P^i}_{\mathrm{SU}}$).
for $j \in N, j \neq i$ **do**
    $\bar{y}^i_j = 0$
    if $c_j > 0$ **then**
        $\bar{x}^i_j = 0$
    else
        $\bar{x}^i_j = u_j$
    end if
end for
if $c_i \leq 0$ **then**
    $\bar{x}^i_i = u_i, \bar{y}^i_i = \frac{r}{u_i}$
else if $c_i > 0, d_i = 0$ **then**
    $\bar{x}^i_i = 1, \bar{y}^i_i = r$
else
    $\eta = \sqrt{\frac{rd_i}{c_i}}$
    if $\eta \in \mathbb{N}$ and $\eta \leq u_i$ **then**
        $\bar{x}^i_i = \eta$
    else if $\eta < 1$ **then**
        $\bar{x}^i_i = 1$
    else if $\eta > u_i$ **then**
        $\bar{x}^i_i = u_i$
    else
        if $c_i \lfloor \eta \rfloor + d_i \frac{r}{\lfloor \eta \rfloor} \leq c_i \lfloor \eta \rfloor + d_i \frac{r}{\lceil \eta \rceil}$ **then**
            $\bar{x}^i_i = \lfloor \eta \rfloor$
        else
            $\bar{x}^i_i = \lceil \eta \rceil$
        end if
    end if
    $\bar{y}^i_i = \frac{r}{\bar{x}^i_i}$
end if

Algorithm 5 Algorithm to solve $\min_{(x,y)\in S^U} c^T x + d^T y$

if One of the components of the vector d is negative **then**
  The problem is unbounded.
else if $c \leq 0, d = 0$ **then**
  $\left(u_1, \frac{r}{u_1}, u_2, 0, \ldots, u_n, 0\right)$ is an extreme point optimal solution.
else$(d \geq 0)$
  for $i = 1, \ldots, n$ **do**
    Use Algorithm 4 to solve the problem ($\mathrm{P^i}_{\mathrm{SU}}$).
    Let $(\bar{x}, \bar{y})^i$ be an optimal solution with optimal value ζ_i.
  end for
  Find the minimum of $\zeta_i, i \in N$. Let $t \in N$ such that $\zeta_t \leq \zeta_i$ for all $i \in N$.
  Then $(\bar{x}, \bar{y})^t$ is an optimal solution of ($\mathrm{P_{SU}}$) with optimal value ζ_t.
end if

Since the running time of the Algorithm 4 is $O(n)$ time, and finding the minimum among $\zeta_i, i \in N$ takes $O(n)$ time, therefore the Algorithm 5 runs in $O(n^2)$ time of the input size.

Appendix C

Additional Proofs

Proposition C.0.1. *Let $j, k \in \mathbb{N}$ with $j \neq k$, then the following two relations*

$$\pi_0 < \sqrt{j(j-1)} < \pi_0 + 1 \text{ and}$$
$$\pi_0 < \sqrt{k(k-1)} < \pi_0 + 1$$

can not hold simultaneously for any non-negative integer π_0.

Proof. Without loss of generality let $k > j$. Note that, it is equivalent to prove that $\sqrt{k(k-1)} - \sqrt{j(j-1)} \geq 1$. This is because, if $\sqrt{k(k-1)} - \sqrt{j(j-1)} \geq 1$ holds, then both the values $\sqrt{k(k-1)}$ and $\sqrt{j(j-1)}$ can not lie between two consecutive integers.

