EXPLOITING MATHEMATICAL STRUCTURE IN OPTIMIZATION PROBLEMS WITH INDICATOR VARIABLES

by

Hyemin Jeon

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Industrial and Systems Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2015

Date of final oral examination: March 19th 2015

The dissertation is approved by the following members of the Final Oral Committee:
Jeffrey T. Linderoth, Professor, Department of Industrial and Systems Engineering
James R. Luedtke, Associate Professor, Department of Industrial and Systems Engineering
Stephen J. Wright. Professor, Department of Computer Sciences
Laura A. McLay, Associate Professor, Department of Industrial and Systems Engineering
Christos T. Maravelias, Associate Professor, Department of Chemical and Biological Engineering

Acknowledgments

This dissertation is a compilation of my research during the Ph.D. study in University of Wisconsin-Madison. However, the first and foremost credit for all the accomplishment should be given to my advisor, Dr. Jeffrey Linderoth. He has guided me with more encouragement and patience than I deserved throughout the stages of graduate school and I offer my sincere gratitude for that. Remaining errors, if any, are my sole responsibility.

I am grateful to Dr. James Luedtke for giving me an excellent introduction to the field of Integer Programming as well as his continued guidance and advice in research. My appreciation also goes to Drs. Andrew Miller, Santanu Dey, Andy Sun, and Burak Kocuk for sharing their insight which was essential for many results presented in the thesis.

I thank my friends and colleagues in the Department of Industrial and Systems Engineering and the Department of Computer Sciences for the enjoyable memories, their support in time of distress, and sharing their experiences and perspective. I also thank my friends in KOTEL for making my life in Madison happier and healthier.

Last but not least, I thank my wonderful family for their unconditional love and support. Special appreciation goes to Joonhoon for being the loving husband he is and for the help in debugging, proofreading, and other chores in work and life.

Contents

List of Tables

Notations

$\mathbb{R}^n$	set of real vectors of size n
$\mathbb{B}^n$	set of vectors of size n with binary(0 or 1) elements
$\mathbb{R}^{m\times n}$	set of real matrices of size $m \times n$
$\mathbb{S}^n$	set of symmetric matrices of size $n \times n$
$\text{diag}(X)$	vector of diagonal elements of a symmetric matrix X, i.e., $\text{diag}(X)_i = X_{ii}$
$[n]$	set of positive integers $\{1, 2, \ldots, n\}$ for some positive integer n
$\text{conv}(P)$	convex hull of set P
$\text{cone}(V)$	cone generated by a set of vectors V
$\text{ext}(P)$	set of extreme points of set P
$\text{int}(P)$	interior of set P
$\dim(P)$	dimension of set P
$\text{rc}(P)$	recession cone of set P
$\text{rec}(f)$	recession function of a function f
$\|P\|$	cardinality of set P
$\|\|x\|\|_0$	number of nonzero elements of vector x
$\|\|x\|\|_2$	Euclidean norm of vector x
e	vector of ones of conformal dimension
e_j	unit vector of conformal dimension where the j^{th} component is 1 and all other components are zeroes.
$\mathbf{0}$	vector of zeroes of conformal dimension

Abstract

In this thesis we investigate two families of mixed-integer quadratic sets with specific structure of binary indicator variables. These two sets appear as substructures in many interesting optimization problems. The focus of the thesis is to obtain good approximations of the convex hull of these sets in order to improve the solution process of the relevant optimization problems. We present in each chapter the result of our analysis of each of these sets or its polyhedral outer-approximations. By exploiting the mathematical structure of the sets, we derive valid inequalities that result in strong convex relaxations. Computational experiments are performed to examine the importance of the inequalities in describing the convex hull of each set. In all cases, we apply our new inequalities to practical optimization problems, and numerical experiments show that the inequalities are useful in improving the bounds in the solution process of the associated optimization problems.

Chapter 1

Introduction

An Integer Program (IP) is an optimization problem where a subset of the variables are restricted to take integer values. These discrete variables are useful in modeling logical constraints, fixed costs for decisions, and finite sets of alternatives. In many interesting applications, the discrete variables are binary *indicator variables* whose value dictates whether a specific constraint is enforced or if sepcific variables are allowed to be non-zero. The main subject of this thesis is to exploit the mathematical structure of optimization problems with indicator variables to develop strong formulations that are helpful in the solution process.

In this chapter, we first explain the structures of two categories of problems investigated in the thesis and introduce some examples of their applications. Theoretical background and review of relevant research are presented in the following sections. A summary of the thesis and contributions concludes the chapter.

1.1 MIQCP with indicator variables

The class of problems we primarily study in this thesis is the 0-1 Mixed Integer Quadratically Constrained Program (MIQCP). It is an optimization problem where the objective function and one or more constraint functions are quadratic, and a subset of the variables are binary indicators involved only in linear terms. Specifically, we consider a problem of the form

$$\begin{aligned}
\min \quad & x^T Q_0 x + c_0^T x && \text{(0-1 MIQCP)}\\
\text{s.t.} \quad & x^T Q_i x + c_i^T x \leqslant d_i \ \forall i \in [k], &&\\
& Ax + Hz \leqslant f, &&\\
& l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [p], && (1.1)\\
& x \in \mathbb{R}^n, z \in \mathbb{B}^p,
\end{aligned}$$

where $Q_0, Q_i \in \mathbb{S}^{n\times n}$, $c_0, c_i \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $H \in \mathbb{R}^{m\times p}$, $f \in \mathbb{R}^m$, and $d_i \in \mathbb{R}\ \forall i \in [k]$. The constraints (1.1) dictate that x_j is either restricted to 0 or allowed to take any value in the interval $[l_j, u_j]$. Binary variables z that enforce these conditions are *indicator variables*, and the continuous variables regulated in this manner such as x are called *semicontinuous variables*. In this thesis, we assume that $-\infty < l_j < u_j < \infty\ \forall j \in [p]$ and focus our attention to the problems where Q_0, Q_i for each i is positive semidefinite.

Let $\mathcal{F}$ denote a mixed integer set that represents the feasible region for a 0-1 MIQCP problem $\mathcal{P}$. Any set $\bar{\mathcal{F}}$ that contains $\mathcal{F}$ is called a *relaxation* of $\mathcal{F}$. The relaxation of $\mathcal{P}$ refers to the optimization problem with the same objective over a relaxation of the feasible region of $\mathcal{P}$. For MIQCP or Mixed Integer Linear Program (MILP), a natural continuous relaxation is to replace the integrality (binary) constraints with appropriate continuous bounds. For example, the constraint $z \in \mathbb{B}^p$ is replaced with $0 \leqslant z_j \leqslant 1, j \in [p]$. We denote this natural relaxation by $\mathcal{R}(\mathcal{F})$.

The feasible region $\mathcal{F}$ of (0-1 MIQCP) is the set of points satisfying the constraints

$$\begin{aligned}
\mathcal{F} := \{(x, z) \in \mathbb{R}^n_+ \times \mathbb{B}^n \mid\ & x^T Q_i x + c_i^T x \leqslant d_i \ \forall i \in [k],\\
& Ax + Hz \leqslant f,\\
& l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [p]\}.
\end{aligned}$$

When Q_o, Q_i are positive semidefinite, the natural relaxation of (0-1 MIQCP) is a convex QCP problem as it is a minimization of a convex function over a convex set $\mathcal{R}(\mathcal{F})$.

By introducing an auxiliary variable, the problem (0-1 MIQCP) can be reformulated without loss of generality to a problem with a linear objective function:

$$\begin{aligned} \min \quad & \nu && \text{(0-1 MIQCP-}l\text{)} \\ \text{s.t.} \quad & \nu \geqslant x^T Q_0 x + c_0^T x, \\ & x^T Q_i x + c_i^T x \leqslant d_i \ \forall i \in [k], \\ & Ax + Hz \leqslant f, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [p], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^p. && (1.2) \end{aligned}$$

The nonlinearity of the objective function is now moved to the constraints. The feasible set $\mathcal{F}_l$ for (0-1 MIQCP-*l*) is defined in a space one dimension higher than $\mathcal{F}$,

$$\begin{aligned} \mathcal{F}_l := \{(\nu, x, z) \in \mathbb{R}^{n+1} \times \mathbb{B}^n \mid \ & \nu \geqslant x^T Q_0 x + c_0^T x, \\ & x^T Q_i x + c_i^T x \leqslant d_i \ \forall i \in [k], \\ & Ax + Hz \leqslant f, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [p]\}. \end{aligned}$$

An important difference between these two equivalent formulations, (0-1 MIQCP) and (0-1 MIQCP-*l*), is that an optimal solution x^* for (0-1 MIQCP-*l*) occurs at an extreme point of the convex hull of the feasible region, whereas the optimal solution for (0-1 MIQCP) may reside in the interior of the convex hull of its feasible region. Thus having the complete description of the convex hull of $\mathcal{F}_l$ allows us to solve a convex continuous optimization problem whose solution also solves (0-1 MIQCP-*l*).

The set $\text{conv}(\mathcal{F}_l)$ is an extremely complicated object not easily amenable to mathematical analysis. Instead, we focus our study on a relaxation of $\mathcal{F}_l$

$$\begin{aligned} \mathcal{T} := \{(\nu, x, z) \in \mathbb{R}^{n+1} \times \mathbb{B}^n \mid \ & \nu \geqslant x^T Q_0 x + c_0^T x, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [p]\}, \end{aligned}$$

where all constraints except for one quadratic constraint and the indicator constraint are ignored. Since $\mathcal{T}$ is a relaxation of $\mathcal{F}_l$, valid inequalities for $\mathcal{T}$ are also valid for $\mathcal{F}_l$. We hope to apply the knowledge gained by studying the structure of $\mathcal{T}$ to solving general 0-1 MIQCP. Note that the feasible region defined by general convex quadratic constraints with indicator variables can be relaxed to the set $\mathcal{T}$ defined above. Specifically, we can introduce an auxiliary variable w_i for each constraint $x^TQ_ix + c_i^Tx \leqslant d_i$ to reformulate it as $w_i \geqslant x^TQ_ix + c_i^Tx,\ w_i \leqslant d_i$. We do not consider the intersection of these relaxations in this study.

Our motivation in studying this relaxation comes from previous studies that have shown that reformulating and generating valid inequalities for important, problem-specific relaxations can be extremely useful when applied to solve the original problem. Crowder, Johnson, and Padberg [18] demonstrated that utilizing the well-understood characteristics of knapsack polytope is very advantageous in the numerical solution of 0-1 MILP problems. Padberg, Van Roy, and Wolsey [49] investigate MIPs with binary variables involved only in indicator constraints, derive classes of facet-defining inequalities for its feasible region, and illustrate using these inequalities to strengthen formulations. For Mixed Integer Nonlinear Program (MINLP), Frangioni and Gentile [28] suggest the technique for strengthening partially reformulated substructure of 0-1 MIQCP and demonstrate its usefulness. Their approach is introduced in more detail in Section 1.5.7.

The problem 0-1 MIQCP is a special case of a 0-1 MINLP where the objective and constraint functions can be general nonlinear functions. Important special cases of 0-1 MIQCP are the 0-1 Mixed Integer Quadratic Program (0-1 MIQP) where $k = 0$, and the 0-1 Mixed Integer Linear Program (0-1 MILP) where $k = p = 0$.

1.2 Cycle substructure with indicator variables

Another problem structure that we investigate arises from the transmission switching problem. In the transmission switching problem, the switching decision on each transmission line is represented by a binary indicator variable. Here we introduce a specific relaxation for this problem and later we provide a more detailed introduction to the DC switching problem in Section 1.3.4 and Chapter 4. The specific relaxation comes from extracting a cycle substructure of the transmission line network. We denote this substructure by a directed cycle (V, C) where $V = [n],\ C = \{(1,2),(2,3),\cdots,(n,1)\}$ and define the set

$$\begin{aligned}\mathcal{C} = \Big\{(x,\theta,z) \in \mathbb{R}^{2n} \times \{0,1\}^n \mid -u_{ij}z_{ij} \leqslant x_{ij} \leqslant u_{ij}z_{ij}\ \forall(i,j) \in C,\\ z_{ij}(\theta_i - \theta_j) = x_{ij}\ \forall(i,j) \in C\Big\}\end{aligned} \tag{1.3}$$

which contains a subset of constraints of the original DC switching problem described in more detail in Chapter 4. The binary variable z_{ij} is an indicator of whether or not the line (i,j) is available. A special characteristic of z_{ij} is that when it enforces one constraint, another constraint is relaxed. If $z_{ij} = 1$, the bounds on flow x_{ij} are relaxed and a non-zero flow is allowed on the line between buses i and j. Additionally, if $z_{ij} = 1$, the flow x_{ij} must satisfy Ohm's law, which states that the power flow on a line is proportional to the potential difference at its endpoints. Otherwise, if $z_{ij} = 0$, the power flow x_{ij} is restricted to 0, but the constraint $\theta_i - \theta_j = x_{ij}$ no longer has to be satisfied.

Writing the constraints in (1.3) as

$$\begin{aligned} x_{12} &= z_{12}(\theta_1 - \theta_2), \\ x_{23} &= z_{23}(\theta_2 - \theta_3), \\ &\cdots, \\ x_{n1} &= z_{n1}(\theta_n - \theta_1), \end{aligned}$$

it becomes easier to see that this system of equations has an interesting attribute that when $z_{ij} = 1 \; \forall (i,j) \in C$, it is implied that

$$\sum_{(i,j)\in C} x_{ij} = (\theta_1 - \theta_2) + (\theta_2 - \theta_3) + \cdots + (\theta_n - \theta_1) = 0.$$

We exploit this characteristic to derive strong valid inequalities for this relaxation in Chapter 4. Note that this approach may be applicable for other problems with similar structure involving network flow models where flow is proportional to the potential difference.

In the remainder of this chapter, we first introduce some examples of applications of convex 0-1 MIQCP in Section 1.3, present brief explanation on the existing algorithms that can be applied to solve general convex 0-1 MIQCPs in Section 1.4, and give background theories and techniques related to solving these problems in Section 1.5. The chapter concludes with remarks on the contribution of the thesis and roadmap for future research in Section 1.6.

1.3 Applications of Optimization with Indicator Variables

In this section, a few of important applications of convex 0-1 MIQCP in various fields and the DC transmission switching problem are introduced.

1.3.1 MIQCP: Portfolio Management

Portfolio Selection is a widely studied problem in finance based on the pioneering work of Markowitz [42] with the goal to construct a portfolio of financial equities that has the lowest possible collective variance/volatility/risk while providing a certain level of expected return. The portfolio is often constructed to closely follow a stock market index as a benchmark, in which case the tracking error with respect to the benchmark, also refered to as active variance, is used as a risk measure instead of portfolio variance itself. The problem of constructing a portfolio with the least tracking error satisfying some constraints is called *Benchmark Tracking*. Due to transaction and contract costs, the number of equities to be included in the portfolio and proportion of each equity is often constrained. There have been many studies on how to formulate and solve the Benchmark Tracking problem including Jansan and van Dijk [51], and the problem can be modeled as a 0-1 MIQP of the following form :

$$\begin{aligned} \min \quad & (x-b)^T Q (x-b) \\ \text{s.t.} \quad & \alpha^T x \geqslant \rho, \\ & e^T x = 1, \\ & e^T z \leqslant \kappa, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n. \end{aligned} \tag{1.4}$$

In this model, variables x represent the proportions of the n individual assets in the portfolio, the vectors $\alpha, b \in \mathbb{R}^n$,respectively, denote the expected return and the proportions of the n assets in the benchmark, and the objective function $(x-b)^T Q(x-b)$ represents the total tracking error defined for a given positive semidefinite covariance matrix Q. The solution to this problem defines a portfolio with the minimum possible tracking error consisting up to κ assets and having an expected return of at least ρ.

1.3.2 MIQCP: Variable Subset Selection for Regression

A fundamental problem in statistics is to describe a relationship between dependent and independent variables refered to as regression. A regression model is estimated from observed or experimentally achieved data, often by building a linear model that describes the data with the minimum sum of squared errors [48]. In multivariate regression with a large number of independent variables, the minimum error is often achieved when all the available independent variables can be utilized. However, there are benefits of having a "simple" explanation for the dependent variable using only a selected subset of independent variables [33, 45]. Some benefits of having a sparse regression model are a lower cost for estimating or predicting, smaller standard errors in coefficient estimation, reduced training and utilization time, and facilitating data visualization. Guyon and Elisseeff [33] also point out that the importance of variable subset selection has become more significant with new domains of applications where the number of variables is huge such as gene selection or text categorization.

Bertsimas and Shioda [11] discuss an algorithm for variable subset selection that relies on a formulation of the problem as QP with a cardinality constraint. The objective of this optimization problem is to find the vector β that minimizes the sum of squared errors $\sum_i (y_i - \beta^T x_i)^2$ for m given data points $(x_i, y_i), x_i \in \mathbb{R}^d, y_i \in \mathbb{R}, i \in [m]$. The problem may be mathematically expressed as

$$\begin{aligned} \min \quad & (y - X\beta)^T (y - X\beta) \\ \text{s.t.} \quad & \|\beta\|_0 \leqslant \kappa, \\ & \beta \in \mathbb{R}^d \end{aligned} \tag{1.5}$$

where $X \in \mathbb{R}^{m \times d}$ whose rows are $x_i, i \in [m]$.

If the independent variables have bounded weights, $l_j \leqslant \beta_j \leqslant u_j \; \forall j \in [d]$, then the problem can be reformulated to a 0-1 MIQCP with binary variables z_j indicating whether or not each β_j is allowed to be positive:

$$\begin{aligned} \min \quad & (y - X\beta)^T (y - X\beta) \\ \text{s.t.} \quad & l_j z_j \leqslant \beta_j \leqslant u_j z_j \;\; j \in [d], \\ & e^T z \leqslant \kappa, \\ & \beta \in \mathbb{R}^d. \end{aligned} \tag{1.6}$$

Typically, $l_j = -M, \; u_j = M$ for a scalar M which needs to be sufficiently large for the solution of (1.6) to be also a solution of (1.5). On the other hand, it is important to choose a value that is not unnecessarily large, as letting M take too large a value results in slow solution process. Bertsimas et. al. [12] suggest an approach to compute these parameters from the given data. The authors also demonstrated in the study that using warm starts in Integer Programming solvers outperforms other state of the art (statistical) methods to solve the subset selection problem. Specifically, their approach solved the problems with m in the $1,000$s and d in the 100s within minutes to provable optimality, and found solutions near optimality for problems with m in the 100s and d in the $1,000$s within minutes.

1.3.3 MIQCP: Sparse Digital Filter Design

A digital filter is a system that processes a digital signal by means of performing arithmetic operations. How the filter mathematically manipulates input signals is specified as the coefficient defined at the filter design stage. As the number of arithmetic operations is the cost-dominating factor for implementation, it is desirable to design a *sparse* filter with few non-zero coefficients. The sparsity of the filter also helps reduce the cost related to hardware and energy consumption [25].

There have been different lines of approach to the sparse filter design problem taking different performance criteria into consideration. The criteria include weighted least-squares in approximating an ideal frequency response, simplified mean squared error, and signal-to-noise ratio. It is known that designing a sparse digital filter with constraints on these three metrics can all be formulated as an optimization problem of the following form:

$$\begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & (x-c)^T Q (x-c) \leqslant \gamma, \\ & x \in \mathbb{R}^n, \end{aligned} \tag{1.7}$$

where $Q \in \mathbb{S}^n$ is positive semidefinite, $c \in \mathbb{R}^n$, and $\gamma \in \mathbb{R}_+$.

This problem can be reformulated as a mixed 0-1 MIQCP of the form

$$\begin{aligned} \min \quad & e^T z \\ \text{s.t.} \quad & (x-c)^T Q (x-c) \leqslant \gamma, \\ & l_j z_i \leqslant x_j \leqslant u_j z_j \ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{Z}^n, \end{aligned}$$

where $l, u \in \mathbb{R}^n$ are lower and upper bounds on x which are selected so that the feasible set of x defined by the quadratic constraints (1.7) is not changed.

1.3.4 DC Transmission Switching

The power flow network is typically depicted in a graph (N, E) where the set of nodes N denote the set of buses and set of arcs E the lines between buses. Variables θ_i for each node $i \in N$ and x_{ij} for each arc $(i, j) \in E$ represent the phase angle of bus i and power flow on line (i, j), respectively. The parameter α_{ij} for each arc $(i, j) \in E$ represent the susceptance of line (i, j). The Direct Current (DC) power flow equation provides a simplification of the physical laws governing the status of components of the power flow network and is written as follows:

$$x_{ij} = \alpha_{ij}(\theta_i - \theta_j) \ \forall (i, j) \in E. \tag{1.8}$$

It is an interesting paradox that the efficiency of the power flow network can be increased by switching some of the lines on or off. The switching decision is typically modeled by introducing binary variables z_{ij} indicating whether or not line (i, j) is used in the network. If the line is switched off, then the flow on the line should be 0. As discussed in Section 1.2, the DC flow equation should not be imposed on the line in this case. This relationship can be modeled by the following two constraints:

$$x_{ij} = \alpha_{ij} z_{ij} (\theta_i - \theta_j) \quad \forall (i, j) \in E, \tag{1.9}$$

$$-u_{ij} z_{ij} \leqslant x_{ij} \leqslant u_{ij} z_{ij} \quad \forall (i, j) \in E. \tag{1.10}$$

The bilinear equation (1.9) can be equivalently formulated using big-M approach as in the work by Fisher, O'Neil, and Ferris [26]. To fully write a formulation for optimal transmission switching, we define the sets, variables, and parameters are as

Sets	$R \subset N$	: set of generator nodes
	$D \subset N$	: set of demand nodes
Variables	x_{ij}	: power flow on arc $(i, j) \in E$
	p_i	: amount of power produced from generator $i \in R$
	θ_i	: voltage angle at node $i \in N$
Parameters	c_i	: unit cost of production at generator $i \in R$
	d_i	: demand at node $i \in D$
	u_{ij}	: capacity on line$(i, j) \in E$
	p_i^{min}/p_i^{max}	: minimum / maximum production at generator $i \in R$.

With these definitions, a formulation of optimal power flow with optimal transmission switching

(OPF-OTS) is of the form:

$$\min \quad \sum_{i \in R} c_i p_i \tag{OPF-OTS}$$

$$\text{s.t.} \quad \sum_{j:(i,j)\in E} x_{ij} - \sum_{j:(j,i)\in E} x_{ij} = \begin{cases} p_i & \forall i \in R \\ d_i & \forall i \in D \\ 0 & \text{otherwise,} \end{cases}$$

$$-u_{ij} z_{ij} \leqslant x_{ij} \leqslant u_{ij} z_{ij} \; \forall (i,j) \in E,$$

$$p_i^{min} \leqslant p_i \leqslant p_i^{max} \; \forall i \in R,$$

$$\alpha_{ij}(\theta_i - \theta_j) - x_{ij} + M(1 - z_{ij}) \geqslant 0 \; \forall (i,j) \in E, \tag{1.11}$$

$$\alpha_{ij}(\theta_i - \theta_j) - x_{ij} - M(1 - z_{ij}) \leqslant 0 \; \forall (i,j) \in E, \tag{1.12}$$

$$\sum_{(i,j)\in E} z_{ij} \leqslant \kappa,$$

$$x_{ij} \in \mathbb{R}, z_{ij} \in \mathbb{B} \; \forall (i,j) \in E, p_i \in \mathbb{R}_+, \theta_i \in \mathbb{R} \; \forall i \in N.$$

The objective of this problem is to find the optimal generating and switching decisions that minimize the total power generation cost satisfying demand where the power flow is governed by the DC flow constraint (1.8). For a scalar M sufficiently large, these constraints are only imposed on the arcs that are switched on.

1.4 Algorithms for solving MINLPs

Some of the most widely used algorithms to solve general convex MINLP are briefly introduced in this section. A more thorough explanation and comparison of algorithms and software for convex MINLP can be found in a review paper of Bonami, Kilinç, and Linderoth [14]. Belotti et al. [9] provide a survey of cutting planes and deterministic methods for convex MINLPs, formulations and heuristics for solving non-convex MINLPs, and a comparison of available solvers for both convex and non-convex MINLPs.

A general MINLP can be expressed in the following algebraic form;

$$\begin{aligned} \min \quad & f(x, z) && \text{(MINLP)} \\ \text{s.t.} \quad & g_j(x, z) \leqslant 0 \; \forall j \in [m], \\ & Ax + Bz \leqslant d, \\ & L \leqslant z \leqslant U, \\ & x \in \mathbb{R}^n, z \in \mathbb{Z}^p. \end{aligned}$$

If the nonlinear functions f, g_j are convex, we say the MINLP is convex. MIQCP and MILP are special cases of MINLP where f, g_j are quadratic and linear, respectively. We assume that the set $X = \{(x, z) \mid Ax + Bz \leqslant d\}$ is bounded and lower bound L and upper bound U on discrete variables z can be deduced for feasible points.

There are different exact methods for solving convex MINLPs. Most commonly applied methods use a branch-and-bound approach, and differ from each other in terms of the subproblems considered in the solution process.

A subproblem often used is the continuous relaxation with restricted bounds. For some lower bound $l \geqslant L$ and upper bound $u \leqslant U$ on discrete variables y, the NLP relaxation *NLPR(l,u)* is defined as

$$\begin{aligned} \min \quad & f(x, y) && \text{(NLPR}(l, u)) \\ \text{s.t.} \quad & g_j(x, y) \leqslant 0 \; \forall j \in [m], \\ & Ax + By \leqslant d, \\ & l \leqslant y \leqslant u, \\ & x \in \mathbb{R}^n, y \in \mathbb{R}^p. \end{aligned}$$

If f, g_j are convex, then *NLPR(l,u)* is tractable. The optimal objective value of NLPR(L, U) provides a lower bound on the optimal objective value of the original MINLP.

Here we briefly introduce the algorithm within which we wish to use the knowledge obtained through our investigation of structured mixed integer sets. NLP-based branch and bound is in principal identical to branch and bound for MILP except that linearity of subproblems is not required. Through enumeration of subproblems, the lower bound ν_L and upper bound ν_U for the optimal objective value ν^* for (MINLP) are updated. The algorithm is summarized in Algorithm 1.1. The set T denotes the set of nodes, where a node is characterized by a pair of vectors (l^i, u^i) representing the subproblem NLPR(l^i, u^i). The optimal objective value and optimal solution of NLPR(l^i, u^i) are denoted by v^i and (x^i, y^i).

Algorithm 1.1: NLP-Based Branch and Bound

---

0. Initialize
Set $z_U = \infty$, $T = \{(L, U)\}$.

---

1. Enumerate
Choose a node (l^i, u^i) and remove it from T. Solve NLPR(l^i, u^i).
if *NLPR(l^i, u^i) is infeasible or* $v^i > v_U$ **then**
    go to step 3.
else if y^i *is integral* **then**
    Set $v_U = v^i$, $(x^*, y^*) = (x^i, y^i)$.
    Delete all nodes from T such that $v_L^j \geqslant v_U$.
    Go to step 3.
else
    go to step 2.

---

2. Divide
Divide the interval (l^i, u^i) into subintervals $(l^{i_1}, u^{i_1}), (l^{i_2}, u^{i_2}), \cdots, (l^{i_k}, u^{i_k})$.
For each $j \in [k]$, create and add a node to T and let $v_L^j = v^j$.
Go to step 1.

---

3. Terminate
if $T = \emptyset$ **then**
    return (x^*, y^*) as optimal solution.
else
    go to step 1.

The efficiency of this algorithm depends on many algorithmic decisions such as node selection or branching rules, cuts added, and the use of heuristics for obtaining a feasible point. Our approach explained in Section 1.1 is applicable in this scheme and will work by strengthening the formulation for subproblem NLPR(l, u). Specifically, we add cuts valid for NLPR(l, u) so that the lower bound it provides is improved, with the goal of improving the entire solution process. If the lower bound v^i improves, the chance of getting node (l^i, u^i) pruned in step *1.* of the algorithm increases, which is likely to lead a smaller amount of enumeration until the termination criterion is satisfied.

1.5 Theoretical Background

In this section, we present some background material on techniques we will use in our investigation of convex 0-1 MIQCP. Our mathematical terminology is essentially standard. See page iv for a complete list of notations used in the thesis.

1.5.1 Theory of Polyhedra

In Chapter 2 and Chapter 4, we will discuss polyhedral relaxations and reformulations of mixed integer nonlinear sets, so here we introduce some common terminology and results related to polyhedron.

Definition 1.1. *A subset of $\mathbb{R}^n$ described by a finite set of linear constraints*

$$P = \{x \in \mathbb{R}^n \mid Ax \leqslant b\}$$

is a polyhedron.

The vast importance of understanding polyhedra with regard to mixed-integer optimization is based on the well known result due to Meyer [44].

Theorem 1.1. *[44] The convex hull of a mixed integer set*

$$X = \{(x, y) \in \mathbb{R}^n_+ \times \mathbb{Z}^p_+ \mid Ax + Gy \leqslant b\}$$

is a polyhedron if A, G, b *are rational.*

It follows from Theorem 1.1 that a Mixed Integer Linear Program(MILP) of the form

$$\max\{cx + dy \mid (x, y) \in X\}$$

can be reformulated as an LP

$$\max\{cx + dy \mid (x, y) \in \text{conv}(X)\}.$$

Therefore, having the complete set of inequalities that describes the convex hull of the feasible set would provide the ideal LP formulation for MIPs. When it is not available, we aim to obtain a good approximation of the convex hull by adding cuts (inequalities) that strengthen the formulation. Knowledge on polyhedra can provide a basis on which we can determine the strength of cuts since the best possible cuts are the ones that are necessary in describing a polyhedron.

Review of relevant results can be found in Wolsey [55] or Conforti, Cornuéjols and Zambelli [17], some of which we introduce below without proof.

A polyhedron may be expressed in two different ways.

Theorem 1.2. *For a set $P \subset \mathbb{R}^n$, the following two conditions are equivalent:*

1. *P is a polyhedron, i.e., there is a matrix A and a vector b such that $P = \{x \in \mathbb{R}^n \mid Ax \leqslant b\}$.*
2. *There are finite sets of vectors $v^1, \dots, v^p, r^1, \dots, r^q$ such that*

$$P = \text{conv}(v^1, \dots, v^p) + \text{cone}(r^1, \dots, r^q).$$

Definition 1.2. *(i) If the maximum number of affinely independent vectors in polyhedron P is $d+1$, the dimension of P is d.*
(ii) The polyhedron P is full dimensional if its dimension is n.

An inequality satisfied by all points in P is called valid for P.

Definition 1.3. *An inequality $\pi x \leqslant \pi_0$ is* valid *for polyhedron P if $\pi\bar{x} \leqslant \pi_0$ $\forall \bar{x} \in P$.*

A polyhedron has faces, and only the highest-dimensional faces are necessary in its inequality description.

Definition 1.4. *(i) F defines a face of the polyhedron P if $F = \{x \in P \mid \pi x = \pi_0\}$ for some valid inequality $\pi x \leqslant \pi_0$ of P.*
(ii) F is a facet of P if F is a face of P and $\dim(F) = \dim(P) - 1$.
(iii) If F is a face of P with $F = \{x \in P \mid \pi x = \pi_0\}$, the valid inequality $\pi x \leqslant \pi_0$ is said to define the face.

Theorem 1.3. *If a polyhedron P is full-dimensional, a valid inequality $\pi x \leqslant \pi_0$ is necessary in the description of P if and only if it defines a facet of P.*

Theorem 1.1 and Theorem 1.3 provide the justification for looking for facets of the convex hull of a mixed integer set as a means of solving an MILP.

1.5.2 Theory of Convex Sets

Recently, Dey and Moran [21] were able to generalize Theorem 1.1 to characterize the convex hull of integer points in other types of closed convex sets besides polyhedra. The characterization relied on the following definition.

Definition 1.5. *[21] The closed convex set* $K \subset \mathbb{R}^n$ *is* thin *if the following condition holds for all* $c \in \mathbb{R}^n$:

$$\min\{c^T x \mid x \in K\} = -\infty \textit{ if and only if there exists } d \in \mathrm{rc}(K) \textit{ such that } c^T d < 0.$$

The following theorem of Dey and Moran states the necessary and sufficient conditions for the convex hull of the integer set $(K \cap \mathbb{Z}^n)$ to be polyhedral. Note that all polyhedral sets are thin convex sets.

Theorem 1.4. *[21] Let* $K \subset \mathbb{R}^n$ *be a closed convex set. If* K *is thin and the recession cone of* K *is a rational polyhedral cone, then* $\mathrm{conv}(K \cap \mathbb{Z}^n)$ *is a polyhedron. Moreover, if* $\mathrm{int}(K) \cap \mathbb{Z}^n \neq \emptyset$ *and* $\mathrm{conv}(K \cap \mathbb{Z}^n)$ *is a polyhedron, then* K *is thin and the recession cone of* K *is a rational polyhedral cone.*

Example 1.1. *Define an unbounded closed convex set*

$$K_1 = \{(x, t, z) \in \mathbb{R}^3 \mid t \geqslant x^2,\ 0 \leqslant x \leqslant z\}$$

with the recession cone $\mathrm{rc}(K_1) = \{(0, d, 0) \mid d \geqslant 0\}$*. Since* K_1 *is thin and* $\mathrm{rc}(K_1)$ *is a rational polyhedral cone, Theorem 1.4 tells us that* $\mathrm{conv}(K_1 \cap \mathbb{Z}^3)$ *is polyhedral.*

It can be easily verified from the graphical representation that

$$\mathrm{conv}(K_1 \cap \mathbb{Z}^3) = \mathrm{conv}\{(0,0,0), (0,0,1), (1,1,1)\} + \mathrm{cone}((0,1,0))$$

and indeed is polyhedral. But the result no longer holds for a mixed integer set with a similar definition and the same recession cone.

