VALID INEQUALITIES AND REFORMULATION TECHNIQUES FOR MIXED INTEGER NONLINEAR PROGRAMMING

by

Sina Modaresi

B.S., Sharif University of Technology, 2010

M.S., University of Pittsburgh, 2012

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Sina Modaresi

It was defended on

November 12, 2015

and approved by

Jayant Rajgopal, Ph.D., Professor, Department of Industrial Engineering

Oleg A. Prokopyev, Ph.D., Associate Professor, Department of Industrial Engineering

Andrew J. Schaefer, Ph.D., Noah Harding Chair and Professor, Department of Computational and Applied Mathematics, Rice University

Juan Pablo Vielma, Ph.D., Richard S. Leghorn Career Development Assistant Professor, Sloan School of Management, Massachusetts Institute of Technology

Dissertation Director: Jayant Rajgopal, Ph.D., Professor, Department of Industrial Engineering

VALID INEQUALITIES AND REFORMULATION TECHNIQUES FOR MIXED INTEGER NONLINEAR PROGRAMMING

Sina Modaresi, PhD

University of Pittsburgh, 2015

One of the most important breakthroughs in the area of Mixed Integer Linear Programming (MILP) is the characterization of the convex hull of specially structured non-convex polyhedral sets in order to develop valid inequalities or cutting planes. Development of strong valid inequalities such as Split cuts, Gomory Mixed Integer (GMI) cuts, and Mixed Integer Rounding (MIR) cuts has resulted in highly effective branch-and-cut algorithms. While such cuts are known to be equivalent, each of their characterizations provides different advantages and insights.

The study of cutting planes for Mixed Integer Nonlinear Programming (MINLP) is still much more limited than that for MILP, since characterizing cuts for MINLP requires the study of the convex hull of a non-convex and non-polyhedral set, which has proven to be significantly harder than the polyhedral case. However, there has been significant work on the computational use of cuts in MINLP. Furthermore, there has recently been a significant interest in extending the associated theoretical results from MILP to the realm of MINLP.

This dissertation is focused on the development of new cuts and extended formulations for Mixed Integer Nonlinear Programs. We study the generalization of split, k-branch split, and intersection cuts from Mixed Integer Linear Programming to the realm of Mixed Integer Nonlinear Programming. Constructing such cuts requires calculating the convex hull of the difference between a convex set and an open set with a simple geometric structure. We introduce two techniques to give precise characterizations of such convex hulls and use them to construct split, k-branch split, and intersection cuts for several classes of non-polyhedral sets. We also study the relation between the introduced cuts and some known classes of cutting planes from MILP. Furthermore, we show how an aggregation technique can be easily extended to characterize the convex hull of sets defined by two quadratic or by a conic quadratic and a quadratic inequality. We also computationally evaluate the performance of the introduced cuts and extended formulations on two classes of MINLP problems.

Keywords: Mixed Integer Linear Programming, Mixed Integer Nonlinear Programming, Valid Inequality, Split Cut, K-branch Split Cut, Gomory Mixed Integer Cut, Mixed Integer Rounding Cut, Intersection Cut, Branch-and-Cut, Extended Formulation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

to my parents, Yousef and Maliheh

This work would not have been possible without the guidance of Dr. Juan Pablo Vielma, my mentor and advisor. I would like to express my deepest gratitude to him, first and foremost, for his constant support, encouragement, professional attitude, constructive criticism, and most importantly, patience.

I would also like to thank the other members of my doctoral committee, Dr. Jayant Rajgopal, Dr. Oleg Prokopyev, and Dr. Andrew Schaefer for their valuable suggestions and comments.

Finally, I am forever indebted to my parents, Yousef and Maliheh, my brother, Sajad, and my sister, Zahra, for their endless love, inspiration, encouragement, and unconditional support.

1.0 INTRODUCTION

An important are of Mixed Integer Linear Programming (MILP) is the characterization of the convex hull of specially structured non-convex polyhedral sets to develop valid inequalities or cutting planes [10, 11, 30, 31, 32, 33, 38, 42, 62, 90]. Development of strong valid inequalities such as Split cuts [32], Gomory Mixed Integer (GMI) cuts [48, 49], and Mixed Integer Rounding (MIR) cuts [66, 74, 75, 93] has resulted in highly effective branch-and-cut algorithms [4, 22, 23, 55, 63]. While such cuts are known to be equivalent [39, 75], each of their characterizations provide different advantages and insights.

The study of cutting planes for Mixed Integer Nonlinear Programming (MINLP) is still much more limited than that for MILP, since characterizing cuts for MINLP requires the study of the convex hull of a non-convex and non-polyhedral set, which has proven to be significantly harder than the polyhedral case. However, there has been significant work on the computational use of cuts in MINLP [25, 28, 44, 59, 87]. Moreover, there has recently been a significant interest in extending the associated theoretical results from MILP to the realm of MINLP [18, 19, 35, 36, 37, 81]. In particular, for the case of Mixed Integer Conic Quadratic Programming (MICQP), there has been a recent surge of theoretical developments [6, 8, 9, 14, 15, 16, 27, 57, 58, 73, 71, 72, 70, 96, 91]. However, most of the known results in the area of MINLP are still limited to very specific sets [54, 86, 88] or to approximations of semi-algebraic sets through Semidefinite Programming (SDP) [45, 61, 76, 77, 78, 79, 80].

While the resulting cuts for MINLP are strong nonlinear inequalities, adding such nonlinear cuts to the continuous relaxation of a MINLP could significantly increase its solution time. Hence there will likely be a strong trade-off between the strength provided by such cuts and their computational cost. It is then unclear if such nonlinear cuts can provide a significant computational advantage over linearization approaches such as those in [25, 59] which do not require explicit cut formulas. However, even in such cases, the developed nonlinear cuts can provide valuable information about the performance of the linearization approaches. For instance, the linearization approaches can sometimes require a large number of iterations to yield a bound improvement similar to that obtained by the associated nonlinear cut. Adding the nonlinear cut provides a simple way to evaluate if the lack of bound improvement is due to lack of strength of the cut or lack of convergence of the linearization approach. Similarly, the availability of explicit formulas of split cuts for quadratic sets proven extremely useful to evaluate the strength of a cutting plane approach based on extended formulations in [71].

This dissertation is focused on the development of new cuts and extended formulations for Mixed Integer Nonlinear Programs. We study the generalization of split, k-branch split, and intersection cuts from Mixed Integer Linear Programming to the realm of Mixed Integer Nonlinear Programming. Constructing such cuts requires calculating the convex hull of the difference between a convex set and an open set with a simple geometric structure. We introduce two techniques to give precise characterizations of such convex hulls and use them to construct split, k-branch split, and intersection cuts for several classes of non-polyhedral sets. We also study the relation between the introduced cuts and some known classes of cutting planes from MILP. Furthermore, we show how an aggregation technique can be easily extended to characterize the convex hull of sets defined by two quadratic or by a conic quadratic and a quadratic inequality. We also computationally evaluate the performance of the introduced cuts and extended formulations on two classes of MINLP problems.

The remainder of this dissertation is organized as follows. In Chapter 3 we study the generalization of split, k-branch split, and intersection cuts from MILP to MINLP. We propose two simple techniques to derive general intersection cuts for several classes of MINLP problems with specific structures. In particular, we give simple formulas for split cuts for essentially all convex sets described by a single conic quadratic inequality. We also give simple formulas for k-branch split cuts and some general intersection cuts for a wide variety of convex quadratic sets.