Also note that, since the function $f(j) = \sqrt{j(j-1)}$ is strictly increasing for $j \in \mathbb{N}$, it is sufficient to prove the result when $k = j+1$, i.e., we show that $\sqrt{j(j+1)} - \sqrt{j(j-1)} \geq 1$. Since we are dealing with positive numbers only, in our following steps of proof, we consider only the positive roots. Now we have for any $j \in \mathbb{N}$,

$$\begin{aligned}
& 4j^2 - 4j + 1 > 4j^2 - 4j \\
\Rightarrow\ & (2j-1)^2 > 4j(j-1) \\
\Rightarrow\ & 2j - 1 > 2\sqrt{j(j-1)} \\
\Rightarrow\ & j^2 + j > 1 + 2\sqrt{j(j-1)} + j^2 - j \\
\Rightarrow\ & j(j+1) > 1 + 2\sqrt{j(j-1)} + j(j-1) \\
\Rightarrow\ & j(j+1) > \left(1 + \sqrt{j(j-1)}\right)^2 \\
\Rightarrow\ & \sqrt{j(j+1)} - \sqrt{j(j-1)} > 1.
\end{aligned}$$

This completes the proof. □

Proposition C.0.2. *Let $k \in \mathbb{N}$ with $k \geq 2$. Consider the positive integers μ_0, μ with $\mu \geq 1$. If $\mu_0 < \mu\sqrt{k(k-1)} < \mu_0 + 1$ then $k - 1 < \frac{\mu_0+1}{\mu} \leq k$ and $k - 1 \leq \frac{\mu_0}{\mu} < k$.*

Proof. Given that $\mu_0 < \mu\sqrt{k(k-1)} < \mu_0 + 1$, which implies $\frac{\mu_0+1}{\mu\sqrt{k(k-1)}} > 1$. Again since $\sqrt{\frac{k-1}{k}} < 1$, we have

$$\begin{aligned}
& \frac{\mu_0 + 1}{\mu\sqrt{k(k-1)}} > \sqrt{\frac{k-1}{k}} \\
\Rightarrow & \frac{\mu_0 + 1}{\mu} > k - 1
\end{aligned}$$

Now, we have to show the other side of the desired inequality. By given condition we have,

$$\begin{aligned}
& \mu\sqrt{k(k-1)} > \mu_0 \\
\Rightarrow & \mu\left\lceil \sqrt{k(k-1)} \right\rceil > \mu_0 \\
\Rightarrow & \mu\left\lceil \sqrt{k(k-1)} \right\rceil \geq \mu_0 + 1 \text{ (since the left hand side is integral)} \\
\Rightarrow & \mu k \geq \mu_0 + 1 \ \left(\text{since } k \geq \left\lceil \sqrt{k(k-1)} \right\rceil\right) \\
\Rightarrow & k \geq \frac{\mu_0 + 1}{\mu}
\end{aligned}$$

Therefore, we have $k - 1 < \frac{\mu_0+1}{\mu} \leq k$.

Since $\frac{\mu_0+1}{\mu} \leq k$, we have $\frac{\mu_0}{\mu} < k$. Now it remains to show $k-1 \leq \frac{\mu_0}{\mu}$. By given relation we have

$$
\begin{aligned}
&\mu_0 + 1 > \mu\sqrt{k(k-1)} \\
\Rightarrow &\mu_0 + 1 > \mu(k-1) \ \left(\text{since } \sqrt{k(k-1)} > k-1\right) \\
\Rightarrow &\mu_0 \geq \mu(k-1) \ \text{(since both sides are integral)} \\
\Rightarrow &\frac{\mu_0}{\mu} \geq k-1
\end{aligned}
$$