Example 1.2. *Define a mixed 0-1 convex set*

$$K_2 = \{(x, t, z) \in \mathbb{R}^2 \times \mathbb{B} \mid t \geqslant x^2,\ 0 \leqslant x \leqslant z\}$$

with the recession cone $\mathrm{rc}(K_2) = \{(0, d, 0) \mid d \geqslant 0\}$*. It can be shown [32] that*

$$\mathrm{conv}(K_2) = \{(x, t, z) \in \mathbb{R}^3 \mid zt \geqslant x^2,\ 0 \leqslant x \leqslant z,\ 0 \leqslant z \leqslant 1\}$$

which is not polyhedral.

Example 1.2 demonstrates that it will not suffice for our analysis to seek facets of the convex hull for the mixed integer sets we study. In interesting recent work, Kilinç-Karzan [39] generalizes the concept of linear inequalities being in a minimal description of a polyhedron to a concept known as $\mathcal{K}$-*minimality* (with respect to cone $\mathcal{K}$) for some specially structured mixed integer conic sets. Yet, for general mixed integer sets with non-polyhedral convex hull, there does not exist a widely-used criterion for determining the necessity of an inequality in describing the convex hull such as the facet-defining quality introduced in Definition 1.4 and Theorem 1.3.

In this work, we instead introduce our own concepts about what it means for valid inequalities to be *useful* for describing a mixed integer set. Let $\mathcal{F}$ denote a mixed integer set representing a feasible region of an MINLP problem,

$$\mathcal{F} = \{x \in \mathbb{R}^{n-p} \times \mathbb{Z}^p \mid f(x) \leqslant 0\}.$$

We assume that $\mathcal{F}$ has a continuous relaxation $\mathcal{R}(\mathcal{F})$ ($\mathcal{F} = \mathcal{R}(\mathcal{F}) \cap (\mathbb{R}^{n-p} \times \mathbb{Z}^p)$) that is closed and convex.

Definition 1.6. *We say an inequality* $h(x) \leqslant 0$ *is* useful *for* $\bar{\mathcal{F}} \supseteq \mathcal{F}$ *if* $h(x) \leqslant 0\ \forall x \in \mathcal{F}$, *but there exists* $\hat{x} \in \bar{\mathcal{F}}$ *such that* $h(\hat{x}) > 0$.

For the NLP-based branch-and-bound algorithm introduced in Section 1.4, the subproblems that gets solved are continuous relaxations with modified bounds or added cuts. If the problem has a linear objective function as in (0-1 MIQCP-*l*), the optimal solutions to the subproblems are obtained at extreme points of the feasible region of relaxations. If the optimal solution to the subproblem is also feasible to the original discrete problem, then the problem is solved. This is always the case when we have a complete description of the convex hull of feasible region, as the extreme points of the convex hull are contained in the original feasible region. If the subproblem solution is not feasible to the original problem, the solution point needs to be separated from the feasible region of the original problem. The purpose of generating a cut (or a valid inequality) is to perform this separation and provide tighter bounds. Definition 1.6 describes precisely the inequalities that are able to serve this purpose for a relaxation $\bar{\mathcal{F}} \supseteq \mathcal{F}$.

The complete description of convex hull is known for some mixed integer sets with specific structure (e.g., K_2 in Example 1.2), but it is generally very difficult to obtain it. However, it is sometimes possible to obtain a description of the convex hull in a higher dimensional space called an *extended formulation*.

Definition 1.7. *For a set* $Q \in \mathbb{R}^n \times \mathbb{Z}^p$, $\bar{Q} \in \mathbb{R}^{n+p+q}$ *is an extended formulation for* Q *if* $\text{proj}_{\mathbb{R}^{n+p}}(\bar{Q}) = \text{conv}(Q)$.

1.5.3 Second Order Cone Programming

The problem (0-1 MIQCP) defined in Section 1.1 with a convex quadratic objective and constraints, (i.e., with matrices $Q, Q_i \succeq 0 \ \forall i \in [k]$,) is a special case of a Mixed Integer Second Order Cone Program (MISOCP). An MISOCP problem can be written as

$$\begin{aligned} \min \quad & c^T x + r^T y && \text{(MISOCP)} \\ \text{s.t.} \quad & \|A_i x + G_i y - b_i\|_2 \leqslant a_i^T x + g_i^T y - h_i, \ i \in [k] \\ & x \in \mathbb{R}^n, y \in \mathbb{Z}^p. \end{aligned}$$

To cast a convex 0-1 MIQCP problem as a 0-1 MISOCP problem, the quadratic constraint

$$x^T Q x + c^T x \leqslant d$$

where $Q \succeq 0$ is transformed to an equivalent second order conic constraint

$$\left\| \begin{matrix} Lx \\ \frac{1+c^T x - d}{2} \end{matrix} \right\|_2 \leqslant \frac{1 - c^T x + d}{2},$$

where L is from the Cholesky factorization of $Q = L^T L$. This can be verified via a simple arithmetic argument. Taking squares of both sides of the inequality and rearranging, we obtain

$$\begin{aligned} & x^T L^T L x + \left(\frac{1 + c^T x - d}{2}\right)^2 \leqslant \left(\frac{1 - c^T x + d}{2}\right)^2 \\ \Leftrightarrow \ & x^T (L^T L) x + \frac{1 + (c^T x)^2 + d^2 + 2c^T x - 2d - 2d(c^T x)}{4} \\ & \qquad - \frac{1 + (c^T x)^2 + d^2 - 2c^T x + 2d - 2d(c^T x)}{4} \leqslant 0 \\ \Leftrightarrow \ & x^T Q x + \frac{4c^T x - 4d}{4} \leqslant 0 \\ \Leftrightarrow \ & x^T Q x + c^T x \leqslant d. \end{aligned}$$

A *rotated* second order cone is represented with a constraint of the form

$$\|x\|^2 \leqslant yz, \ y, z \geqslant 0 \tag{1.13}$$

for some vector $x \in \mathbb{R}^n$. The inequality (1.13) can be equivalently written as the standard form of conic quadratic constraint

$$\left\| \begin{matrix} 2x \\ y - z \end{matrix} \right\|_2 \leqslant y + z. \tag{1.14}$$

The second-order cone is often referred to as the *Lorentz* cone defined as

$$\mathcal{L}^m = \left\{ (x_1, \cdots, x_{m-1}, x_m) \in \mathbb{R}^m \mid x_m \geqslant \|(x_1, \cdots, x_{m-1})\|_2 \right\}.$$

It is often useful to transform a convex quadratic constraint to a conic quadratic constraint as efficient algorithms are available for solving SOCPs. We take the definition suggested by Ben-Tal and Nemirovski [10] for sets that are represented by conic quadratic inequalities.

Definition 1.8. *A set* $X \in \mathbb{R}^n$ *is Second Order Cone Representable (SOC-Representable) if there exists a system of finitely many inequalities of the form* $A_j x + H_j u - b_j \in \mathcal{L}^{m_j}$ *where* $x \in \mathbb{R}^n$, u *are additional variables, and* $\mathcal{L}^{m_j}$ *is the Lorentz cone of the appropriate dimension such that*

$$x \in X \Leftrightarrow \exists u \text{ such that } A_j x + H_j u - b_j \in \mathcal{L}^{m_j} \; \forall j \in [k].$$

Software packages that are able to solve SOCP include commercial codes CPLEX, MOSEK, Gurobi, and academic codes SeDuMi, and CVX. A brief overview of some of the solvers can be found in the survey of Pólik [50]. Later in Chapter 3 we derive valid inequalities that are conic quadratic, which allows for the use of above SOCP solvers for computations.

Ben-Tal and Nemirovski [10] introduce various types of inequalities that can be reformulated to conic quadratic constraints. Aktürk, Atamtürk, and Gürel [2] introduce a conic quadratic reformulation of the machine-job assignment problem that can be used for other 0-1 convex problems, and demonstrate its effectiveness for solving the problem to optimality. Rather recently, there has been research on the techniques of strengthening the formulation for MISOCP some of which are briefly introduced in Section 1.5.6.

1.5.4 Disjunctive Programming

A disjunctive set is a union of sets. Disjunctive programming is optimization over disjunctive sets. The concept was developed for the union of polyhedra in the seventies by Balas. The notations and terminology in this section is mainly according to [7].

The following theorem describes the extended formulation of convex hull of a disjunctive set consisting of union of polyhedra.

Theorem 1.5. *[6] Given polyhedra* $P_i = \{x \in \mathbb{R}^n \mid A^i x \geqslant b^i\} \neq \emptyset$, $i \in \mathcal{Q}$, *the closed convex hull of* $\bigcup_{i\in\mathcal{Q}} P_i$ *is the set of* $x \in \mathbb{R}^n$ *for which there exist vectors* $(y^i, y_0^i) \in \mathbb{R}^{n+1}$, $i \in \mathcal{Q}$ *satisfying*

$$\begin{aligned}
&x - \sum_{i\in\mathcal{Q}} y^i = 0,\\
&A^i y^i - b^i y_0^i \geqslant 0,\\
&y_0^i \geqslant 0,\ i \in \mathcal{Q},\\
&\sum_{i\in\mathcal{Q}} y_0^i = 1.
\end{aligned} \tag{1.15}$$

In particular, denoting by $P_\mathcal{Q} := \operatorname{clconv} \bigcup_{i\in\mathcal{Q}} P_i$ *the closed convex hull of* $\bigcup_{i\in\mathcal{Q}} P_i$ *and by* $\mathscr{P}$ *the set of vectors* $(x, \{y^i, y_0^i\}_{i\in\mathcal{Q}})$ *satisfying 1.15,*

(i) if x^* *is an extreme point of* $P_\mathcal{Q}$*, then* $(\bar{x}, \{\bar{y}^i, \bar{y}_0^i\})_{i\in\mathcal{Q}}$ *is an extreme point of* $\mathscr{P}$*, with* $\bar{x} = x^*$, $(\bar{y}^k, \bar{y}_0^k) = (x^*, 1)$ *for some* $k \in \mathcal{Q}$*, and* $(\bar{y}^i, \bar{y}_0^i) = (0, 0)\ \forall i \in \mathcal{Q} \setminus \{k\}$*.*

(ii) if $(\bar{x}, \{\bar{y}^i, \bar{y}_0^i\})_{i\in\mathcal{Q}}$ *is an extreme point of* $\mathscr{P}$*, then* $\bar{y}^k = \bar{x} = x^*$ *and* $\bar{y}_0^k = 1$ *for some* $k \in \mathcal{Q}$, $(\bar{y}^i, \bar{y}_0^i) = (0, 0)\ \forall i \in \mathcal{Q} \setminus \{k\}$*, and* x^* *is an extreme point of* $P_\mathcal{Q}$*.*

This representation requires a number of variables proportional to the number of disjunctions $|\mathcal{Q}|$. Therefore, it is necessary that a tractable number of disjunctions are defined for this to be useful in practice. If, for instance, a disjunctive set is defined for each of all integral feasible point for 0-1 MIP with p binary variables, the number of variables in the extended representation will be linear in 2^p, resulting in a computationally disadvantageous formulation.

Important for our work on non-polyhedral sets in Chapter 3 is the generalization of this result to disjunctive convex sets by Ceria and Soares [15]. Their work gives a starting point of using perspective function to obtain strong formulation for MINLPs, and is used as a basis for the reformulation technique introduced in Section 1.5.7.

The perspective of a closed convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\tilde{f} : \mathbb{R}^{n+1} \to \mathbb{R} \cup \{+\infty\}$ is the function

$$\tilde{f}(\lambda, x) = \begin{cases} \lambda f(x/\lambda) & \text{if } \lambda > 0 \\ +\infty & \text{otherwise} \end{cases}$$

It is known that the perspective function of a closed convex function is convex [35].

The main result of Ceria and Soares is stated in the following theorem.

Theorem 1.6. *[15] For* $t \in T$*, let* $G^t : \mathbb{R}^n \to \mathbb{R}^{m_t}$ *be a vector-valued function with the property that the corresponding sets*

$$K^t = \{x \in \mathbb{R}^t \mid G^t(x) \leqslant 0\}$$

are nonempty, closed and convex. Let $K = \cup_{t \in T} K^t$ *be bounded below or above. Then* $x \in \text{conv}(K)$ *if and only if the following system is feasible:*

$$x = \sum_{t \in T} x^t;\ \sum_{t \in T} \lambda^t = 1;\ \text{cl}(\tilde{G}^t)(\lambda_t, x^t) \leqslant 0,\ \lambda_t \geqslant 0\ \forall t \in T,$$

where $\text{cl}(\tilde{G})(\lambda, x)$ *denotes the closure of the perspective function of* G.

The same work provides the characterization of the nontrivial constraint:

$$\text{cl}(\tilde{G}^t)(\lambda_t, x^t) \leqslant 0 \ \Leftrightarrow \begin{cases} \lambda_t G^t(x^t/\lambda_t) \leqslant 0 & \text{if } \lambda > 0, \\ \text{rec}(G^t)(x^t) \leqslant 0 & \text{if } \lambda = 0. \end{cases}$$

Theorem 1.6 provides an extended formulation for the convex hull of the union of convex sets. Similar to the linear case, this result is not always readily applicable to solving discrete optimization problems as it may require a large number of additional variables in the problem. Furthermore, the nonlinear functions $\text{cl}(\hat{G}^t)(\lambda_t, x^t)$ can present additional numerical difficulties.

In the later chapters, we present continuous formulations for 0-1 MIQCP that approximates the convex hull in a space with slightly higher dimension than the space of original variables, and both of the results from [6] and [15] serve as references for determining how close the approximation is to the convex hull.

1.5.5 Lifting

Lifting of inequalities typically refers to the extension of inequalities that are valid for a lower-dimensional restricted subset to obtain inequalities valid for the higher dimensional original set. The restricted set is generally obtained by fixing a subset of variables, and a valid inequality involving the remaining variables is "lifted" to involve all the original variables. The conventional form of lifting is linear; a valid linear inequality is extended by adding linear terms involving the variables previously fixed. In case of MIP, linear lifting provides strong formulations for many classes of problems since all the inequalities required for an ideal LP formulation (the convex hull of the feasible region) are linear as we know from Theorem 1.1. The technique of lifting plays a major role in our effort to generate cutting planes for (0-1 MIQCP) in chapters 2 and 3.

We give a short description of the lifting procedure in the case of linear inequalities, a large part of which is taken from [31] and [3]. Define a mixed 0-1 knapsack set

$$K = \{(x, y) \in \mathbb{B}^n \times \mathbb{R}^m_+ \mid a^T x + g^T y \leqslant b\}$$

where $a \in \mathbb{R}^n$, $g \in \mathbb{R}^m, b \in \mathbb{R}$. For a general vector $v \in \mathbb{R}^n$, let v_I denote the subvector of v consisting of elements v_j, $\forall j \in I$ for some index set $I \subset [n]$. For some partition (L, U, R) of $[n]$, define a nonempty subset of K as

$$K(L, U, R) = \{(x, y) \in \mathbb{B}^{|R|} \times \mathbb{R}^m_+ \mid a_R^T x_R + g^T y \leqslant d\},$$

where $d = b - |U|$. Note that $K(L, U, R)$ is obtained from K by fixing $x_j = 0 \ \forall j \in L$ and $x_j = u_j \ \forall j \in U$.

Given a valid inequality for $K(L, U, R)$,

$$\pi_R^T x_R + \sigma^T y \leqslant \pi_0, \tag{1.16}$$

we aim to construct a valid inequality for K of the form

$$\pi_R^T x_R + \pi_L^T x_L + \pi_U^T (1 - x_U) + \sigma^T y \leqslant \pi_0. \tag{1.17}$$

The *lifting function* for the inequality (1.16) $\Phi : \mathbb{R} \mapsto \mathbb{R} \cup \{\infty\}$ is defined as

$$\Phi(\alpha) = \pi_0 - \max\{\pi_R^T x_R + \sigma^T y \mid a_R^T x_R + g^T y \leqslant d - \alpha,\ x_R \in \mathbb{B}^{|R|},\ y \in \mathbb{R}^m_+\}.$$

The involved optimization problem, called the *lifting problem*, is a useful tool in proving the validity of the extended inequality as it determines the maximum value of the slack in inequality (1.16) if α is subtracted from the right-hand-side of the knapsack constraint from K. Specifically, it is easy to see that (1.17) is valid for K if and only if $\pi_L^T x_L + \pi_U^T (1 - x_T) \leqslant \Phi(a_L^T x_L + a_U^T (x_U - 1)) \ \forall (x, y) \in K$. We use this same idea numerous times in chapters 2 and 3.

The lifting can be executed sequentially for one variable at a time or simultaneously for all variables $x_j \ \forall j \in L \cup U$. For sequential lifting, we start by choosing a variable x_k. If $k \in L$, the lifted inequality will be of the form

$$\pi_R^T x_R + \pi_k x_k + \sigma^T y \leqslant \pi_0$$

which is valid for $K(L \backslash \{k\}, U, R \cup \{k\})$ if and only if $\pi_k x_k \leqslant \Phi(a_k x_k) \Leftrightarrow \pi_k \leqslant \Phi(a_k)$. If $k \in U$, the lifted inequality will be of the form

$$\pi_R^T x_R + \pi_k (1 - x_k) + \sigma^T y \leqslant \pi_0$$

which is valid for $K(L, U \backslash \{k\}, R \cup \{k\})$ if and only if $\pi_k (1 - x_k) \leqslant \Phi(a_k (x_k - 1)) \Leftrightarrow \pi_k \leqslant \Phi(-a_k)$.

Note that the lifting function will change after each step of sequential lifting, and different lifting problems need to be solved while lifting each variables in $L \cup U$. Therefore, the coefficient of a lifted variable x_l depends on its order in the sequence if lifted during sequential lifting. However, Wolsey [56] showed that if the lifting function Φ is superadditive, then it is possible to obtain coefficient for all variables using the same lifting function regardless of the sequence of lifting.

For (1.17), if Φ is superadditive, then

$$\pi_R^T x_R + \sum_{i \in L} \Phi(a_i) x_L + \sum_{i \in U} \Phi(-a_i)(1 - x_U) + \sigma^T y \leqslant \pi_0. \tag{1.18}$$

is valid. Note that even if Φ is not superadditive, a superadditive lower bound $\phi : \mathbb{R} \mapsto \mathbb{R} \cup \{\infty\}$ of Φ can be used in place of Φ for sequence independent lifting.

This concept has been effectively used in strengthening formulation of a variety of sets by deriving strong valid inequalities, especially for the ones involving binary variables. Examples of application of this technique can be found in many studies including the work of Gu, Nemhauser and Savelsbergh [30].

1.5.6 Cutting Planes, Theory and Algorithms for MINLPs

Our work will derive inequalities for specific structures found in MIQCPs. However, many other researchers have adapted general inequalities from MILP to be applicable for MINLPs. Cezik and Iyengar [16] demonstrated that Gomory mixed integer cuts can be used when the nonlinear constraint set can be represented with conic inequalities. Atamtürk and Narayanan generalize the theory of lifting to integer programming with conic constraints (Conic Integer Programming) [4] and extend mixed integer rounding procedure to SOCPs [5]. The lift and project procedure for 0-1 MINLPs was derived by Stubbs and Mehrotra [52] based on earlier work of Balas, Ceria and Cornuéjols [8]. Kilinç, Linderoth and Luedtke [37] introduce a procedure for generating disjunctive inequalities for convex MINLPs, and report that the use of this inequality within a branch-and-cut solver reduces solution time for many instances. Kilinç, Linderoth, Luedtke and Miller [38] make use of the information from the strong branching technique for MINLP in deriving valid inequalities. The authors demonstrate that this approach can significantly improve existing algorithms for MINLP. Modaresi, Kilinç, and Vielma [46] generalize the split and intersection cuts and provide simple formulas for split cuts for convex sets with single quadratic constraint and for intersection cuts for various convex quadratic sets. Dong and Linderoth [22] investigate valid inequalities for a set involving xx^T for $x \in [0,1]^n$ and indicator binary variables, which appears in a relaxation of MIQP. The authors derive different classes of inequalities by lifting the inequalities valid for a related set without binary variable, and show that one of the classes of inequalities describe the convex hull for relevant convex 0-1 quadratic set and also generalizes the perspective constraints.

Extension of theory on polyhedra from MILP includes the work of Dadush, Dey, and Vielma [19], [20] that generalize the earlier studies on rational polytope and show that the split closure of a strictly convex body and the Chvátal-Gomory closure of a compact convex set are both rational polytopes.

There also have been studies on the solution process of MISOCP. Vielma, Ahmed, and Nemhauser [54] present a linear programming based branch and bound algorithm that uses a polyhedral relaxation of conic quadratic constraints. Drewes and Pokutta [23] suggest a cutting plane framework based on the generalized Benders cut for 0-1 SOCPs where continuous and binary variables are solely coupled in the conic constraints. Drewes and Ulbrich [24] introduce a subgradient-based outerapproximation approach to solve MISOCPs and also generalizes the convergence result for differentiable MINLP to subdifferentiable constraint functions.

1.5.7 Perspective Reformulation

A useful technique to get a strong relaxation for some classes of convex MINLPs with indicator variables is the perspective reformulation introduced by Günluk and Linderoth[32]. In particular, they study the convex hull description of sets where some of the variables are restricted to assume fixed values or to be contained in a convex set depending on whether the indicator variables are turned on or off. The following lemma provides an important ingredient of our subsequent work presented in Chapter 2 and Chapter 3.

Lemma 1.1. *[32] Define the sets*

$$\begin{aligned} W^0(\bar{x}) &= \{(x,z) \in \mathbb{R}^{n+1} \mid x = \bar{x}, z = 0\} \\ W^1(\bar{x}) &= \{(x,z) \in \mathbb{R}^{n+1} \mid f_i(x) \leqslant 0 \ \forall i \in [t], u \geqslant x - \bar{x} \geqslant l, z = 1\} \\ W^- &= \{(x,z) \in \mathbb{R}^{n+1} \mid z f_i(x/z) \leqslant 0 \ \forall i \in I, uz \geqslant x \geqslant lz, 1 \geqslant z \geqslant 0\} \end{aligned}$$

where $\bar{x} \in \mathbb{R}^n$ *and* $u, l \in \mathbb{R}^n_+$. *If* W^1 *is convex, then* $\operatorname{conv}(W^0 \cup W^1) = W^- \cup W^0 = \operatorname{cl}(W^-)$.

The name *perspective reformulation* comes from the fact that the formulation for W^- uses the perspective function. Lemma 1.1 can be viewed as an adaptation of Theorem 1.6 to special indicator-induced convex sets.

An earlier work that is closely related to perspective reformulation comes from Frangioni and Gentile [27]. They consider MINLPs with one indicator variable of the form

$$\begin{aligned} \min \quad & f(x) + cz \\ \text{s.t.} \quad & Ax \leqslant bz, \\ & x \in \mathbb{R}^n, z \in \mathbb{B} \end{aligned}$$

where the polyhedron $\{x \mid Ax \leqslant b\}$ is bounded, f is a closed convex function that is finite on $\{x \mid Ax \leqslant b\}$, and $f(0) = 0$.

Then Frangioni and Gentile show that for an equivalent MINLP

$$\begin{aligned} \min \quad & \nu \\ \text{s.t.} \quad & \nu \geqslant f(x) + cz, \\ & Ax \leqslant bz, \\ & x \in \mathbb{R}^n, z \in \mathbb{B}, \end{aligned}$$

the *perspective cut*

$$\nu \geqslant f(\bar{x}) + c + s^T(x - \bar{x}) + (c + f(\bar{x}) - s^T\bar{x})(z - 1) \tag{1.19}$$

for $s \in \partial f(\bar{x})$ is a valid inequality. The inequality (1.19) is derived based on a first-order analysis of the convex envelope of the original objective function $f(x) + cz$. This inequality is valid for any $\bar{x}$ such that $A\bar{x} \leqslant b$ and defines maximal face for the epigraph of convex envelope of $f + cz$ of dimension at least one. Günlük and Linderoth [32] demonstrate that (1.19) can be derived as a linear outerapproximation of the convex inequality in the definition of the perspective reformulation.

1.5.8 Strengthened Reformulations

Both perspective cuts and the perspective reformulation are developed for problems with semicontinuous variables regulated by indicator variable(s) where the semicontinuous variables appear *separably* in convex functions. Therefore it cannot be directly applied for the problems of different structure in terms of interaction between continuous and binary variables. Frangioni and Gentile in their successive work [28] suggest how to utilize perspective cuts for problems where the semicontinuous variables appear together in non-separable convex quadratic function:

An example is the problem

$$\begin{aligned} \min \quad & x^T Q x + q^T x + c^T z, \\ \text{s.t.} \quad & Ax + Hy \geqslant b, \\ & l_i z_i \leqslant x_i \leqslant u_i z_i \ \forall i \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n \end{aligned}$$

We demonstrate how to apply their approach for the problem of our interest:

$$\begin{aligned} \min \quad & \nu \\ \text{s.t.} \quad & \nu \geqslant x^T Q x + c^T x, \\ & Ax + Hz \leqslant f, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n. \end{aligned} \tag{0-1 MIQCP-l}$$

Note that for simplicity of formulation we consider the case where all variables x are semi-continuous, i.e., $p = n$, and leave out quadratic constraints other than the original objective. The relaxation approach explained below can be separately applied to each of the omitted inequalities.

If Q is a diagonal matrix, the quadratic function can be written as $\sum_{i \in [n]} Q_{ii}(x_i^2 + c_i x_i)$ and the perspective cuts or perspective reformulation can be directly applied to generate a strong formulation. The perspective cut for this case at $(\bar{x}, \bar{z})$ for each $i \in [n]$ can be written as

$$\nu_i \geqslant (2Q_{ii}\bar{x}_i + c_i)x_i - Q_{ii}\bar{x}_i^2 z_i \tag{1.20}$$

In fact, an extended formulation of the convex hull of the substructure of this problem is known.

Lemma 1.2. *[53] Define the set* X *as*

$$X = \{(\nu, x, z) \in \mathbb{R}^{1+n} \times \mathbb{B}^n \mid \nu \geqslant \sum_{j \in [n]} q_j x_j^2,\ l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [n]\},$$

and an extended formulation is $\bar{X}$

$$\bar{X} = \{(\nu, x, t, z) \in \mathbb{R}^{1+2n} \times \mathbb{B}^n \mid \nu \geqslant \sum_{j \in [n]} q_j t_j,\ t_j \geqslant x_j^2,\ l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [n]\}. \tag{1.21}$$

The convex hull of $\bar{X}$ is

$$\operatorname{conv}(\bar{X}) = \{(v, x, t, z) \in \mathbb{R}^{1+2n} \times [0,1]^n \mid v \geqslant \sum_{j \in [n]} q_j t_j,$$
$$t_j z_j \geqslant x_j^2,\ l_j z_j \leqslant x_j \leqslant u_j z_j\ \forall j \in [n]\}. \tag{1.22}$$

The inequalities

$$t_j z_j \geqslant x_j^2\ \forall j \in [n]$$

come from the (separable) application of the perspective reformulation. For a general positive semidefinite matrix Q, we can create an equivalent formulation with the matrix split into a diagonal matrix and a remainder matrix. Specifically, with any nonnegative diagonal matrix D such that $Q - D \succeq 0$, (0-1 MIQCP-*l*) can be equivalently formulated as

$$\begin{aligned} \min\quad & w \\ \text{s.t.}\quad & w \geqslant x^T(Q - D)x + x^T D x + c^T x, \\ & Ax + Hz \leqslant f, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j\ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n. \end{aligned}$$

For a substructure of this problem

$$\{(v, x, t, z) \in \mathbb{R}^{1+2n} \times \mathbb{B}^n \mid v \geqslant x^T D x,\ l_j z_j \leqslant x_j \leqslant u_j z_j\ \forall j \in [n]\},$$

both the convexification of the extended formulation of Lemma 1.2, or the inequalities (1.20) can be directly applied to construct a stronger formulation.

To satisfy $Q - D \succeq 0$ and thus maintain the convexity of the problem, Frangioni and Gentile suggest two different choices of D. The first one is to set $D = \lambda_{min} I$ where λ_{min} is the minimum eigenvalue of Q. The second approach aims to get 'larger' D so that more improvement on the (root) lower bound can be achieved by adding perspective cuts. To maximize the total extracted amount while keeping the remainder matrix positive semidefinite, one can solve the following Semidefinite Program (SDP):

$$\max \Big\{ \sum_{i=1}^{n} d_i \mid Q - \sum_{i=1}^{n} d_i (e_i e_i^T) \succeq 0,\ d \geqslant 0 \Big\}. \tag{1.23}$$

Computations of Frangioni and Gentile [28] demonstrates convincingly that even though the perspective cuts were applied only to a subset with simple structure, the reformulation, relaxation and convexification techniques help improve the root node bound for the unit commitment problem and minimum variance portfolio problem.

We suggest a generalization of this approach to handle nonseparability. Consider a more general decomposition of $Q = R + B$ such that $B, R \succeq 0$, which allows for the reformulation of (0-1 MIQCP-*l*) to

$$\begin{aligned} \min \quad & \nu \\ \text{s.t.} \quad & \nu \geqslant x^T(Q - B)x + x^T Bx + c^T x, \\ & Ax + Hz \leqslant f, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n. \end{aligned} \tag{0-1 MIQCP-r}$$

We then propose to *diagonalize* the extracted matrix $x^T Bx$ using the Cholesky factorization $B = LL^T$. By introducing a set of new variables y and letting $y = L^T x$, we get $x^T Bx = x^T LL^T x = y^T y$ and (0-1 MIQCP-r) is equivalently formulated as

$$\begin{aligned} \min \quad & \nu \\ \text{s.t.} \quad & \nu \geqslant x^T(Q - B)x + \sum_{j \in [n]} t_j + c^T x, \\ & t_j \geqslant y_j^2 \ \forall j \in [n], \\ & y = L^T x, \\ & Ax + Hz \leqslant f, \\ & l_j z_j \leqslant x_j \leqslant u_j z_j \ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n. \end{aligned}$$

As in the diagonal decomposition, we create a relaxation of the feasible region of the reformulated problem by taking only the inequalities involving variables y and the lower and upper bound constraints on variables x. Substituting the variables x with $L^{-T}y$, we obtain a lower dimensional set

$$S_0 = \{(y, t, z) \in \mathbb{R}^{2n} \times \mathbb{B}^n \mid t_j \geqslant y_j^2,\ l_j z_j \leqslant [L^{-T}y]_j \leqslant u_j z_j \ \forall j \in [n]\} \tag{1.24}$$

and aim to find a good approximation to $\text{conv}(S_0)$.

The set S_0 is very similar to $\bar{X}$ defined by (1.21) as all of the quadratic constraints are represented by separable functions. Yet the perspective cuts (1.20) or Lemma 1.2 cannot be immediately applied to approximate $\text{conv}(S)$ because the indicator variables z now have a more complicated interaction with the continuous variables y through constraint $l_j z_j \leqslant [L^{-T}y]_j \leqslant u_j z_j \; \forall j \in [n]$. Even relaxing the quadratic constraints $t_j \geqslant y_j^2$, the set S_0 is equivalent to the intersection of knapsack constraints for which no good characterization is known [43]. Therefore we concentrate our analysis on a low dimensional case of S_0 denoted by S:

$$\begin{aligned} S = \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{B}^2 \mid t_1 \geqslant y_1^2,\ t_2 \geqslant y_2^2,& \\ 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant z_1,& \\ 0 \leqslant \qquad\quad a_{22}y_2 \leqslant z_2\ \}& \end{aligned} \tag{1.25}$$

where the matrix

$$L^{-T} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Note that we use $l_j = 0, u_j = 1 \; \forall j \in [n]$ without loss of generality as we can scale the variables for any $-\infty < l_j < u_j < +\infty$.

Note that although S is a restricted set, what we learn about $\text{conv}(S)$ can still be utilized in constructing a good formulation for (0-1 MIQCP-*l*) by taking the decomposition $Q = R + B$ for a 2×2 block diagonal matrix $B \succeq 0$. Then the Cholesky factor L and L^{-T} are also block diagonal which makes the constraint $l_j z_j \leqslant [L^{-T}y]_j \leqslant u_j z_j \; \forall j \in [n]$ partially separable in a sense that the i^{th} block L_i interacts with only variables $y_{2i},\ y_{2i+1}$. Therefore, any strong formulation for S can be applied separately for substructures involving each 2×2 block. We discuss this approach more fully in Chapter 2. The diagonal decomposition is a special case of block diagonal decomposition. Our goal is to obtain a formulation for S that results in improving the bounds in the solution process for problems defined as (0-1 MIQCP).