In Chapter 4 we study split cuts and extended formulations for MICQP. In particular, we study the relation between Conic MIR (CMIR) cuts introduced by Atamtürk and Narayanan

[9] and nonlinear split cuts for a class of MICQP problems. We also study an extended formulation for such a class of MICQP and illustrate how the power of an extended formulation can improve the strength of a cutting plane procedure in MINLP.

In Chapter 5 we consider an aggregation technique introduced by Yıldıran [94] to study the convex hull of regions defined by two quadratic or by a conic quadratic and a quadratic inequality. Yıldıran [94] shows how to characterize the convex hull of sets defined by two quadratics using Linear Matrix Inequalities (LMI). We show how this aggregation technique can be easily extended to yield valid conic quadratic inequalities for the convex hull of sets defined by two quadratic or by a conic quadratic and a quadratic inequality. We also show that in many cases under additional assumptions, these valid inequalities characterize the convex hull exactly.

In Chapter 6 we computationally evaluate the performance of the introduced linear and nonlinear cuts and extended formulations on two classes of MINLP problems (Closest Vector Problem and Mean-variance Capital Budgeting). We compare the strength of the nonlinear cuts added to the original formulations versus the linear cuts added to an extended formulation.

Finally, Chapter 7 concludes the discussion summarizing the contributions of this dissertation.

2.0 NOTATION AND PRELIMINARIES

We use the following notation throughout the dissertation. Let $e^i \in \mathbb{R}^n$ be the i-th unit vector, $0_n \in \mathbb{R}^n$ be the zero vector, $I \in \mathbb{R}^{n \times n}$ be the identity matrix where n is an appropriate dimension that we omit if evident from the context, and $\mathbb{S}^n$ denote symmetric matrices with n rows and columns. For $a \in \mathbb{R}$ we let $(a)^+ := \max\{0, a\}$ and $\lfloor a \rfloor := \max\{k \in \mathbb{Z} : k \leq a\}$. We also let $\|x\|_2 := \sqrt{\sum_{i=1}^n x_i^2}$ denote the Euclidean norm of $x \in \mathbb{R}^n$ and $|x| \in \mathbb{R}^n$ be the vector whose components are the absolute value of the components of $x \in \mathbb{R}^n$. In addition, we let $\|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ denote the p-norm of a given vector $x \in \mathbb{R}^n$ and for a vector $v \in \mathbb{R}^n$, we let the projection onto its span be $P_v := \frac{vv^T}{\|v\|_2^2}$ and onto its orthogonal complement be $P_v^\perp := I - \frac{vv^T}{\|v\|_2^2}$. We also let $\{\pi_i\}_{i=1}^k \subseteq \mathbb{R}^n \setminus \{0_n\}$ be an arbitrary set of vectors, and not necessarily a sequence of vectors. For a matrix $\mathcal{P}$, we let $\pi(\mathcal{P})$ denote the number of negative eigenvalues of $\mathcal{P}$ and $\text{null}(\mathcal{P})$ denote its null space. For a set $S \subseteq \mathbb{R}^n$, we let $\text{int}(S)$ be its interior, $\text{bd}(S)$ be its boundary, $\text{conv}(S)$ be its convex hull, $\overline{\text{conv}}(S)$ be the closure of its convex hull, $\text{aff}(S)$ be its affine hull, $\text{lin}(S) := \{d \in \mathbb{R}^n : x + \lambda d \in S \text{ for all } x \in S \text{ and } \lambda \in \mathbb{R}\}$ be its lineality space, and S_∞ be its recession cone. For a function $G : \mathbb{R}^n \to \mathbb{R}$ we let $\text{epi}(G) := \{(x, t) \in \mathbb{R}^{n+1} : G(x) \leq t\}$ be its epigraph, $\text{gr}(G) := \{(x, t) \in \mathbb{R}^{n+1} : G(x) = t\}$ be its graph, and $\text{hyp}(G) := \{(x, t) \in \mathbb{R}^{n+1} : G(x) \geq t\}$ be its hypograph. In addition, we let the second-order cone (a.k.a. Lorentz cone) be the epigraph of the Euclidean norm defined as $\{(x, t) \in \mathbb{R}^{n+1} : \|x\|_2 \leq t\}$. Finally, we let $[n] := \{1, \ldots, n\}$.

3.0 INTERSECTION CUTS FOR NONLINEAR INTEGER PROGRAMMING: CONVEXIFICATION TECHNIQUES FOR STRUCTURED SETS

One of the most important breakthroughs in the area of Mixed Integer Linear Programming (MILP) is the development of strong valid inequalities or cutting planes such as split and intersection cuts. However, the study of cuts for Mixed Integer Nonlinear Programming (MINLP) is still much more limited than that for MILP. Most of the known results in this area are limited to very specific sets [54, 86, 88] or to approximations of semi-algebraic sets through Semidefinite Programming (SDP) [45, 61, 76, 77, 78, 79, 80]. While some precise SDP representations of the convex hulls of semi-algebraic sets exist [50, 51, 52, 85], these require the use of auxiliary variables. Such higher dimensional, extended, or lifted representations are extremely powerful. However, there are theoretical and computational reasons to want representations in the original space and/or in the same class as the original set (e.g. representations that do not jump from quadratic basic semi-algebraic to SDP). We refer to characterizations that satisfy both these requirements as *projected* and *class preserving*. Projected and class preserving are in general incompatible (e.g. the convex hull of the basic semi-algebraic set $\{x \in \mathbb{R}^2 : (x_1^2 - x_2)x_1 \geq 0,\, x_2 \geq 0\}$ has no projected basic semi-algebraic representation, but has a lifted basic semi-algebraic representation [24]). Furthermore, even giving an algebraic characterization of the boundary of the convex hull of a variety [82, 83] or giving a projected SDP representation of the convex hull of certain varieties and quadratic semi-algebraic sets [84, 94, 95] requires very complex techniques from algebraic geometry. All such issues make extending MILP cutting planes to the MINLP setting extremely challenging. To alleviate such challenges, in this chapter we concentrate on the extension of split, k-branch split, and other intersection cuts to the MINLP setting [10, 32, 38, 48, 49, 62].

Split, k-branch split, and intersection cuts for MILP can all be obtained by taking the convex hull of the difference between a convex set and a set with a simple geometric structure. This characterization allows for a straightforward extension of the cuts to the MINLP setting. However, this conceptual extension does not provide a practical construction procedure for the cuts. For this reason, we follow the approach of the simple, but extremely powerful Mixed Integer Rounding (MIR) cut [66, 74, 75, 93]. The MIR procedure can be used to generate every split cut for a MILP and, together with the closely related Gomory Mixed Integer (GMI) cut procedure [48, 49], yields the most effective cutting plane approach for general MILP [22, 23]. In particular, one version of the MIR procedure shows that every split cut can be constructed through a simple two step procedure. The first step is the construction of a canonical cut known as the *simple* or *basic* MIR. This cut is obtained by taking the convex hull of the difference between two simple convex sets in $\mathbb{R}^2$, both of which are described by two linear inequalities. The second step simply uses linear transformations to obtain all split cuts from the basic MIR. In this chapter we show that a similar approach can be used to construct a wide range of intersection cuts. More specifically, we show how two very simple techniques can be used to construct projected class preserving characterizations of the convex hull of difference between certain canonical sets. The techniques we consider are only tailored to the geometric structure of these canonical sets and do not require the sets to have any additional algebraic properties (e.g. being quadratic, basic semi-algebraic, etc.). Thanks to this, the resulting characterizations are quite general, but give simple closed form expressions. While the canonical sets are somewhat specific, we can also use affine transformations to obtain more general cuts. In particular, these techniques can be used to construct split cuts for essentially all convex sets described by a single conic quadratic inequality, and to extend k-branch split and general intersection cuts to a wide variety of quadratic sets of interest to trust region and lattice problems. In both cases, the only algebraic property of the quadratic sets needed for the construction is the symmetry of the Euclidean norm. This suggests that the techniques could be useful to construct cuts for additional classes of sets by only exploiting similar basic properties.