Therefore, we have $k-1 \leq \frac{\mu_0}{\mu} < k$. □

Publications and Conference Talks

- **Rahman H.**, Mahajan A. : *Facets of a Mixed-Integer Bilinear Covering Set With Bounds on Variables.* Under second round of review in *Journal of Global Optimization (Springer).*
- **Rahman H.**, Mahajan A. : *On The Facet Defining Inequalities of the Mixed-Integer Bilinear Covering Set.* Under review in *Mathematical Methods of Operations Research (Springer)*
- Ramamurthy A. **Rahman H.**, Mahajan A. : *A Branch-and-Cut Algorithm to solve Pattern Minimization Problem (PMP). To be communicated.*
- **Rahman H.**, Mahajan A. : *On the split-rank of the facet defining inequalities of mixed-integer bilinear covering set. International Symposium on Combinatorial Optimization (ISCO 2018), Marrakesh, Morocco, April 2018*
- **Rahman H.**, Mahajan A. : *Facets of a Mixed-Integer Bilinear Covering Set With Bounds on Variables. International Symposium on Operations Research and Game Theory: Modeling and Computation, New Delhi, January 2018*
- **Rahman H.**, Mahajan A. : *On the split-rank and disjunctions for the facet defining inequalities for mixed-integer bilinear covering set. International Conference on 'Advancing Frontiers in Operational Research: Towards a Sustainable World' (AFOR 2017), Calcutta, December 2017*
- **Rahman H.**, Mahajan A. : *An Efficient Algorithm to Solve a Linear Function Over Mixed-Integer Bilinear Covering Set. International Conference on Analytics in Operational Research, Delhi NCR, December 2016*

Bibliography

[1] FICO Xpress Solver. https://www.fico.com/en/products/fico-xpress-solver.

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

[3] CPLEX 12.1 User's Manual, 2016. IBM ILOG Inc. Mountain View, CA.

[4] K. Abhishek, S. Leyffer, and J. T. Linderoth. FilMINT - An Outer Approximation Based Solver for Nonlinear Mixed Integer Programs. INFORMS Journal On Computing, 22(4):555 – 567, 2010.

[5] K. Anderson and A. Jensen. Intersection cuts for mixed integer conic quadratic sets, in : M. Goemans, J. Correa (Eds). IPCO in : LNCS, 7808:37 – 48, 2013.

[6] K. M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. Journal of Global Optimization, 43:471 – 484, 2009.

[7] K. M. Anstreicher. On convex relaxations for quadratically constrained quadratic programming. Mathematical Programming, B 136:233 – 251, 2012.

[8] K.M. Anstreicher and S. Burer. Computable representations for convex hulls of low-dimensional quadratic forms. Mathematical Programming, B 124:33 – 43, 2010.

[9] MOSEK ApS. The MOSEK optimization toolbox for MATLAB manual. Version 8.1., 2017.

[10] A. Atamtürk and V. Narayanan. Conic mixed-integer rounding cuts. Mathematical Programming, A 122:1 – 20, 2010.

[11] Alper Atamtürk. On the facets of the mixed–integer knapsack polyhedron. Mathematical Programming, 98(1):145–175, 2003.

[12] E. Balas. Disjunctive programming. Annals of Discrete Mathematics, 5:3 – 51, 1979.

[13] E. Balas. Disjunctive programming: properties of the convex hull of feasible points. Discrete Applied Mathematics, 58(1-3):3 – 44, 1998.

[14] E. Balas, S. Ceria, and G. Cornuéjols. A Lift-and-Prject Cutting Plane Algorithm for Mixed 0-1 Programs. Mathematical Programming, 58:295 – 324, 1993.

[15] X. Bao, N. V. Sahinidis, and M. Tawarmalani. Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons. Mathematical Programming, B 129:129 – 157, 2011.

[16] P. Belotti. Design of Telecommunication Networks with Shared Protection, 2009. Available at: www.minlp.org/library/problem/index.php?i=51.

[17] P. Belotti. Disjunctive cuts for nonconvex MINLP. In : Lee and J. and Leyffer and S.(eds.) Mixed Integer Nonlinear Programming. The IMA Volumes in Mathematics and its Applications, volume 154, pages : 117 - 144. Springer, 2012.

[18] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan. Mixed-integer nonlinear optimization. Acta Numerica, 22:1 – 131, 2013.

[19] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Waechter. Branching and bounds tightening techniques for nonconvex MINLP. Optimization Methods and Software, 24 : 4-5:597 – 634, 2009.