1.6 Contributions

In Chapter 1, we explained our motivation to investigate mixed-integer sets with indicator variables that arise from 0-1 MIQCP and the transmission switching problem. These classes of optimization problems appear in numerous applications from various fields such as the portfolio problem in financial engineering, subset selection for regression in statistics and electricity network design problems. The goal of this thesis is to develop strong formulations of substructures appearing in these problems that can improve the solution process by providing better bounds in enumeration algorithms. Previous research and theories related to the subject were briefly reviewed. The review includes the existing techniques that were directly applied in our attempt to construct the relaxation for 0-1 MIQCP and different approach suggested by other researchers. What makes our approach unique is the reformulation process using the Cholesky factorization of the Hessian. This results in a feasible set defined by separable quadratic constraints involving new variables and more complicated linear constraints involving indicator variables. The methods we develop in Chapter 2 and Chapter 3 are applicable to this reformulated feasible set.

For 0-1 MIQCP, we focus on a low-dimensional case with 2 indicator variables and start the investigation with a polyhedral outer-approximation of the problem in Chapter 2. Five different types of valid inequalities are derived for this MILP mainly using techniques such as lifting, variable substitution, constraint aggregation, and the perspective reformulation. These inequalities are shown to be facet-defining for the convex hull of the mixed integer set under some mild conditions. We conduct computational experiments to examine how well the resulting formulation approximates the convex hull and to infer which facets are more important in describing the convex hull. By solving optimization problems with random linear objective function over the set defined by our formulation, it is revealed that we do not have the complete description of the convex hull, but the formulation provides exact solutions for 90 % of the instances. We also demonstrate through computation on the minimum variance portfolio problem how to use the results in the low dimensional case to general problems of larger size by a 2×2 block-diagonal matrix extraction.

In Chapter 3, based on what was developed in the previous chapter, four types of quadratic valid inequalities are derived for the low dimensional case. The techniques applied to derive them are very similar to what were used in Chapter 2. All of the inequalities derived are shown to be second order cone representable. Since there does not exist a theoretical tool to determine the necessity of an inequality in describing a non-polyhedral set whereas facet defining quality is well-defined for polyhedra, we examine the importance of these inequalities through computational experiments. Complete characterization of the set of extreme points of the natural relaxation and the set of extreme points of the convex hull are given. We also specify which extreme points of the natural relaxation are violated by each of the inequalities we introduce. Computational experiments on the same instances of the minimum variance portfolio problem from Chapter 2 are performed.

In Chapter 4, we investigate the substructure of the DC transmission switching problem. As for 0-1 MIQCP, we define a relaxation by including a subset of constraints for the original

DC transmission switching problem. This subset of constraints is selected by finding a specific subgraph in the form of a directed cycle of the underlying transmission line network. We exploit the attribute of the relaxation to derived valid linear inequalities for the relaxation and prove that they are facet-defining for the relaxation. It was also shown by our collaborators that these are the only non-trivial inequalities necessary to describe the convex hull of a lower-dimensional set obtained by projecting out the voltage angle variable from our relaxation. Computational experiments are performed to demonstrate the usefulness of the inequalities in a branch-and-cut scheme on IEEE instances.

The following list summarizes the main result and contributions of this thesis from each chapter.

In Chapter 1:

- We motivate the study of mixed-integer sets with specific structure involving indicator variable and provide a review of related theoretical results and tools that are helpful in the investigation.
- For 0-1 MIQCP, we introduce a unique approach to reformulate its relaxation through Choesky factorization to obtain separability of quadratic constraints.
- For DC transmission switching problem, we introduce our approach to construct a relaxation by extracting a subgraph from the transmission line network.

In Chapter 2:

- For a polyhedral outer-approximation $P(A, B)$ of a low-dimensional (transformed) relaxation of 0-1 MIQCP, we derive 5 classes of valid inequalities and prove that they are facet-defining under some mild conditions.
- We measure the importance of each facet via shooting experiment and present computational evidence that these 5 classes of inequalities do not provide complete characterization of $\text{conv}(P(A, B))$.
- We describe a framework to utilize the result obtained for a low-dimensional relaxation in a solution process for a general problem of larger dimension by reformulation through matrix decomposition. This was also demonstrated in computational experiments on the minimum variance portfolio optimization.

In Chapter 3:

- We define a low-dimensional (transformed) relaxation of 0-1 MIQCP with quadratic constraint, S, and characterize the set of extreme points of its simple continuous relaxation $\mathcal{R}(S)$ and $\text{conv}(S)$.
- We derive 4 classes of nonlinear inequalities valid for S and prove that they are all second-order cone representable.
- Computational evidence is provided to show that these 4 inequalities do not provide complete characterization of $\text{conv}(S)$.
- We measure the impact of the inequalities by i) analyzing which extreme points of $\mathcal{R}(S)$ are cut off by them and at which extreme points of $\text{conv}(S)$ they are tight, and ii) conducting shooting experiments.
- Computations were performed on the same instances of the minimum variance portfolio optimization from Chapter 2 to demonstrate the ability of the valid inequalities to improve bounds.

In Chapter 4:

- We construct a useful relaxation $\mathcal{C}$ of the DC transmission switching problem based on the physical intuition of Kirchoff's voltage law. the relaxation is based on a directed cycle.
- We derive a class of valid inequalities for the relaxation $\mathcal{C}$ and prove that they are also facet-defining for $\mathcal{C}$.
- We investigate the separation problem and describe a closed form solution to its subproblem formulated as a knapsack problem.
- Computations were performed on IEEE instances of the DC transmission switching problem to demonstrate the usefulness of the inequalities in the solution process.

Chapter 2

Polyhedral Outer-approximation of S

2.1 Introduction

With the reformulation strategy explained in Section 1.5.8 in mind, the focus of this chapter is the mixed integer set

$$S_0 = \{(y, t, z) \in \mathbb{R}^{2n} \times \mathbb{B}^n \mid t_j \geqslant y_j^2,\ l_j z_j \leqslant [L^{-T} y]_j \leqslant u_j z_j\ \forall j \in [n]\}. \tag{2.1}$$

As described in Section 1.5.8, the set S_0 can be used in a relaxation of the feasible region of the following reformulation of 0-1 MIQCP:

$$\begin{aligned}
\min\quad & v \\
\text{s.t.}\quad & v \geqslant x^T(Q - B)x + \sum_{i \in [n]} t_i + c^T x, \\
& t_i \geqslant y_i^2\ \forall i \in [n], \\
& y = L^T x, \\
& x^T Q_i x + c_i^T x \leqslant d_i\ \forall i \in [k], \\
& Ax + Hz \leqslant f, \\
& l_j z_j \leqslant x_j \leqslant u_j z_j\ \forall j \in [p], \\
& x \in \mathbb{R}^n, z \in \mathbb{B}^p.
\end{aligned}$$

We begin with an analysis of the lower-dimensional set defined as

$$\begin{aligned} S = \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{B}^2 \mid t_1 \geqslant y_1^2, t_2 \geqslant y_2^2, \\ 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant z_1, \\ 0 \leqslant \qquad a_{22}y_2 \leqslant z_2 \}. \end{aligned} \tag{2.2}$$

Note that the lower and upper bounds on the variables y_1, y_2 can be deduced from the linear inequalities. Specifically, $y_1 \in [l_1, u_1], y_2 \in [l_2, u_2]\ \forall y_1, y_2 \in S_0$ where

$$\begin{aligned} l_1 :=& \min\{0, -\frac{a_{12}}{a_{11}}\frac{1}{a_{22}}\}, \\ u_2 :=& \max\{\frac{1}{a_{11}}, \frac{1}{a_{11}} - \frac{a_{12}}{a_{11}}\frac{1}{a_{22}}\}, \\ l_2 :=& 0, \\ u_2 :=& \frac{1}{a_{22}}. \end{aligned}$$

For ease of notation, we ignore the dependence of the set S on the parameters a_{11}, a_{12}, and a_{22} when we refer to the general structure. However, when we explicitly write the dependence, e.g., to address specific instances, we refer to the set as $S(a_{11}, a_{12}, a_{22})$.

In this chapter, we investigate a polyhedral outer-approximation of S. The relaxation can be constructed using tangential lines of the quadratic constraints and made to be arbitrarily close to S by increasing the number of these lines. We denote the polyhedral set by

$$P(A,B) = \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{B}^2 \mid 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant z_1, \tag{2.3a}$$

$$0 \leqslant \qquad a_{22}y_2 \leqslant z_2, \tag{2.3b}$$

$$t_1 \geqslant 2\alpha y_1 - \alpha^2\ \forall \alpha \in A, \tag{2.3c}$$

$$t_2 \geqslant 2\beta y_2 - \beta^2\ \forall \beta \in B\ \} \tag{2.3d}$$

where $A = \{\alpha_1, \alpha_2, \cdots, \alpha_m\} \subset [l_1, u_1]$ and $B = \{\beta_1, \beta_2, \cdots, \beta_p\} \subset [l_2, u_2]$ are the index sets of the points at which we take linearizations for $t_1 \geqslant y_1^2$ and $t_2 \geqslant y_2^2$, respectively. Without loss of generality, assume that $\alpha_i < \alpha_{i+1}\ \forall \alpha_i, \alpha_{i+1} \in A$, and $\beta_j < \beta_{j+1}\ \forall \beta_j, \beta_{j+1} \in B$. We define parameters σ and δ for A and B to be the minimum distance between consecutive lineariation points:

$$\sigma := \min_{i=1,\dots,m-1} \{\alpha_{i+1} - \alpha_i\} \tag{2.4}$$

$$\delta := \min_{j=1,\dots,p-1} \{\beta_{j+1} - \beta_j\}. \tag{2.5}$$

These parameters are used specifically in the facet proofs in section Section 2.2.

2.2 Valid Inequalities for $P(A, B)$

Five methods of deriving strong inequalities for $P(A, B)$ are presented in this section. For simplicity of notation, we define a constant $r = \frac{a_{12}}{a_{11}}$.

The first facet-defining inequality is the perspective cut of Frangioni and Gentile [27].

Proposition 2.1. *For each* $\bar{\beta} \in B$*, the inequality*

$$t_2 \geqslant 2\bar{\beta}y_2 - \bar{\beta}^2 z_2. \tag{Ineq L1}$$

is valid for $P(A, B)$*.*

Proof. Since $z_2 = 0$ implies $y_2 = 0$, $t_2 \geqslant 2\bar{\beta}y_2 - \bar{\beta}^2$ for any $\bar{\beta} \in B$ (equation (2.3d) in the description of $P(A, B)$) can be strengthened through the perspective reformulation introduced in Section 1.5.7 to (Ineq L1). □

Proposition 2.2. *The inequality* (Ineq L1) *defines a facet of* $\operatorname{conv}(P(A, B))$*.*

Proof. To prove that (Ineq L1) is a facet of $\operatorname{conv}(P(A, B))$, we present the following 6 affinely independent points that are contained in $P(A, B)$ and satisfy $t_2 = 2\bar{\beta}y_2 - \bar{\beta}^2 z_2$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}\frac{1}{a_{11}}\\0\\(\frac{1}{a_{11}})^2\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-r\bar{\beta}\\\bar{\beta}\\r^2\bar{\beta}^2\\\bar{\beta}^2\\0\\1\end{bmatrix}, \begin{bmatrix}0\\\bar{\beta}-\frac{\delta}{2}\\0\\\bar{\beta}(\bar{\beta}-\delta)\\1\\1\end{bmatrix}.$$

Note that δ defined in (2.5) is a parameter representing the minimum distance between values in B. It is obvious that the first five points are in $P(A, B)$. For the sixth point to be feasible, the condition $\bar{\beta}(\bar{\beta} - \delta) \geqslant 2\beta(\bar{\beta} - \frac{\delta}{2})\ \forall \beta \in B$ must be true. The condition holds since $2\bar{\beta}y_2 - \bar{\beta}^2 \geqslant 2\beta y_2 - \beta^2\ \forall \beta, \bar{\beta} \in B$ if $y \in [\bar{\beta} - \frac{\delta}{2}, \bar{\beta} + \frac{\delta}{2}]$. □

The next three inequalities are derived by fixing and lifting the variable z_1 with the perspective reformulation. The inequalities differ from each other in the substitution made between the variables y_1 and y_2.

Proposition 2.3. *For each* $\alpha_k \in A$*, define*

$$L_1(\alpha_k) = \begin{cases} 0 & \text{if } 0 \leqslant \alpha_k \leqslant -\frac{r}{a_{22}} \text{ or } 0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}},\ r > 0 \\ \frac{2r(\alpha_{p1} - \alpha_k)}{a_{22}} + \alpha_{p1}^2 - \alpha_k^2 & \text{if } -\frac{r}{a_{22}} \leqslant \alpha_k,\ r < 0 \\ -\frac{2\alpha_k}{a_{11}} & \text{if } \frac{1}{a_{11}} - \frac{r}{a_{22}} \leqslant \alpha_k \leqslant 0 \\ -\frac{2\alpha_{p2}}{a_{11}} + \frac{2r(\alpha_{p2} - \alpha_k)}{a_{22}} + \alpha_{p2}^2 - \alpha_k^2 & \text{if } \alpha_k \leqslant \frac{1}{a_{11}} - \frac{r}{a_{22}} \leqslant 0,\ r > 0 \\ -\frac{2r\alpha_k}{a_{22}} - \alpha_k^2 & \text{if } \alpha_k \leqslant 0,\ \frac{1}{a_{11}} - \frac{r}{a_{22}} \geqslant 0 \end{cases} \tag{2.6}$$

where α_{p1} *and* α_{p2} *are parameters that can be uniquely chosen by finding elements in* A *satisfying* $\frac{\alpha_{p1}+\alpha_{p1-1}}{2} \leqslant -\frac{r}{a_{22}} \leqslant \frac{\alpha_{p1}+\alpha_{p1+1}}{2}$ *and* $\frac{\alpha_{p2}+\alpha_{p2-1}}{2} \leqslant \frac{1}{a_{11}} - \frac{r}{a_{22}} \leqslant \frac{\alpha_{p2}+\alpha_{p2+1}}{2}$.

Then the inequality

$$-2r\alpha_k y_2 - t_1 - L_1(\alpha_k) z_1 - \alpha_k^2 z_2 \leqslant 0 \tag{Ineq L2}$$

is valid for $P(A, B)$.

Proof. If $z_1 = 0$, (2.3a) implies $y_1 + r y_2 = 0$ where $r = \frac{a_{12}}{a_{11}}$. Substituting $y_1 = -r y_2$ in the inequality $t_1 \geqslant 2\alpha_k y_1 - \alpha_k^2$ for some $\alpha_k \in A$ gives $t_1 \geqslant -2r\alpha_k y_2 - \alpha_k^2$. Applying the perspective reformulation leads to a strengthened inequality $-2r\alpha_k y_2 - t_1 - \alpha_k^2 z_2 \leqslant 0$ that is valid in the case that $z_1 = 0$.

We can maximally lift this inequality back in z_1 to $-2r\alpha_k y_2 - t_1 - \alpha_k^2 z_2 \leqslant L_1(\alpha_k) z_1$ by solving

$$L_1(\alpha_k) = \max\{-2r\alpha_k y_2 - t_1 - \alpha_k^2 z_2 \mid (y, t, z) \in P(A, B)\}. \tag{2.7}$$

We need only consider the case $z_1 = 1$, and solve the two cases of (2.7) where $z_2 = 0$ and 1. If $z_1 = 1, z_2 = 0$, then (2.3b) implies that $y_2 = 0$, (2.7) reduces to

$$\max\{-t_1 \mid 0 \leqslant y_1 \leqslant 1/a_{11}, t_1 \geqslant 2\alpha y_1 - \alpha^2\ \forall \alpha \in A, t_1 \geqslant 0\}$$

and the optimal value in this case 0.

When $z_1 = z_2 = 1$, (2.7) is written as follows:

$$
\begin{aligned}
L_1(\alpha_k) = \max\ & -2r\alpha_k y_2 - t_1 - \alpha_k^2 && \text{(Lift1-P)}\\
\text{s.t. } & a_{11}y_1 + a_{12}y_2 \geqslant 0, && (\lambda_1)\\
& a_{11}y_1 + a_{12}y_2 \leqslant 1, && (\lambda_2)\\
& a_{22}y_2 \geqslant 0, && (\lambda_3)\\
& a_{22}y_2 \leqslant 1, && (\lambda_4)\\
& 2\alpha y_1 - t_1 \leqslant \alpha^2\ \forall \alpha \in A, && (\mu_i)\\
& 2\beta y_2 - t_2 \leqslant \beta^2\ \forall \beta \in B, && (\gamma_j)\\
& t_1, t_2 \geqslant 0.
\end{aligned}
$$

The variables in parentheses (λ, μ, γ) are the dual variables associated with each constraint. We characterize optimal solutions to (Lift1-P) by examining its dual linear program involving these variables:

$$
\begin{aligned}
\min\ & \lambda_2 + \lambda_4 + \sum_{i=1}^{m} \alpha_i^2 \mu_i + \sum_{j=1}^{n} \beta_j \gamma_j - \alpha_k^2 && \text{(Lift1-D)}\\
\text{s.t. } & -a_{11}\lambda_1 + a_{11}\lambda_2 + \sum_{i=1}^{m} 2\alpha_i \mu_i = 0,\\
& -a_{12}\lambda_1 + a_{12}\lambda_2 - a_{22}\lambda_3 + a_{22}\lambda_4 + \sum_{j=1}^{n} 2\beta_j \gamma_j = -2r\alpha_k,\\
& -\sum_{i=1}^{m} \mu_i \geqslant -1,\\
& -\sum_{j=1}^{n} \gamma_j \geqslant 0,\\
& \lambda, \mu, \gamma \geqslant 0.
\end{aligned}
$$

The optimal solution to (Lift1-P) can be characterized by describing feasible points $(\hat{y}_1, \hat{y}_2, \hat{t}_1, \hat{t}_2)$ to the primal (Lift1-P) and $(\hat{\lambda}, \hat{\mu}, \hat{\gamma})$ to dual (Lift1-D), showing that these points have the same objective value. Different points are characterized for different cases defined depending on the instances of $P(A, B)$ and $\alpha_k \in A$. As $A \subset [l_1, u_1]$, we know that if $a_{12} < 0$, then $\alpha_k \in [0, 1/a_{11} - r/a_{22}]$, and if $a_{12} > 0$, then $\alpha_k \in [-r/a_{22}, 1/a_{11}]$. We partition these intervals and consider two types of instances where $a_{12} > 0$, ($0 \leqslant a_{12} \leqslant a_{22}$ and $a_{12} > a_{22}$) to define the following 6 different cases:

- Case 1: $a_{12} < 0,\ \alpha_k \in [0, -r/a_{22}]$
- Case 2: $a_{12} < 0,\ \alpha_k \in [-r/a_{22}, 1/a_{11} - r/a_{22}]$
- Case 3: $0 \leqslant a_{12} \leqslant a_{22},\ \alpha_k \in [-r/a_{22}, 0]$
- Case 4: $0 \leqslant a_{12} \leqslant a_{22}$ or $a_{12} > a_{22},\ \alpha_k \in [0, 1/a_{11}]$
- Case 5: $a_{12} > a_{22},\ \alpha_k \in [-r/a_{22}, 1/a_{11} - r/a_{22}]$
- Case 6: $a_{12} > a_{22},\ \alpha_k \in [1/a_{11} - r/a_{22}, 0]$

The feasible points for (Lift1-P) and (Lift1-D) and the common (optimal) objective value are described for each of these cases. Note that if the objective value for these points is less than 0, the optimal objective value for (2.7) is 0 which is obtained when $z_2 = 0$.

- Case 1:
  $(\alpha_k, -\frac{\alpha_k}{r}, \alpha_k^2, (\frac{\alpha_k}{r})^2)$ is feasible for (Lift1-P) and $((\frac{2\alpha_k}{a_{11}}, 0, 0, 0), e_k, \mathbf{0})$ is feasible for (Lift1-D) with optimal objective value 0.

- Case 2:
  $(-\frac{r}{a_{22}}, \frac{1}{a_{22}}, -\frac{2r\alpha_{p1}}{a_{22}} - \alpha_{p1}^2, (\frac{1}{a_{22}})^2))$ is feasible for (Lift1-P) and $((\frac{2\alpha_{p1}}{a_{11}}, 0, 0, \frac{2r(\alpha_{p1}-\alpha_k)}{a_{22}}), e_p, \mathbf{0})$ is feasible for (Lift1-D) with optimal objective value $\frac{2r(\alpha_{p1}-\alpha_k)}{a_{22}} + \alpha_{p1}^2 - \alpha_k^2$. Note that if $\alpha_{p1} = \alpha_k$, the resulting inequality coincides with Case 1.

- Case 3:
  $(0, \frac{1}{a_{22}}, 0, (\frac{1}{a_{22}})^2)$ is feasible for (Lift1-P) and $((0, 0, 0, -\frac{2r\alpha_k}{a_{22}}), \mathbf{0}, \mathbf{0})$is feasible for (Lift1-D) with optimal objective value $-\frac{2r\alpha_k}{a_{22}} - \alpha_k^2$. Note that this value is nonnegative since

  $$-\frac{2r\alpha_k}{a_{22}} - \alpha_k^2 = -\alpha_k(\alpha_k + \frac{2r}{a_{22}}) = -\alpha_k(\alpha_k + \frac{r}{a_{22}} + \frac{r}{a_{22}})$$

  where $\alpha_k \leqslant 0$ and $r > 0$ by assumption and $\alpha_k + \frac{r}{a_{22}} \geqslant 0$ since $\alpha_k \in A \subset [-\frac{r}{a_{22}}, \frac{1}{a_{11}}]$.

- Case 4:
  $(0, 0, 0, 0)$ is feasible for (Lift1-P) and $((0, 0, \frac{2r\alpha_k}{a_{22}}, 0), \mathbf{0}, \mathbf{0})$is feasible for (Lift1-D) with optimal objective value $-\alpha_k^2$. As this is less than 0, the optimal objective value to the original lifting problem involving z_2 is 0 as in Case 1.

- Case 5:
$(\frac{1}{a_{11}} - \frac{r}{a_{22}}, \frac{1}{a_{22}}, 2\alpha_{p2}(\frac{1}{a_{11}} - \frac{r}{a_{22}}) - \alpha_{p2}^2, (\frac{1}{a_{22}})^2)$ is feasible for (Lift1-P) and $((0, -\frac{2\alpha_{p2}}{a_{22}}, 0, \frac{2r(\alpha_{p2}-\alpha_k)}{a_{22}}), e_p, \mathbf{0})$ is feasible for (Lift1-D) with optimal objective value $-\frac{2\alpha_{p2}}{a_{11}} + \frac{2r(\alpha_{p2}-\alpha_k)}{a_{22}} + \alpha_{p2}^2 - \alpha_k^2$. If $\alpha_{p2} = \alpha_k$, the resulting inequality coincides with Case 6.

- Case 6:
$(\alpha_k, \frac{1}{a_{12}} - \frac{\alpha_k}{r}, \alpha_k^2, (\frac{1}{a_{12}} - \frac{\alpha_k}{r})^2)$ is feasible for (Lift1-P) and $((0, -\frac{2\alpha_k}{a_{11}}, 0, 0), e_k, \mathbf{0})$is feasible for (Lift1-D) with optimal objective value $-\frac{2\alpha_k}{a_{11}}$.

We have characterized the lifting coefficient $L_1(\alpha_k)$ as (2.6) by providing a closed form solution to (2.7) in all cases, which establishes the validity of (Ineq L2) for $P(A, B)$. □

Proposition 2.4. *The inequality* (Ineq L2) *defines a facet of* $\text{conv}(P(A, B))$.

Proof. To prove the claim, we give 6 affinely independent points in $P(A, B)$ that satisfy

$$-2r\alpha_k y_2 - t_1 - L_1(\alpha_k)z_1 - \alpha_k^2 = 0$$

for each of the 6 cases defined in the proof of Proposition 2.3. Note that σ is a parameter representing the minimum distance between values in A defined as (2.4).

- Case 1 or Case 4: ($a_{12} < 0, \alpha_k \in [0, -r/a_{22}]$ or $a_{12} \geqslant 0, \alpha_k \in [0, 1/a_{11}]$)
The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $-2r\alpha_k y_2 - t_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{a_{11}} \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_k \\ -\frac{\alpha_k}{r} \\ \alpha_k^2 \\ (\frac{\alpha_k}{r})^2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \alpha_k - \frac{\sigma}{2} \\ -\frac{1}{r}(\alpha_k - \frac{\sigma}{2}) \\ \alpha_k(\alpha_k - \sigma) \\ (\frac{1}{r}(\alpha_k - \frac{\sigma}{2}))^2 \\ 0 \\ 1 \end{bmatrix}.$$

- Case 2: ($a_{12} < 0,\ \alpha_k \in [-r/a_{22}, 1/a_{11} - r/a_{22}]$)
For any ϵ such that $-\frac{r}{a_{22}} - \epsilon \in [\frac{\alpha_{p1}+\alpha_{p1-1}}{2}, \frac{\alpha_{p1}+\alpha_{p1+1}}{2}]$, the following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $-2r\alpha_k y_2 - t_1 - (\frac{2r(\alpha_{p1}-\alpha_k)}{a_{22}} + \alpha_{p1}^2 -$

$\alpha_k^2)z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\-\frac{\alpha_k}{r}\\\alpha_k^2\\(\frac{\alpha_k}{r})^2\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k-\frac{\sigma}{2}\\-\frac{1}{r}(\alpha_k-\frac{\sigma}{2})\\\alpha_k(\alpha_k-\sigma)\\(\frac{1}{r}(\alpha_k-\frac{\sigma}{2}))^2\\0\\1\end{bmatrix}, \begin{bmatrix}-\frac{r}{a_{22}}\\\frac{1}{a_{22}}\\-\frac{2r\alpha_{p1}}{a_{22}}-\alpha_{p1}^2\\(\frac{1}{a_{22}})^2\\1\\1\end{bmatrix}, \begin{bmatrix}-\frac{r}{a_{22}}-\epsilon\\\frac{1}{a_{22}}\\-\frac{2r\alpha_{p1}}{a_{22}}-\alpha_{p1}^2\\(\frac{1}{a_{22}})^2\\1\\1\end{bmatrix}.$$

- Case 3: $(0 \leqslant a_{12} \leqslant a_{22},\ \alpha_k \in [-r/a_{22}, 0])$
  The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $-2r\alpha_k y_2 - t_1 - (-\frac{2r\alpha_k}{a_{22}} - \alpha_k^2)z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\-\frac{\alpha_k}{r}\\\alpha_k^2\\(\frac{\alpha_k}{r})^2\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k-\frac{\sigma}{2}\\-\frac{1}{r}(\alpha_k-\frac{\sigma}{2})\\\alpha_k(\alpha_k-\sigma)\\(\frac{1}{r}(\alpha_k-\frac{\sigma}{2}))^2\\0\\1\end{bmatrix}, \begin{bmatrix}0\\\frac{1}{a_{22}}\\0\\(\frac{1}{a_{22}})^2\\1\\1\end{bmatrix}, \begin{bmatrix}-\frac{\sigma}{2}\\\frac{1}{a_{22}}\\0\\(\frac{1}{a_{22}})^2\\1\\1\end{bmatrix}.$$

- Case 5: $(a_{12} > a_{22},\ \alpha_k \in [-r/a_{22}, 1/a_{11} - r/a_{22}])$
  For any ϵ such that $\frac{1}{a_{11}} - \frac{r}{a_{22}} - \epsilon \in [\frac{\alpha_{p2}+\alpha_{p2-1}}{2}, \frac{\alpha_{p2}+\alpha_{p2+1}}{2}]$, the following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $-2r\alpha_k y_2 - t_1 - (-\frac{2\alpha_{p2}}{a_{11}} + \frac{2r(\alpha_{p2}-\alpha_k)}{a_{22}} + \alpha_{p2}^2 - \alpha_k^2)z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\-\frac{\alpha_k}{r}\\\alpha_k^2\\(\frac{\alpha_k}{r})^2\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k-\frac{\sigma}{2}\\-\frac{1}{r}(\alpha_k-\frac{\sigma}{2})\\\alpha_k(\alpha_k-\sigma)\\(\frac{1}{r}(\alpha_k-\frac{\sigma}{2}))^2\\0\\1\end{bmatrix}, \begin{bmatrix}\frac{1}{a_{11}}-\frac{r}{a_{22}}\\\frac{1}{a_{22}}\\2\alpha_{p2}(\frac{1}{a_{11}}-\frac{r}{a_{22}})-\alpha_{p2}^2\\(\frac{1}{a_{22}})^2\\1\\1\end{bmatrix}, \begin{bmatrix}\frac{1}{a_{11}}-\frac{r}{a_{22}}-\epsilon\\\frac{1}{a_{22}}\\2\alpha_{p2}(\frac{1}{a_{11}}-\frac{r}{a_{22}})-\alpha_{p2}^2\\(\frac{1}{a_{22}})^2\\1\\1\end{bmatrix}.$$

- Case 6: $(a_{12} > a_{22},\ \alpha_k \in [1/a_{11} - r/a_{22}, 0])$
  The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy

$-2r\alpha_k y_2 - t_1 - (-\frac{2\alpha_k}{a_{11}})z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\-\frac{\alpha_k}{r}\\\alpha_k^2\\(\frac{\alpha_k}{r})^2\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k-\frac{\sigma}{2}\\-\frac{1}{r}(\alpha_k-\frac{\sigma}{2})\\\alpha_k(\alpha_k-\sigma)\\(\frac{1}{r}(\alpha_k-\frac{\sigma}{2}))^2\\0\\1\end{bmatrix}, \begin{bmatrix}0\\-\frac{\alpha_k}{r}+\frac{1}{a_{12}}\\\alpha_k^2\\(-\frac{\alpha_k}{r}+\frac{1}{a_{12}})^2\\1\\1\end{bmatrix}, \begin{bmatrix}0\\-\frac{\alpha_k}{2r}+\frac{1}{a_{12}}\\0\\(-\frac{\alpha_k}{2r}+\frac{1}{a_{12}})^2\\1\\1\end{bmatrix}.$$

□

We can apply a similar lifting procedure without substituting the y variables to obtain the following valid inequality.

Proposition 2.5. *For each* $\alpha_k \in A$*, define*

$$L_3(\alpha_k) = \begin{cases} 0 & \text{if } \alpha_k \leqslant 0 \\ \alpha_k^2 & \text{if } 0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}} \\ \frac{2\alpha_k - 2\alpha_q}{a_{11}} + \alpha_q^2 & \text{if } \alpha_k \geqslant \frac{1}{a_{11}} \end{cases} \tag{2.8}$$

where $\alpha_q \in A$ *can be uniquely chosen by finding an element in* A *that satisfies* $\frac{\alpha_q+\alpha_{q-1}}{2} \leqslant \frac{1}{a_{11}} \leqslant \frac{\alpha_q+\alpha_{q+1}}{2}$*. Then the inequality*

$$2\alpha_k y_1 - t_1 - L_3(\alpha_k)z_1 - \alpha_k^2 z_2 \leqslant 0 \tag{Ineq L3}$$

is valid for $P(A, B)$*.*

Proof. We start from a valid inequality $2\alpha_k y_1 - t_1 - \alpha_k^2 \leqslant 0$ for some $\alpha_k \in A$ in the description of $P(A, B)$. Under the restriction $z_1 = 0$, $z_2 = 0$ implies $y_1 = 0$ due to (2.3a), so we can apply the perspective reformulation to tighten this inequality. As a result, we obtain the inequality $2\alpha_k y_1 - t_1 - \alpha_k^2 z_2 \leqslant 0$, which is valid if $z_1 = 0$. Then we lift this in z_1 to

$$2\alpha_k y_1 - t_1 - \alpha_k^2 z_2 \leqslant L_3(\alpha_k)z_1. \tag{2.9}$$

The minimal lifting coefficient is attained by solving

$$L_3(\alpha_k) = \max\{2\alpha_k y_1 - t_1 - \alpha_k^2 z_2 \mid (y, t, z) \in S\}. \tag{2.10}$$

We need only consider the case $z_1 = 1$. If $z_1 = z_2 = 1$, the objective function of (2.10) reduces to $2\alpha_k y_1 - t_1 - \alpha_k^2$, which is nonnegative by (2.3c) in the description of $P(A, B)$. A feasible solution $y_1 = \alpha_k, t_1 = \alpha_k^2$ has the objective value 0. Therefore, the optimal objective value for (2.10) in this case is 0. When $z_1 = 1, z_2 = 0$, (2.10) is written as follows:

$$\begin{aligned} L_3(\alpha_k) = \max\ & 2\alpha_k y_1 - t_1 && \text{(Lift3-P)} \\ \text{s.t. } & a_{11} y_1 \geqslant 0, && (\lambda_1) \\ & a_{11} y_1 \leqslant 1, && (\lambda_2) \\ & 2\alpha y_1 - t_1 \leqslant \alpha^2 \ \forall \alpha \in A, && (\mu_i) \\ & t_1 \geqslant 0. \end{aligned}$$

Its optimal solution can be characterized by examining its dual linear program:

$$\begin{aligned} \min\ & \lambda_2 + \sum_{i=1}^{m} \alpha_i^2 \mu_i && \text{(Lift3-D)} \\ \text{s.t. } & -a_{11}\lambda_1 + a_{11}\lambda_2 + \sum_{i=1}^{m} 2\alpha_i \mu_i = 2\alpha_k, \\ & -\sum_{i=1}^{m} \mu_i \geqslant -1, \\ & \lambda, \mu \geqslant 0. \end{aligned}$$

By describing feasible points $(\hat{y}_1, \hat{t}_1)$ to primal (Lift3-P) and $(\hat{\lambda}, \hat{\mu})$ to dual (Lift3-D) with common optimal objective value and comparing it with 0, we obtain closed form solutions to (2.10). If these points have objective value less than 0, the optimal objective value of (2.10) is 0 obtained when $z_2 = 0$. Three possible cases of (Ineq L3) are considered based on the value of α_k: $\alpha_k \leqslant 0$, $0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}}$, and $\alpha_k \geqslant \frac{1}{a_{11}}$.