Constructing such cuts requires calculating the convex hull of the difference between a convex set and an open set with a simple geometric structure. We introduce two techniques to give precise characterizations of such convex hulls and use them to construct split, k-branch split, and intersection cuts for several classes of non-polyhedral sets. In particular, we give simple formulas for split cuts for essentially all convex sets described by a single conic quadratic inequality. We also give simple formulas for k-branch split cuts and some general intersection cuts for a wide variety of convex quadratic sets.

The rest of this chapter is organized as follows. We begin with Section 3.1 where we introduce some notation and review some known results. Section 3.2 then introduces an interpolation technique that can be used to construct split and k-branch split cuts for many classes of sets. Then, in Section 3.3 we use the interpolation technique to characterize intersection cuts for conic quadratic sets. Finally, Section 3.4 introduces an aggregation technique that can be used to construct a wide array of general intersection cuts. In both Sections 3.2 and 3.4, we first present the basic principles behind the techniques in a simple, but abstract setting, and then utilize them to construct more specific cuts to illustrate their power and limitations.

3.1 NOTATION, KNOWN RESULTS AND OTHER PRELIMINARIES

In addition to the notation introduced in Chapter 2, we use the following notation and definitions in this chapter.

Definition 1 (Intersection, Split, k-branch Split, and t-inclusive Split Cuts)**.** *Let $B \subseteq \mathbb{R}^n$ be a closed convex set that we refer to as the* base set*, $F \subseteq \mathbb{R}^n$ be a closed set that we refer to as the* forbidden set*, and $g : \mathbb{R}^n \to \mathbb{R}$ be an arbitrary function. We say inequality $g(x) \leq 0$ is an* intersection cut *for B and F if $\overline{\operatorname{conv}}\,(B \setminus \operatorname{int}(F)) \subseteq \{x \in \mathbb{R}^n : g(x) \leq 0\}$ and g is convex.*

We let a split *be a set of the form $\left\{x \in \mathbb{R}^n : \pi^T x \in [\pi_0, \pi_1]\right\}$ for some $\pi \in \mathbb{R}^n \setminus \{0_n\}$ and $\pi_0, \pi_1 \in \mathbb{R}$ such that $\pi_0 < \pi_1$. If F is a split, we say that the associated intersection cut is a* split cut*. Besides, if F is a split with $\pi = e^i$ for some $i \in [n]$, we refer to F as an* elementary split *and to the the associated split cut as an* elementary split cut*.*

We let a k-branch split *be a set of the form* $\bigcup_{i=1}^{k}\left\{x \in \mathbb{R}^n : \pi_0^i \leq \pi_i^T x \leq \pi_1^i\right\}$ *for some* $\{\pi_i\}_{i=1}^{k} \subseteq \mathbb{R}^n \setminus \{0_n\}$, $\pi_0^i, \pi_1^i \in \mathbb{R}$ *such that* $\pi_0^i < \pi_1^i$ *for all* $i \in [k]$*. If* F *is a k-branch split, we say that the associated intersection cut is a* k-branch split cut*.*

When considering epigraphical sets of the form $B = \{(x,t) \in \mathbb{R}^{n+1} : G(x) \leq t\}$ *for some closed convex function* $G(x)$*, we often assume that* F *is a cylinder whose axis lies along* t *(i.e.,* F *is of the form* $S \times \mathbb{R}$ *for some* $S \subseteq \mathbb{R}^n$*). For instance, if* F *is a split, we have* $F = \left\{(x,t) \in \mathbb{R}^{n+1} : \pi^T x \in [\pi_0, \pi_1]\right\}$*. However, in some cases, we consider a split that includes* t *and we refer to such a split as a* t-inclusive split*. More specifically, we let a t-inclusive split be a set of the form* $\left\{(x,t) \in \mathbb{R}^{n+1} : \pi^T x + \hat{\pi} t \in [\pi_0, \pi_1]\right\}$ *for some* $(\pi, \hat{\pi}) \in \mathbb{R}^{n+1}$ *such that* $\hat{\pi} \neq 0$[1]*, and* π_0*,* $\pi_1 \in \mathbb{R}$ *such that* $\pi_0 < \pi_1$*. If* F *is a t-inclusive split, we say that the associated intersection cut is a* t-inclusive split cut*.*

We mostly restrict to the cases in which $\operatorname{conv}(B \setminus \operatorname{int}(F))$ *is closed, so for notational convenience, we let* $\overline{B} := \operatorname{conv}(B \setminus \operatorname{int}(F))$ *when* F *is evident from the context.*

We note that the term intersection cut was introduced by Balas [10] for the case in which B is a translated simplicial cone, F is the euclidean ball or a cylinder of a lower dimensional euclidean ball, and the unique vertex of B is in $\operatorname{int}(F)$. In this setting, we have that $\operatorname{conv}(B \setminus \operatorname{int}(F))$ is closed and can be described by adding a single linear inequality to B. Furthermore, this single linear inequality has a simple formula dependent on the *intersections* of the extreme rays of B with F. While we do not always have such intersection formulas for other classes of sets, we continue to use the term intersection cut in the more general setting and avoid any additional qualifiers for simplicity. In particular, we do not use the term *generalized intersection cut* as it has already been used for the case of polyhedral B and F and in conjunction with an improved cut generation procedure for MILP [12]. The term split cut was introduced by Cook, Kannan and Schrijver [32], and their original definition directly generalizes to non-polyhedral sets as in Definition 1. The term k-branch split cut was introduced by Li and Richard [62]; 2-branch split cuts are also called cross cuts in Dash, Dey and Günlük [38]. These definitions also directly generalize to non-polyhedral sets as in Definition 1.

[1] We allow $\pi = 0_n$ to consider disjunctions that only affect t.

The interest of intersection cuts for MILP and MINLP arises from the fact that if $\mathrm{int}(F)\cap(\mathbb{Z}^p\times\mathbb{R}^q)=\emptyset$, an intersection cut for B and F is valid for $\overline{\mathrm{conv}}\,(B\cap(\mathbb{Z}^p\times\mathbb{R}^q))$. Hence, intersection cuts can be used to strengthen the continuous relaxation of MILP and MINLP problems.

Intersection cuts are particularly attractive in the MILP setting, since they can be quite strong and can easily be constructed. They were extensively studied when they were first proposed in the 1970s [10, 48, 49] and have recently received renewed interest [31, 42]. Part of the relative simplicity and effectiveness of intersection cuts for MILP stems from two basic facts. The first one is that in the MILP setting, B is a polyhedron (i.e., the continuous relaxation of a MILP is an LP). The second one is the fact that every convex set F such that $\mathrm{int}(F)\cap\mathbb{Z}^n=\emptyset$ (usually denoted a *lattice free convex set*) and that is maximal with respect to inclusion for this property is also a polyhedron [64]. Restricting both B and F to be (convex) polyhedra give intersection cuts for MILP several useful properties. For instance, if B and F are polyhedra, then $\overline{\mathrm{conv}}\,(B\setminus\mathrm{int}\,(F))$ is a polyhedron [42]. Hence, in the MILP setting, we can restrict our attention to linear intersection cuts. In particular, if F is a split and B is a polyhedron, then all linear intersection cuts for B and F can be constructed from simplicial relaxations of B and hence have simple formulas [5, 40, 89]. As discussed in the introduction, GMI cuts [48, 49] and MIR cuts [66, 74, 75, 93] are two versions of these formulas. For more information on the ongoing efforts to duplicate this effectiveness for other lattice free polyhedra, we refer the reader to [31, 42]. In this context, we note that $\mathrm{conv}\,(B\setminus\mathrm{int}\,(F))$ can fail to be closed even if B and F are polyhedra and F is not a split (e.g. consider $B=\{x\in\mathbb{R}^2 : x_2\geq 0\}$ and $F=\{x\in\mathbb{R}^2 : x_2\leq 1,\, x_1+x_2\leq 1\}$). However, $\mathrm{conv}\,(B\setminus\mathrm{int}\,(F))$ is closed in the polyhedral case if F is convex and full-dimensional and the recession cone of F is a linear subspace [7].