[20] A. Ben-Tal and A.S. Nemirovski. Lectures on modern convex optimization: Analysis, algorithms and engineering applications. SIAM, 2001.

[21] H. Y. Benson and Ü. Saglam. Mixed-Integer Second-Order Cone Programming: A Survey. INFORMS TutORials in Operations Research, pages 13 – 36, 2014.

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

[23] E. Birgin and J. Martínez. Practical Augmented Lagrangian Methods for Constrained Optimization. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014.

[24] E.G. Birgin and J.M. Martínez. On the application of an Augmented Lagrangian algorithm to some portfolio problems. EURO Journal on Computational Optimization, 4(1):79–92, 2016.

[25] P. Bonami. Lift-and-project cuts for mixed integer convex programs, in : O. Günlük, G. J. Woeginger (Eds). IPCO, in : LNCS, 6655:52 – 64, 2011.

[26] P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, and A. Wächter. An algorithmic framework for convex mixed integer nonlinear programs. Discrete Optimization, 5:186 – 204, 2010.

[27] B. Borchers. A C Library for Semidefinite Programming. Optimization Methods and Software, 1(11):613 – 623, 1999.

[28] J. Borwein and A.S. Lewis. Convex Analysis and Nonlinear Optimization. Springer-Verlag New York, second edition, 2006.

[29] S. Boyd and L. Vandenberghe. Convex Optimization. Cambridge University Press, Cambridge., 2004.

[30] S. Burer. On the copositive representation of binary and continuous non-convex continuous programs. Mathematical Programming, A 120:479 – 495, 2009.

[31] S. Burer and J. Chen. Relaxing the optimality conditions of box QP. Computational Optimization and Applications, 48(3):653 – 673, 2011.

[32] R. H. Byrd, J. Nocedal, and R. A. Waltz. Knitro: An Integrated Package for Nonlinear Optimization, pages 35–59. Springer US, 2006.

[33] M. T. Çezik and G. Iyengar. Cuts for mixed 0-1 conic programming. Mathematical Programming, A 104:179 – 202, 2005.

[34] V. Chvátal. Linear Programming. W.H. Freeman, 1983.

[35] M. Conforti, G. Cornuéjols, and G. Zambelli. Polyhedral approaches to mixed integer linear programming. In : 50 Years of Integer Programming 1958 - 2008. Springer, 2010.

[36] M. Conforti, G. Cornuéjols, and G. Zambelli. Integer Programming. Springer, 2014.

[37] S. A. Cook. The complexity of theorem-proving procedures. Proceedings of the Third Annual ACM Symposium on the Theory of Computing, ACM, New York:151 – 158, 1971.

[38] W. Cook, R. Kannan, and A. Schrijver. Chvátal closures for mixed integer programming problems. Mathematical Programming, 47:155 – 174, 1990.

[39] G. Cornuéjols. Valid inequalities for mixed integer linear programs. Mathematical Programming, B 112:3 – 44, 2008.

[40] D. Dadush, S. S. Dey, and J. P. Vielma. The split closure of a strictly convex body. Operations Research Letters, 39:121 – 126, 2011.

[41] R. J. Dakin. A tree search algorithm for mixed-integer programming problems. Computer Journal, 8:250 – 255, 1965.

[42] G. B. Dantzig. Maximization of a linear function of variables subject to linear inequalities, In T.C. Koopmans, ed. Activity Analysis of Production and Allocation, Cowles Commission Monograph, No. 13, Wiley, New York:339 – 347, 1951.

[43] S. Dash, N. B. Dobbs, O. Günlük, T. J. Nowicki, and G. M. Swirszcz. Lattice-free sets, multi-branch split disjunctions, and mixed-integer programming. Mathematical Programming, A 145:483 – 508, 2014.

[44] S. Dey. A note on the split rank of intersection cuts. Mathematical Programming, A 130:107 – 124, 2011.