- If $\alpha_k \leqslant 0$:
  $(0, 0)$ is feasible for (Lift3-P) and $((-\frac{2\alpha_k}{a_{11}}, 0), \mathbf{0})$ is feasible for (Lift3-D) with optimal objective value 0.

- If $0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}}$:
  (α_k, α_k^2) is feasible for (Lift3-P) and $((0, 0), e_k)$ is feasible for (Lift3-D) with optimal objective value α_k^2.

- If $\alpha_k \geqslant \frac{1}{a_{11}}$:

$(\frac{1}{a_{11}}, \frac{2\alpha_q}{a_{11}} - \alpha_q^2)$ is feasible for (Lift3-P) and $((0, -\frac{2\alpha_k - 2\alpha_q}{a_{11}})), e_p)$ is feasible for (Lift3-D) with optimal objective value $\frac{2\alpha_k - 2\alpha_q}{a_{11}} + \alpha_q^2$.

By definition of $L_3(\alpha_k)$, (2.9) is valid for $P(A, B)$. By explicitly characterizing $L_3(\alpha_k)$ as (2.8), the proof is complete. □

Proposition 2.6. *The inequality* (Ineq L3) *defines a facet of* $\text{conv}(P(A, B))$ *if* i) $\alpha_k < 0$*, or* ii) $0 \leqslant \alpha_k < \frac{1}{a_{11}}$ *and* $0 \leqslant -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$*, or* iii) $\alpha_k \geqslant \frac{1}{a_{11}}$*,* $0 \leqslant -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$*, and* $\exists \epsilon : \frac{2\alpha_q}{a_{11}} - \alpha_q^2 - 2\alpha_k \epsilon \geqslant \frac{2\alpha}{a_{11}} - \alpha^2 \; \forall \alpha \in A$.

Proof. To prove the claim, we provide 6 affinely independent points in $P(A, B)$ that satisfies

$$2\alpha_k y_1 - t_1 - L_3(\alpha_k) z_1 - \alpha_k^2 z_2 = 0$$

for each of 3 cases.

- If $\alpha_k < 0$:
  The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $2\alpha_k y_1 - t_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_k \\ -\frac{1}{r}\alpha_k \\ \alpha_k^2 \\ (\frac{\alpha_k}{r})^2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \alpha_k + \frac{\sigma}{2} \\ -\frac{1}{r}(\alpha_k + \frac{\sigma}{2}) \\ \alpha_k(\alpha_k - \sigma) \\ \frac{1}{r^2}(\alpha_k + \frac{\sigma}{2})^2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \alpha_k \\ \frac{1}{r}(\frac{1}{a_{11}} - \alpha_k) \\ \alpha_k^2 \\ \frac{1}{r^2}(\frac{1}{a_{11}} - \alpha_k)^2 \\ 1 \\ 1 \end{bmatrix}.$$

- If $0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}}$ and $0 \leqslant -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$:
  The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_k \\ 0 \\ \alpha_k^2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_k - \frac{\sigma}{2} \\ 0 \\ \alpha_k(\alpha_k - \sigma) \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_k \\ -\frac{1}{r}\alpha_k \\ \alpha_k^2 \\ (\frac{\alpha_k}{r})^2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \alpha_k - \frac{\sigma}{2} \\ -\frac{1}{r}(\alpha_k - \frac{\sigma}{2}) \\ \alpha_k(\alpha_k - \sigma) \\ \frac{1}{r^2}(\alpha_k - \frac{\sigma}{2})^2 \\ 0 \\ 1 \end{bmatrix}.$$

- If $\alpha_k > \frac{1}{a_{11}}$, $0 \leqslant -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$, and $\exists \epsilon : \frac{2\alpha_q}{a_{11}} - \alpha_q^2 - 2\alpha_k\epsilon \geqslant \frac{2\alpha}{a_{11}} - \alpha^2 \ \forall \alpha \in A$:
  The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A, B)$ and satisfy $2\alpha_k y_1 - t_1 - (\frac{2\alpha_k - 2\alpha_q}{a_{11}} + \alpha_q^2)z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\frac{1}{a_{11}}\\0\\\frac{2\alpha_q}{a_{11}} - \alpha_q^2\\0\\1\\0\end{bmatrix}, \begin{bmatrix}\frac{1}{a_{11}} - \epsilon\\0\\\frac{2\alpha_q}{a_{11}} - \alpha_q^2 - 2\alpha_k\epsilon\\0\\1\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\-\frac{1}{r}\alpha_k\\\alpha_k^2\\(\frac{\alpha_k}{r})^2\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k - \frac{\sigma}{2}\\-\frac{1}{r}(\alpha_k - \frac{\sigma}{2})\\\alpha_k(\alpha_k - \sigma)\\\frac{1}{r^2}(\alpha_k - \frac{\sigma}{2})^2\\0\\1\end{bmatrix}.$$

□

Fixing and lifting the variable z_2 and applying the perspective reformulation, we obtain the next inequality.

Proposition 2.7. *For each $\alpha_k \in A$, define*

$$L_4(\alpha_k) = \begin{cases} 0 & \text{if } -\frac{\alpha_k}{r} < 0 \\ \alpha_k^2 & \text{if } 0 \leqslant -\frac{\alpha_k}{r} < \frac{1}{a_{22}} \\ \frac{2r(\alpha_s - \alpha_k)}{a_{22}} + \alpha_s^2 & \text{if } -\frac{\alpha_k}{r} \geqslant \frac{1}{a_{22}} \end{cases} \tag{2.11}$$

where $\alpha_s \in A$ can be uniquely chosen from A by finding an element satisfying $\frac{\alpha_s + \alpha_{s-1}}{2} \leqslant -\frac{r}{a_{22}} \leqslant \frac{\alpha_s + \alpha_{s+1}}{2}$. Then the inequality

$$2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 - L_4(\alpha_k) z_2 \leqslant 0 \tag{Ineq L4}$$

is valid for $P(A, B)$.

Proof. If $z_2 = 0$, it is implied that $y_2 = 0$, so in this case $z_1 = 0$ implies $y_1 = 0$ by (2.3a). This allows the application of the perspective reformulation to $t_1 \geqslant 2\alpha_k y_1 - \alpha_k^2$ for any $\alpha_k \in A$ in the description of $P(A, B)$ to obtain tighter inequality $2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 \leqslant 0$ valid in the case $z_2 = 0$. We lift it back in z_2 to

$$2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 \leqslant L_4(\alpha_k) z_2 \tag{2.12}$$

by solving

$$L_4(\alpha_k) = \max\{2\alpha y_1 - t_1 - \alpha^2 z_1 \mid (y,t,z) \in P(A,B)\}. \tag{2.13}$$

We need only consider the case $z_2 = 1$. If $z_1 = z_2 = 1$, (2.13) reduces to

$$\max\{2\alpha_k y_1 - t_1 - \alpha_k^2 \mid (y,t,z) \in P(A,B)\},$$

for which $y_1 = \alpha_k, t_1 = \alpha_k^2$ is feasible with the optimal objective value 0.

If $z_1 = 0, z_2 = 1$, (2.13) is written as follows:

$$\begin{aligned} \max\ & 2\alpha_k y_1 - t_1 \\ \text{s.t.}\ & -\frac{1}{r} y_1 \geqslant 0, \\ & -\frac{1}{r} y_1 \leqslant \frac{1}{a_{22}}, \\ & 2\alpha y_1 - t_1 \leqslant \alpha^2 \ \forall \alpha \in A, \\ & t_1 \geqslant 0. \end{aligned} \tag{Lift4-P}$$

We use its dual linear program written below to characterize the optimal objective value for (2.13):

$$\begin{aligned} \min\ & \frac{1}{a_{22}} \lambda_2 + \sum_{i=1}^{m} \alpha_i^2 \mu_i \\ \text{s.t.}\ & \frac{1}{r}\lambda_1 - \frac{1}{r}\lambda_2 + \sum_{i=1}^{m} 2\alpha_i \mu_i = 2\alpha_k, \\ & -\sum_{i=1}^{m} \mu_i \geqslant -1, \\ & \lambda, \mu \geqslant 0. \end{aligned} \tag{Lift4-D}$$

The solution to (2.13) can be characterized by describing feasible points $(\hat{y}_1, \hat{t}_1)$ to the primal (Lift4-P) and $(\hat{\lambda}, \hat{\mu})$ to the dual (Lift4-D) with common optimal objective value and comparing this value with 0, which was the objective value obtained when $z_1 = z_2 = 1$. The characterization is done for three different cases defined depending on the values of a_{11}, a_{12}, a_{22}, and α_k.

- If $-\frac{\alpha_k}{r} \leqslant 0:$
  $(0,0)$ is feasible for (Lift4-P) and $((2r\alpha_k, 0), \mathbf{0})$ is feasible for (Lift4-D) with optimal objective

value 0.

- If $0 \leqslant -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$:
  (α_k, α_k^2) is feasible for (Lift4-P) and $((0,0), e_k)$ is feasible for (Lift4-D) with optimal objective value α_k^2.

- If $-\frac{\alpha_k}{r} \geqslant \frac{1}{a_{22}}$:
  $(-\frac{r}{a_{22}}, -\frac{2r\alpha_s}{a_{22}} - \alpha_s^2)$ is feasible for (Lift3-P) and $((0, 2r(\alpha_s - \alpha_k)), e_p)$ is feasible for (Lift3-D) with optimal objective value $\frac{2r(\alpha_s-\alpha_k)}{a_{22}} + \alpha_s^2$.

By definition of $L_4(\alpha_k)$, (2.12) is valid for $P(A,B)$. Therefore, characterizing the optimal objective value to (2.13) as $L_4(\alpha_k)$ defined in (2.11) suffices to establish validity. □

Proposition 2.8. *The inequality* (Ineq L4) *defines a facet of* $\text{conv}(P(A,B))$ *if* i) $-\frac{\alpha_k}{r} \leqslant 0$ *or* ii) $0 < -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$ *and* $0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}}$.

Proof. We provide 6 affinely independent points in $P(A,B)$ that satisfy $2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 - L_4(\alpha_k) z_2 = 0$ for each of the two conditions in the claim.

1. If $-\frac{\alpha_k}{r} \leqslant 0$:
   The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A,B)$ and satisfy $2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\0\\\alpha_k^2\\0\\1\\0\end{bmatrix}, \begin{bmatrix}\alpha_k - \frac{\sigma}{2}\\(\frac{1}{a_{12}} - \frac{\alpha_k}{r})\\\alpha_k(\alpha_k - \sigma)\\(\frac{1}{a_{12}} - \frac{\alpha_k}{r})^2\\1\\1\end{bmatrix}, \begin{bmatrix}0\\\frac{1}{a_{22}}\\0\\0\\0\\1\end{bmatrix}.$$

2. If $0 < -\frac{\alpha_k}{r} \leqslant \frac{1}{a_{22}}$ and $0 \leqslant \alpha_k \leqslant \frac{1}{a_{11}}$:
   The following 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ are in $P(A,B)$ and satisfy

$2\alpha_k y_1 - t_1 - \alpha_k^2 z_1 - \alpha_k^2 z_2 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}\alpha_k\\-\frac{\alpha_k}{r}\\\alpha_k^2\\\frac{\alpha_k^2}{r^2}\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k - \frac{\sigma}{2}\\-\frac{1}{r}(\alpha_k - \frac{\sigma}{2})\\\alpha_k(\alpha_k - \sigma)\\\frac{1}{r^2}(\alpha_k - \frac{\sigma}{2})^2\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha_k\\0\\\alpha_k^2\\0\\1\\0\end{bmatrix}, \begin{bmatrix}\alpha_k - \frac{\sigma}{2}\\0\\\alpha_k(\alpha_k - \sigma)\\0\\1\\0\end{bmatrix}.$$

Note that under this condition, inequality (Ineq L4) coincides with the second case of inequality (Ineq L3).

□

The last valid inequality we introduce is derived using constraint aggregation and the perspective reformulation.

Proposition 2.9. *If* $a_{12} > 0$, *the inequality*

$$(a_{11}y_1 + a_{12}y_2) - \frac{a_{11}}{2\alpha}t_1 - \frac{a_{12}}{2\beta}t_2 - (\frac{\alpha a_{11}}{2} + \frac{\beta a_{12}}{2})z_1 \leqslant 0. \qquad \text{(Ineq L5)}$$

where $0 < \alpha \in A,\ 0 < \beta \in B$ *is valid for* $P(A, B)$.

Proof. We show validity for two values that z_1 takes. If $z_1 = 0$, (2.3a) implies that $a_{11}y_1 + a_{12}y_2 = 0$, thus (Ineq L5) reduces to $\frac{a_{11}}{2\alpha}t_1 + \frac{a_{12}}{2\beta}t_2 \geqslant 0$. This is valid since t_1, t_2 are nonnegative variables and $a_{11}, a_{12}, \alpha, \beta$ are all assumed to be positive. If $z_1 = 1$, (2.3a) is equivalent to $\frac{a_{11}}{2\alpha} \times (t_1 \geqslant 2\alpha y_1 - \alpha^2) + \frac{a_{12}}{2\beta} \times (t_2 \geqslant 2\beta y_2 - \beta^2)$. As this is a linear combination of two valid inequalities with positive weights $\frac{a_{11}}{2\alpha}$ and $\frac{a_{12}}{2\beta}$, it is valid for $P(A, B)$. □

Proposition 2.10. *The inequality* (Ineq L5) *defines a facet of* $\text{conv}(P(A, B))$.

Proof. To prove claim, we present 6 affinely independent points $(y_1, y_2, t_1, t_2, z_1, z_2)$ in $P(A, B)$

that satisfy $(a_{11}y_1 + a_{12}y_2) - \frac{a_{11}}{2\alpha}t_1 - \frac{a_{12}}{2\beta}t_2 - (\frac{\alpha a_{11}}{2} + \frac{\beta a_{12}}{2})z_1 = 0$:

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix}, \begin{bmatrix}\alpha\\\beta\\\alpha^2\\\beta^2\\1\\1\end{bmatrix}, \begin{bmatrix}\alpha-\frac{\sigma}{2}\\\beta\\\alpha(\alpha-\sigma)\\\beta^2\\1\\1\end{bmatrix}, \begin{bmatrix}\alpha\\\beta-\frac{\delta}{2}\\\alpha^2\\\beta(\beta-\delta)\\1\\1\end{bmatrix}, \begin{bmatrix}-\frac{\sigma}{2}\\\frac{\sigma}{2r}\\0\\0\\0\\1\end{bmatrix}.$$

□

We use the following example to demonstrate the usefulness of the derived inequalities.

Example 2.1. *For the following instance of* S*,*

$$\begin{aligned}\hat{S} = \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{B}^2 \mid\ & 0 \leqslant 3y_1 + y_2 \leqslant z_1,\\ & 0 \leqslant \quad 2y_2 \leqslant z_2,\\ & t_1 \geqslant y_1^2,\ t_2 \geqslant y_2^2 \quad\},\end{aligned}$$

we construct a polyhedral outerapproximation with tangential lines of the quadratic inequalities at lower and upper bounds of variables y_1, y_2*. The resulting polyhedron is written as*

$$\begin{aligned}\hat{P}(A,B) = \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{B}^2 \mid\ & 0 \leqslant 3y_1 + y_2 \leqslant z_1,\\ & 0 \leqslant \quad 2y_2 \leqslant z_2,\\ & t_1 \geqslant -\frac{1}{3}y_1 - \frac{1}{36},\\ & t_1 \geqslant \frac{2}{3}y_1 - \frac{1}{9},\\ & t_2 \geqslant y_2 - \frac{1}{4} \quad\}\end{aligned}$$

where $A = \{-\frac{1}{6}, \frac{1}{3}\}$*,* $B = \{0, \frac{1}{2}\}$*.*

It can be shown that the only nontrivial facet-defining inequalities for $\text{conv}(P_1)$ *are*

$$4y_2 - 4t_2 - z_2 \leqslant 0, \tag{f1-1}$$

$$4y_2 - 36t_1 - z_1 - z_2 \leqslant 0, \tag{f1-2}$$

$$-12y_1 - 36t_1 - z_2 \leqslant 0, \tag{f1-3}$$

$$6y_1 - 9t_1 - z_1 \leqslant 0, \tag{f1-4}$$

$$12y_1 + 4y_2 - 18t_1 - 4t_2 - 3z_1 \leqslant 0. \tag{f1-5}$$

All five facets can be generated by using each of the linear inequalities we derived. Specific inequalities and parameters are given below.

- (Ineq L1) *with* $\bar{\beta} = \frac{1}{2}$ *generates facet* (f1-1).
- (Ineq L2) *with* $\alpha_k = -\frac{1}{6},\ L_1(\alpha_k) = 1$ *generates facet* (f1-2).
- (Ineq L3) *with* $\alpha_k = -\frac{1}{6},\ L_3(\alpha_k) = 0$ *generates facet* (f1-3).
- (Ineq L4) *with* $\alpha_k = \frac{1}{3},\ L_4(\alpha_k) = 0$ *generates facet* (f1-4).
- (Ineq L5) *with* $\alpha = \frac{1}{3},\ \beta = \frac{1}{2}$ *generates facet* (f1-5).

2.3 Block-diagonal Decomposition

Recall from the discussion in Section 1.5.8 that our focus on the lower dimensional set S is based on our approach to handle nonseparability by decomposing $Q = R + B$, where B is a 2×2 block diagonal matrix. The choice of $B \succeq 0$ is not unique. Mimicking the earlier work of Frangioni and Gentile [28], we introduce a way of constructing a non-overlapping block diagonal $B \in \mathbb{R}^{n \times n}$ by solving an SDP whose objective is to maximize the amount extracted. We assume that n is an even number and formulate the SDP as follows:

$$\begin{aligned}
\max \quad & \sum_{j=1}^{n} r\mu_j + \sum_{j \in \{1,3,5,\cdots,n-1\}} s\lambda_j \\
\text{s.t.} \quad & Q - \sum_{j=1}^{n} \mu_j e_j e_j^T - \sum_{j \in \{1,3,5,\ldots n-1\}} \lambda_j (e_j e_{j+1}^T + e_{j+1} e_j^T) \succeq 0, \\
& \mu_j e_j e_j^T + \mu_{j+1} e_{j+1} e_{j+1}^T + \lambda_j (e_j e_{j+1}^T + e_{j+1} e_j^T) \succeq 0 \ \forall i \in \{1,3,5,\cdots,n-1\}.
\end{aligned} \tag{2.14}$$

Variables μ_j and λ_j represent the diagonal and off-diagonal elements of a symmetric matrix B, respectively. Different matrices B can be obtained for different value of objective weights r and s. The effectiveness of the relaxation depends on which matrix B is used, but it is still not clear what the best strategy to construct B is. Similar consideration was taken for obtaining the best diagonal decomposition by Zheng, Sun and Li [57]. They formulated an SDP to find the diagonal matrix D which results in the tightest possible continuous relaxation when the perspective reformulation is applied. Their formulation results in a larger dimensional SDP compared to the formulation (1.23) used by Frangioni and Gentile, but provides a tighter root bound. Dong and Linderoth [22] obtain similar result but give computational evidence that extracting from diagonal to optimize root bound is not always the best strategy. Our focus is not on the specific extraction, but on whether or not performing a non-diagonal extraction to reformulate can be useful. For that reason, we do not perform extensive analysis on the effect of optimizing root bound, or on effect of choice of the objective coefficient r and s from (2.14). In all our experiments in the thesis we use $r = 5, s = 2$.

2.4 Computational Experiments

2.4.1 Comparing Relaxations of $P(A, B)$

Let the simple continuous relaxation of $P(A, B)$ be denoted by

$$\begin{aligned}\mathcal{R}(P(A,B)) := \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times [0,1]^2 \mid & 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant z_1,\\ & 0 \leqslant \qquad\quad a_{22}y_2 \leqslant z_2,\\ & t_1 \geqslant 2\alpha y_1 - \alpha^2 \ \forall \alpha \in A,\\ & t_2 \geqslant 2\beta y_2 - \beta^2 \ \forall \beta \in B\ \},\end{aligned}$$

and a stronger relaxation obtained by adding the 5 classes of valid inequalities derived in Section 2.2 by

$$\mathcal{T}(P(A,B)) := \{(y,t,z) \in \mathcal{R}(P(A,B)) \mid \text{(Ineq L1)} - \text{(Ineq L5)}\}.$$

Since $\text{conv}(P(A,B))$ is the tightest possible relaxation, we know that $P(A,B) \subseteq \text{conv}(P(A,B)) \subseteq \mathcal{T}(P(A,B)) \subseteq \mathcal{R}(P(A,B))$. What is not clear is how different these sets are from each other. In particular, we would like to tell whether or not $\mathcal{T}(P(A,B)) = \text{conv}(P(A,B))$. In this section, we perform a simple numerical experiment to estimate the difference between the sets which also shows that $\mathcal{T}(P(A,B)) \neq \text{conv}(P(A,B))$.

We generated random instances of $P(A, B)$ along with random vectors $c \in \mathbb{R}^6$ and solved

$$\omega_H = \min_{(y,t,z) \in \text{proj}_{(y,t,z)} H} c^T(y, t, z) \tag{2.15}$$

for each $H \in \{P(A, B), \text{conv}(P(A, B)), \mathcal{T}(P(A, B)), \mathcal{R}(P(A, B))\}$. For $H = \text{conv}(P(A, B))$, the extended description given in Theorem 1.5 on disjunctive sets is used. This is a result by Balas [6] and brief introduction can be found in Section 1.5.4. Note that we know by Theorem 1.1 (introduced in Section 1.5.1) that $\omega^* := \omega_{P(A,B)} = \omega_{\text{conv}(P(A,B))}$. We test how useful the facet-defining inequalities are in describing $\text{conv}(P(A, B))$ by measuring their ability to close the gap between the MILP solution ω^* and the natural relaxation solution $\omega_{\mathcal{R}(P(A,B)}$. To do that, we define the Gap Closure(GC) in percentage(%) as below:

$$\text{GC} := 100 \times \Big(1 - \frac{\omega_{\mathcal{T}(P(A,B))} - \omega^*}{\omega_{\mathcal{R}(P(A,B))} - \omega^*}\Big).$$

This value represents the proportion of the gap reduced by the inequalities derived in Section 2.2 out of the original gap obtained by solving the simple continuous relaxation.

We created random instances of $P(A, B)$ and (2.15) in the following manner. Random matrices of the form $\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}$ where $a_{11}, a_{22} > 0$ were generated by drawing each element a_{11}, a_{12}, and a_{22} from uniform distributions of integers in the range $[1, 10]$, $[-10, 10]$, $[1, 10]$, respectively. If the generated $a_{12} = 0$, the draw for a_{12} was repeated until we obtain a nonzero value. The objective coefficients c were also drawn from uniform distributions: $c_1, c_2 \sim \mathcal{U}(-100, 10)$, $c_3, c_4 \sim \mathcal{U}(0, 100)$ and $c_5, c_6 \sim \mathcal{U}(-10, 100)$. The coefficients c_3, c_4 on t are chosen to be always nonnegative as otherwise (2.15) will be unbounded. The draws for c were repeated until we obtain an instance of (2.15) satisfying $\omega_{\mathcal{R}(P(A,B))} - \omega^* \geqslant 10^{-3}$, since otherwise there is no gap to close. For each random instance of $P(A, B)$, we compute the lower and upper bounds of y_1 (l_1, l_2) and $y_2(l_2, u_2)$. The set of linearization points A and B to be used for equations (2.3c) and (2.3d) in the description of $P(A, B)$ were constructed to contain 40 equally-spaced points in the intervals $[l_1, u_1]$ and $[l_2, u_2]$, respectively, including the bounds. For each $\alpha \in A$ and $\beta \in B$, all applicable inequalities from Section 2.2 were generated and added to formulate $\mathcal{T}(P(A, B))$.

A hundred random instances were generated by this procedure and (2.15) was solved for each instance. The average gap closure for all instances was 94.52 %. The gap was completely closed (i.e., GC was 100 %) for 90 instances. For the other 10 instances, the remaining gap ranges between 0.1 % and 100 %. The fact that there exist instances for which GC is not 100 % proves that $\mathcal{T}(P(A, B)) \neq \text{conv}(P(A, B))$. Therefore, the inequlaities (Ineq L1) - (Ineq L5) are not the only nontrivial inequalities necessary in the description of $\text{conv}(P(A, B))$.

2.4.2 Shooting Experiment

In Section 2.2, we provided conditions under which each of the valid inequalities for $P(A, B)$ were facet-defining. In this section, we measure the importance of each of the facets with shooting experiment. To our knowledge, it was first proposed by Kuhn [41] as a means to identify new facets of the TSP polytope. Random rays from a fixed point in the interior of the polytope are extended until they intersect facets of the polytope and these facets are recorded. Later Gomory, Johnson, and Evans [29] reported a similar experiment conducted on the cyclic group polyhedra where the random shooting rays originate from outside of the polyhedra and the number of rays intersected for each facet is recorded. The facets hit by the highest percentage of rays are considered most important. Hunsaker [36] found that the facets whose "hit percentage" is large correlate strongly with empirical usefulness of the inequalities.

We conduct a shooting experiment on random instances of $\mathcal{T}(P(A, B))$ to determine which inequalities are more important in optimization. A point $d_0 \in \mathbb{R}^6$ in the interior of $\mathcal{T}(P(A, B))$ is selected, and random rays in direction $d \in \mathbb{R}^6$ are shoot from d_0 until they hit a facet. A shooting for a specific direction $\hat{d}$ corresponds to solving the following LP:

$$\begin{aligned} \max \quad & s \\ \text{s.t.} \quad & (d_0 + sd) \in \mathcal{T}(P(A, B)), \\ & s \in \mathbb{R}_+. \end{aligned} \tag{2.16}$$

Random instances of $\mathcal{T}(P(A, B))$ were generated by using the exactly same procedure described in Section 2.4.1. For each randomly generated instance, d_0 was defined as $(1/4a_{11}, 1/4a_{22}, 1/(8a_{11}^2), 1/(8a_{22}^2), 0.5, 0.5)$ to be sure that none of the inequalities in the description of $\mathcal{R}(P(A, B))$ is tight at the origin of shooting. Each element of the shooting direction d was drawn from a uniform distribution. Specifically, $d_1, d_2, d_5, d_6 \sim \mathcal{U}(-1, 1)$ and $d_3, d_4 \sim \mathcal{U}(0, 1)$. The coefficients d_3, d_4 for variables t_1, t_2 are chosen to be nonnegative as otherwise (2.16) is unbounded. Thirty instances of $\mathcal{T}(P(A, B))$ were generated and 500 randomly generated d were used to solve (2.16) for each of them.

Table 2.1 summarizes the result of the shooting experiment. The entries denote the number of random rays (out of 500) for which the inequalities are tight at the optimal solution of (2.16). The columns represent the inequalities in the description of $\mathcal{R}(P(A, B))$ and (Ineq L1) - (Ineq L5). For numerical purposes, we say the inequality $f(x) \leqslant 0$ is *tight* at $\hat{x}$ if $f(\hat{x}) < 10^{-6}$.

inst.	$\mathcal{R}(P(A,B))$	(Ineq L1)	(Ineq L2)	(Ineq L3)	(Ineq L4)	(Ineq L5)
1	4	68	0	427	1	0
2	71	390	5	13	13	0
3	116	299	8	42	35	0
4	66	361	0	63	9	0
5	0	31	0	0	469	0
6	19	461	0	8	12	0
7	0	368	0	0	131	0
8	500	0	0	0	0	0
9	76	423	0	0	0	0
10	0	31	0	1	467	1
11	0	14	0	486	0	0
12	0	171	0	0	328	1
13	14	34	0	450	2	0
14	0	2	0	0	498	0
15	0	2	0	0	498	0
16	246	82	116	48	6	0
17	10	314	0	176	0	0
18	8	487	0	0	1	0
19	2	487	0	0	0	0
20	1	442	0	1	55	0
21	4	399	0	0	96	0
22	138	361	0	0	0	0
23	0	48	0	0	452	0
24	6	485	0	8	0	0
25	1	246	0	253	0	0
26	10	478	0	6	5	0
27	0	60	0	0	440	0
28	500	0	0	0	0	0
29	0	147	0	0	353	0
30	0	4	0	496	0	0
Average	59.73	223.17	4.30	82.60	129.03	0.07

Table 2.1: Shooting experiment for $\mathcal{T}(P(A,B))$

All inequalities derived in Section 2.2 were hit at least once. On the average, inequalities (Ineq L1), (Ineq L3) and (Ineq L4) were hit by most number of rays.

2.5 Application

In this section, we apply the cutting planes and decomposition technique introduced in Sections 2.2 and 2.3 to application problems. Our goal is to demonstrate the strength of the resulting formulation compared to Frangioni and Gentile's approach [27].

We first define the general relaxation formulation for the two methods. Let $\mathcal{F}$ denote the feasible region of (0-1 MIQCP-*l*):

$$\begin{aligned}\mathcal{F} = \{(v, x, z) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{B}^n \mid\ & v \geqslant x^T Q_0 x + c_0^T x,\\ & Ax + Hz \leqslant f,\\ & 0 \leqslant x_j \leqslant z_j\ \forall j \in [n],\\ & x \in \mathbb{R}^n, z \in \mathbb{B}^p.\}\end{aligned}$$

Let $\mathcal{R}_D(\mathcal{F})$ denote the reformulation of continuous relaxation of $\mathcal{F}$ resulting from the diagonal decomposition suggested by Frangioni and Gentile with $Q = R + D$, $D = \text{diag}(d)$. Let $\mathcal{R}_B(\mathcal{F})$ denote the reformulation from the 2×2 block diagonal decomposition $Q = R + B$:

$$\begin{aligned}\mathcal{R}_D(\mathcal{F}) = \{(x, z, v, w) \in\ \mathbb{R}_+^n \times [0,1]^n \times \mathbb{R}_+^{1+n} \mid\ & v \geqslant x^T R x + d^T w,\\ & w_j \geqslant x_j^2\ \forall j \in [n],\\ & Ax + Hz \leqslant f,\\ & 0 \leqslant x_j \leqslant z_j\ \forall j \in [n]\ \},\end{aligned}$$

$$\begin{aligned}\mathcal{R}_B(\mathcal{F}) = \{(x, y, z, v, t) \in\ \mathbb{R}_+^n \times \mathbb{R}^n \times [0,1]^n \times \mathbb{R}_+^{1+n} \mid & v \geqslant x^T R x + e^T t,\\ & t_j \geqslant y_j^2\ \forall j \in [n],\\ & Ax + Hz \leqslant f,\\ & L^T x = y,\\ & 0 \leqslant x_j \leqslant z_j\ \forall j \in [n].\ \}\end{aligned}$$

Let $\mathcal{T}_D^{\mathcal{O}}(\mathcal{F})$ denote the formulation obtained by applying perspective cuts to a polyhedral outer-approximation of $\mathcal{R}_D(\mathcal{F})$. We denote by $\mathcal{T}_B^{\mathcal{O}}(\mathcal{F})$ the formulation obtained by adding the inequalities (Ineq L1) - (Ineq L5) from Section 2.2 to a polyhedral outer-approximation of $\mathcal{R}_B(\mathcal{F})$. Assuming n is an even number, we obtain $n/2$ blocks and handle each of them separately to obtain polyhedral relaxations of the form (2.3). Specifically, we construct the 6-dimensional polyhedral set denoted by $P^i(A^i, B^i, L_{ii}^{-T}, L_{i,i+1}^{-T}, L_{i+1,i+1}^{-T})$ for each $i \in \{1, 3, \cdots, n-1\}$. The sets A^i, B^i are the set of linearization points chosen between the lower and upper bounds on variables y_i and y_{i+1}, respectively, and L is the Cholesky factor satisfying $B = LL^T$.