In the MINLP setting, there has been significant work on the computational use of linear split cuts [25, 28, 87, 44, 59]. From the theoretical side, we know that if F is a split, then $\mathrm{conv}\,(B\setminus\mathrm{int}\,(F))$ is closed even if B is not polyhedral [37]. With respect to formulas for intersection cuts, there has been some progress in the description of split cuts for quadratic sets in [8, 9, 37, 14]. Dadush et al. [37] show that, if B is an ellipsoid and F is a split, then $\mathrm{conv}\,(B\setminus\mathrm{int}\,(F))$ can be described by intersecting B with either a linear half space, an affine transformation of the second-order cone, or an ellipsoidal cylinder. In addition, they give simple closed form expressions for all these linear and nonlinear split cuts. Independently, [14] studies split cuts for more general quadratic sets, but only for splits in which $\{x \in B : \pi^T x = \pi_0\}$ and $\{x \in B : \pi^T x = \pi_1\}$ are bounded. They give a procedure to find the associated split cuts, but do not give closed form expressions for them. Finally, [8, 9] give a simple formula for an elementary split cut for the standard three dimensional second-order cone. While [14] develops a procedure to construct split cuts through a detailed algebraic analysis of quadratic constraints developed in [15], [8, 9, 37] give formulas for split cuts through simple geometric arguments. As we have recently shown at the MIP 2012 Workshop, these geometric techniques can be extended to additional quadratic and basic semi-algebraic sets [69]. In this paper we show that the principles behind these geometric arguments can be abstracted from the semi-algebraic setting to develop split and k-branch split cut formulas for a wider class of specially structured convex sets. This abstraction greatly simplifies the proofs and can be used to construct split cuts for essentially all convex sets described by a single conic quadratic inequality through simple linear algebra arguments. In addition to studying split and k-branch split cuts, we show how a commonly used aggregation technique can be used to develop formulas for general nonlinear intersection cuts for the case in which B and F are both non-polyhedral, but share a common structure. While a non-polyhedral F is not necessary in the MINLP settings (it still should be sufficient to consider maximal lattice free convex sets, which are polyhedral), they could still provide an advantage and are important in other settings such as trust region problems [18, 19, 79] and lattice problems [26, 71]. We finally note that similar results for the quadratic case have recently been independently developed in [6]. We discuss the relation between the results in [6] and our work at the end of Section 3.3.2.

To describe our approach, we use the following additional definition.

Definition 2. *Let $B \subseteq \mathbb{R}^n$ be a closed convex set, $F \subseteq \mathbb{R}^n$ be a closed set, and $g : \mathbb{R}^n \to \mathbb{R}$ be an arbitrary function. We say inequality $g(x) \leq 0$ is a:*

- valid cut *if $\overline{B} \subseteq \{x \in \mathbb{R}^n : g(x) \leq 0\}$,*
- binding valid cut *if it is valid and $\{x \in B \setminus \operatorname{int}(F) : g(x) = 0\} \neq \emptyset$, and*

- sufficient cut, *if* $\{x \in B : g(x) \leq 0\} \subseteq \overline{B}$.

Binding valid cuts correspond to valid cuts that cannot be improved by translations, and sufficient cuts are those that are violated by any point of B outside $\overline{B}$. We can show that a convex cut that is sufficient and valid is enough to describe $\overline{B}$ together with the original constraints defining B. Our approach to generating such cuts will be to construct cuts that are binding and valid by design, and that have simple structures from which sufficiency can easily be proven.

3.2 INTERSECTION CUTS THROUGH INTERPOLATION

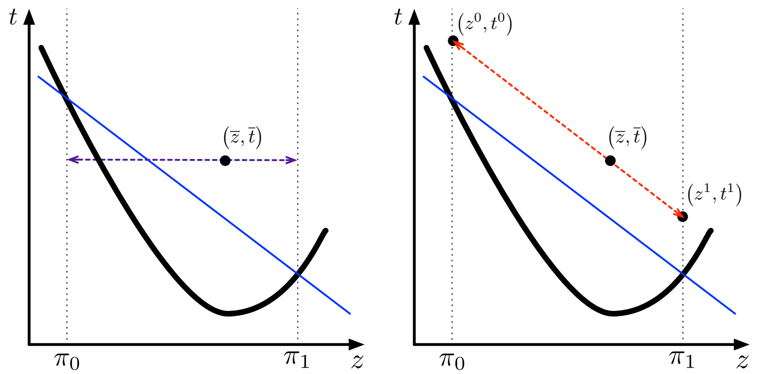
In this section we consider the case in which the base set is either the epigraph, lower level set, or a section of the epigraph of a convex function and the forbidden set corresponds to a split, t-inclusive split, or a k-branch split. Our cut construction approach is based on a simple interpolation technique that can be more naturally explained for splits and epigraphs of specially structured functions. For this reason, we begin with such a case and then consider special cases of non-epigraphical sets and discuss the limits of the interpolation technique. While the structures for which the technique yields simple formulas are quite specific, we can consider broader classes by considering affine transformations. In Section 3.3 we illustrate the power of this approach by showing how the interpolation technique yields formulas for intersection cuts for convex quadratic sets.

3.2.1 Split Cuts for Epigraphical Sets

Let $G : \mathbb{R} \to \mathbb{R}$ be a closed convex function and let F be an elementary split associated with $\pi = e^1$. Then $\overline{\operatorname{epi}(G)} = \operatorname{epi}(G) \cap \operatorname{epi}(J)$ for

$$J(z) = \frac{G(\pi_1) - G(\pi_0)}{\pi_1 - \pi_0} z + \frac{\pi_1 G(\pi_0) - \pi_0 G(\pi_1)}{\pi_1 - \pi_0}. \tag{3.1}$$

This is illustrated in Figure 1, where the graph of G is given by the thick black curve and the graph of J is depicted by the thin blue line. Indeed, since J is a linear function and hence



(a) Naive friends construction. (b) Friends by following the slope.

Figure 1: Interpolation Technique for Univariate Functions.

$\text{epi}(G) \cap \text{epi}(J)$ convex, it is enough to show that $J(z) \leq t$ is a valid and sufficient cut. We can check that $J(z) \leq t$ is a binding valid cut because J is the (affine) linear interpolation of G through $z = \pi_0$ and $z = \pi_1$. Convexity of G then implies that this interpolation is below G for $z \notin (\pi_0, \pi_1)$.