[45] S. Dey and Q. Louveaux. Split Rank of Triangle and Quadrilateral Inequalities. Mathematics of Operations Research, 36(3):432 – 461, 2011.

[46] S. S. Dey, A. Santana, and Y. Wang. New SOCP relaxation and branching rule for bipartite bilinear programs. Optimization and Engineering (accepted), 2018.

[47] N. J. Driebeek. An algorithm for the solution of mixed-integer programming problems. Management Science, 12:576 – 587, 1965.

[48] M. A. Duran and I. E. Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Mathematical Programming, 30:307 – 339, 1986.

[49] F. Eisenbrand and G. Shmonin. Carathéodory bounds on integer cones. Operations Research Letters, 34:564 – 568, 2006.

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

[51] R. Fletcher. Practical Methods of Optimization. Wiley, 2013.

[52] R. Fletcher and S. Leyffer. Solving mixed integer nonlinear programs by outer approximation. Mathematical Programming, 66:327–349, 1994.

[53] R. Fletcher and S. Leyffer. Usermanual for filterSQP. https://wiki.mcs.anl.gov/leyffer/images/5/58/SQP_manual.pdf, 1998.

[54] C.A. Floudas. Global optimziation in design and control of chemical process systems. Journal of Process Control, 10:125 – 134, 2001.

[55] J. Forrest and R. Lougee-Heimerl. CBC (Coin-or Branch-and-Cut) Solver. https://projects.coin-or.org/Cbc.

[56] J. Forrest, M. Saltzman, L. Hafer, and J. Hall. COIN-OR Linear Program Solver (CLP). https://projects.coin-or.org/Clp.

[57] A. Frangioni and C. Gentile. Perspective cuts for a class of convex 01 mixed integer programs. Mathematical Programming, A 106:225 – 236, 2006.

[58] T. Gau and G. Wascher. CUTGEN1: A Problem Generator for the Standard One-dimensional Cutting Stock Problem. European Journal of Operational Research, 84:572 – 579, 1995.

[59] M.X. Goemans and T. Rothvoß. Polynomiality for bin packing with a constant number of item types. https://arxiv.org/abs/1307.5108, 2013.

[60] R. E. Gomory. Solving linear programming problems in integers. In: Bellman, M.H.R. (ed.) Combinatorial Analysis. Proceedings of Symposia in Applied Mathematics, 10:269 – 308, 1960.

[61] R. E. Gomory. An algorithm for integer solutions to linear programs, In : Recent Advances in Mathematical Programming. R.L. Graves & P. Wolfe, eds., McGraw-Hill, NewYork, 1963.

[62] R. E. Gomory. Some polyhedra related to combinatorial problems. Linear Algebra and its Apllications, 2:451 – 558, 1969.

[63] I.E. Grossmann. Systematic Methods of Chemical Process Design. Prentice Hall, Uppder Saddle River (NJ), 1997.

[64] M. Grötschel, L. Lovász, and A. Schrijver. The Ellipsoid Methods and its Consequences in Combinatorial Optimization. Combinatorica, 1:169 – 197, 1981.

[65] Martin Grötschel and Manfred W. Padberg. On the symmetric travelling salesman problem II: Lifting theorems and facets. Mathematical Programming, 16(1):281–302, 1979.

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

[67] R. Horst and P.M. Pardalos. Handbook of Global Optimization. Kluwer Academic Publishers, 1995.

[68] R. Horst and H. Tuy. Global Optimization: Deterministic Approaches. Springer Verlag, Berlin, third edition, 1996.

[69] M. Hoskins, R. Masson, G. G. Melançon, J.E. Mendoza, C. Meyer, and L.M. Rousseau. The PrePack Optimization Problem, In: Simonis H. (eds) Integration of AI and OR Techniques in Constraint Programming. CPAIOR 2014. Lecture Notes in Computer Science. Springer, Cham, pages 136 – 143, 2014.