Mathematically, these formulations are written as

$$\begin{aligned}\mathcal{T}_D^{\mathcal{O}}(\mathcal{F}) = \{(x, z, v, w) \in \ & \mathbb{R}_+^n \times [0,1]^n \times \mathbb{R}_+^{1+n} \mid \\ & v \geqslant x^T R x + d^T w, \\ & w_j \geqslant 2\eta_{k_j} x_j - \eta_{k_j}^2 z_j \ \forall j \in [n], k_j \in K^j, \\ & x_j \leqslant z_j \ \forall j \in [n] \ \}\end{aligned}$$

where K^j is the set of points where the outer-approximations of $w_j \geqslant x_j^2$ is taken for each $j \in [n]$, and

$$\begin{aligned}\mathcal{T}_B^{\mathcal{O}}(\mathcal{F}) = \{(x, y, z, v, t) \in \ & \mathbb{R}_+^n \times \mathbb{R}^n \times [0,1]^n \times \mathbb{R}_+^{1+n} \mid \\ & v \geqslant x^T R x + e^T t, \\ & t_i \geqslant 2\alpha_i y_i - \alpha_i^2 \ \forall \alpha_i \in A^i, i \in [n], i \text{ odd }, \\ & t_i \geqslant 2\beta_i y_i - \beta_i^2 \ \forall \beta_i \in B^i, i \in [n], i \text{ even }, \\ & (y, t, z) \in \mathcal{T}(P^i(A^i, B^i, L_{ii}^{-T}, L_{i,i+1}^{-T}, L_{i+1,i+1}^{-T})) \ \forall i \in [n], i \text{ odd }, \\ & L^T x = y, \\ & 0 \leqslant x_j \leqslant z_j \ \forall j \in [n] \ \}\end{aligned}$$

for some choice of sets of linearization points A^i, B^i for the $(\frac{i+1}{2})^{\text{th}}$ pair of y variables.

In all following computational experiments, the set of linearization points are defined to contain 40 equally-spaced points between lower and upper bounds of the associated y variables (for A^i, B^i) or x variables (for K^j) including the bounds. As in Section 2.4.1, for each $\alpha \in A^i$ and $\beta \in B^i$, all applicable inequalities of the form (Ineq L1) - (Ineq L5) were generated and added to the definition of the formulation $\mathcal{T}(P^i(A^i, B^i, L_{ii}^{-T}, L_{i,i+1}^{-T}, L_{i+1,i+1}^{-T}))$.

2.5.1 Minimum Variance Portfolio Problem

We consider the portfolio optimization problem where the goal is to produce a financial portfolio with the minimum possible variance satisfying a set of constraints. We use the formulation similar to (1.4).

$$\begin{aligned} \nu^* = \min \quad & x^T Q x \\ \text{s.t.} \quad & e^T x = 1, \\ & e^T z \leqslant K, \\ & \alpha^T x \geqslant \rho, \\ & 0 \leqslant x_j \leqslant z_j \; j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n \end{aligned} \tag{2.17}$$

We created instances of (2.17) randomly in the following manner. The parameters α_j were drawn from uniform distributions in the range $[-0.02, 0.5]$, and the parameter ρ from a uniform distribution in the range $[0, 0.02]$. A random, positive semidefinite matrix Q is constructed as

$$Q = A^T A + \text{diag}(\zeta) + \sum_{j=1}^{n/2} \gamma_j (e_{2j-1} + e_{2j})(e_{2j-1} + e_{2j})^T$$

where $A \in \mathbb{R}^{n \times n}, \zeta \in \mathbb{R}^n, \gamma \in \mathbb{R}^{n/2}$. Each element of A is drawn uniformly at random from the range $[-5, 5]$, each ζ_j is a uniform random variable in the range $[0, \bar{\zeta}]$ and γ_j is a uniform random variable in the range $[0, \bar{\gamma}]$. The third addend in Q is a 2×2 block-diagonal matrix whose j^{th} block has all 4 entries equal to $\gamma_j, j = 1, 2, \cdots, n/2$. By increasing $\bar{\zeta}$, the resulting matrix Q becomes more diagonally dominant. By increasing $\bar{\gamma}$, we can make Q more "block"-diagonally dominant. We would expect the block extraction technique with the facet-defining inequalities (Ineq L1)-(Ineq L5) to perform best if $\bar{\gamma}$ is large. Note that by construction we have $Q \succeq 0$. We created three families of random instances. In the first family, $\bar{\zeta} = \bar{\gamma} = 0$; in the second family, $\bar{\zeta} = 100, \bar{\gamma} = 0$; and in the third family $\bar{\zeta} = 0, \bar{\gamma} = 100$.

To precisely define the relaxations, let

$$P := \{(x, z) \in [0, 1]^n \times [0, 1]^n \mid e^T x = 1, e^T z \leqslant K, \alpha^T x \geqslant \rho, x_j \leqslant z_j \forall j \in [n]\}$$

be the continuous relaxation of the feasible region of (3.40). The natural continuous relaxation obtains root objective value

$$\nu_R := \min_{x,z} \{x^T Q x \mid (x, z) \in P\}.$$

To create a relaxation based on a diagonal matrix-splitting, we solve the SDP (1.23) to obtain an optimal solution d, and we let $Q = R + \text{diag}(d)$. The strengthened relaxation obtains root objective value

$$\nu_D := \min_{x,z,\nu,w} \Big\{ \nu \mid \quad \nu \geqslant x^T R x + e^T w$$
$$(x,z) \in P$$
$$w_j \geqslant 2\eta_{k_j} x_j - \eta_{k_j}^2 z_j \; \forall j \in [n], k_j \in K^j \Big\}.$$

To create a relaxation based on a block-diagonal matrix splitting, we solve the SDP (2.14) to obtain an optimal solution solution B, and we let $Q = R + B, withB = LL^T$. The strengthened relaxation employing our new facet-defining inequalities obtains root objective value

$$\nu_B := \min_{x,z,\nu,t,y} \Big\{ \nu \mid \quad \nu \geqslant x^T R x + e^T t$$
$$(x,z) \in P$$
$$L^T y = x$$
$$(y,t,z) \in \mathcal{T}(P(A^j, B^j, L_{jj}^{-T}, L_{j,j+1}^{-T}, L_{j+1,j+1}^{-T})) \; \forall j \in \{1, 3, \cdots, n-1\} \Big\}.$$

Note the dependence of the feasible region, specifically the sets $P(A^j, B^j)$, on the inverse of the Cholesky factor of B.

We also create a relaxation based on performing the same block-diagonal matrix splitting used to compute ν_B, but we replace $\mathcal{T}(P(A,B))$ (our approximation of $\text{conv}(P(A,B))$), with $\text{conv}(P(A,B))$ using the extended formulation described in Theorem 1.5. Specifically, we compute the value

$$\nu_{\text{conv}} := \min_{x,z,\nu,t,y} \Big\{ \nu \mid \quad \nu \geqslant x^T R x + e^T t$$
$$(x,z) \in P$$
$$L^T y = x$$
$$(y,t,z) \in \text{conv}(P(A^j, B^j, L_{jj}^{-T}, L_{j,j+1}^{-T}, L_{j+1,j+1}^{-T})) \; \forall j \in \{1, 3, \cdots, n-1\} \Big\}.$$

Relaxations are compared with respect to their root optimality gap

$$\text{Gap}(\nu_?) := \frac{\nu_? - \nu^*}{\nu^*},$$

where $\nu_?$ is one of ν_R, ν_D, ν_B, or ν_{conv} depending on which relaxation is being evaluated.

Tables 2.2, 2.3, and 2.4 show the average root optimality gaps for the four different relaxation methods over the 20 randomly generated instances of each problem size.

n	Gap(ν_R)	Gap(ν_D)	Gap(ν_B)	Gap(ν_{conv})
10	39.02	38.56	38.20	38.00
20	33.47	33.17	33.46	33.46
30	27.98	27.49	27.93	27.85
40	19.00	18.60	19.00	19.00

Table 2.2: Average Gap(%) for Min Variance Portfolio Problem with ($\bar{\zeta} = 0, \bar{\gamma} = 0$)

n	Gap(ν_R)	Gap(ν_D)	Gap(ν_B)	Gap(ν_{conv})
10	40.09	12.22	18.73	10.91
20	34.72	11.76	21.94	13.48
30	35.73	9.38	24.68	11.70
40	35.73	10.71	25.22	12.70

Table 2.3: Average Gap(%) for Min Variance Portfolio Problem with ($\bar{\zeta} = 100, \bar{\gamma} = 0$)

n	Gap(ν_R)	Gap(ν_D)	Gap(ν_B)	Gap(ν_{conv})
10	25.42	20.39	15.80	12.36
20	21.65	17.46	15.01	13.84
30	24.73	17.83	16.57	15.61
40	30.71	15.80	22.00	20.78

Table 2.4: Average Gap(%) for Min Variance Portfolio Problem with ($\bar{\zeta} = 0, \bar{\gamma} = 100$)

Table 2.2 shows that for a matrix Q that is not diagonally or block-diagonally dominant, ($\bar{\zeta} = \bar{\gamma} = 0$) using perspective cuts or the facet-defining inequalities from Section 2.2 makes little contribution to tightening the gap at the root node. However, as shown in Table 2.4 and Table 2.3, the root gap was improved significantly for the other two families of instances. When Q is diagonally dominant, ($\bar{\zeta} = 100, \bar{\gamma} = 0$) using diagonal extraction to add perspective cuts works best. The value Gap(z_D) < Gap(z_{conv}) for some instances in this case, which contradicts our expectation that the reverse would be true, since block-diagonal extraction is a generalization of diagonal extraction. However, as discussed in Section 2.3, the root gap depends on the extracted matrix and we do not focus on obtaining a specific extraction that optimizes the root bound. The relaxation from block-diagonal extraction is the strongest when Q is block-diagonally dominant. ($\bar{\zeta} = 0, \bar{\gamma} = 100$)

2.6 Summary

In this chapter, we investigate the polyhedral outer-approximation of S, denoted by $P(A, B)$. This polyhedron is constructed by replacing the quadratic constraints in the description of S with their linearizations taken at a number of break points. We drive 5 classes of valid inequalities for $P(A, B)$ by using techniques such as lifting, variable substitution, constraint aggregation, and the perspective reformulation. It was shown that they are facet-defining under some mild conditions, and the importance of each facet was measured through shooting experiment. Computational experiment provide evidence that these inequalities do not suffice in the description of $\text{conv}(P(A, B))$. Although the inequalities are only applicable for low-dimensional relaxations, they can be helpful in a solution process for a problem of larger dimension. We decribe how to utilize them for each low-dimensional substructure of an optimization problem obtained by matrix decomposition, and demonstrate this framework in computations on the minimum variance portfolio optimization problems of size $n \in \{10, 20.30, 40\}$.

Chapter 3

Quadratic Cutting Surfaces for S

3.1 Introduction

In this chapter, we direct our attention back to the nonlinear set

$$S_0 = \{(y, t, z) \in \mathbb{R}^{2n} \times \mathbb{B}^n \mid t_j \geqslant y_j^2,\ l_j z_j \leqslant [L^{-T} y]_j \leqslant u_j z_j \ \forall j \in [n]\} \tag{3.1}$$

that relaxes the feasible region of a reformulation of 0-1 MIQCP. As in Chapter 2, our investigation focuses on the low-dimensional set

$$\begin{aligned} S = \{(y, t, z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{B}^2 \mid t_1 &\geqslant y_1^2, && (3.2a)\\ t_2 &\geqslant y_2^2, && (3.2b)\\ 0 &\leqslant a_{11} y_1 + a_{12} y_2 \leqslant z_1, && (3.2c)\\ 0 &\leqslant \qquad\quad a_{22} y_2 \leqslant z_2 \}. && (3.2d) \end{aligned}$$

We first characterize in Section 3.1.1 the extreme points of $\mathrm{conv}(S)$ and the extreme points of its natural relaxation. The derivation procedure for the valid inequalities is explained in Section 3.2 where their relationship with the linear inequalities from Chapter 2 is discussed. The results of computational experiments to demonstrate the strength of the formulation that uses the new nonlinear inequalities is reported in sections 3.3 and 3.4.

3.1.1 Extreme Points Characterization

We denote the continuous relaxation of S by

$$\begin{aligned}\mathcal{R}(S) := \{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times \mathbb{R}^2 \mid\ & t_1 \geqslant y_1^2, t_2 \geqslant y_2^2,\\ & 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant z_1,\\ & 0 \leqslant \qquad\quad a_{22}y_2 \leqslant z_2,\\ & 0 \leqslant z_1 \leqslant 1,\ 0 \leqslant z_2 \leqslant 1 \quad \}.\end{aligned} \tag{3.3}$$

Recall from Section 1.5.8 that

$$L^{-T} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

so we may assume $a_{11}, a_{22} > 0$.

S is a union of 4 subsets in which (z_1, z_2) take values among $(0,0), (0,1), (1,0)$, and $(1,1)$. These subsets are all defined by the same set of quadratic inequalities and only differ in the polyhedral feasible region for y_1, y_2. We let $Q_{\bar{z}_1\bar{z}_2}$ denote the polyhedron containing y_1, y_2 for each subset:

$$\begin{aligned}Q_{\bar{z}_1\bar{z}_2} = \{(y_1,y_2) \in \mathbb{R}^2 \mid\ & 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant \bar{z}_1,\\ & 0 \leqslant \qquad\quad a_{22}y_2 \leqslant \bar{z}_2\ \}\end{aligned}$$

for $(\bar{z}_1, \bar{z}_2) \in \{(0,0), (0,1), (1,0), (1,1)\}$.

As explained in Section 1.5.1, we will determine the usefulness of nonlinear valid inequalities based on whether they strengthen the relaxation by cutting off extreme points of the natural relaxation. The following lemma and theorem provide the characterization of the points in $\text{ext}(\mathcal{R}(S))$.

Lemma 3.1. *If* $p = (\hat{y}_1, \hat{y}_2, \hat{t}_1, \hat{t}_2, \hat{z}_1, \hat{z}_2) \in \text{ext}(\mathcal{R}(S))$, *then* $\hat{t}_1 = \hat{y}_1^2, \hat{t}_2 = \hat{y}_2^2$.

Proof. If $p^* = (y_1^*, y_2^*, t_1^*, t_2^*, z_1^*, z_2^*)$ is an optimal solution to the optimization problem

$$\min_{p \in \mathcal{R}(S)} \{c^T p\}$$

with $c > 0$, then it must be that $t_1^* = (y_1^*)^2, t_2^* = (y_2^*)^2$. If not, a solution with strictly smaller objective function value would exist. □

Theorem 3.1. *Define the following 4 infinite families of points parametrized by $\alpha, \beta \in \mathbb{R}$:*

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad 1, \quad 1 \quad) \tag{3.4}$$

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad 1, \quad a_{22}\beta \quad) \tag{3.5}$$

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad a_{11}\alpha + a_{12}\beta, \quad 1 \quad) \tag{3.6}$$

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad a_{11}\alpha + a_{12}\beta, \quad a_{22}\beta \quad). \tag{3.7}$$

The set of extreme points of $\mathcal{R}(S)$ is defined as

$$\text{ext}(\mathcal{R}(S)) = \{(3.4), (3.5), (3.6), (3.7) \mid (\alpha, \beta) \in Q_{11}\}.$$

Proof. ($\supseteq$) Suppose $p = (\bar{y}_1, \bar{y}_2, \bar{t}_1, \bar{t}_2, \bar{z}_1, \bar{z}_2)$ is one of the forms (3.4) - (3.7), but not an extreme point. Then there must exist 2 points distinct from p in $\mathcal{R}(S)$ which we denote by

$$\begin{aligned} u =& (y_{1u}, y_{2u}, t_{1u}, t_{2u}, z_{1u}, z_{2u}), \\ v =& (y_{1v}, y_{2v}, t_{1v}, t_{2v}, z_{1v}, z_{2v}) \end{aligned}$$

such that $(1-\lambda)u + \lambda v = p$ for some $\lambda \in (0, 1)$.

This relationship and the definition of $\mathcal{R}(S)$ imply that

$$\begin{aligned} \bar{y}_1 &= \alpha = (1-\lambda)y_{1u} + \lambda y_{1v} \\ \bar{t}_1 &= \alpha^2 = (1-\lambda)t_{1u} + \lambda t_{1v} \\ & t_{1u} \geqslant y_{1u}^2, \ t_{1v} \geqslant y_{1v}^2. \end{aligned}$$

This system of equations and inequalities implies that

$$\begin{aligned} \bar{t}_1 = \alpha^2 =& (1-\lambda)t_{1u} + \lambda t_{1v} \\ \geqslant& (1-\lambda)y_{1u}^2 + \lambda y_{1v}^2 \\ \geqslant& ((1-\lambda)y_{1u} + \lambda y_{1v})^2 \\ =& \bar{y}_1^2. \end{aligned}$$

Due to the strict convexity of function $f(y) = y^2$, the last inequality holds at equality only if $y_{1u} = y_{1v}$. It follows that $y_{1u} = y_{1v} = \alpha$, $t_{1u} = t_{1v} = \alpha^2$. The same argument applies to the system of equations and inequalities involving $\bar{y}_2, y_{2u}, y_{2v}, \bar{t}_2, t_{2u}, t_{2v}$ and implies that

$y_{2u} = y_{2v} = \beta,\ t_{2u} = t_{2v} = \beta^2$. Then the definition of $\mathcal{R}(S)$ implies that

$$\begin{aligned} 0 \leqslant a_{11}\alpha + a_{12}\beta &\leqslant \hat{z}_1 \leqslant 1, \\ 0 \leqslant \qquad\quad a_{22}\beta &\leqslant \hat{z}_2 \leqslant 1, \end{aligned}$$

and thus for a point p of the form (3.4),(3.5),(3.6),or (3.7), any solution to the equation $(1-\lambda)u + \lambda v = p,\ \lambda \in (0,1)$ satisfies $z_{1u} = z_{1v} = \hat{z}_1,\ z_{2u} = z_{2v} = \hat{z}_2$. We have established that $p = u = v$, which contradicts the assumption that p is not an extreme point. Therefore any point in $\mathcal{R}(S)$ of the given form is an extreme point.

($\subseteq$) Suppose there exists an extreme point $p = (\bar{y}_1, \bar{y}_2, \bar{t}_1, \bar{t}_2, \bar{z}_1, \bar{z}_2)$ that cannot be written in the form of (3.4),(3.5),(3.6), or (3.7). From Lemma 3.1, we know p can be written as $(\alpha, \beta, \alpha^2, \beta^2, \hat{z}_1, \hat{z}_2)$. Since $p \in \mathcal{R}(S)$, it satisfies

$$\begin{aligned} 0 \leqslant a_{11}\alpha + a_{12}\beta &\leqslant \bar{z}_1 \leqslant 1 \\ 0 \leqslant \qquad\quad a_{22}\beta &\leqslant \bar{z}_2 \leqslant 1 \end{aligned}$$

As we assumed $\bar{z}_1 \notin \{a_{11}\alpha + a_{12}\beta, 1\}$, $a_{11}\alpha + a_{12}\beta < \bar{z}_1 < 1$. Then we can take two points in $\mathcal{R}(S)$

$$\begin{aligned} u =& (\alpha, \beta, \alpha^2, \beta^2, a_{11}\alpha + a_{12}\beta, \bar{z}_2), \\ v =& (\alpha, \beta, \alpha^2, \beta^2, 1, \bar{z}_2) \end{aligned}$$

that satisfy $p = (1-\lambda)u + \lambda v$ where

$$\lambda = \frac{\bar{z}_1 - (a_{11}\alpha + a_{12}\beta)}{1 - (a_{11}\alpha + a_{12}\beta)} \in (0, 1).$$

This contradicts the assumption that p is extreme, and thus we conclude that $\bar{z}_1 \in \{a_{11}\alpha + a_{12}\beta, 1\}$. A similar argument shows that $\bar{z}_2 \in \{a_{22}\beta, 1\}$, proving that any extreme point can be written in the forms (3.4),(3.5),(3.6), or (3.7). □

Note that the points in the 4 families (3.4),(3.5),(3.6), and (3.7) may not be distinct from each other depending on the values of $a_{22}\beta$ and $a_{12}\alpha + a_{12}\beta$. As there are infinitely many of these combinations, this family of points constitutes an uncountably infinite set of extreme points.

With a similar argument, we also characterize $\text{ext}(\text{conv}(S))$. Not surprisingly, $\text{conv}(S)$ is not a polyhedron.

Theorem 3.2. *Define the following 4 infinite families of points* $(\bar{y}_1, \bar{y}_2, \bar{t}_1, \bar{t}_2, \bar{z}_1, \bar{z}_2)$ *parametrized by* $\alpha, \beta \in \mathbb{R}$*:*

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad 1, \quad 1 \quad) \tag{3.8}$$

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad 1, \quad 0 \quad) \tag{3.9}$$

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad 0, \quad 1 \quad) \tag{3.10}$$

$$(\quad \alpha, \quad \beta, \quad \alpha^2, \quad \beta^2, \quad 0, \quad 0 \quad). \tag{3.11}$$

The set of extreme points of $\text{conv}(S)$ *is defined as*

$$\begin{aligned}\text{ext}(\text{conv}(S)) = \mathcal{E} :=& \{(3.8) \mid (\alpha, \beta) \in Q_{11}\} \cup \{(3.9) \mid (\alpha, \beta) \in Q_{10}\} \\ & \cup \{(3.10) \mid (\alpha, \beta) \in Q_{01}\} \cup \{(3.11) \mid (\alpha, \beta) \in Q_{00}\}.\end{aligned}$$

Proof. ($\supseteq$) Suppose for contradiction that there exists a point $p = (\bar{y}_1, \bar{y}_2, \bar{t}_1, \bar{t}_2, \bar{z}_1, \bar{z}_2)$ which is in the set $\mathcal{E}$ defined above, but which not an extreme point. Then there exist two distinct points in $\text{conv}(S)$ which we denote by

$$\begin{aligned}u =& (y_{1u}, y_{2u}, t_{1u}, t_{2u}, z_{1u}, z_{2u}), \\ v =& (y_{1v}, y_{2v}, t_{1v}, t_{2v}, z_{1v}, z_{2v})\end{aligned}$$

such that $(1 - \lambda)u + \lambda v = p$ for some $\lambda \in (0, 1)$. Since $(\hat{z}_1, \hat{z}_2) \in \mathbb{B}^2$ and $z_{iu}, z_{iv} \in [0, 1],\ i = 1, 2$, any solution to this equation satisfies $z_{iu} = z_{iv} = \hat{z}_i,\ i = 1, 2$. Applying the same argument from the proof of Theorem 3.1, we can deduce that $y_{iu} = y_{iv} = \bar{y}_i,\ t_{iu} = t_{iv} = \bar{t}_i = \bar{y}_i^2,\ i = 1, 2$. This results in $u = v = p$, which contradicts the assumption that p is not extreme.

($\subseteq$) Suppose for contradiction that there exists a point $p \in \text{ext}(\text{conv}(S))$ that cannot be written in the forms (3.8)-(3.11). First note that $(\bar{z}_1, \bar{z}_2) \in \mathbb{B}^2$ is a necessary condition for p to be extreme. If $(\bar{z}_1, \bar{z}_2) \notin \mathbb{B}^2$, then $p \in \text{conv}(S) \backslash S$ which means that p is written as a convex combination of two or more points in S and by definition not an extreme point. The condition $(\bar{y}_1, \bar{y}_2) \in Q_{\bar{z}_1 \bar{z}_2}$ follows from feasibility requirement. Therefore, that p is not of one of the forms (3.8) - (3.11) means $\bar{t}_i > \bar{y}_i^2$ for $i = 1$ or 2. Let $\bar{t}_i = \bar{y}_i^2 + \epsilon_i$ for $\epsilon_i > 0,\ i = 1, 2$ and take

$$\begin{aligned}u =& (y_{1u}, y_{2u}, \bar{y}_1^2 + 2\epsilon_1, \bar{y}_2^2 + 2\epsilon_2, z_{1u}, z_{2u}), \\ v =& (y_{1v}, y_{2v}, \bar{y}_1^2, \bar{y}_2^2, z_{1v}, z_{2v}).\end{aligned}$$

Then $p = 0.5u + 0.5v$, so P is not extreme, contradicting the assumption. □

3.2 Valid and Useful Inequalities for S

Motivated by an intuitive conjecture that facets of $P(A, B)$ would correspond to linearization of nonlinear inequalities that describe the boundaries of $\text{conv}(S)$, we take a similar approach to generate strong quadratic inequalities valid for S. Specifically, procedures such as fixing the value of a variable followed by substitution, (nonlinear) lifting, and strengthening the inequalities via the perspective reformulation were applied to produce convex nonlinear inequalities valid for S. The usefulness of these inequalities is illustrated in Section 3.3.2 by showing when the points in $\text{ext}(\mathcal{R}(S))$ (characterized in Theorem 3.1) are cut off by the inequalities.

The first valid inequality is a result of direct application of perspective reformulation to the variables t_2, y_2 and z_2.

Proposition 3.1.

$$z_2 t_2 \geqslant y_2^2. \tag{Ineq 1}$$

is a valid inequality for S.

Proof. If $z_2 = 1$, then (Ineq 1) is the same as (3.2d) in the description of S. If $z_2 = 0$, then (3.2d) implies that $y_2 = 0$, so (Ineq 1) is valid in this case as well. □

Note that (Ineq 1) is in the form of a rotated second order cone, since $z_2, t_2 \geqslant 0$.

The second valid inequality is obtained by fixing the variable $z_1 = 0$, substituting variable y_1 with $-ry_2$, applying the perspective reformulation on t_1, y_2 and z_2 and lifting for validity when $z_1 = 1$.

Proposition 3.2. *The inequality*

$$C_1 z_1 z_2 + z_2 t_1 \geqslant r^2 y_2^2 \tag{Ineq 2}$$

where

$$C_1 = \begin{cases} 0 & \text{if } a_{12} < 0 \\ \frac{r^2}{a_{22}^2} & \text{if } 0 \leqslant a_{12} \leqslant a_{22} \\ \frac{r^2}{a_{22}^2} - (\frac{1}{a_{11}} - r\frac{1}{a_{22}})^2 & \text{if } a_{12} > a_{22} \end{cases}$$

is valid for S.

Proof. When $z_2 = 0$, (Ineq 2) is valid since both left and right-hand side of the inequality reduce to 0 as (3.2d) implies that $y_2 = 0$. If $z_2 = 1$ and $z_1 = 0$, then (3.2c) implies that $y_1 = -ry_2$. By substitution, (Ineq 2) is equivalent to (3.2a) and therefore valid in this case. We need only prove that (Ineq 2) is valid when $z_1 = z_2 = 1$, which is true if $C_1 \geqslant r^2y_2^2 - t_1$. Thus the minimum lifting coefficient is attained by solving

$$\begin{aligned} \max\ & r^2y_2^2 - y_1^2 \\ \text{s.t. } 0 \leqslant\ & a_{11}y_1 + a_{12}y_2 \leqslant 1, \\ 0 \leqslant\ & \qquad\qquad a_{22}y_2 \leqslant 1. \end{aligned} \tag{3.12}$$

which is a nonconvex optimization problem. Note that in the objective of the lifting problem (3.12), y_1^2 replaces t_1 as $t_1 = y_1^2$ at an optimal solution. Instead of using an optimality certificate to identify a closed form solution to the problem as in Chapter 2, we exploit the fact that the objective function is separable to obtain a closed-form solution. Define

$$y_1^*(y_2) = \underset{y_1}{\operatorname{argmax}}\{r^2y_2^2 - y_1^2 \mid 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant 1\}.$$

Observing that for a fixed y_2 this optimization problem is equivalent to

$$\begin{aligned} & \underset{y_1}{\operatorname{argmin}}\{|y_1| \mid 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant 1\} \\ = & \underset{y_1}{\operatorname{argmin}}\{y_1 \mid -ry_2 \leqslant y_1 \leqslant \frac{1}{a_{11}} - ry_2\} \end{aligned} \tag{3.13}$$

and $y_2 \geqslant 0\ \forall y_2 \in S$, we characterize $y_1^*(y_2)$ as

$$y_1^*(y_2) = \begin{cases} -ry_2 & \text{if } r < 0 \\ 0 & \text{if } r > 0,\ \frac{1}{a_{11}} - ry_2 \geqslant 0 \\ \frac{1}{a_{11}} - ry_2 & \text{if } r > 0,\ \frac{1}{a_{11}} - ry_2 < 0. \end{cases}$$

Here we see that $y_1^*(y_2)$ is unique. So now the lifting problem (3.12) becomes an optimization problem over only y_2.

$$C_1 = \max_{y_2}\left\{r^2y_2^2 - (y_1^*(y_2))^2 \mid 0 \leqslant a_{11}y_1^*(y_2) + a_{12}y_2 \leqslant 1, 0 \leqslant a_{22}y_2 \leqslant 1\right\}. \tag{3.14}$$

Recall that $r = \frac{a_{12}}{a_{11}}$ and as $a_{11} > 0$, $a_{12} > 0 \Leftrightarrow r > 0$. If $a_{12} < 0$, then substituting $y_1^*(y_2) = -ry_2$

into (3.14) gives

$$C_1 = \max_{y_2}\{0 \mid 0 \leqslant a_{22}y_2 \leqslant 1\} = 0.$$

If $a_{12} > 0$, then we must consider two subcases depending on the value of $y_1^*(y_2)$. Specifically, in the case $y_2 \leqslant 1/a_{12}$, we substitute $y_1^*(y_2) = 0$ into (3.14), and in the case $y_2 \leqslant 1/a_{12}$, we substitute $y_1^*(y_2) = \frac{1}{a_{11}} - ry_2$ into (3.14). Performing these substitutions and simplifying the expressions yields

$$\begin{aligned} C_1 = \max\Big\{ &\max_{y_2 \leqslant 1/a_{12}} \{r^2 y_2^2 \mid 0 \leqslant a_{12}y_2 \leqslant 1, 0 \leqslant a_{22}y_2 \leqslant 1\}, \\ &\max_{y_2 > 1/a_{12}} \{\frac{r^2}{a_{22}^2} - (\frac{1}{a_{11}} - r\frac{1}{a_{22}})^2 \mid 0 \leqslant a_{22}y_2 \leqslant 1\}\Big\}. \end{aligned} \tag{3.15}$$

We can solve each of the two optimization problems in the definition of C_1 in closed form. If $a_{12} \leqslant a_{22}$, the first maximization problem in (3.15) has optimal value $(a_{12}/(a_{11}a_{22}))^2$, and the second optimization problem is infeasible. Thus $C_1 = r^2/a_{22}^2$ when $a_{12} \leqslant a_{22}$. If $a_{12} \geqslant a_{22}$, the first maximization problem in (3.15) has optimal value $1/a_{11}^2$, and the second maximization problem in (3.15) has optimal value $\frac{r^2}{a_{22}^2} - (\frac{1}{a_{11}} - r\frac{1}{a_{22}})^2 = (1/a_{11}^2)(2a_{12}/a_{22} - 1)$. Since $(2a_{12}/a_{22} - 1) \geqslant 1$ when $a_{12} \geqslant a_{22}$, we know that $C_1 = (1/a_{11}^2)(2a_{12}/a_{22} - 1)$ when $a_{12} \geqslant a_{22}$. Having established the proper value of C_1 in all cases, the proof is complete. □

Note that since $C_1 \geqslant 0$ in all cases, (Ineq 2) can be written in the form of a rotated second order cone constraint:

$$(C_1 z_1 + t_1)z_2 \geqslant (ry_2)^2.$$

A similar lifting procedure involving variable z_2 was used to derive the following inequality.

Proposition 3.3. *The inequality*

$$C_2 z_1 z_2 + a_{11}^2 z_1 t_1 \geqslant (a_{11}y_1 + a_{12}y_2)^2 \tag{Ineq 3}$$

where

$$C_2 = \begin{cases} 0 & \text{if } a_{12} < 0 \\ -(\frac{a_{12}}{a_{22}})^2 + \frac{2a_{12}}{a_{22}} & \text{if } 0 \leqslant a_{12} \leqslant a_{22} \\ 1 & \text{if } a_{12} > a_{22} \end{cases}$$

is valid for S.

Proof. If $z_1 = 0$, then (Ineq 3) reduces to $0 \geqslant (a_{12}y_1 + a_{12}y_2)^2$, which is implied by (3.2c) in the description of S. If $z_1 = 1, z_2 = 0$, then (Ineq 3) is $a_{11}^2 t_1 \geqslant (a_{11}y_1)^2$, which is equivalent to (3.2a).