To show that the cut is sufficient, we need to show that any point $(\overline{z}, \overline{t}) \in \text{epi}(G)$ that satisfies the cut is in $\overline{\text{epi}(G)}$. To achieve this, we can find two points (z^0, t^0) and (z^1, t^1) in $\text{epi}(G)$ such that $z^0 \leq \pi_0$, $z^1 \geq \pi_1$, and $(\overline{z}, \overline{t}) \in \text{conv}\left(\{(z^0, t^0), (z^1, t^1)\}\right)$. Following [41], we will denote these points the *friends* of $(\overline{z}, \overline{t})$. One naive way to construct the friends is to wiggle $(\overline{z}, \overline{t})$ by decreasing and increasing $\overline{z}$ until it reaches π_0 and π_1, respectively. However, as illustrated in Figure 1(a), this can result in one of the friends falling outside $\text{epi}(G)$. Fortunately, as illustrated in Figure 1(b), we can always wiggle by following the slope of the cut J to assure that the friends are in $\text{epi}(J)$. Correctness (i.e., containment of the friends in $\text{epi}(G)$) then follows by noting that $J(z) = G(z)$ at $z = \pi_0$ and $z = \pi_1$, since $J(z) \leq t$ is a binding valid cut. This two-stage procedure of binding validity through interpolation and sufficiency through friends can be formalized for general closed convex sets as follows.

Proposition 1. *Let $B \subseteq \mathbb{R}^n$ be a closed convex set and $F \subseteq \mathbb{R}^n$ be closed. If $C \subseteq \mathbb{R}^n$ is a closed convex set such that*

$$B \cap \operatorname{bd}(F) \quad = \quad C \cap \operatorname{bd}(F) \tag{3.2a}$$

$$B \setminus \operatorname{int}(F) \quad \subseteq \quad C \setminus \operatorname{int}(F), \tag{3.2b}$$

and if

$$\text{for all } \bar{x} \in C \cap \operatorname{int}(F) \text{ there exists a finite set } \Gamma \subseteq C \cap \operatorname{bd}(F) \text{ such that } \bar{x} \in \operatorname{conv}(\Gamma), \tag{3.3}$$

then

$$\overline{B} = B \cap C. \tag{3.4}$$

Proof. We have that

$$B \setminus \operatorname{int}(F) \subseteq B \cap C \subseteq \overline{B}, \tag{3.5}$$

where the first containment comes from (3.2b) and the last from (3.3) and (3.2a). The result follows by taking convex hull in (3.5) and noting that $B \cap C$ is convex because both B and C are convex. □

Note that if F is a split, we can always consider Γ containing exactly two points (e.g Figure 1 and Propositions 2 and 4), while larger sets Γ might be necessary for other forbidden sets (e.g. Proposition 7). Our general approach to use Proposition 1 is to construct a convex function that yields binding valid cuts (i.e., satisfies (3.2)) and to use its specific geometric structure to construct friends for sufficiency. We now consider two structures in which the appropriate interpolation can easily be constructed once we identify the interpolation's general form. The geometric structures of the resulting cuts yield two friends construction techniques. The first technique generalizes the univariate argument in Figure 1(b) by noting that following the slope of J is equivalent to moving in $\operatorname{lin}(\operatorname{epi}(J))$. The second technique constructs the friends by moving in a ray contained in an appropriately constructed cone. These techniques are described in detail in Sections 3.2.1.1 and 3.2.1.2 respectively.

3.2.1.1 Separable Functions Let G be a separable function of the form $G(z,y) = f(z) + g(y)$ with $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^p \to \mathbb{R}$ closed convex functions, and let F be an elementary split associated with $\pi = e^1$. Analogous to (3.1), we can simply interpolate G parametrically on y to obtain

$$J(z,y) = \frac{G(\pi_1,y) - G(\pi_0,y)}{\pi_1 - \pi_0} z + \frac{\pi_1 G(\pi_0,y) - \pi_0 G(\pi_1,y)}{\pi_1 - \pi_0}. \tag{3.6}$$

In this case, the interpolation simplifies to $J(z,y) = \frac{f(\pi_1)-f(\pi_0)}{\pi_1-\pi_0} z + \frac{\pi_1 f(\pi_0) - \pi_0 f(\pi_1)}{\pi_1 - \pi_0} + g(y)$, which is convex on (z,y) and linear on z. Our original univariate argument follows through directly and we get $\overline{\text{epi}\,(G)} = \text{epi}\,(G) \cap \text{epi}\,(J)$. To illustrate this, consider $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $G(z,y) = z^2 + y^2$ and let F be the elementary split associated with $\pi = e^1$, $\pi_0 = -10$, and $\pi_1 = 1$. Constructing a parametric linear interpolation as in (3.6) yields

$$J(z,y) = \frac{1-100}{11} z + \frac{(100+y^2) + 10\,(1+y^2)}{11} = -9z + 10 + y^2.$$

Function J is convex on (z,y), linear on z, and satisfies the conditions of Proposition 1. We can thus conclude that it yields the associated split cut. In contrast, if we consider the non-elementary split $\pi = (1,1)^T$ with the previous choices of π_0 and π_1 on the same function G, we need to proceed with more care. In particular, the parametric interpolation (3.6) cannot be directly applied since the disjunction affects both z and y. However, we can construct the split cut by exploiting the fact that G can be represented as

$$G(z,y) = \frac{(z+y)^2}{2} + \frac{(z-y)^2}{2} = \frac{\left(\pi^T(z,y)\right)^2}{2} + \frac{\left(h^T(z,y)\right)^2}{2}, \tag{3.7}$$

where $h = (1,-1)^T$ is orthogonal to π. If we let $\tilde{z} = \pi^T(z,y)$, $\tilde{y} = h^T(z,y)$, $\tilde{\pi} = (1,0)$, $\tilde{\pi}_0 = -10$, $\tilde{\pi}_1 = 1$, and $\tilde{G}\,(\tilde{z},\tilde{y}) = \tilde{z}^2/2 + \tilde{y}^2/2$, we revert to the elementary case where we can apply the parametric interpolation (3.6) to obtain the split cut

$$\tilde{J}\,(\tilde{z},\tilde{y}) = \frac{\tilde{G}\,(\tilde{\pi}_1,\tilde{y}) - \tilde{G}\,(\tilde{\pi}_0,\tilde{y})}{\tilde{\pi}_1 - \tilde{\pi}_0}\tilde{z} + \frac{\tilde{\pi}_1 \tilde{G}\,(\tilde{\pi}_0,\tilde{y}) - \tilde{\pi}_0 \tilde{G}\,(\tilde{\pi}_1,\tilde{y})}{\tilde{\pi}_1 - \tilde{\pi}_0} = \frac{-9\tilde{z} + 10 + \tilde{y}^2}{2}. \tag{3.8}$$

We can then recover the split cut in the original (z,y) space by replacing the definitions of $\tilde{z}$ and $\tilde{y}$. The same procedure can be used for any separable function that is of, or can be converted to, the form $G(x) = f(\pi^T x) + g(P_\pi^\perp x)$ where $g : \mathbb{R}^n \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ are closed convex functions ($P_\pi^\perp x$ plays the same role as $h^T(z,y)$ in (3.7)).

To formally prove this, we first show how the friends construction procedure of Figure 1(b) can be extended to a general closed convex set C by considering properties of $\mathrm{lin}\,(C)$.

Proposition 2. *Let $F \subseteq \mathbb{R}^n$ be a split and $C \subseteq \mathbb{R}^n$ be a closed convex set. If there exists $u \in \mathrm{lin}\,(C)$ such that $\pi^T u \neq 0$, then condition (3.3) in Proposition 1 is satisfied.*

Proof. Let $\bar{x} \in C$ such that $\pi^T \bar{x} \in (\pi_0, \pi_1)$ and $u \in \mathrm{lin}\,(C)$ such that $\pi^T u \neq 0$. Also let $x^i := \bar{x} + \lambda_i u$ for $i \in \{0, 1\}$, where $\lambda_i = \frac{\pi_i - \pi^T \bar{x}}{\pi^T u}$, and let $\beta \in (0, 1)$ be such that $\pi^T \bar{x} = \beta \pi_0 + (1 - \beta)\, \pi_1$. Because $u \in \mathrm{lin}\,(C)$ and since $\pi^T x^i = \pi_i$, we have $x^i \in C \cap \mathrm{bd}\,(F)$ for $i \in \{0, 1\}$. The result then follows by noting that $\bar{x} = \beta x^0 + (1 - \beta)\, x^1$. □

Using Propositions 1 and 2 we obtain the following split cut formula for separable functions.