[70] Gurobi Optimization Inc. Gurobi optimizer reference manual, 2016. http://www.gurobi.com.

[71] R. G. Jeroslow. There cannot be any algorithm for integer programming with quadratic constraints. Operations Research, 21(1):221 – 224, 1973.

[72] J. Kallrath. Solving planning and design problems in the process industry using mixed integer and global optimization. Annals of Operations Research, 140:339 – 373, 2005.

[73] J. Kallrath, S. Rebennack, J. Kallrath, and R. Kusche. Solving real-world cutting stock-problems in the paper industry: Mathematical approaches, experience and challenges. European Journal of Operational Research, 238(1):374 – 389, 2014.

[74] N. Karmarkar. A new polynomial time algorithm for linear programming. Combinatorica, 4:375 – 395, 1984.

[75] G. M. Katta and N. K. Santosh. Some NP-Complete Problems in Quadratic and Nonlinear Programming. Mathematical Programming, 39:117 – 129, 1987.

[76] L. G. Khachiyan. A polynomial time algorithm for linear programming. Soviet Mathematics Doklady, 20:191 – 194, 1979.

[77] M. R. Kılınç, J. T. Linderoth, and J. Luedtke. Effective Separation of Disjunctive Cuts for Convex Mixed Integer Nonlinear Programs. Technical Report, University of Wisconsin, Madison, 2010.

[78] A. H. Land and A. G. Doig. An automatic method for solving discrete programming problems. Econometrica, 28:497 – 520, 1960.

[79] H. W. Lenstra. Integer programming with a fixed number of variables. Mathematics of Operations Research, 8:538 – 548, 1983.

[80] A. N. Letchford. Integer quadratic quasi-polyhedra. Integer Programming and Combinatorial Optimization (IPCO), 2010.

[81] L. Liberti, C. Lavor, and N. Maculan. A branch-and-bound algorithm for the molecular distance geometry problem. International Transactions in Operational Research, 15:1 – 17, 2008.

[82] J. Luedtke, M. Namazifar, and J. Linderoth. Some results on the strength of relaxations of multilinear functions. Mathematical Programming, B 136:325–351, 2012.

[83] A. Mahajan, S. Leyffer, J. Linderoth, J. Luedtke, and T. Munson. MINOTAUR: A mixed-integer nonlinear optimization toolkit. ANL/MCS-P8010-0817, Argonne National Lab, 2017.

[84] A. Martin, M. Möller, and S. Moritz. Mixed integer models for the stationary case of gas network optimization. IEEE Transaction on Power Systems, 12:444 – 455, 1997.

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

[86] S. Mitchell, A. Kean, A. Mason, M. O'Sullivan, and A. Phillips. PuLP. https://www.coin-or.org/PuLP/, 2011.

[87] S. Modaresi, M. R. Kılınç, and J. P. Vielma. Split cuts and extended formulations for Mixed Integer Conic Quadratic Programming. Operations Research Letters, 43:10 – 15, 2015.

[88] S. Modaresi, M. R. Kılınç, and J. P. Vielma. Intersection cuts for nonlinear integer programming: convexification techniques for structured sets. Mathematical Programming, A 155:575 – 611, 2016.

[89] J. Momoh, R. Koessler, M. Bond, B. Stott, D. Sun, A. Papalexopoulos, and P. Ristanovic. Challenges to optimal power flow. IEEE Transaction on Power Systems, 12:444 – 455, 1997.

[90] G. L. Nemhauser and L. A. Wolsey. Integer and Combinatorial Optimization. Wiley, 1988.

[91] G. L. Nemhauser and L. A. Wolsey. A recursive procedure to generate all cuts for 0-1 mixed-integer programs. Mathematical Programming, 46:379 – 390, 1990.

[92] Y. Nesterov and A. Nemirovskii. Interior-Point Polynomial Algorithms in Convex Programming. SIAM, 1994.