Thus, we need only prove that (Ineq 3) is valid for the case $z_1 = z_2 = 1$, which is true if

$$\begin{aligned} C_2 &\geqslant \max\{(a_{11}y_1 + a_{12}y_2)^2 - a_{11}^2 t_1 \mid 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant 1, 0 \leqslant a_{22}y_2 \leqslant 1, t_1 \geqslant y_1^2\} \\ &= \max\{(a_{11}y_1 + a_{12}y_2)^2 - a_{11}^2 y_1^2 \mid 0 \leqslant a_{11}y_1 + a_{12}y_2 \leqslant 1, 0 \leqslant a_{22}y_2 \leqslant 1\}. \end{aligned}$$

Defining $Y = a_{11}y_1 + a_{12}y_2$, we can rewrite this as

$$C_2 = \max\{Y^2 - a_{11}^2 y_1^2 \mid 0 \leqslant Y \leqslant 1, 0 \leqslant \frac{a_{22}}{a_{12}}Y - \frac{a_{11}a_{22}}{a_{12}}y_1 \leqslant 1\}. \tag{3.16}$$

We again can exploit separability to derive a closed-form solution to this problem. Let

$$\begin{aligned} y_1^*(Y) &:= \underset{y_1}{\operatorname{argmax}}\{Y^2 - a_{11}^2 y_1^2 \mid -\frac{a_{22}}{a_{12}}Y \leqslant \frac{-a_{11}a_{22}}{a_{12}}y_1 \leqslant 1 - \frac{a_{22}}{a_{12}}Y\} \\ &= \underset{y_1}{\operatorname{argmin}}\{|y_1| \mid -\frac{a_{22}}{a_{12}}Y \leqslant \frac{-a_{11}a_{22}}{a_{12}}y_1 \leqslant 1 - \frac{a_{22}}{a_{12}}Y\}. \end{aligned}$$

The definition of C_2 in (3.16) can be replaced by a 1-dimensional optimization problem

$$C_2 = \max_Y\{Y^2 - a_{11}^2(y_1^*(Y))^2 \mid 0 \leqslant Y \leqslant 1, 0 \leqslant \frac{a_{22}}{a_{12}}Y - \frac{a_{11}a_{22}}{a_{12}}y_1^*(Y) \leqslant 1\} \tag{3.17}$$

assuming that $y_1^*(Y)$ is a singleton $\forall Y$. If $a_{12} \leqslant 0$, then

$$y_1^*(Y) = \underset{y_1}{\operatorname{argmin}}\{|y_1| \mid Y/a_{11} \leqslant y_1 \leqslant Y/a_{11} - a_{12}/(a_{11}a_{22})\} = Y/a_{11},$$

since $Y/a_{11} \geqslant 0$. Substituting this into (3.17), we get that

$$C_2 = \max_Y\{0 \mid 0 \leqslant Y \leqslant 1\} = 0.$$

If $a_{12} > 0$, then

$$\begin{aligned} y_1^*(Y) &= \underset{y_1}{\operatorname{argmin}}\{|y_1| \mid Y/a_{11} \geqslant y_1 \geqslant Y/a_{11} - a_{12}/(a_{11}a_{22})\} \\ &= \begin{cases} 0 & \text{if } Y \leqslant a_{12}/a_{22} \\ Y/a_{11} - a_{12}/(a_{11}a_{22}) & \text{if } Y \geqslant a_{12}/a_{22}. \end{cases} \end{aligned}$$

$$C_2 = \max\left\{ \max_{Y \leqslant a_{12}/a_{22}} \{Y^2 \mid 0 \leqslant Y \leqslant 1, 0 \leqslant (a_{22}/a_{12})Y \leqslant 1\}, \right. \tag{3.18}$$
$$\left. \max_{Y > a_{12}/a_{22}} \{2a_{12}Y/a_{22} - (a_{12}/a_{22})^2 \mid 0 \leqslant Y \leqslant 1\} \right\}$$

If $a_{12} \geqslant a_{22}$, the first maximization problem in (3.18) has optimal value 1, and the second optimization problem is infeasible. So $C_2 = 1$ in the case that $a_{12} > 0, a_{12} \geqslant a_{22}$. If $a_{12} \leqslant a_{22}$, the first maximization problem in (3.18) has optimal value $(a_{12}/a_{22})^2$, and the second maximization problem in (3.18) has optimal value $2a_{12}/a_{22} - (a_{12}/a_{22})^2$. Since

$$2a_{12}/a_{22} - (a_{12}/a_{22})^2 - (a_{12}/a_{22})^2 = \frac{2a_{12}}{a_{22}}\left(1 - \frac{a_{12}}{a_{22}}\right) \geqslant 0$$

when $a_{22} > a_{12}$, we can conclude that $C_2 = 2a_{12}/a_{22} - (a_{12}/a_{22})^2$ in this case. Having established the value of C_2 in all cases, the proof is complete. □

Since if $0 \leqslant a_{12} \leqslant a_{22}$, we have that $C_2 = (a_{12}/a_{22})(2 - a_{12}/a_{22}) \geqslant 0$, (Ineq 3) can be written as the rotated second order cone constraint

$$(C_2 z_2 + a_{11}^2 t_1) z_1 \geqslant (a_{11} y_1 + a_{12} y_2)^2.$$

The final valid inequality is equivalent to applying the perspective reformulation to the variables t_1, y_1 and $(z_1 + z_2)$ which is applicable because $z_1 + z_2 = 0$ implies $y_1 = 0$.

Proposition 3.4.

$$t_1(z_1 + z_2) \geqslant y_1^2. \tag{Ineq 4}$$

is a valid inequality for S.

Proof. We verify the validity for each value that z_1, z_2 take.

1. $(z_1 = 0,\ z_2 = 0)$
   (Ineq 4) simply reduces to $0 \geqslant 0$.

2. $(z_1 = 0,\ z_2 = 1$ or $z_1 = 1,\ z_2 = 0)$
   (Ineq 4) is equivalent to (3.2a) in the description of S.

3. $(z_1 = 1,\ z_2 = 1)$
   (Ineq 4) is equivalent to $2t_1 \geqslant y_1^2$, which is dominated by (3.2a) in the description of S.

□

Note that the feasible region for (Ineq 4) is convex since it is clearly in rotated second-order cone form.

Adding the valid inequalities (Ineq 1)-(Ineq 4) to the continuous relaxation of S results in a strengthened relaxation of S, which we denote as

$$
\begin{aligned}
\mathcal{T}(S) = \Big\{(y,t,z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \mid\ & t_1 \geqslant y_1^2, t_2 \geqslant y_2^2, \\
& z_2 t_2 \geqslant y_2^2 \\
& C_1 z_1 z_2 + z_2 t_1 \geqslant \frac{a_{12}^2}{a_{11}^2} y_2^2 \\
& C_2 z_1 z_2 + a_{11}^2 z_1 t_1 \geqslant (a_{11} y_1 + a_{12} y_2)^2 \\
& t_1 (z_1 + z_2) \geqslant y_1^2 \\
& 0 \leqslant a_{11} y_1 + a_{12} y_2 \leqslant z_1, \\
& 0 \leqslant \qquad\quad a_{22} y_2 \leqslant z_2, \\
& 0 \leqslant z_1 \leqslant 1,\ 0 \leqslant z_2 \leqslant 1 \quad \Big\}.
\end{aligned}
$$

3.2.1 Relationship between the linear and nonlinear inequalities

The derivation technique for the strong inequalities for linear outer-approximation $P(A,B)$ and the nonlinear inequalities for S are very similar in principle, and the resulting sets of inequalities are closely related to each other. Specifically, we found that the linearization of some of the nonlinear inequalities at particular points exactly coincides with the facet-defining linear inequalities for the polyhedral outer-approximation introduced in Chapter 2.

For example, the first valid inequality $z_2 t_2 \geqslant y_2^2$ (Ineq 1) for S can be linearized to generate the valid inequality $t_2 \geqslant 2\bar{\beta} y_2 - \bar{\beta}^2 z_2$ (Ineq L1) from Chapter 2. In general, the linear outer-approximation of a convex inequality of the form $f(x) \leqslant 0$ at $x = \bar{x}$ is the following inequality:

$$\nabla f(\bar{x})^T (x - \bar{x}) + f(\bar{x}) \leqslant 0.$$

Writing (Ineq 1) as $f(y_2, t_2, z_2) = y_2^2 - t_2 z_2 \leqslant 0$, its outer-approximation at the point $(\bar{y}_2, \bar{t}_2, \bar{z}_2)$ is

Nonlinear Ineq		Linearization Point $(\bar{y}_1, \bar{y}_2, \bar{t}_1, \bar{t}_2, \bar{z}_1, \bar{z}_2)$	Linear Ineq	Condition
(Ineq 1)		$(\bar{y}_1, \beta, \bar{t}_1, \beta^2, \bar{z}_1, 1)$	(Ineq L1)	$\beta \in B$
(Ineq 2)	$r < 0$	$(\bar{y}_1, -\alpha_k/r, \alpha_k^2, \bar{t}_2, \bar{z}_1, 1)$	(Ineq L2)	$0 \leqslant \alpha_k \leqslant 1/a_{11}$
	$0 < a_{12} \leqslant a_{22}$	$(\bar{y}_1, 1/a_{22}, \alpha_k^2, \bar{t}_2, 0, 1)$	(Ineq L2)	$\alpha_k = -r/a_{22} < 0$
	$a_{12} > a_{22}$	$(\bar{y}_1, -\alpha_k/r, \alpha_k^2, \bar{t}_2, 0, 1)$	(Ineq L2)	$a_{22} = a_{12}, 0 \leqslant \alpha_k \leqslant 1/a_{11}$
(Ineq 3)		-	-	-
(Ineq 4)		$(\alpha_k, \bar{y}_2, \alpha_k^2, \bar{t}_2, 1, 0)$	(Ineq L3)	$0 \leqslant \alpha_k \leqslant 1/a_{11}$
		$(\alpha_k, \bar{y}_2, \alpha_k^2, \bar{t}_2, 0, 1)$	(Ineq L4)	$0 \leqslant -\alpha_k/r \leqslant 1/a_{22}$

Table 3.1: Relationship between nonlinear and linear inequalities

given by

$$\nabla f(\bar{y}_2, \bar{t}_2, \bar{z}_2)^T \begin{bmatrix} y_2 - \bar{y}_2 \\ t_2 - \bar{t}_2 \\ z_2 - \bar{z}_2 \end{bmatrix} + f(\bar{y}_2, \bar{t}_2, \bar{z}_2)$$

$$= \begin{bmatrix} 2\bar{y}_2 \\ -\hat{z}_2 \\ -\hat{t}_2 \end{bmatrix}^T \begin{bmatrix} y_2 - \bar{y}_2 \\ t_2 - \bar{t}_2 \\ z_2 - \bar{z}_2 \end{bmatrix} + f(\bar{y}_2, \bar{t}_2, \bar{z}_2)$$

$$= 2\bar{y}_2 y_2 - \bar{z}_2 t_2 - \bar{t}_2 z_2 - \bar{y}_2^2 + \bar{t}_2 \bar{z}_2 \leqslant 0.$$

At the point $(\bar{y}_2, \bar{t}_2, \bar{z}_2) = (\bar{\beta}, \bar{\beta}^2, 1)$, this linear inequality is $2\bar{\beta} y_2 - t_2 - \bar{\beta}^2 \leqslant 0$, which is precisely (Ineq L1).

We can similarly write the outer-approximation of each nonlinear inequality we derived in Section 3.2 at $(\bar{y}_1, \bar{y}_2, \bar{t}_1, \bar{t}_2, \bar{z}_1, \bar{z}_2)$.

$$\text{(Ineq 2)} \rightarrow 2r^2\bar{y}_2 y_2 - \bar{z}_2 t_1 - C_1 \bar{z}_2 z_1 - (C_1 \bar{z}_1 + \bar{t}_1) z_2 - r^2 \bar{y}_2^2 + \bar{t}_1 \bar{z}_2 + C_1 \bar{z}_1 \bar{z}_2 \leqslant 0$$

$$\text{(Ineq 3)} \rightarrow (2a_{11}^2 \bar{y}_1 + 2a_{11}a_{12}\bar{y}_2) y_1 + (2a_{11}a_{12}\bar{y}_1 + 2a_{12}^2 \bar{y}_2) y2 - a_{11}^2 \bar{z}_1 t_1$$
$$- (a_{11}^2 \bar{t}_1 + C_2 \bar{z}_2) z_1 - C_2 \bar{z}_1 z_2 - (a_{11}\bar{y}_1 + a_{12}\bar{y}_2)^2 + a_{11}^2 \bar{t}_1 \bar{z}_1 + C_2 \bar{z}_1 \bar{z}_2 \leqslant 0$$

$$\text{(Ineq 4)} \rightarrow 2\bar{y}_1 y_1 - (\bar{z}_1 + \bar{z}_2) t_1 - \bar{t}_1 z_1 - \bar{t}_1 z_2 - \bar{y}_1^2 + \bar{t}_1 \bar{z}_1 + \bar{t}_1 \bar{z}_2 \leqslant 0$$

Table 3.1 summarizes the conditions under which these inequalities are equivalent to the linear valid inequalities for $P(A, B)$ derived in Chapter 2.

3.3 Measuring the Impact of Inequalities

In mixed integer linear programming, it is known that the convex hull of a mixed integer linear set defined by rational inequalities is a polyhedron [44]. Theorem 3.2 established that $\text{conv}(S)$ is not a polyhedron, so we have no nice way of demonstrating that the inequalities we derived are necessary in the description of $\text{conv}(S)$, like demonstrating facets of a polyhedron. The theory of mixed integer nonlinear sets is not fully-developed, but Kilinç-Karzan [39] has done some work to extend the notion of the necessity of inequalities in the description of mixed-integer nonlinear sets. In this section, we take a different approach to providing evidence of the usefulness of the conic inequalities (Ineq 1)-(Ineq 4). First, we provide a short empirical experiment to demonstrate that adding the inequalities to the continuous relaxation $\mathcal{R}(S)$ does *not* provide a complete description of $\text{conv}(S)$, i.e. $\mathcal{T}(S) \neq \text{conv}(S)$. The next two subsections provide evidence that the inequalities (Ineq 1)-(Ineq 4) may be useful by characterizing which fractional extreme points of the continuous relaxation are cut off by each of the inequalities and by characterizing at which points of $\text{conv}(S)$ each of the inequalities is tight. The section also contains results of a "shooting experiment," wherein we estimate the size of each surface of $\mathcal{T}(S)$.

3.3.1 Proportion of Fractional Extreme Points

Our first question about the strength of the relaxation $\mathcal{T}(S)$ is whether it is the complete characterization of $\text{conv}(S)$, i.e., whether the conic inequalities (Ineq 1) - (Ineq 4) are the only nontrivial inequalities necessary in the description of $\text{conv}(S)$. We show that $\mathcal{T}(S) \neq \text{conv}(S)$ by estimating the probability that $\mathcal{T}(S)$ has fractional extreme points through computational experiments. As Theorem 3.2 has shown that all extreme points of $\text{conv}(S)$ are integral in the z_1, z_2 components, this probability should be zero if $\mathcal{T}(S) = \text{conv}(S)$.

The extreme points of $\mathcal{T}(S)$ were obtained by solving

$$\min_{(y,t,z)\in\mathcal{T}(S)} \alpha^T(y,t,z) \tag{3.19}$$

where $\alpha \in \mathbb{R}^6$. The coefficients $\alpha_1, \alpha_2, \alpha_5, \alpha_6$ were drawn from a uniform distribution in $[-1, 1]$, and the coefficients for the t variables α_3, α_4 were drawn from a uniform distribution in $[0, 1]$ since they need to be non-negative to ensure that (3.19) is not unbounded.

Thirty random instances were generated and 50 random linear objectives were tried on each instance of S. The same experiment was repeated using $\mathcal{R}(S)$ for comparison. Table 3.2 summarizes the proportion of fractional extreme points out of all extreme points obtained by solving (3.19) for different cases.

Instance Classes	$\mathcal{R}(S)$	$\mathcal{T}(S)$
A ($a_{12} < 0$)	0.064	0.084
B ($0 \leqslant a_{12} \leqslant a_{22}$)	0.052	0.14
C ($a_{22} < a_{12}$)	0.048	0.092

Table 3.2: Probability of having fractional extreme points

The probabilities associated with $\mathcal{T}(S)$ are not zero, which implies $\mathcal{T}(S) \neq \text{conv}(S)$. There does not seem to be apparent difference between the classes of instances in the proportion of fractional extreme points. Note that $\mathcal{T}(S)$ appears to be more likely to have fractional extreme point than $\mathcal{R}(S)$, which means the valid inequalities we applied cut off extreme points while generating more fractional extreme points.

3.3.2 Extreme points cut off by the inequalities

We investigate the usefulness of the inequalities by examining which of the points of $\text{ext}(\mathcal{R}(S))$ are cut off by each of the inequalities introduced in Section 3.2. Table 3.3 demonstrates whether the fractional extreme points in each of the families (3.5), (3.6), and (3.7) are cut off by inequalities (Ineq 1) - (Ineq 4), or the condition under which they are cut off. In the table, it is assumed that $0 < a_{11}\alpha + a_{12}\beta < 1$ and $0 < a_{22}\beta < 1$ so that points represented by (3.5) - (3.7) are fractional. The entries of the table means either the points are always cut off (O), are never cut off (X), or are cut off under some condition given in the equations referenced.

The analysis demonstrates that all 4 inequalities are able to eliminate some fractional extreme points of the simple relaxation $\mathcal{R}(S)$. Equation Ineq 1 eliminates all extreme points with fractional z_2 value.

3.3.3 Tightness of inequalities

Another property of inequalities that may point toward their usefulness in strengthening the continuous relaxation of a mixed integer nonlinear set S is whether or not the inequality is tight at extreme points of the convex hull of S. In Table 3.4 we examine for each class of extreme points of $\text{conv}(S)$ whether or not each inequality (Ineq 1) - (Ineq 4) holds with equality at that point. In the table, the notation A denotes that the inequality is always tight for all extreme points of the associated type. The notation X denotes that the inequality is never tight for that class of extreme points. If the inequality is tight for some of the extreme points in the class, then the conditions under which the extreme point is tight is given by the equation number listed in the table.

Inequality		Extreme Points		
		(3.5)	(3.6)	(3.7)
(Ineq 1)		O	X	O
(Ineq 2)	$a_{12} < 0$	(3.20)	X	(3.20)
	$0 < a_{12} \leqslant a_{22}$	X	(3.21)	(3.22)
	$a_{12} > a_{22}$	(3.23)	(3.24)	(3.25)
(Ineq 3)	$a_{12} < 0$	X	(3.26)	(3.26)
	$0 < a_{12} \leqslant a_{22}$	(3.27)	X	(3.28)
	$a_{12} > a_{22}$	(3.29)	(3.30)	(3.31)
(Ineq 4)		X	X	(3.32)

$$a_{22}\alpha^2 - (\frac{a_{12}}{a_{11}})^2\beta < 0 \tag{3.20}$$

$$(\frac{r}{a_{22}})^2(a_{11}\alpha + a_{12}\beta) + \alpha^2 - r^2\beta^2 < 0 \tag{3.21}$$

$$(\frac{r}{a_{22}})^2(a_{11}\alpha + a_{12}\beta) + \alpha^2 - \frac{r^2}{a_{22}}\beta^2 < 0 \tag{3.22}$$

$$\frac{r^2}{a_{22}^2} - (\frac{1}{a_{11}} - \frac{r}{a_{22}})^2 + \alpha^2 - (\frac{r}{a_{22}})^2\beta < 0 \tag{3.23}$$

$$(\frac{r^2}{a_{22}^2} - (\frac{1}{a_{11}} - \frac{r}{a_{22}})^2)(a_{11}\alpha + a_{12}\beta) + \alpha^2 - r^2\beta^2 < 0 \tag{3.24}$$

$$(\frac{r^2}{a_{22}^2} - (\frac{1}{a_{11}} - \frac{r}{a_{22}})^2)(a_{11}\alpha + a_{12}\beta) + \alpha^2 - \frac{r^2}{a_{22}}\beta < 0 \tag{3.25}$$

$$a_{11}^2\alpha^2 - a_{11}\alpha - a_{12}\beta < 0 \tag{3.26}$$

$$a_{22}\beta + a_{11}^2\alpha^2 - (a_{11}\alpha + a_{12}\beta)^2 < 0 \tag{3.27}$$

$$a_{11}^2\alpha^2 - a_{11}\alpha - a_{12}\beta + a_{22}\beta < 0 \tag{3.28}$$

$$(\frac{2a_{12}}{a_{22}} - (\frac{a_{12}}{a_{22}})^2)a_{22}\beta + a_{11}^2\alpha^2 - (a_{11}\alpha + a_{12}\beta)^2 < 0 \tag{3.29}$$

$$(\frac{2a_{12}}{a_{22}} - (\frac{a_{12}}{a_{22}})^2) + a_{11}^2\alpha^2 - (a_{11}\alpha + a_{12}\beta) < 0 \tag{3.30}$$

$$(\frac{2a_{12}}{a_{22}} - (\frac{a_{12}}{a_{22}})^2)a_{22}\beta + a_{11}^2\alpha^2 - (a_{11}\alpha + a_{12}\beta) < 0 \tag{3.31}$$

$$a_{11}\alpha + a_{12}\beta + a_{22}\beta < 1 \tag{3.32}$$

Table 3.3: Extreme points cut off

Inequality		Extreme Points			
		(3.8)	(3.9)	(3.10)	(3.11)
(Ineq 1)		A	A	A	A
(Ineq 2)	$a_{12} < 0$	A	A	(3.33)	(3.34)
	$0 < a_{12} \leqslant a_{22}$	A	A	(3.33)	(3.35)
	$a_{12} > a_{22}$	A	A	(3.33)	(3.36)
(Ineq 3)	$a_{12} < 0$	A	A	A	(3.33)
	$0 < a_{12} \leqslant a_{22}$	A	A	A	(3.37)
	$a_{12} > a_{22}$	A	A	A	(3.38)
(Ineq 4)		A	A	A	X

$$\beta = 0 \tag{3.33}$$

$$a_{11}\alpha + a_{12}\beta = 0 \text{ or } a_{11}\alpha - a_{12}\beta = 0 \tag{3.34}$$

$$\alpha = 0 \text{ and } \beta = \frac{1}{a_{22}} \tag{3.35}$$

$$\frac{1}{a_{11}^2}\left(\frac{2a_{12}}{a_{22}} - 1\right) + \alpha^2 = \left(\frac{a_{12}}{a_{11}}\right)^2 \tag{3.36}$$

$$-\left(\frac{a_{12}}{a_{22}}\right)^2 + \frac{2a_{12}}{a_{22}} + a_{11}^2\alpha^2 = (a_{11}\alpha + a_{12}\beta)^2 \tag{3.37}$$

$$\alpha = 0 \text{ and } \beta = \frac{1}{a_{12}} \tag{3.38}$$

Table 3.4: Tightness of inequalities at extreme points of $\text{conv}(S)$

3.3.4 Shooting Experiment

In this section, we present the result of shooting experiments for $\mathcal{T}(S)$ which is analogous to what we conduct for $\mathcal{T}(P(A, B))$ in Section 2.4.2. We would like to understand the importance of each of the surfaces in our relaxation $\mathcal{T}(S)$ by estimating the likelihood of hitting that surface first if traversing a random direction from a fixed point in the interior of $\text{conv}(S)$.

A shooting with ray $r \in \mathbb{R}^6$ originating from a point $\hat{p} = (\hat{y}_1, \hat{y}_2, \hat{t}_1, \hat{t}_2, \hat{z}_1, \hat{z}_2)$ corresponds to solving the following conic programming problem:

$$
\begin{aligned}
\max_{y,t,z,d} \quad & d \\
\text{s.t.} \quad & (y,t,z) = (\hat{p} + rd) \\
& (y,t,z) \in \mathcal{T}(S) \\
& d \in \mathbb{R}_+.
\end{aligned} \tag{3.39}
$$

By selecting $\hat{p}$ as a convex combination of the extreme points of $\text{conv}(S)$ (characterized in Theorem (3.2)), we can be sure that $\hat{p} \in \text{conv}(S)$. In an attempt to approximate the centroid of $\text{conv}(S)$, we consider all possible combinations of the lower and upper bounds on variables y_1, y_2 as the values of α, β in each of the 4 cases (3.8) - (3.11) to obtain the following 9 points:

$$
\begin{aligned}
p_1 &= (0,0,0,0,1,1) \\
p_2 &= (1/a_{11}, 0, (1/a_{11})^2, 0, 1, 1) \\
p_3 &= (-a_{12}/a_{11}a_{22}, 1/a_{22}, (a_{12}/a_{11}a_{22})^2, (1/a_{22})^2, 1, 1) \\
p_4 &= (1/a_{11} - a_{12}/a_{11}a_{22}, 1/a_{22}, (1/a_{11} - a_{12}/a_{11}a_{22})^2, (1/a_{22})^2, 1, 1) \\
\hline
p_5 &= (0,0,0,0,0,1) \\
p_6 &= (-a_{12}/a_{11}a_{22}, 1/a_{22}, (a_{12}/a_{11}a_{22})^2, (1/a_{22})^2, 0, 1) \\
\hline
p_7 &= (0,0,0,0,1,0) \\
p_8 &= (1/a_{11}, 0, (1/a_{11})^2, 0, 1, 0) \\
\hline
p_9 &= (0,0,0,0,0,0).
\end{aligned}
$$

As the starting point for our shooting experiment we use

$$
\hat{p} = \frac{1}{9}\sum_{j=1}^{9} p_j.
$$

We choose the ray $r \in \mathbb{R}^6$ to be uniformly distributed over an appropriate hemisphere in $r \in \mathbb{R}^6$. Since $\text{conv}(S)$ is unbounded in the directions $(0,0,1,0,0,0)$ and $(0,0,0,1,0,0)$, we are not interested in rays r that have positive components in the t directions ($r_3 > 0$ or $r_4 > 0$). We generate the direction by first generating a vector r from the multivariate normal distribution $r \sim \mathcal{N}(0, I)$, thus ensuring that r is uniformly distributed on the unit sphere in $\mathbb{R}^6$. If $r_3 > 0$ or $r_4 > 0$, we simply flip the sign for that component of the vector. For each random instance of S, the shooting problem (3.39) was solved 1,000 times using different directions r chosen in this manner.

Table (3.5) summarizes the result of the shooting experiment for 10 different sets $S(a_{11}, a_{12}, a_{22})$, characterized by the values of the coefficients (a_{11}, a_{12}, a_{22}). Entries in the table denote the number of times out of 1,000 the inequalities (Ineq 1) - (Ineq 4), and the original inequalities in the description of $\mathcal{R}(S)$ are tight at the optimal solution of shooting problem. For numerical purposes, we define the inequality $f(x) \leqslant 0$ to be *tight* at $\hat{x}$ if $\|f(\hat{x})\| < 10^{-6}$.

(a_{11}, a_{12}, a_{22})	(Ineq 1)	(Ineq 2)	(Ineq 3)	(Ineq 4)	$\mathcal{R}(S)$
$(5, -9, 6)$	838	2	7	0	153
$(9, -9, 4)$	560	37	0	0	403
$(5, -9, 3)$	668	9	0	0	323
$(7, 5, 5)$	616	0	0	0	384
$(8, 2, 9)$	662	0	0	0	338
$(3, 3, 9)$	949	0	0	0	51
$(1, 2, 8)$	963	0	0	0	37
$(8, 2, 6)$	446	0	0	0	554
$(5, 4, 1)$	171	0	2	1	826
$(3, 5, 1)$	254	1	0	0	745

Table 3.5: Shooting Experiment

The results of the experiment show convincingly that the most important nonlinear inequalities are the perspective inequalities (Ineq 1). However, each of the inequalities is sometimes useful in strengthening the set $\mathcal{R}(S)$. For all 10,000 instances of (3.39), the optimal solution value lies in $\text{conv}(S)$.

3.4 Computations

This section demonstrates how to utilize the useful valid inequalities from Section 3.2 to a general instance of 0-1 MIQCP. As in the previous Chapter, we use the 2×2 block diagonal decomposition strategy $Q = R + B$ introduced in Section 2.3. We compare the strength of the resulting formulation with the ones obtained by diagonal decomposition $Q = R + D$, $D = \text{diag}(d)$.

We first define the general relaxation formulation for the two methods. Let $\mathcal{F}$ denote the feasible region of (0-1 MIQCP-*l*):

$$\begin{aligned}\mathcal{F} = \{(\nu, x, z) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{B}^n \mid\ & \nu \geqslant x^T Q_0 x + c_0^T x, \\ & Ax + Hz \leqslant f, \\ & 0 \leqslant x_j \leqslant z_j\ \forall j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^p\}.\end{aligned}$$

We denote the formulation of a relaxation of $\mathcal{F}$ obtained by diagonal decomposition and applying quadratic perspective inequalities by $\mathcal{T}_D(\mathcal{F})$, the formulation obtained by block diagonal decomposition and applying inequalities (Ineq 1) - (Ineq 4) by $\mathcal{T}_B(\mathcal{F})$, respectively. The two formulations can be written as follows:

$$\begin{aligned}\mathcal{T}_D(\mathcal{F}) = \{(x, z, \nu, w) \in\ \mathbb{R}_+^n \times [0,1]^n \times \mathbb{R}_+^{1+n} \mid & \\ & \nu \geqslant x^T R x + d^T w, \\ & Ax + Hz \leqslant f, \\ & w_j z_j \geqslant x_j^2\ \forall j \in [n], \\ & x_j \leqslant z_j\ \forall j \in [n]\ \},\end{aligned}$$

$$\begin{aligned}\mathcal{T}_B(\mathcal{F}) = \{(x, y, z, \nu, t) \in\ \mathbb{R}_+^n \times \mathbb{R}^n \times [0,1]^n \times \mathbb{R}_+^{1+n} \mid & \\ & \nu \geqslant x^T R x + e^T t, \\ & Ax + Hz \leqslant f, \\ & t_j \geqslant y_j^2\ \forall j \in [n], \\ & (y_i, y_{i+1}) \in \mathcal{T}(S)\ \forall i \in [n], i \text{ odd }, \\ & L^T x = y, \\ & 0 \leqslant x_j \leqslant z_j\ \forall j \in [n]\ \}.\end{aligned}$$

3.4.1 Minimum Variance Portfolio Problem

To demonstrate the potential of the block diagonal extraction and the strong conic inequalities introduced in Section 3.2, we report the results of an experiment computing root node bounds for small instances of a minimum variance portfolio problem. The problem can be stated mathematically as

$$\begin{aligned} z^* := \min \quad & x^\mathsf{T} Q x \\ \text{s.t.} \quad & e^\mathsf{T} x = 1, \\ & e^\mathsf{T} z \leqslant K, \\ & \alpha^\mathsf{T} x \geqslant \rho, \\ & 0 \leqslant x_j \leqslant z_j \ j \in [n], \\ & x \in \mathbb{R}^n, z \in \mathbb{B}^n. \end{aligned} \tag{3.40}$$

We use the same 3 families of randomly generated instances from Section 2.5.1 where Q was written as

$$Q = A^\mathsf{T} A + \operatorname{diag}(\zeta) + \sum_{j=1}^{n/2} \gamma_j (e_{2j-1} + e_{2j})(e_{2j-1} + e_{2j})^\mathsf{T}.$$

Recall from Section 2.5.1 that each element of $A \in \mathbb{R}^{n \times n}$ were drawn from discrete uniform distribution within $[-5, 5]$ and each element of $\zeta \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}^{n/2}$ were drawn from continuous uniform distributions within ranges $[0, \bar{\zeta}]$ and $[0, \bar{\gamma}]$, respectively. As with the facet-defining inequalities for $\operatorname{conv}(P(A, B))$, we would expect that block extraction technique with the conic inequlaities (Ineq 1)-(Ineq 4) to perform best if Q is more "block"-diagonally dominant, i.e., when $\bar{\gamma}$ is large.

In each family, we solved 20 random instances for each size $n \in \{10, 20, 30, 40\}$ to compare the strength of different relaxations. For each instance, we set the cardinality constraints to have value $K = \lfloor 2n/5 \rfloor$. To precisely define the relaxations, let

$$P := \{(x, z) \in [0, 1]^n \times [0, 1]^n \mid e^\mathsf{T} x = 1, e^\mathsf{T} z \leqslant K, \alpha^\mathsf{T} x \geqslant \rho, x_j \leqslant z_j \forall j \in [n]\}$$

be the continuous relaxation of the feasible region of (3.40). The natural continuous relaxation obtains root objective value

$$z_R := \min_{x,z} \{x^\mathsf{T} Q x \mid (x, z) \in P\}.$$

To create a relaxation based on a diagonal matrix-splitting, we solve the SDP (1.23) to obtain an optimal solution d, and we let $Q = R + \operatorname{diag}(d)$. The strengthened relaxation obtains root objective value

$$z_D := \min_{x,z,\nu,t} \Big\{ \nu \mid \quad \nu \geqslant x^T R x + e^T t$$
$$(x,z) \in P$$
$$z_j t_j \geqslant d_j x_j^2 \ \forall j \in [n] \Big\}.$$

To create a relaxation based on a block-diagonal matrix splitting, we solve the SDP (2.14) to obtain an optimal solution solution B, and we let $Q = R + B$, with $B = LL^T$. The strengthened relaxation employing our new conic inequalities obtains root objective value

$$z_B := \min_{x,z,\nu,t,y} \Big\{ \nu \mid \quad \nu \geqslant x^T R x + e^T t$$
$$(x,z) \in P$$
$$L^T y = x$$
$$(y,t,z) \in \mathcal{T}(S(L_{jj}^{-T}, L_{j,j+1}^{-T}, L_{j+1,j+1}^{-T})) \ \forall j \in \{1,3,\cdots,n-1\} \Big\}.$$

Note the dependence of the feasible region, specifically the sets S, on the inverse of the Cholesky factor of B.