Proposition 3. *Let F be a split, $g : \mathbb{R}^n \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ be closed convex functions,*

$$S_{g,f} := \left\{ (x, t) \in \mathbb{R}^{n+1} : g(P_\pi^\perp x) + f(\pi^T x) \leq t \right\},$$

$a = \frac{f(\pi_1) - f(\pi_0)}{\pi_1 - \pi_0}$, and $b = \frac{\pi_1 f(\pi_0) - \pi_0 f(\pi_1)}{\pi_1 - \pi_0}$. Then $\overline{S_{g,f}} = S_{g,f} \cap C$, where

$$C = \left\{ (x, t) \in \mathbb{R}^{n+1} \;:\; g(P_\pi^\perp x) + a \pi^T x + b \leq t \right\}.$$

Proof. Interpolation condition (3.2) holds by the definition of a and b and convexity of f. Friends condition (3.3) follows from Proposition 2 by noting that $u = \left(\pi, a \left\|\pi\right\|_2^2\right) \in \mathrm{lin}\,(C)$ and $(\pi, 0)^T u \neq 0$. The result then follows from Proposition 1. □

3.2.1.2 Non-separable Positive Homogeneous Functions Proposition 1 can also be used to construct cuts for some non-separable functions, but as illustrated in the following example, we need slightly more complicated interpolations. Consider $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $G(z, y) = \sqrt{z^2 + y^2}$ and let F be the elementary split associated with $\pi = e^1$, $\pi_0 = -10$, and $\pi_1 = 1$. Constructing a parametric *linear* interpolation as in (3.6) yields

$$J_L(z, y) = \frac{10\sqrt{1+y^2} + \sqrt{100+y^2} + z\left(\sqrt{1+y^2} - \sqrt{100+y^2}\right)}{11}. \tag{3.9}$$

The associated cut is certainly valid, binding, and sufficient for $\overline{\text{epi}\,(G)}$ (we can always find friends by wiggling z toward π_0 and π_1, and using t to correct by following the slope of J_L for fixed y). However, while J is linear with respect to z, it is not convex with respect to y. We hence cannot use Proposition 1 for this interpolation. Fortunately, we can construct an alternative interpolation given by

$$J_C(z, y) = \sqrt{\left(\frac{20 - 9z}{11}\right)^2 + y^2} \tag{3.10}$$

that is convex on (z, y). This function is not linear on z for fixed y, but we can still show it satisfies the interpolation condition (3.2) by noting that $\left(\frac{20-9z}{11}\right)^2 \leq z^2$ for any $z \notin (\pi_0, \pi_1)$ and that equality holds for $z \in \{\pi_0, \pi_1\}$. This is illustrated in Figure 2 for $y = -4$ where the graphs of G, J_C, and J_L are given by the thick black curve, the thin blue curve, and the dash-dotted green line, respectively. The figure shows that $J_C(z, y) \leq t$ is a nonlinear binding valid cut, but is strictly weaker than $J_L(z, y) \leq t$. While J_C yields a weaker cut than J_L, J_C is in fact the strongest *convex* function that satisfies the interpolation condition (3.2) and we can show that $\overline{\text{epi}(G)} = \text{epi}(G) \cap \text{epi}(J_C)$. However, for the point $\left(\overline{z}, \overline{y}, \overline{t}\right) \in \text{epi}\,(J_C) \cap \text{int}\,(F)$ with $\overline{y} = -4$ depicted in Figure 2, the friends construction cannot be done by wiggling in a direction that leaves $\overline{y}$ fixed to -4. In other words, there are points in $\hat{H} := \{(z, y, t) \in \mathbb{R}^3 \,:\, y = -4\}$ that do not have friends in $\hat{H}$. We can construct friends by wiggling in a direction that does change $\overline{y}$, but since $\text{lin}(\text{epi}(J_C)) = \{0\}$, such a direction cannot be directly obtained from Proposition 2. Fortunately, the general idea of Proposition 2 can be adapted to obtain a variant that directly reveals an appropriate direction.

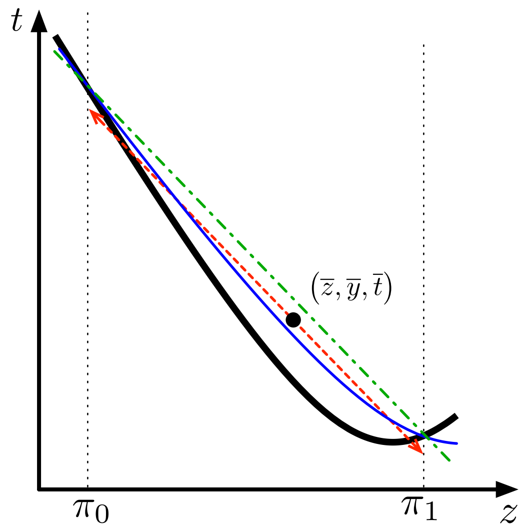


Figure 2: Nonlinear Interpolation for Non-separable Functions.

The variant of Proposition 2 that we need, exploits a different geometric characteristic of $\mathrm{epi}\,(J_C)$ through the generalization of a technique used in [8, 9]. The required geometric characteristic is given by the following definition.

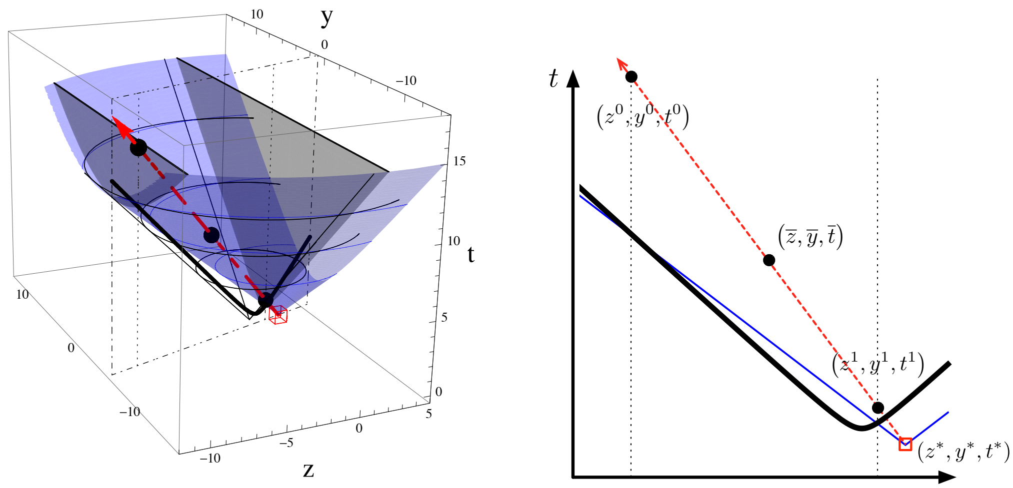
Definition 3. *Let $C \subseteq \mathbb{R}^n$ be a closed convex set. We say C is a* translated cone *or* conic set *if there exists $x^* \in C$ such that $C - x^*$ is a convex cone. We refer to such x^* as* an apex *of C, noting that it is not necessarily unique (e.g. a half space is a conic set whose apex is not unique).*

One can check that $\mathrm{epi}\,(J_C)$ is a conic set with the unique apex $(z^*, y^*, t^*) = (20/9, 0, 0)$. Hence, because $(\bar{z}, \overline{y}, \bar{t}) \in \mathrm{epi}\,(J_C)$, we have that the ray

$$R := \{(z^*, y^*, t^*) + \alpha\left((\bar{z}, \overline{y}, \bar{t}) - (z^*, y^*, t^*)\right) \,:\, \alpha \geq 0\} \subseteq \mathrm{epi}\,(J_C)\,. \tag{3.11}$$

Furthermore, because $z^* > \pi_1$ and $\bar{z} \in (\pi_0, \pi_1)$, there exists $\alpha_i > 0$ such that $z^* + \alpha_i(\bar{z} - z^*) = \pi_i$ for each $i \in \{0, 1\}$. Therefore the friends of $(\bar{z}, \overline{y}, \bar{t})$ are given by $(z^i, y^i, t^i) := (z^*, y^*, t^*) + \alpha_i\left((\bar{z}, \overline{y}, \bar{t}) - (z^*, y^*, t^*)\right)$ for $i \in \{0, 1\}$.