[93] A. Neumaier. Molecular modeling of proteins and mathematical prediction of protein structure. SIAM Review, 39:407 – 460, 1997.

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

[95] J. Nocedal and S. J. Wright. Numerical Optimization. Springer, 2006.

[96] J. H. Owen and S. Mehrotra. A disjunctive cutting plane procedure for general mixed integer linear programs. Mathematical Programming, A 89:437 – 448, 2001.

[97] I. Quesada and I. E. Grossmann. An LP/NLP based branch-and-bound algorithm for convex MINLP optimization problems. Computers & Chemical Engineering, 16:937–947, 1992.

[98] A. Ramamurthy. Lower bound for cutting stock problem and pattern minimization problem, 2018. M.Tech. Thesis, Indian Institute of Technology Bombay.

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

[100] R. T. Rockafellar. Lagrange Multipliers and Optimality. SIAM Review, 35 (2):183 – 238, 1993.

[101] N. V. Sahinidis. Global Optimization and Constraint Satisfaction: The Branch-and-Reduce Approach, in the book Global Optimization and Constraint Satisfaction, pages 1–16. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[102] A. Saxena, P. Bonami, and J. Lee. Disjunctive Cuts for Non-convex Mixed Integer Quadratically Constrained Programs. In: A. Lodi, A. Panconesi, G. Rinaldi (eds) Integer Programming and Combinatorial Optimization. IPCO. Lecture Notes in Computer Science, vol 5035. Springer, Berlin, Heidelberg, 2008.

[103] N. Z. Shor. Dual quadratic estimates in polynomial and boolean programming. Annals of Operations Research, 25:163 – 168, 1990.

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

[105] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for Optimization over Symmetric Cones. Optimization Methods and Software, 11(12):625 – 653, 1999.

[106] M. Tawarmalani., J. P. Richard, and K. Chung. Strong valid inequalities for orthogonal disjunctions and bilinear covering sets. Mathematical Programming, B 124:481 – 512, 2010.

[107] M. Tawarmalani and N. V. Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. Mathematical Programming, A 99:563 – 591, 2004.

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

[109] S. Umetani, M. Yagiura, and T. Ibaraki. One dimensional cutting stock problem to minimize the number of different patterns. European Journal of Operational Research, 146:388 – 402, 2003.

[110] F. Vanderbeck. Exact Algorithm for Minimising the Number of Setups in the One-Dimensional Cutting Stock Problem. Operations Research, 48(6):915 – 926, 2000.

[111] A. Wächter and L.T. Biegler. On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming. Mathematical Programming, A 106:25 – 57, 2006.

[112] H. Wolkowicz. Semidefinite and lagrangian relaxations for hard combinatorial problems, In: Proceedings of the 19th IFIP TC7 Conference on System Modelling and Optimization. Kluwer, B.V., Deventer, 35 (2):269 – 310, 2000.

[113] L. A. Wolsey. Integer Programming. John Wiley & Sons INC, 1998.

[114] K. Wolter and D. Steffy. Solving Constraint Integer Programs (SCIP). https://scip.zib.de/.

[115] W. Xie and N. V. Sahinidis. A branch-and-bound algorithm for the continuous facility layout problem. Comput. Chem. Eng., 32:1016 – 1028, 2008.

[116] Cui Yaodong, Cheng Zhong, and Yao Yi. Pattern-set generation algorithm for the one-dimensional cutting stock problem with setup cost. European Journal of Operational Research, 243(2):540–546, 2015.

[117] F. You and I.E. Grossmann. Mixed-integer nonlinear programming models and algorithms for supply chain design with stochastic inventory management, 2009. Available at: http://www.minlp.org/library/problem/index.php?i=30.

[118] D. B. Yudin and A. S. Nemirovski. Informational complexity and effective methods of solution of convex extremal problems. Economics and Mathematical Methods, 12:357 – 369, 1976.