We also create a relaxation based on performing the same block-diagonal matrix splitting used to compute z_B, but we replace $\mathcal{T}(S)$ (our approximation of $\text{conv}(S)$), with $\text{conv}(S)$ using the extended formulation described in Theorem 1.6. Specifically, we compute the value

$$z_{\text{conv}} := \min_{x,z,\nu,t,y} \Big\{ \nu \mid \quad \nu \geqslant x^T R x + e^T t$$
$$(x,z) \in P$$
$$L^T y = x$$
$$(y,t,z) \in \text{conv}(S(L_{jj}^{-T}, L_{j,j+1}^{-T}, L_{j+1,j+1}^{-T})) \ \forall j \in \{1,3,\cdots,n-1\} \Big\}.$$

Relaxations are compared with respect to their root optimality gap

$$\text{Gap}(z_?) := \frac{z_? - z^*}{z^*},$$

where $z_?$ is one of z_R, z_D, z_B, or z_{conv} depending on which relaxation is being evaluated.

Tables 3.6, 3.7, and 3.8 show the average root optimality gaps for the four different relaxation methods over the 20 randomly generated instances of each problem size.

n	Gap(z_R)	Gap(z_D)	Gap(z_B)	Gap(z_{conv})
10	38.47	38.26	37.55	37.44
20	34.04	34.03	34.03	34.03
30	24.19	24.10	24.08	24.06
40	18.43	18.43	18.43	18.43

Table 3.6: Average Gap(%) for Min Variance Portfolio Problem with ($\bar{\zeta} = 0, \bar{\gamma} = 0$)

n	Gap(z_R)	Gap(z_D)	Gap(z_B)	Gap(z_{conv})
10	39.68	12.07	12.25	10.69
20	32.05	12.27	15.01	13.40
30	31.93	8.00	11.90	9.75
40	29.34	7.98	11.34	9.27

Table 3.7: Average Gap(%) for Min Variance Portfolio Problem ($\bar{\zeta} = 100, \bar{\gamma} = 0$)

n	Gap(z_R)	Gap(z_D)	Gap(z_B)	Gap(z_{conv})
10	24.72	20.33	13.20	11.80
20	19.99	17.22	12.41	11.76
30	18.84	17.36	11.98	11.53
40	17.27	16.07	12.72	12.39

Table 3.8: Average Gap(%) for Min Variance Portfolio Problem ($\bar{\zeta} = 0, \bar{\gamma} = 100$)

Table 3.6 shows that strengthening the relaxation using the perspective reformulation or the new conic inequalities appears to have little or no impact for the instances with $\bar{\zeta} = 0, \bar{\gamma} = 0$. However, Tables 3.7 and 3.8 show that there are significant root gap improvements for instances in the families where Q contains elements specifically added to the diagonal or off-diagonal. The diagonal extraction method works best for instances that are the most diagonally dominant ($\bar{\zeta} = 100, \bar{\gamma} = 0$). For these instances, we even observe Gap(z_D) $<$ Gap(z_{conv}). Since the 2×2 block extraction technique is a generalization of the diagonal extraction technique, we would expect the reverse. However, the final root gap depends on the extracted matrix D or B. We placed arbitrary weights ($s = 5, r = 2$) on extracting diagonal and off-diagonal elements in our extraction SDP (2.14) which cannot guarantee that Gap(z_D) $>$ Gap(z_{conv}). In Table 3.8, clearly the relaxations from block-diagonal extraction (z_B and z_{conv}) are the strongest. This meshes with our intuition, since we specifically studied and strengthened the sets arising from this reformulation. We are encouraged that the values for Gap(z_B) are close to Gap(z_{conv}). This indicates that the new conic inequalities (Ineq 1)-(Ineq 4) do a good job approximating $\text{conv}(S)$, at least in the direction of the objective function. Overall, we conclude that the block-extraction technique and proposed conic inequalities may be effective for instances where Q is close to block-diagonal.

3.5 Summary

In this chapter, we direct our attention back to the nonlinear set S. We begin by completely characterizing the set of extreme points for two important relaxations of S, the simple continuous relaxation $\mathcal{R}(S)$ and $\text{conv}(S)$. As an immediate consequence of this anlysis, it is established that $\text{conv}(S)$ is not a polyhedron. Using the similar techniques used in Chapter 2, 4 classes of valid quadratic inequalities for S are derived. They are all shown to be second-order cone representable, which allows us to use conic programming solvers to optimize over the strengthened relaxation. We show that the inequalities are closely related to the linear inequalities from Chapter 2 as we can exactly generate some of the facet of $P(A, B)$ by linearizing the quadratic inequalities at specific points. As we do not have a nice way of demonstrating if an inequality is necessary for describing a convex hull of a non-polyhedral set, we measure the impact of our inequalities through extreme point analysis and computations including shooting experiment. Specifically, we have characterized which extreme points of $\mathcal{R}(S)$ violate each inequality and at which extreme points of $\text{conv}(S)$ the inequalities are tight. The result of computational experiment implies that we still do not have the complete description of $\text{conv}(S)$. We conduct additional computation on the same instances of the minimum variance portfolio optimization problem from Chapter 2 to demonstrate the ability of the derived inequalities to improve bounds.

Chapter 4

Valid Inequalities for DC Transmission Switching

4.1 Introduction

Optimization problems arising from the design and operation of electric power systems have been a very active topic of research. The types of optimization problems include optimal power flow, unit commitment, optimal transmission switching, pricing, and others [47]. Many of these optimization problems involve the non-convex, nonlinear power flow equations that model the steady-state power flow of Alternating Current (AC).

The power flow network is typically depicted in a graph $G = (N, E)$ where the set of nodes N denote the set of buses and set of arcs E the lines between buses. Variables v_i and θ_i for each node $i \in N$ represent the voltage magnitude and phase angle of bus i, respectively. Variables x_{ij}, y_{ij} represent the active and reactive power flow on line (i, j) and p_i, q_i represent the active and reactive power at bus i. Parameters α_{ij} and g_{ij} for each arc $(i, j) \in E$ represent the susceptance and conductance of line (i, j).

The physical laws dictating the status of these components are written as:

$$\begin{aligned} x_{ij} &= g_{ij}v_i^2 - g_{ij}v_iv_j\cos(\theta_i - \theta_j) - \alpha_{ij}v_iv_j\sin(\theta_i - \theta_j) \\ y_{ij} &= -\alpha_{ij}v_i^2 + \alpha_{ij}v_iv_j\cos(\theta_i - \theta_j) - g_{ij}v_iv_j\sin(\theta_i - \theta_j) \end{aligned} \quad \forall (i, j) \in E \tag{4.1}$$

$$\begin{aligned} p_i &= \textstyle\sum_{j:(i,j)\in E} x_{ij} \\ q_i &= \textstyle\sum_{j:(i,j)\in E} y_{ij} \end{aligned} \quad \forall i \in N. \tag{4.2}$$

As a means of creating a computationally tractable approximation of the non-convex set defined by the power flow equations, the Direct Current (DC) power flow model is often used. The simplification is based on several assumptions [34]:

- The phase angle difference $\theta_i - \theta_j \approx 0$, which implies $\cos(\theta_i - \theta_j) \approx 1$ and $\sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$.
- The voltage magnitude $v_i \approx 1$.
- The relative power y_{ij} is ignored.

With these assumptions, the first two addends for the active power x_{ij} in (4.1) cancels. The resulting DC model reduces equations (4.1) from the AC power flow model to the linear equations

$$x_{ij} = -\alpha_{ij}(\theta_i - \theta_j)\ \forall (i,j) \in E. \tag{4.3}$$

Recall from Section 1.3.4 that for the purpose of describing the mathematical formulations for optimization problems arising from transmission line network, we define the relevant sets, variables, and parameters as follows:

Sets	$R \subset N$	: set of generator nodes
	$D \subset N$	: set of demand nodes
Variables	x_{ij}	: power flow on arc $(i,j) \in E$
	p_i	: amount of power produced from generator $i \in R$
	θ_i	: voltage angle at node $i \in N$
Parameters	c_i	: unit cost of production at generator $i \in R$
	d_i	: demand at node $i \in D$
	u_{ij}	: capacity on line$(i,j) \in E$
	p_i^{min}/p_i^{max}	: minimum / maximum production at generator $i \in R$.

A classical power network optimization problem is the economic dispatch problem where the objective is to determine the generator operation decision that minimizes the total generator cost while satisfying certain level of demands. With the DC power flow model, the economic dispatch problem (ED) for power network $G = (N, E)$ can be written as:

$$
\begin{aligned}
\min \quad & \sum_{i \in R} c_i p_i && \text{(ED)}\\
\text{s.t.} \quad & \sum_{j:(i,j)\in E} x_{ij} - \sum_{j:(j,i)\in E} x_{ij} = \begin{cases} p_i & \forall i \in R \\ d_i & \forall i \in D \\ 0 & \text{otherwise} \end{cases}, \\
& x_{ij} = \alpha_{ij}(\theta_i - \theta_j) \; \forall (i,j) \in E && (4.4)\\
& -u_{ij} \leqslant x_{ij} \leqslant u_{ij} \; \forall (i,j) \in E, && (4.5)\\
& p_i^{min} \leqslant p_i \leqslant p_i^{max} \; \forall i \in R, \\
& x_{ij} \in \mathbb{R} \; \forall (i,j) \in E, p_i \in \mathbb{R}_+, \theta_i \in \mathbb{R} \; \forall i \in N.
\end{aligned}
$$

In this problem, the only operation decision is how much power should be produced at each generater $i \in R$, represented by variable p_i. Once the decision is made, the flows x_{ij} on each line (i, j) are determined according to the DC power flow equation and subject to the line capacity U_{ij}.

By introducing the decision of switching some of the lines on or off, the efficiency of the power flow network can be improved. Binary variables z_{ij} are introduced to represent the switching decision, indicating whether the line (i, j) is used in the network ($z_{ij} = 1$) or not ($z_{ij} = 0$). If the line is switched off, then the flow on the line should be 0 and the DC flow equation should not be imposed on the line. This relationship can be modeled by modifying the constraints (4.4) and (4.5) as follows:

$$x_{ij} = \alpha_{ij} z_{ij}(\theta_i - \theta_j) \quad \forall (i,j) \in E, \tag{4.6}$$

$$-U_{ij} z_{ij} \leqslant x_{ij} \leqslant U_{ij} z_{ij} \quad \forall (i,j) \in E. \tag{4.7}$$

The bilinear equation (4.6) can be equivalently formulated using big-M approach as in the work by Fisher, O'Neil, and Ferris [26]. Their formulation of optimal power flow with optimal transmission switching (OPF-OTS) is of the form

$$
\begin{aligned}
\min \quad & \sum_{i \in R} c_i p_i && \text{(OPF-OTS)} \\
\text{s.t.} \quad & \sum_{j:(i,j)\in E} x_{ij} - \sum_{j:(j,i)\in E} x_{ij} = \begin{cases} p_i & \forall i \in R \\ d_i & \forall i \in D \\ 0 & \text{otherwise,} \end{cases} \\
& -U_{ij} z_{ij} \leqslant x_{ij} \leqslant U_{ij} z_{ij} \; \forall (i,j) \in E, \\
& p_i^{min} \leqslant p_i \leqslant p_i^{max} \; \forall i \in R, \\
& \alpha_{ij}(\theta_i - \theta_j) - x_{ij} + M(1 - z_{ij}) \geqslant 0 \; \forall (i,j) \in E, && (4.8) \\
& \alpha_{ij}(\theta_i - \theta_j) - x_{ij} - M(1 - z_{ij}) \leqslant 0 \; \forall (i,j) \in E, && (4.9) \\
& \sum_{(i,j)\in E} z_{ij} \leqslant \kappa, \\
& x_{ij} \in \mathbb{R}, z_{ij} \in \mathbb{B} \; \forall (i,j) \in E, p_i \in \mathbb{R}_+, \theta_i \in \mathbb{R} \; \forall i \in N.
\end{aligned}
$$

The objective of this problem is to find the optimal generating and switching decisions that minimize the total power generation cost satisfying demand where the power flow is governed by the DC flow model (1.8). If line (i,j) is switched on, ($z_{ij} = 1$) equations (4.8) and (4.9) dictate that the DC power flow constraint (4.4) is imposed on this line. Otherwise if $z_{ij} = 0$, we need the scalar M to be sufficiently large so that

$$-M \leqslant \alpha_{ij}(\theta_i - \theta_j) \leqslant M$$

is valid for the feasible region of (ED).

We propose to consider the set defined by a subset of the constraints for a subgraph of G. Specifically, we extract a directed cycle from the network and consider only the DC powerflow equations (4.6) and the bounds on flow on each arc (4.7), ignoring the flow balance equation and the generator capacity constraints. As in our approach for 0-1 MIQCP from Chapter 2 and Chapter 3, we aim to improve the solution process for optimization problems in power flow network such as (OPF-OTS) that contain this relaxation as a substructure.

We repeat the definition of the relaxation of focus from Section 1.2

$$
\begin{aligned}
\mathcal{C} = \Big\{ (x, \theta, z) \in \mathbb{R}^{2n} \times \{0,1\}^n \;\mid\; & -u_{ij} z_{ij} \leqslant x_{ij} \leqslant u_{ij} z_{ij} \; \forall (i,j) \in C, \\
& z_{ij}(\theta_i - \theta_j) = x_{ij} \; \forall (i,j) \in C \Big\}, \qquad (4.10)
\end{aligned}
$$

where (V, C) is a directed cycle with $V = [n]$, $C = \{(i, i+1) \mid \forall i \in [n-1]\} \cup \{(n, 1)\}$. As $\mathcal{C}$ is the union of 2^n polyhedra, clconv$(\mathcal{C})$ is a polyhedron by Theorem 1.5 due to Balas [6] from Section 1.5.4. For any additive function $\mathcal{M}$ of subset of arcs $S \subseteq C$, we define $\mathcal{M}(S) := \sum_{e \in S} M_e$. A class of strong valid inequalities for $\mathcal{C}$ are introduced in Section 4.2 that are also shown to be facet-defining. We formulate and study the separation problem for these inequalities in Section 4.3 and use the result to design heuristics for separation. Computational result on the DC switching problem is reported in Section 4.4.

4.2 Valid inequalities for clconv($\mathcal{C}$)

In this section, we present a series of lemmata and theorems to characterize the extreme rays of $\mathcal{C}$, introduce strong valid inequalities for $\mathcal{C}$, and prove that these inequalities are facet-defining for the closure of its convex hull clconv$(\mathcal{C})$.

Lemma 4.1. *The extreme rays of* clconv$(\mathcal{C})$ *are* $(\mathbf{0}, e_i, \mathbf{0}),\ (\mathbf{0}, -e_i, \mathbf{0})\ \forall i \in [n]$.

Proof. Consider two points in $\mathcal{C}$; $(\bar{x}, \bar{\theta}, \bar{z})$ and $(\mathbf{0}, \frac{\alpha}{1-\lambda} e_i, \mathbf{0})$ for some $\alpha \in \mathbb{R}_+,\ i \in [n]$ and $\lambda \in (0, 1)$. A convex combination of these points,

$$\lambda(\bar{x}, \bar{\theta}, \bar{z}) + (1-\lambda)(\mathbf{0}, \frac{\alpha}{1-\lambda} e_i, \mathbf{0}) = (\lambda\bar{x}, \lambda\bar{\theta} + \alpha e_i, \lambda\bar{z})$$

is contained in conv$(\mathcal{C})$. By constructing a sequence of scalars $\{\lambda_j\}_{j=1}^{\infty}$ such that $\lambda_j \in (0, 1)\ \forall j$ and $\lim_{j\to\infty} \lambda_j = 1$, (e.g., $\lambda_j = 1 - \frac{1}{j}$) we can construct a sequence of points contained in conv$(\mathcal{C})$ of the form $\{(\lambda_j\bar{x}, \lambda_j\bar{\theta} + \alpha e_i, \lambda_j\bar{z})\}_{j=1}^{\infty}$. Therefore, the limit of this sequence, $(\bar{x}, \bar{\theta} + \alpha e_i, \bar{z})$, is contained in clconv$(\mathcal{C})$. We have established that for any point $(\bar{x}, \bar{\theta}, \bar{z}) \in \mathcal{C}$,

$$(\bar{x}, \bar{\theta}, \bar{z}) + \alpha(\mathbf{0}, e_i, \mathbf{0}) \in \text{clconv}(\mathcal{C})\ \forall i \in [n], \forall \alpha \in \mathbb{R}_+.$$

A similar argument shows that $(\mathbf{0}, -e_i, \mathbf{0})$ is also a ray of clconv$(\mathcal{C})$ $\forall i \in [n]$. □

Corollary 4.1. *Any inequality valid for* clconv$(\mathcal{C})$ *of the form*

$$\sum_{a \in C} \lambda_a x_a + \sum_{j \in [n]} \gamma_j \theta_j + \sum_{a \in C} \mu_a z_a \leqslant D, \tag{4.11}$$

has $\gamma_j = 0\ \forall j \in [n]$.

Proof. Suppose that $\gamma_i \neq 0$ for some $i \in [n]$ and let $(\bar{x}, \bar{\theta}, \bar{z}) \in \text{clconv}(\mathcal{C})$ satisfy (4.11). Then we can find a point of the form $(\bar{x}, \bar{\theta} + \alpha(\mathbf{0}, e_i, \mathbf{0}), \bar{z})$ or $(\bar{x}, \bar{\theta} + \alpha(\mathbf{0}, -e_i, \mathbf{0}), \bar{z})$ which is also contained in $\text{clconv}(\mathcal{C})$ by Lemma 4.1, but violates (4.11) for α sufficiently large. □

The following theorem states the main result of this section.

Theorem 4.2. *For* $S \subseteq C$ *such that* $u(C \setminus S) < u(S)$*, the inequalities*

$$x(S) + \sum_{a \in C} \beta_a^S z_a \leqslant b^S, \tag{4.12}$$

$$-x(S) + \sum_{a \in C} \beta_a^S z_a \leqslant b^S \tag{4.13}$$

are valid for $\mathcal{C}$*, where*

$$\beta_a^S = u(S \setminus a) - u(C \setminus S) \quad \forall a \in C, \tag{4.14}$$

$$b^S = (n-1)(2u(S) - u(C)). \tag{4.15}$$

Proof. Let $T \subseteq C$ be a subset of variables that are fixed to 0. The inequalities (4.12) and (4.13) are valid if

$$x(S \setminus T) \leqslant b^S - \beta^S(C \setminus T) \quad \forall T \subseteq C, \tag{4.16}$$

$$-x(S \setminus T) \leqslant b^S - \beta^S(C \setminus T) \quad \forall T \subseteq C \tag{4.17}$$

hold, respectively. In the case $T = \emptyset$, $x(C) = \sum_{(i,j) \in C} (\theta_i - \theta_j) = 0$, so that the left hand side of (4.16) is $x(S) = -x(C \setminus S)$ and the right hand side of (4.16) is

$$\begin{aligned} b^S - \beta^S(C) &= (n-1)(2u(S) - u(C)) - \sum_{a \in C} (u(S \setminus a) - u(C \setminus S)) \\ &= (n-1)(2u(S) - u(C)) - (n-1)u(S) + n(u(C) - u(S)) \\ &= u(C \setminus S) \end{aligned}$$

which proves the validity of (4.12). Similarly, when $T = \emptyset$, $x(C) = x(S) + x(C \setminus S) = 0$ implies that the left hand side of the the inequality (4.13) is $-x(S) = x(C \setminus S)$. Thus (4.13) reduces to $x(C \setminus S) \leqslant u(C \setminus S)$ which is valid.

We can show that the inequality is valid for $|T| \geqslant 1$ by examining the right hand side of (4.16) and (4.17):

$$\begin{aligned}
b^S - \beta^S(C \setminus T) =& (n-1)(2u(S) - u(C)) - \sum_{a \in C \setminus T} u(S \setminus a) + \sum_{a \in C \setminus T} u(C \setminus S) \\
=& (n-1)(2u(S) - u(C)) - \sum_{a \in C} u(S \setminus a) \\
& + \sum_{a \in T} u(S \setminus a) + |C \setminus T|(u(C) - u(S)) \\
=& (n-1)(2u(S) - u(C)) - (n-1)u(S) \\
& + |T|u(S) - u(S \cap T) + (|T| - n)(u(S) - u(C)) \\
=& (n-1)(u(S) - u(C)) + |T|u(S) - u(S) \\
& + u(S \setminus T) + (|T| - n)(u(S) - u(C)) \\
=& (|T| - 1)(2u(S) - u(C)) + u(S \setminus T)
\end{aligned}$$

Since the assumption $u(C \setminus S) < u(S)$ implies $u(S) > \frac{1}{2}u(C)$, this shows that $b^S - \beta^S(C \setminus T) \geqslant u(S \setminus T)\ \forall T : |T| \geqslant 1$ and both (4.16), (4.17) hold.

□

Theorem 4.3. *The inequalities* (4.12), (4.13) *for any* $S \subseteq C$ *satisfying* $u(C \setminus S) < u(S)$ *are facet-defining for set* $\mathcal{C}$.

Proof. The proof we present here addresses only (4.12). A very similar argument can establish that (4.13) is facet-defining as well. Let F denote the face of $\mathcal{C}$ defined by (4.12) and G a face defined by an arbitrary valid inequality for $\mathcal{C}$.

$$\begin{aligned}
F :=& \Big\{(x, \theta, z) \in \mathcal{C} \mid x(S) + \sum_{a \in C} \beta^S_a z_a = b^S\Big\} \\
G :=& \Big\{(x, \theta, z) \in \mathbb{R}^3 \mid \sum_{a \in C} \lambda_a x_a + \sum_{j \in [n]} \gamma_j \theta_j + \sum_{a \in C} \mu_a z_a = D\Big\}.
\end{aligned}$$

To prove the theorem, we will show that if $F \subseteq G$, then the coefficients defining G, $(\lambda, \gamma, \mu, D)$, must be a scalar multiple of the coefficients defining F, $(\sum_{i \in S} e_i, \mathbf{0}, \beta^S, b^S)$. Note that it follows from Theorem 4.1 that $\gamma = \mathbf{0}$.

First observe that a point in F satisfies the following conditions.

$$x(S) = u(C \setminus S) \text{ if } z_a = 1\ \forall a \in C \tag{4.18}$$

$$x(S) = u(S \setminus \hat{a}) \text{ if } z_a = 1\ \forall a \in C \setminus \hat{a}, z_{\hat{a}} = 0. \tag{4.19}$$

If $z_a = 1\ \forall a \in C$,

$$\begin{aligned} x(S) &= b^S - \sum_{a\in C} \beta_a^S z_a \\ \Leftrightarrow x(S) &= (n-1)(2u(S) - u(C)) - \sum_{a\in C} (u(S \setminus a) - u(C \setminus S)) \\ &= (n-1)(2u(S) - u(C)) - (n-1)u(S) + nu(C \setminus S) \\ &= u(C \setminus S). \end{aligned}$$

Note that if $x(S) = u(C \setminus S)$, then it is implied that $x_a = -u_a\ \forall a \in C \setminus S$ in order for $(x, \theta, z) \in \mathcal{C}$ as it needs to satisfy $x(C) = 0$.

If $z_a = 1\ \forall a \in C \setminus \hat{a}$ and $z_{\hat{a}} = 0$, we consider whether or not $\hat{a} \in S$. If $\hat{a} \in S$,

$$\begin{aligned} x(S) &= b^S - \sum_{a\in C} \beta_a^S z_a \\ \Leftrightarrow \sum_{a\in S\setminus \hat{a}} x_a &= (n-1)(2u(S) - u(C)) - \sum_{a\in C\setminus \hat{a}} (u(S \setminus a) - u(C \setminus S)) \\ &= (n-1)(2u(S) - u(C)) - (n-1)u(S) + u(S) - u_{\hat{a}} - (n-1)(u(C \setminus S)) \\ &= u(S) - u_{\hat{a}} \end{aligned}$$

otherwise if $\hat{a} \in C \setminus S$,

$$\begin{aligned} x(S) &= b^S - \sum_{a\in C} \beta_a^S z_a \\ \Leftrightarrow x(S) &= (n-1)(2u(S) - u(C)) - \sum_{a\in C\setminus \hat{a}} (u(S \setminus a) - u(C \setminus S)) \\ &= (n-1)(2u(S) - u(C)) - (n-1)u(S) + u(S) - (n-1)(u(C \setminus S)) \\ &= u(S). \end{aligned}$$

Define a parameter $\rho := \frac{u(C\setminus S)}{u(S)} \in [0, 1)$. We consider the following points in F satisfying conditions (4.18) and (4.19). A point $\chi = (\bar{x}, \bar{\theta}, \bar{z})$ satisfying

$$\bar{x}_a = \begin{cases} \rho u_a & \text{if } a \in S \\ -u_a & \text{if } a \in C \setminus S \end{cases}$$
$$\bar{\theta}_i = \begin{cases} 0 & \text{if } i = 1 \\ -\bar{x}_{12} & \text{if } i = 2 \\ \bar{\theta}_{i-1} - \bar{x}_{i-1,i} & \text{if } 2 < i \leqslant n \end{cases}$$
$$\bar{z}_a = 1 \; \forall a \in C,$$

and a class of n points $\xi^{\hat{a}} = (\hat{x}, \hat{\theta}, \hat{z})$ satisfying

$$\hat{x}_a = \begin{cases} u_a & \text{if } a \in C \setminus \hat{a} \\ 0 & \text{if } a = \hat{a} \end{cases}$$
$$\hat{\theta}_i = \begin{cases} 0 & \text{if } (i, i+1) = \hat{a} \\ \sum_{a \in C} \hat{x}_a & \text{if } (i-1, i) = \hat{a} \\ \hat{\theta}_{i-1} - \hat{x}_{i-1,i} & \text{otherwise} \end{cases}$$
$$\hat{z}_a = \begin{cases} 0 & \text{if } i = \hat{a} \\ 1 & \text{if } a \in C \setminus \hat{a}. \end{cases}$$

We first show that $\lambda_a = 0 \, \forall a \in C \setminus S$. For any $k \in C \setminus S$, we can find a point $\xi^{\hat{a}} - (\epsilon e_k, \Delta, \mathbf{0}) \in F$ for some $\hat{a} \neq k$, $\epsilon \in \mathbb{R}$ sufficiently small and $\Delta \in \mathbb{R}^n$. By the assumption $F \subseteq G$, both $\xi^{\hat{a}}$ and $\xi^{\hat{a}} - (\epsilon e_k, \Delta, \mathbf{0}) \in G$, which implies

$$\sum_{a \in C} \lambda_a u_a - \lambda_{\hat{a}} u_{\hat{a}} + \sum_{a \in C} \mu_a - \mu_{\hat{a}} = D \tag{4.20}$$
$$\sum_{a \in C} \lambda_a u_a - \lambda_{\hat{a}} u_{\hat{a}} - \lambda_k \epsilon + \sum_{a \in C} \mu_a - \mu_{\hat{a}} = D. \tag{4.21}$$

Subtracting the equations, (4.20) − (4.21) gives $\lambda_k \epsilon = 0$. Since this relationship holds for any $k \in C \setminus S$, we have

$$\lambda_a = 0 \; \forall a \in C \setminus S. \tag{4.22}$$

Next we will show that λ_a are equal $\forall a \in S$. If $|S| = 1$, there is nothing to show. Otherwise, for any pair of arcs $k, l \in S$, we can find a point $\chi - (\epsilon e_k - \epsilon e_l, \Delta, \mathbf{0}) \in F$ for $\epsilon \in \mathbb{R}$ sufficiently small and some $\Delta \in \mathbb{R}^n$ and obtain the equations

$$\sum_{a \in C} \lambda_a \bar{x}_a + \sum_{a \in C} \mu_a = D \tag{4.23}$$

$$\sum_{a \in C} \lambda_a \bar{x}_a - \lambda_k \epsilon + \lambda_l \epsilon + \sum_{a \in C} \mu_a = D. \tag{4.24}$$

Subtracting the equations, (4.23) − (4.24) gives $\lambda_k \epsilon - \lambda_l \epsilon = 0$. Since $k, l \in S$ were arbitrary, every coefficients $\lambda_a,\ a \in S$ are equal to each other. For some scalar Λ, let

$$\lambda_a = \Lambda \ \forall a \in S. \tag{4.25}$$

The remainder of the proof shows that μ_a is a multiple of β_a^S defined in (4.14), and D is a multiple of b^S defined in (4.15). For any pair of arcs $k, l \in C$, $\xi^k, \xi^l \in F \subseteq G$ results in the equations

$$\sum_{a \in C} \lambda_a u_a - \lambda_k u_k + \sum_{a \in C} \mu_a - \mu_k = D \tag{4.26}$$

$$\sum_{a \in C} \lambda_a u_a - \lambda_l u_l + \sum_{a \in C} \mu_a - \mu_l = D \tag{4.27}$$

and the subtraction (4.26) − (4.27) implies $\lambda_k - \lambda_l + \mu_k - \mu_l = 0$. Due to (4.22) and (4.25), this can be written as the following 4 equations depending on whether or not $k, l \in S$:

$$\Lambda(u_k - u_l) = -(\mu_k - \mu_l), \quad k, l \in S \tag{4.28}$$

$$\Lambda u_k = -(\mu_k - \mu_l), \quad k \in S,\ l \in C \setminus S \tag{4.29}$$

$$-\Lambda u_l = -(\mu_k - \mu_l), \quad k \in C \setminus S, l \in S \tag{4.30}$$

$$0 = -(\mu_k - \mu_l), \quad k, l \in C \setminus S \tag{4.31}$$

Adding (4.28) and (4.29) for all choices of $l \in C$, and for a specific $k \in S$, we obtain

$$n\Lambda u_k - \Lambda u(S) = -n\mu_k + \sum_{a \in C} \mu_a \ \forall k \in S. \tag{4.32}$$

Similarly, we add (4.30) and (4.31) for all choices of $a_l \in C$ for a specific $k \in C \setminus S$ to obtain

$$-\Lambda u(S) = -n\mu_k + \sum_{a \in C} \mu_a \ \forall k \in C \setminus S. \tag{4.33}$$

The equations (4.32) and (4.33) provide expressions for μ_k $\forall k \in C$ in terms of $\Lambda, u(S), u_k$, and $\sum_{a\in C} \mu_a$. To determine μ_a, we first obtain the description for $\sum_{a\in C} \mu_a$ by considering the following equations. First we obtain by adding (4.20) for all $\hat{a} \in C$ and substituting λ with $\Lambda \sum_{a\in S} e_i$,

$$(n-1)\Lambda u(S) + (n-1) \sum_{a\in C} \mu_a = nD. \tag{4.34}$$

By substituting λ with $\Lambda \sum_{a\in S} e_i$ and ρ with $\frac{u(C\setminus S)}{u(S)}$ in (4.23), we obtain

$$\Lambda \rho u(S) + \sum_{a\in C} \mu_a \tag{4.35}$$

$$= \Lambda u(C \setminus S) + \sum_{a\in C} \mu_a = D. \tag{4.36}$$

The subtraction $n \times$ (4.36) $-$ (4.34) gives

$$\begin{aligned}
\sum_{a\in C} \mu_a =& (n-1)\Lambda u(S) - n\Lambda u(C \setminus S) \\
=& (n-1)\Lambda u(S) - n\Lambda(u(C) - u(S)) \\
=& \Lambda((2n-1)u(S) - nu(C)).
\end{aligned} \tag{4.37}$$

Plugging (4.37) into (4.32) results in

$$\begin{aligned}
\mu_k =& \frac{1}{n}(\sum_{a\in C} \mu_a + \Lambda u(S) - n\Lambda u_k) \\
=& \frac{1}{n}\Lambda(2nu(S) - nu(C) - nu_k) \\
=& \Lambda(u(S \setminus k) - u(C \setminus S)) \; \forall k \in S,
\end{aligned}$$

and plugging (4.37) into (4.33) gives

$$\begin{aligned}
\mu_k =& \frac{1}{n}(\sum_{a\in C} \mu_a + \Lambda u(S)) \\
=& \frac{1}{n}\Lambda(2nu(S) - nu(C)) \\
=& \Lambda(u(S \setminus k) - u(C \setminus S)) \; \forall k \in C \setminus S.
\end{aligned}$$

So we have shown that $\mu_a = \Lambda\beta_a^S \; \forall a \in C$. Finally, plugging (4.37) into (4.36), we obtain

$$\begin{aligned} D =& \Lambda u(C \setminus S) + \sum_{a \in C} \mu_a \\ =& \Lambda(u(C) - u(S) + (2n-1)u(S) - nu(C)) \\ =& \Lambda(n-1)(2u(S) - u(C)), \end{aligned}$$

which establishes $D = \Lambda b^S$. We have shown that if $F \subseteq G$, then

$$(\lambda, \gamma, \mu, D) = \Lambda(\sum_{i \in S} e_i, \mathbf{0}, \beta^S, b^S) \text{ for some } \Lambda \in \mathbb{R},$$

which establishes that (4.12) is facet-defining for $\text{clconv}(\mathcal{C})$. □

In a collaborative work with Kocuk, Dey, and Sun, we also have shown that the inequalities (4.12), (4.13) characterize the convex hull of $\text{proj}_{(x,z)}(\mathcal{C})$.