Figure 3 illustrates the ray-based friends construction for $(\bar{z}, \overline{y}, \bar{t})$ with $\overline{y} = -4$. Figure 3(a) shows the construction in the (z, y, t) space, while Figure 3(b) shows the section



(a) Construction in the (z, y, t) space. (b) Construction in the hyperplane $\tilde{H}$.

Figure 3: Friends Construction for Non-separable Positive Homogeneous Functions.

obtained by intersecting Figure 3(a) with the hyperplane $\tilde{H} := \operatorname{aff}(R \cup \{(0, 0, 1)\})$, for the ray R given in (3.11). The intersection of $\tilde{H}$ with the bounding box is depicted by the dash-dotted line in Figure 3(a). The graph of G is given by a black wire-frame in Figure 3(a), while the intersection of this graph with $\tilde{H}$ is given by the thick black curve in both figures. Meanwhile, the graph of J_C is depicted by the blue shaded region in Figure 3(a) and by a thin blue curve in Figure 3(b). The figures also depict (z^i, y^i, t^i) for $i \in \{0, 1\}$ and $(\bar{z}, \bar{y}, \bar{t})$ as black dots and (z^*, y^*, t^*) as a red box. In addition, the intersection of $z = \pi_i$ for $i \in \{0, 1\}$ with the epigraphs of both G and J_C are depicted in Figure 3(a) by the gray shaded regions. The intersection of $z = \pi_i$ for $i \in \{0, 1\}$ with $\tilde{H}$ are depicted in both figures by dotted lines. Finally, ray R is depicted in both figures as a red dashed arrow. Note that $\tilde{H}$ is tilted in the (z, y) space precisely to contain (z^*, y^*, t^*) and $(\bar{z}, \bar{y}, \bar{t})$. Noting that $y^* \neq \bar{y}$ we have that, unlike $\hat{H}$, $\tilde{H}$ allows the variation of y. Furthermore, while $(\bar{z}, \bar{y}, \bar{t}) \in \hat{H} \cap \tilde{H}$ might not have friends in $\hat{H}$, Figure 3 shows that it does have friends in $\tilde{H}$. Similarly to Proposition 2, the above construction can be extended to general convex sets as follows.

Proposition 4. *Let $F \subseteq \mathbb{R}^n$ be a split. If $C \subseteq \mathbb{R}^n$ is a conic set with apex $x^* \in \mathbb{R}^n$ such that $\pi^T x^* \notin (\pi_0, \pi_1)$, then condition (3.3) in Proposition 1 is satisfied.*

Proof. Let $\bar{x} \in C$ such that $\pi^T \bar{x} \in (\pi_0, \pi_1)$. Note that since x^* is the apex of C, all points on the ray $R := \{x^* + \alpha(\bar{x} - x^*) : \alpha \in \mathbb{R}_+\}$ belong to C. Let the intersections of R with the hyperplanes $\pi^T x = \pi_0$ and $\pi^T x = \pi_1$ be x^0 and x^1, respectively. Such points are obtained from R by setting $\alpha_i = \frac{\pi_i - \pi^T x^*}{\pi^T \bar{x} - \pi^T x^*}$, for $i \in \{0, 1\}$. We have $x^i \in C \cap \operatorname{bd}(F)$ for $i \in \{0, 1\}$, since $\pi^T x^i = \pi_i$ and $R \subseteq C$. Note that $\bar{x}$ is obtained from R by setting $\alpha = 1$. If $\alpha_0 < 1 < \alpha_1$ or $\alpha_1 < 1 < \alpha_0$, then there exists $\beta \in (0, 1)$ such that $\bar{x} = \beta x^0 + (1 - \beta) x^1$. Seeing that $\pi^T \bar{x} \in (\pi_0, \pi_1)$ and $\pi^T x^* \notin (\pi_0, \pi_1)$, one can check $\alpha_0 < 1 < \alpha_1$ or $\alpha_1 < 1 < \alpha_0$. □

Note that Propositions 2 and 4 ask for very different requirements on C. In Proposition 2, we only need to have a direction $u \in \operatorname{lin}(C)$ such that $\pi^T u \neq 0$. In such a case, C always defines a non-pointed region (i.e., C contains a line). On the other hand, as illustrated by (3.10), the sets C for which Proposition 4 is applicable are usually pointed (i.e. C has at least one extreme point). However, pointedness is not a requirement in Proposition 4 (e.g. half-spaces are conic sets). The real price of Proposition 4 over Proposition 2 is requiring C to be conic, which is a much more global requirement than asking for the lineality space of C to contain a non-orthogonal direction to π. However, both propositions are needed to construct split cuts for positive homogeneous functions. To see this, consider the same function $G(z, y) = \sqrt{z^2 + y^2}$ for which (3.10) yields a split cut, but instead consider the split $z \in [-1, 1]$. For this case, we can check that $\overline{\operatorname{epi}(G)} = \operatorname{epi}(G) \cap \operatorname{epi}(J_D)$ for $J_D(z, y) = \sqrt{1 + y^2}$, which does not have a conic epigraph. However, $(1, 0, 0) \in \operatorname{lin}(\operatorname{epi}(J_C))$ and hence Proposition 2 is applicable. This dichotomy between a non-pointed and a conic (and potentially pointed) cut will be a common occurrence that we highlight further when characterizing intersection cuts for conic quadratic sets in Section 3.3.

While Propositions 2 and 4 can be used to prove sufficiency of the split cuts for positive homogeneous functions, such cuts first have to be constructed with an appropriate interpolation technique. Fortunately, both interpolations of $G(z, y) = \sqrt{z^2 + y^2}$ (conic and non-pointed) can be generalized to functions based on p-norms by using the following simple lemma.

Lemma 1. *Let $p \in \mathbb{N}$, $\pi_0, \pi_1 \in \mathbb{R}$ such that $\pi_0 < \pi_1$, $l \in \mathbb{R}$, $a = \frac{(|l|^p + |\pi_1|^p)^{1/p} - (|l|^p + |\pi_0|^p)^{1/p}}{\pi_1 - \pi_0}$, and $b = \frac{\pi_1(|l|^p + |\pi_0|^p)^{1/p} - \pi_0(|l|^p + |\pi_1|^p)^{1/p}}{\pi_1 - \pi_0}$.*

- *If $s \in \{\pi_0, \pi_1\}$, then $|as+b|^p = |s|^p + |l|^p$ and*
- *if $s \notin (\pi_0, \pi_1)$, then $|as+b|^p \leq |s|^p + |l|^p$.*

Proof. We show the equivalent version of the lemma given by

1. If $s \in \{\pi_0, \pi_1\}$, then $|as+b| = (|s|^p + |l|^p)^{1/p}$ and
2. if $s \notin (\pi_0, \pi_1)$, then $|as+b| \leq (|s|^p + |l|^p)^{1/p}$.