Acknowledgments

First of all, I thank the Almighty Allah for giving me the strength, motivation, capability and patience in the entire academic journey and to accomplish the highest academic degree, Doctor of Philosophy.

The Masters and Ph.D. life in IIT Bombay was full of excitement, failures, hopes, complaints, breakthrough and of course fun, and it is a great experience for me. I am privileged to work in such a supportive environment in the beautiful campus as well as in the city of Mumbai. In my entire journey, many people supported, accompanied and mentored me in a different ways. I take this opportunity to acknowledge and thank all of them who were an integrated part of this beautiful journey.

I would like to express my deepest gratitude to my Ph.D. supervisor Prof. Ashutosh Mahajan (Ashu). I am extremely grateful and indebted to him for his continuous support, encouragement, advice and friendly behavior throughout my entire Ph.D. journey. He introduced me to the challenging field of Global Optimization and pushed me towards both the theoretical and computational directions. I am fortunate to have such an inspiring mentor, without whom this thesis would not have been possible. Also, I am privileged to be the first Ph.D. student guided by him.

I would like to thank Prof. Vishnu Narayanan for his valuable suggestions and comments which helped me to improve my Ph.D. work, and also being a part of my Research Progress Committee (RPC). I thank our faculty advisor Prof. K. S. Mallikarjuna Rao for his valuable guidance throughout the journey. I would like to give my heartiest thanks to the other faculty members of the Department of Industrial Engineering and Operations Research - Prof. N. Hemachandra, Prof. Narayan Rangaraj, Prof. Jayendran Venkateswaran, Prof. Veeraruna Kavitha, Prof. P. G. Awate, Prof. Manjesh Hanawal for teaching me, giving me the necessary skill-sets and helping me whenever needed to achieve my goals. I also thank Prof. S. A. Soman for being a part of my RPC. I am blessed to have had professors like them. I also thank Prof. Akshay Gupte (Clemson University, USA) for suggesting me different papers related to my research work.

I would like to thank IEOR office staff, Abasaheb Molavane, Pramod Pawar and Siddarth Salve for their assistance with the office related works and timely help.

I thank my departmental friends/seniors/juniors Arko Chatterjee, Venkateswarao K, Umakanta Pattanayak, Puja Sahu, Siddhartha Paul, Sandhya Tripathi, Arun Verma, Tejas Ghorpade, Indrajit Saha, Prashant Palkar, G. Chandramouli, Sambit Mohanty, Vanessa Beddoe, Khushboo Agarwal, Manish Punjabi, Arun Ramamurthy, Aanchal Panwar (Billi), Shraman Bhaduri, Sitadri Roy etc. for their company, help with little things related to both academic and non-academic, having fun during the entire journey.

I appreciate the friendship of lovely and amazing people in IIT Bombay. I specially thank Ms Anuradha Tewari (Anu) for our unforgettable memories. I thank Siva Muthuprakash, Pragati Chaturvedi (night tea partner), Majid Hassan, Archana Sinha, Tanushree Roy, Subinoy Roy, Aniket Kundu, Shaukat Ali, Jawaid Alam, Mohd Zafar etc. for being with me in all the laughs, the pains, the fight-backs, the struggles which kept me going.

I also would like to thank my friends from the Mumbai city, Manjari Singh and Lakshmi Ahuja, specially for our weekend fun. I thank Mukul Goel for introducing me with such lovely people in the city of Mumbai.

I also thank my childhood friend Kamal Ezaz for his constant support from our childhood.

I am grateful to my parents Azmal Hossain (father) and Najiba Bibi (mother) and other family members for their sacrifices and support for my studies. A special thanks to my brother-in-law Saidul Islam for his continuous support and encouragement which helped me to accomplish my journey and I am truly blessed to have a person like him in my life.

Lastly, I thank our Ministry of Human Resource and Development for the financial support I got for my Ph.D. research work.