Theorem 4.4. *[40]*

$$\text{conv}(\text{proj}_{(x,z)}(\mathcal{C})) = \left\{(x,z) \mid (4.12), (4.13), -u_{ij}z_{ij} \leqslant x_{ij} \leqslant u_{ij}z_{ij}, z_{ij} \leqslant 1 \; \forall (i,j) \in C\right\}$$

For proof of the theorem and more discussion on complexity, separation, and some computational result, we refer the readers to Kocuk et. al. [40]

4.3 The Separation Problem

Given an arbitrary instance of (OPF-OTS) based on network $G = (N, E)$, it may be possible to generate exponentially many inequalities of the form (4.12) and (4.13) since the number of cycles in G may be exponentially large. To make the inequalities useful in the solution process, we need a procedure to generate specific inequalities that can separate a given fractional solution to a continuous relaxation for (OPF-OTS). The procedure comprises two different aspects: the first is to detect directed cycles in the underlying network G from which we construct relaxations $\mathcal{C}$ defined in (4.10), and the second is to identify the inequalities of the form (4.12) or (4.13) from each cycle (V, C) that are violated by the current solution. Given a fractional point $(\hat{x}, \hat{z})$, the separation problem for (4.12) is written as

$$\max_{C \in E: C \text{ is a cycle}} \left\{ \max_{S \subseteq C: 2u(S) \geqslant u(C)} \{\hat{x}(S) + \sum_{a \in C} \beta_a^S \hat{z}_a - b^S\} \right\}, \tag{4.38}$$

an optimization problem whose objective is to maximize the amount of violation by $(\hat{x}, \hat{z})$.

Instead of considering all directed cycles contained in G, we specify a characteristic of a cycle C that is useful in separating a fractional point $(\hat{x}, \hat{z})$. The following proposition provides a necessary condition for C to generate inequalities (4.12) or (4.13) violated by this point. Given a cycle C and a point $(\hat{x}, \hat{z})$, we first define a parameter $K_C(\hat{z})$ as

$$K_C(\hat{z}) := 1 - \sum_{a \in C} (1 - \hat{z}_a). \tag{4.39}$$

For each subset $S \subseteq C$ satisfying $2u(S) > u(C \setminus S)$, we denote by $\text{viol}(S)$ the amount of violation for inequality (4.12) by $(\hat{x}, \hat{z})$:

$$\text{viol}(S) := \hat{x}(S) + \sum_{a \in C} \beta_a^S \hat{z}_a - b^S. \tag{4.40}$$

Proposition 4.1. *Given* $(\hat{x}, \hat{z})$ *for a cycle* C*, if* $K_C(\hat{z}) \leqslant 0$*, then the inequalities* (4.12), (4.13) *are not violated.*

Proof. Consider the inequality (4.12) for some $S \subseteq C$ satisfying $u(S) - u(C \setminus S) > 0$. The amount of violation for (4.12) by $(\hat{x}, \hat{z})$ can be written using $K_C(\hat{z})$ by rearrangement:

$$\begin{aligned} \text{viol}(S) &= \sum_{a \in S} \hat{x}_a + \sum_{a \in C} (u(S \setminus a) - u(C \setminus S))\hat{z}_a - (n-1)(2u(S) - u(C)) \\ &= \sum_{a \in S} (\hat{x}_a - u_a \hat{z}_a) + \sum_{a \in C} (2u(S) - u(C))\hat{z}_a - (n-1)(2u(S) - u(C)) \\ &= \sum_{a \in S} (\hat{x}_a - u_a \hat{z}_a) + (2u(S) - u(C))K_C(\hat{z}). \end{aligned} \tag{4.41}$$

Since $\hat{x}_a - u_a \hat{z}_a \leqslant 0 \; \forall a \in C$ from feasibility and $2u(S) - u(C) = u(S) - u(C \setminus S) > 0$ by assumption, this value cannot be positive if $K_C(\hat{z}) \leqslant 0$. A similar argument establishes that this condition is also necessary for (4.13). □

This result allows us to select and consider only the cycles with $K_C(\hat{z}) > 0$, or equivalently, $\sum_{a \in C} \hat{z}_a > n - 1$.

Given a cycle C, and a solution $(\hat{x}, \hat{z}) \in \mathbb{R}_+^n \times [0, 1]^n$, the separation problem for inequalities (4.12) is equivalent to finding a subset of arcs S that maximizes the violation. This is the inner optimization problem from (4.38):

$$\max_{S \subseteq C : 2u(S) \geqslant u(C)} \{\hat{x}(S) + \sum_{a \in C} \beta_a^S \hat{z}_a - b^S\}. \tag{4.42}$$

We introduce binary variables $y \in \{0,1\}^n$ to represent the decision whether or not each arc $a \in C$ is contained in S and use $\text{viol}(S)$ rearranged as (4.41) to formulate (4.42) as the following Integer Program:

$$\begin{aligned} \max_{y \in \{0,1\}^n} \quad & \sum_{a \in C} (\hat{x}_a - u_a \hat{z}_a) y_a + (\sum_{a \in C} 2u_a y_a - u(C)) K_C(\hat{z}) \\ \text{s.t.} \quad & \sum_{a \in C} u_a y_a \geqslant \frac{1}{2} u(C). \end{aligned} \tag{4.43}$$

Let $\hat{v} = \hat{x}_a - u_a \hat{z}_a + 2u_a K_C(\hat{z})$. Then by rearrangement,the separation problem is written as a knapsack problem:

$$\nu = \max_{y \in \{0,1\}^n} \Big\{ \sum_{a \in C} \hat{v}_a y_a \;\mid\; \sum_{a \in C} u_a y_a \geqslant \frac{1}{2} u(C) \Big\} \tag{4.44}$$

If $\nu - u(C) K_C(\hat{z}) > 0$ and the optimal solution to (4.44) is y^*, then the inequality generated using $S = \{a \in C \mid y_a^* = 1\}$ is violated.

Due to the special structure of the knapsack problem (4.44), it is possible to describe its optimal solution in closed form. Define the set of arcs

$$S_C^* := \{a \in C \mid \hat{x}_a - u_a \hat{z}_a + 2u_a K_C(\hat{z}) > 0\}. \tag{4.45}$$

The following proposition shows that S_C^* is the only solution to the separation problem (4.42) that we should consider. For ease of presentation we define a parameter Δ for S as

$$\Delta(S) := u(S) - u(C \setminus S).$$

Proposition 4.2. *Assume* $K_C(\hat{z}) > 0$*. If there is any* $S \subseteq C$ *with* $\Delta(S) > 0$ *and* $\text{viol}(S) > 0$*, then the separation problem* (4.44) *is solved by* S_C^**.*

Proof. Let T^* denote the optimal solution of (4.42). First note by construction S_C^* contains every $a \in C$ that has a positive contribution to the objective $\text{viol}(S)$. Therefore, $T^* \supseteq S_C^*$ as otherwise we can find an alternative solution with improved objective by adding the element $a \in S_C^* \setminus T^*$ to T^*, which will also satisfy the constraint $u(S) \geqslant \frac{1}{2} u(C)$.

Suppose that $\Delta(S_C^*) > 0$, so that S_C^* is a feasible solution to (4.44). Then $T^* = S_C^*$ as adding any more element in $C \setminus S_C^*$ to S_C^* would decrease the objective value. Suppose that S_C^* is infeasible, i.e., $\Delta(S_C^*) \leqslant 0$. Then

$$\text{viol}(S_C^*) = \sum_{a \in S} (\hat{x}_a - u_a \hat{z}_a) + \Delta(S_C^*) K_C(\hat{z}) \leqslant 0$$

because the first term $\hat{x}_a \leqslant u_a \hat{z}_a$ is nonpositive to satisfy the line capacity constraint, and the second term is nonpositive by assumption. We have $T^* \supset S_C^*$, which implies $\text{viol}(T^*) \leqslant \text{viol}(S_C^*) \leqslant 0$ since elements $a \notin S_C^*$ contribute a non-positive amount to the objective. Thus in the case $\Delta(S_C^*) < 0$, there is no $S \subseteq C$ with $\Delta(S) > 0$ and $\text{viol}(S) > 0$. □

We have shown in Proposition 4.1 that $K_C(\hat{z}) > 0$ is a necessary condition for finding violated inequality from cycle C. Therefore, the result of Proposition 4.2 implies that a violated inequality exists for a given cycle C with $K_C(\hat{z}) > 0$ if and only if:

$$\sum_{a \in C} \Big(\max\{0, \hat{x}_a - u_a \hat{z}_a + 2u_a K_C(\hat{z})\} - u_a K_C(\hat{z}) \Big) > 0. \tag{4.46}$$

The separation reduces to a search for a cycle C satisfying this condition and $K_C(\hat{z}) > 0$.

Although the analysis we conduct here may provide useful ideas on how to efficiently find violated inequalities, it remains an open question to establish the complexity of exactly solving the separation problem defined in (4.38). We conjecture that it is a hard problem, as the separation problem appears to have similarities to the longest path problem.

4.4 Computation

4.4.1 Algorithm

In this section, we explain our algorithmic framework for applying the inequalities (4.12) and (4.13) as cuts in solving the DC transmission switching problem instances. Algorithm 4.1 provides a sketch of the basic framework we use to generate cuts. Description of some details of each step of the algorithm follows.

Let $\mathcal{F}$ denote the feasible region of (OPF-OTS). The first continuous relaxation we solve to obtain solution $(\hat{x}, \hat{z})$ and lower bound $\hat{L}$ is the natural relaxation $\mathcal{R}(\mathcal{F})$ obtained by replacing the constraints $z \in \{0,1\}^n$ with $z \in [0,1]^n$. Observe that in the description of $\mathcal{F}$, there is no incentive to keep $\hat{z}$ from taking the largest possible value, 1.0. This is a disadvantage since separating the solution $(\hat{x}, \hat{z})$ with an unnecessarily large value of $\hat{z}$ may not improve the lower bound of the

Algorithm 4.1: Cut Generation

Initialize: cutoff = true, $r = 0$.
while *cutoff = true and* $r < maxR$ **do**
  Solve a continuous relaxation to obtain solution $(\bar{x}, \bar{z})$ and the lower bound $\hat{L}$.
  Construct the collection $\mathcal{B}$ of directed cycles C satisfying $K_C(\bar{x}) > 0$.
  for *every cycle* $C \in \mathcal{B}$ **do**
    Find S^*_C defined in (4.45).
    if $u(S^*_C) > \frac{1}{2}u(C)$ **then**
      Generate cuts and add them to formulation.
    else
      cutoff = false.
    $r \leftarrow r + 1$.

optimal objective value at all. Our inequalities always have positive coefficients on at least some of the binary variables $z_{ij}, (i,j) \in C$ by definition of the coefficients β^S_{ij}. Therefore, the amount of violation by $(\hat{x}, \hat{z})$ will reduce if the values of the binary variables z_{ij} decrease for $(i,j) \in C$ such that $\beta^S_{ij} > 0$. If the solution $(\hat{x}', \hat{z}')$ to the updated formulation has $\hat{z}' \leqslant \hat{z}$ to satisfy the added cuts but has the same flow on the lines, ($\hat{x}' = \hat{x}$) the objective function value will remain the same. To avoid this problem, we obtain an alternative solution $(\bar{x}, \bar{z})$ by solving

$$(\bar{x}, \bar{z}) = \operatorname*{argmin}_{(x,z)} \Big\{ \sum_{(i,j) \in E} z_{ij} \mid \sum_{i \in R} c_i p_i = \hat{L}, (x, z, p, \theta) \in \mathcal{R}(\mathcal{F}) \Big\}. \tag{4.47}$$

We look for cycles with $K_C(\bar{z}) > 0$ instead of $K_C(\hat{z}) > 0$, hoping to find cycles that are more useful in improving the lower bound.

To find the cycles at each iteration within the *while* loop in Algorithm 4.1, a simple heuristic based on the Depth First Search (DFS) was implemented. We start the search from each node $i \in N$ satisfying $\sum_{(i,j) \in E} z_{ij} > 1 + 10^{-3}$, as otherwise the resulting cycle is not likely to satisfy the necessary condition for violation given in Proposition 4.1. In each round of cut generation, we collect up to 300 cycles. ($|\mathcal{B}| \leqslant 300$)

We iterate through the set of cycles collected to generate inequalities. For each cycle satisfying $u(S^*_C) > \frac{1}{2}u(C)$, we generate at least one inequality. The first inequality is generated by using S^*_C:

$$x(S^*_C) + \sum_{a \in C} \beta_a^{S^*_C} z_a \leqslant b^{S^*_C}.$$

The fact that S^*_C solves the separation subproblem (4.42) does not mean that it provides the only violated inequality from cycle C. For example, for any $a \in S^*_C$ satisfying

$$u(S^*_C \setminus a) > \frac{1}{2}u(C),$$

we can generate a violated inequality using $S^*_C \setminus a$. For any $a \in C \setminus S^*_C$, if

$$\text{viol}(S^*_C) + \hat{x}_a - u_a\hat{z}_a + 2u_a K_C(\hat{z}) > 0,$$

then $S^*_C \cup a$ also generates a violated inequality which is valid since $u(S^*_C) > \frac{1}{2}u(C)$ implies $u(S^*_C \cup a) > \frac{1}{2}u(C)$. We examine this possibility and add all the inequalities we find that are violated by the current solution $(\hat{x}, \hat{z})$ or the alternative solution $(\bar{x}, \bar{z})$.

In the following sections, we demonstrate the usefulness of our inequalities by applying the algorithm described above in the solution process for some instances of DC transmission switching problem. We use the Concert Technology libraries for C++ of CPLEX 12.5 [1] to solve the problems and implement our algorithm. In all the computational result we present, the cut generation algorithm described in Algorithm 4.1 was applied only at the root node within a Callback structure.

4.4.2 DC transmission switching

In this section, we solve the DC transmission switching problems on some variants of the 118-bus instance introduced by Blumsack [13], which has 118 buses ($|N| = 118$), 186 lines ($|E| = 186$), and 19 generators ($|R| = 19$). The instances from [13] were modified by generating a discrete random variable that follows a uniform distribution in the range $[0, 15]$ for each of the demand nodes and adding the random number to the demand. We generated 15 instances by this modification and solved them with and without applying our cut generation algorithm. Since using a Callback suppresses some default features of CPLEX (e.g., dynamic search) and we do not know its entire effects, we used an empty Callback that does not contain any computation for fair comparison. The problems were solved using traditional branch and cut strategy (instead of dynamic search) using a single thread with the time limit of 1 hour.

We first solved the instances with three types of Callbacks, the empty Callback, cut generation Callback up to 20 rounds, and cut generation Callback up to 200 rounds, with all default CPLEX cuts turned off. Table 4.1 summarizes the result for each case. The entries in each row represent the average across 15 instances for the initial optimality gap at the root node, the ending optimality gap at the root node, solution time, number of nodes explored, number of cuts added to the formulation, and number of unsolved instances out of 15.

	Empty Callback	Cut Callback (*maxR* = 20)	Cut Callback (*maxR*=200)
Init RG(%)	21.22	21.22	21.22
End RG(%)	21.22	17.26	17.25
Time(s)	1178.67	921.15	303.37
# Nodes	1118548	1062023	450812
# Cuts	-	516	2244
# Unsolved	1	0	0

Table 4.1: Impact of Inequalities for DC OTS without CPLEX Cuts

Using the cut Callback with *maxR* = 200 performed best with respect to all criteria. On the average, applying our cut generation algorithm helped close up to 3.97% of optimality gap at the root node. It significantly improved the speed of solution process, reducing the average solution time by 21.85 % when *maxR* = 20 and by 74.26 % when *maxR* = 200. The number of explored nodes was reduced by 5.05 % and 59.70 % by the two cut Callbacks. We were able to solve all 15 instances to provable optimality within 1 hour using the cuts, including the instance that was not solved within the time limit without the cuts. The remaining gap for the unsolved instance after the time limit with the empty Callback was 9.04 %.

4.5 Summary

In Chapter 4, we discuss a strong relaxation for the cycle relaxation $\mathcal{C}$ of the DC transmission switching problem. A class of valid linear inequalities were derived for $\mathcal{C}$. We know that $\text{clconv}(\mathcal{C})$ is a polyhedron unlike $\text{conv}(S)$, which allows us to use the theoretical tool to determine if our inequalities are necessary in the description of $\text{clconv}(\mathcal{C})$. We prove that the inequalities are in fact facet-defining for $\text{clconv}(\mathcal{C})$. In a collaborative work with Kocuk, Dey, and Sun, it was also shown that these inequalities are the only non-trivial inequalities in the description of the convex hull of a lower-dimensional set obtained by projecting out a subset of variables. We also discuss the separation problem for the inequalities, which involves two subproblems: finding directed cycles in a graph and finding violated inequalities from each cycle. Due to the characteristic of $\mathcal{C}$, we can provide a closed-form solution to the second subproblem formulated as a knapsack problem. Computational experiment on IEEE instances were conducted to demonstrate the usefulness of the derived inequalities.

Chapter 5

Conclusion

The main subjects of investigation in this thesis are two classes of mixed-integer quadratic sets that have the specific structure of binary indicator variables. The first set, denoted by S, is constructed by reformulating a low-dimensional relaxation for 0-1 MIQCP via a Cholesky factorization. The second set, denoted by $\mathcal{C}$, is a relaxation for the DC transmission switching problem based on a directed cycle extracted from the underlying transmission line network and Kirchoff's voltage law.

In Chapter 1, we define the sets S, $\mathcal{C}$ and introduce various applications where the sets appear as substructures. We briefly introduce the algorithmic framework within which our findings can be utilized to improve the efficiency of the solution process. A review of previous studies that help our investigation is provided. Each of the following three chapters of the thesis present the main result of our investigation for three different sets, a polyhedral outer-approximation of S, S, and $\mathcal{C}$, respectively.

Chapter 2 investigates the polyhedral outer-approximation of S, denoted by $P(A, B)$. This polyhedron is constructed by replacing the quadratic constraints in the description of S with their linearizations taken at a number of break points. We derive 5 classes of valid inequalities for $P(A, B)$ by using techniques such as lifting, variable substitution, constraint aggregation, and the perspective reformulation. It was shown that they are facet-defining under some mild conditions, and the importance of each facet was measured through a shooting experiment. Computational experiments provide evidence that these inequalities do not describe $\text{conv}(P(A, B))$. Although the inequalities are only applicable for low-dimensional relaxations, they can be helpful in a solution process for a problem of larger dimension. We decribe how to utilize them for each low-dimensional substructure of an optimization problem obtained by matrix decomposition, and demonstrate this framework in computations on the minimum variance portfolio optimization problems.

In Chapter 3, we direct our attention back to the nonlinear set S. We begin by completely characterizing the set of extreme points for two important relaxations of S, the simple continuous relaxation $\mathcal{R}(S)$ and $\text{conv}(S)$. As an immediate consequence of this anlysis, it is established that $\text{conv}(S)$ is not a polyhedron. Using similar techniques used in Chapter 2, 4 classes of valid quadratic inequalities for S are derived. They are all shown to be second-order cone representable, which allows us to use conic programming solvers to optimize over the strengthened relaxation. We show that the inequalities are closely related to the linear inequalities from Chapter 2 as we can exactly generate some of the facet of $P(A, B)$ by linearizing the quadratic inequalities at specific points. As we do not have a nice way of demonstrating if an inequality is necessary for describing a convex hull of a non-polyhedral set, we measure the impact of our inequalities through extreme point analysis and computations including shooting experiment. Specifically, we have characterized which extreme points of $\mathcal{R}(S)$ violate each inequality and at which extreme points of $\text{conv}(S)$ the inequalities are tight. The results of our computational experiments imply that we do not have the complete description of $\text{conv}(S)$. We conduct additional computation on the same instances of the minimum variance portfolio optimization problem from Chapter 2 to demonstrate the ability of the derived inequalities to improve bounds.

Finally, in Chapter 4, we discuss the cycle relaxation $\mathcal{C}$ of the DC transmission switching problem. A class of valid linear inequalities were derived for $\mathcal{C}$. We know that $\text{clconv}(\mathcal{C})$ is a polyhedron unlike $\text{conv}(S)$, which allows us to use the theoretical tool to determine if our inequalities are necessary in the description of $\text{clconv}(\mathcal{C})$. We prove that the inequalities are in fact facet-defining for $\text{clconv}(\mathcal{C})$. In a collaborative work with Kocuk, Dey, and Sun, it was also shown that these inequalities are the only non-trivial inequalities in the description of the convex hull of a lower-dimensional set obtained by projecting out a subset of variables. We also discuss the separation problem for the inequalities, which involves two subproblems: finding directed cycles in a graph and finding violated inequalities from each cycle. Due to the characteristic of $\mathcal{C}$, we can provide a closed-form solution to the second subproblem formulated as a knapsack problem. Computational experiment on IEEE instances were conducted to demonstrate the usefulness of the derived inequalities.

We conclude the thesis with remarks on the questions that remain open on the subjects. In chapters 2 and 3, we have established that the valid inequalities we derive are not the only ones necessary to describe $\text{conv}(P(A, B))$ or $\text{conv}(S)$. Identifying the missing inequalities to obtain complete characterizations of their convex hulls would be a meaningful contribution. Also, it would be interesting to see if it is useful to extend the result to generate strong relaxations for a larger dimensional set. On the computational side, we chose not to further investigate the strategy for extracting a block-diagonal matrix. An extensive analysis or experiments on the SDP we solve to perform matrix decomposition may shed light on how to extract small matrices so that the impact of our approach is maximized in the solution process. The algorithmic framework we applied in Chapter 4 is based on very simple heuristics, and we impose upper bounds on the number of cycles and inequalities that we handle to avoid computation time getting too long. It may be possible to design more efficient heuristics that utilize the result of our analysis of the separation problem presented in the thesis. Additional research goals would be to establish the complexity of separation for our inequalities, and to understand how to combine flow balance considerations with Kirchoff's voltage law to construct a stronger relaxation for the DC switching problem. We hope that these studies will spur additional research in to these and other important related sets.

Appendix A

Supplemental Proofs

A.1 Affine independence of points used in proof of Proposition 2.2

We have claimed in the proof of Proposition 2.2 that the following points are affinely independent.

$$\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}\frac{1}{a_{11}}\\0\\(\frac{1}{a_{11}})^2\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-r\bar{\beta}\\\bar{\beta}\\r^2\bar{\beta}^2\\\bar{\beta}^2\\0\\1\end{bmatrix}, \begin{bmatrix}0\\\bar{\beta}-\frac{\delta}{2}\\0\\\bar{\beta}(\bar{\beta}-\delta)\\1\\1\end{bmatrix}.$$

Denote these points by $p_0, p_1, \cdots, p_5$. To prove the claim, we show that the five points $(p_5 - p_0), (p_4 - p_0), \cdots, (p_1 - p_0)$ are linearly independent. Note that p_0 is the origin and that these five points are linearly independent if and only if $\sum_{i=1}^{5} \lambda_i p_i = \mathbf{0}$ implies $\lambda_i = 0,\ i = 1, 2, \cdots, 5$.

The first equation is written as the following system of equations:

$$\frac{1}{a_{11}}\lambda_3 - r\bar{\beta}\lambda_4 = 0 \tag{A.1}$$

$$\bar{\beta}\lambda_4 + (\bar{\beta} - \frac{\delta}{2})\lambda_5 = 0 \tag{A.2}$$

$$\lambda_2 + (\frac{1}{a_{11}})^2\lambda_3 + r^2\bar{\beta}^2\lambda_4 = 0 \tag{A.3}$$

$$\bar{\beta}^2\lambda_4 + \bar{\beta}(\bar{\beta} - \delta)\lambda_5 = 0 \tag{A.4}$$

$$\lambda_1 + \lambda_3 + \lambda_5 = 0 \tag{A.5}$$

$$\lambda_4 + \lambda_5 = 0 \tag{A.6}$$

We solve this system for λ. Taking the subtraction (A.6) $\times\bar{\beta}$ − (A.2), we obtain $\frac{\delta}{2}\lambda_5 = 0$. By (A.6), $\lambda_5 = 0$ implies $\lambda_4 = 0$, which in turn implies that $\lambda_3 = 0$ by (A.1). Plugging in $\lambda_3 = 0, \lambda_4 = 0$ to (A.3) and $\lambda_3 = 0, \lambda_5 = 0$ to (A.5), we obtain $\lambda_2 = \lambda_1 = 0$. We have established that $\lambda = \mathbf{0}$, which proves that $p_0, p_1, \cdots, p_5$ are affinely independent. The affine independence of all sets of points used in the proofs of Propositions in Chapter 2 can be established by similar exercise.

Bibliography

[1] *User's Manual for CPLEX Version 12.4*. IBM, 2011.

[2] S. Aktürk, A. Atamtürk, and S. Gürel. A strong conic quadratic reformulation for machine-job assignment with controllable processing times. *Operations Research Letters*, 37:187–191, 2009.

[3] A. Atamtürk. Cover and pack inequalities for (mixed) integer programming. *Annals of Operations Research*, 139:21–38, 2005.

[4] A. Atamtürk and V. Narayanan. Lifting for conic mixed integer programming. *Mathematical Programming Series A*, 126:351–363, 2007.

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

[6] E. Balas. Disjunctive programming: properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89:3–44, 1998.

[7] E. Balas and M. Perregaard. Lift-and-project for mixed 0-1 programming: recent progress. *Discrete Applied Mathematics*, 123:129–154, 2002.

[8] E. Balas, S. Ceria, and G. Cornuejols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming*, 58:295–324, 1993.

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

[10] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization*. MPS-SIAM Series on Optimization. SIAM, 2001.

[11] D. Bertsimas and R. Shioda. Algorithms for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43:1–22, 2009.

[12] D. Bertsimas, A. King, and R. Mazumder. Best subset selection via a modern optimization lens. In *The 2014-2015 Philip McCord Morse Lecture*. Informs Annual Meeting, 2014.

[13] S.A. Blumsack. *Network Topologies and Transmission Investment Under Electric-Industry Restructuring*. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, May 2006.

[14] P. Bonami, M. Kilinç, and J. Linderoth. Algorithms and software for convex mixed integer nonlinear programs. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, volume 154 of *The IMA Volumnes in Mathematics and its Applications*, pages 1–39. Springer, 2012.

[15] S. Ceria and J. Soares. Convex programming for disjunctive optimization. *Mathematical Programming*, 86:595–614, 1999.

[16] M. T. Cezik and G. Iyengar. Cuts for mixed 0-1 conic programming. *Mathematical Programming*, pages 179–202, 2005.

[17] M. Conforti, G. Cornuéjols, and G. Zambelli. Polyhedral approaches to mixed integer linear programming. In *50 Years of Integer Programming 1958-2008*, pages 343–385. Springer, 2010.

[18] H. Crowder, E.L. Johnson, and M. Padberg. Solving large-scale zero-one linear programming problems. *Operations Research*, 31, 1983.

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

[20] J.P. Vielma D. Dadush, S.S. Dey. On the Chvátal-Gomory closure of a compact convex set. *Mathematical Programming (to appear)*, 2013.

[21] S.S. Dey and D.A. Moran R. Some properties of convex hulls of integer points contained in general convex sets. *Mathematical Programming Series A*, 141:507–526, 2013.

[22] H. Dong and J. Linderoth. On valid inequalities for quadratic programming with continuous variables with binary indicators. In M. Goemans and J. Correa, editors, *16th International Conference on Integer Programming and Combinatorial Optimization*, volume 7801 of *Lecture Notes in Computer Science*, pages 169–180, 2013.

[23] S. Drewes and S. Pokutta. Cutting-planes for weakly-coupled 0/1 second order cone programs. *Electronic Notes in Discrete Mathematics*, 36:735–742, 2010.

[24] S. Drewes and S. Ulbrich. Subgradient based outer approximation for mixed integer second order cone programming. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, volume 154 of *The IMA Volumnes in Mathematics and its Applications*, pages 41–59. Springer, 2012.

[25] D.Wei, C.K.Sestok, and A.V.Oppenheim. Sparse filter design under a quadratic constraint: Low-complexity algorithms. *IEEE Transactions on Signal Processing*, 61:857–870, 2013.

[26] E.B. Fisher, R.P. O'Neil, and M.C. Ferris. Optimal transmission switching. *IEEE Transactions on Power Systems*, 23(3):1346–1355, 2008.

[27] A. Frangioni and C. Gentile. Perspective cuts for a class of convex 0-1 mixed integer programs. *Mathematical Programming*, 106:225–236, 2006.

[28] A. Frangioni and C. Gentile. SDP diagonalizations and perspective cuts for a class of nonseparable miqp. *Operations Research Letters*, 35:181–185, 2007.

[29] R.E. Gomory, E.L. Johnson, and L. Evans. Corner polyhedra and their connection with cutting planes. *Mathematical Programming Series B*, 96:321–329, 2003.

[30] Z. Gu, G.L. Nemhauser, and M.W.P. Savelsbergh. Lifted flow cover inequalities for mixed 0-1 integer programs. *Mathematical Programming*, 85:439–467, 1999.

[31] Z. Gu, G.L. Nemhauser, and M.W.P. Savelsbergh. Sequence independent lifting in mixed integer programming. *Journal of Combinatorial Optimization*, 4:109–129, 2000.

[32] O. Günlük and J. Linderoth. Perspective relaxation of mixed integer nonlinear programs with indicator variables. *Mathematical Programming Series B*, 2010.

[33] I. Guyon and A. Elisseeff. An introduction to variables and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.

[34] H.L. Hijazi, C. Coffrin, and P. Van Hentenryck. Convex quadratic relaxations of mixed-integer nonlinear programs in power systems. Technical report, NICTA, 2013. URL `http://www.optimization-online.org/DB_HTML/2013/09/4057.html`.

[35] J. B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, Berlin, Germany, 1993.

[36] Brady Hunsaker. *Measuring Facets of Polyhedra to Predict Usefulness in Branch-and-cut Algorithms*. PhD thesis, Georgia Institute of Technology, 2003.

[37] M. Kilinç, J. Linderoth, and J. Luedtke. Effective separation of disjunctive cuts for convex mixed integer nonlinear programs. Technical Report 1681, Computer Science Department, University of Wisconsin, 2010.

[38] M. Kilinç, J. Linderoth, J. Luedtke, and A. Miller. Strong branching inequalities for convex mixed integer nonlinear programs. *Computational Optimization and Applications*, 59:639–665, 2014.

[39] F. Kilinç-Karzan. On minimal valid inequalities for mixed integer conic programs. `http://www.optimization-online.org/DB_HTML/2013/06/3936.html`, 2013.

[40] B. Kocuk, H. Jeon, S. Dey, J. Linderoth, J. Luedtke, and A. Sun. Cycle-based formulation and valid inequalities for the dc transmission switching problem. *Operations Research*, 2014. submitted.

[41] H. W. Kuhn. On the origin of the Hungarian method. In J. K. Lenstra, A. H. G. Rinnooy Kan, and A. Schrijver, editors, *History of Mathematical Programming: A Collection of Personal Reminiscences*, pages 77–81, Amsterdam, 1991. Elsevier. Postscript of the article.

[42] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7:77–91, 1952.

[43] A. Martin and R. Weismantel. *The Intersection of Knapsack Polyhedra and Extensions*, volume 1412, pages 243–256. Springer, 1998.

[44] R.R. Meyer. On the existence of optimal solutions to integer and mixed integer programming problems. *Mathmatical Programming*, 7:223–235, 1974.

[45] A.J. Miller. Selection of subsets of regression variables. *Journal of Royal Statistical Society. Series A*, 147:389–425, 1984.

[46] S. Modaresi, M. Kilinç, and J.P. Vielma. Intersection cuts for nonlinear integer programming: convexification techniques for structured sets. 2013. Submitted for publication.

[47] J.A. Momoh. *Electric Power System Applications of Optimization*. Power Engineering(Willis). CRC Press, 2 edition, 2001.

[48] J. Neter, W. Wasserman, and M.H. Kutner. *Applied Linear Regression Models*. McGraw-Hill, 2 edition, 1989.

[49] M.W. Padberg, T.J. Van Roy, and L.A. Wolsey. Valid inequalities for fixed charge problems. *Operations Research*, 33, 1985.

[50] I. Pólik. *Conic Optimization Software*. John Wiley & Sons, Inc., 2010. ISBN 9780470400531. doi: 10.1002/9780470400531.eorms0182. URL `http://dx.doi.org/10.1002/9780470400531.eorms0182`.

[51] R.Jansen and R. van Dijk. Optimal benchmark tracking with small portfoilo. *Journal of Portfolio Management*, 28:33–39, 2002.

[52] R. Stubbs and S. Mehrotra. A branch-and-cut method for 0-1 mixed convex programming. *Mathematical Programming*, pages 515–532, 1999.

[53] R. A. Stubbs. *Branch-and-Cut Methods for Mixed 0-1 Convex Programming*. PhD thesis, Northwestern University, December 1996.

[54] J.P. Vielma, S. Ahmed, and G. Nemhauser. A lifted linear programming branch-and-bound algorithm for mixed integer conic quadratic programs. *INFORMS Journal on Computing*, 20: 438–450, 2008.

[55] L. A. Wolsey. *Integer Programming*. Wiley-Interscience publication, 1998.

[56] L.A. Wolsey. Valid inequalities and superadditivity for 0-1 integer programs. *Mathematics of Operations Research*, 2(1):66–77, 1977.

[57] X. Zheng, X. Sun, and D. Li. Improving the performance of miqp solvers for quadratic programs with cardinality and minimum threshold constraints: A semidefinite program approach, 2010. URL `http://www.optimization-online.org/DB_HTML/2010/11/2797.html`.