Let $f(s) := as + b$ and $g(s) := (|s|^p + |l|^p)^{1/p}$. By definition of a and b we have that $f(\pi_i) = g(\pi_i)$ for $i \in \{0,1\}$. Indeed, $f(s)$ is the (affine) linear interpolation of $g(s)$ through $z = \pi_0$ and $z = \pi_1$. Convexity of $g(s)$ then implies $f(s) \leq g(s)$ for all $s \notin (\pi_0, \pi_1)$. If $|\pi_0| = |\pi_1|$, then $|as+b| = f(s)$ and the result follows directly. If $|\pi_0| \neq |\pi_1|$, one can check that $|as+b| = f(s)$ for $s \in [\pi_0, \pi_1]$ and hence (1) holds. For (2) it suffices to show that $-as - b \leq g(s)$ for all $s \in \mathbb{R}$. To show this we first assume $a > 0$ and hence $\pi_1 > 0$ (case $a < 0$ is analogous). Because $f(s)$ is affine and $f(\pi_i) = g(\pi_i)$ for $i \in \{0,1\}$, by a sub-differential version of the mean value theorem we have that there exists $\bar{s} \in (\pi_0, \pi_1)$ such that $a \in \partial g(\bar{s})$. Then, by symmetry of $g(s)$ and its convexity, we have that $g(s) \geq g(-\bar{s}) - a(s + \bar{s}) = -as + g(-\bar{s}) - a\bar{s}$ for $s \in \mathbb{R}$. The result then follows by noting that $g(-\bar{s}) - a\bar{s} \geq -b$ for all $\bar{s} \in (\pi_0, \pi_1)$ because $g(s) - as \geq 0$ for all $s \in \mathbb{R}$ and $-b \leq 0$. □

Using this lemma we can construct split cuts for epigraphs of a wide range of positive homogeneous convex functions and their sections (i.e. the epigraphs of such positive homogeneous functions after a variable is fixed to a constant).

Proposition 5. *Let F be a split, $\beta, l \in \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$ be a positive homogeneous closed convex function, a and b as in Lemma 1, and*

$$H_{p,g} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \left(g\left(P_\pi^\perp x\right)^p + \left|\beta \pi^T x\right|^p + |\beta l|^p \right)^{1/p} \leq t \right\}.$$

Then $\overline{H_{p,g}} = H_{p,g} \cap C$, where $C = \left\{ (x,t) \in \mathbb{R}^{n+1} : \left(g\left(P_\pi^\perp x\right)^p + \left|\beta\left(a\pi^T x + b\right)\right|^p \right)^{1/p} \leq t \right\}$.

Proof. Interpolation condition (3.2) holds by the definition of a and b and Lemma 1. If $|\pi_0| = |\pi_1|$, then $(\pi, 0) \in \operatorname{lin}(C)$ and friends condition (3.3) follows from Propositions 2. If $|\pi_0| \neq |\pi_1|$, then C is a conic set with apex $(x^*, t^*) = \left(\frac{-b}{a\|\pi\|_2^2}\pi, 0\right)$. Furthermore,

$$(\pi, 0)^T (x^*, t^*) = \pi^T x^* = \pi_1 + (|l|^p + |\pi_1|^p)^{1/p} \rho = \pi_0 + (|l|^p + |\pi_0|^p)^{1/p} \rho,$$

where $\rho = \frac{\pi_0 - \pi_1}{(|l|^p + |\pi_1|^p)^{1/p} - (|l|^p + |\pi_0|^p)^{1/p}}$. If $|\pi_1| < |\pi_0|$, then $\pi^T x^* \geq \pi_1$ and if $|\pi_1| > |\pi_0|$, then $\pi^T x^* \leq \pi_0$. Therefore, friends condition (3.3) follows from Proposition 4. The result then follows from Proposition 1. □

The following direct corollary of Proposition 5 yields simplified formulas for split cuts when $l = 0$ and $H_{p,g}$ is the epigraph of a positive homogeneous convex function.

Corollary 1. *Let F be a split, $\beta \in \mathbb{R}$, $p \in \mathbb{N}$, $g : \mathbb{R}^n \to \mathbb{R}$ be a positive homogeneous closed convex function, $a = \frac{\pi_0 + \pi_1}{\pi_1 - \pi_0}$, $b = -\frac{2\pi_1\pi_0}{\pi_1 - \pi_0}$, and*

$$C_{p,g} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \left(g\left(P_\pi^\perp x\right)^p + \left|\beta \pi^T x\right|^p \right)^{1/p} \leq t \right\}.$$

If $0 \notin (\pi_0, \pi_1)$, then $\overline{C_{p,g}} = C_{p,g}$. Otherwise, $\overline{C_{p,g}} = C_{p,g} \cap C$, where

$$C = \left\{ (x,t) \in \mathbb{R}^{n+1} : \left(g\left(P_\pi^\perp x\right)^p + \left|\beta\left(a\pi^T x + b\right)\right|^p \right)^{1/p} \leq t \right\}.$$

In particular, if g is a p-norm and the splits are elementary, Corollary 1 further specializes as follows.

Corollary 2. *Let F be an elementary split associated with $\pi = e^k$, $K_p := \{(x,t) \in \mathbb{R}^{n+1} : \|x\|_p \leq t\}$, a and b as in Corollary 1, and $\widehat{A} := I - e^k e^{kT}$. If $0 \notin (\pi_0, \pi_1)$, then $\overline{K_p} = K_p$. Otherwise, $\overline{K_p} = K_p \cap C$, where*

$$C = \left\{ (x,t) \in \mathbb{R}^{n+1} \;:\; \left\| \left(\widehat{A} + a e^k e^{kT} \right) x + b e^k \right\|_p \leq t \right\}.$$

Proof. Direct from Corollary 1 by noting that $K_p = \left\{ (x,t) \in \mathbb{R}^{n+1} \;:\; \left(\left\| \widehat{A} x \right\|_p^p + |x_k|^p \right)^{1/p} \leq t \right\}$, $C = \left\{ (x,t) \in \mathbb{R}^{n+1} \;:\; \left(\left\| \widehat{A} x \right\|_p^p + |a x_k + b|^p \right)^{1/p} \leq t \right\}$, and seeing that $\widehat{A} = P_\pi^\perp$. □

3.2.2 Split Cuts For Level Sets

The interpolation technique can also be applied to some non-epigraphical sets. This is illustrated in the following proposition.

Proposition 6. *Let F be a split, $g : \mathbb{R}^n \to \mathbb{R}$ be a positive homogeneous convex function, $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a closed convex function such that $f(\pi_0), f(\pi_1) \leq 0$, $a = \frac{f(\pi_1)-f(\pi_0)}{\pi_1-\pi_0}$, $b = \frac{\pi_1 f(\pi_0)-\pi_0 f(\pi_1)}{\pi_1-\pi_0}$ and*

$$L_{g,f} := \left\{x \in \mathbb{R}^n : g(P_\pi^\perp x) + f(\pi^T x) \leq 0\right\}.$$

Then $\overline{L_{g,f}} = L_{g,f} \cap C$, where $C = \left\{x \in \mathbb{R}^n \,:\, g(P_\pi^\perp x) + a\pi^T x + b \leq 0\right\}$.

Proof. Interpolation condition (3.2) holds by the definition of a and b and convexity of f. If $f(\pi_0) = f(\pi_1)$, then $\pi \in \operatorname{lin}(C)$ and friends condition (3.3) follows from Proposition 2. If $f(\pi_0) \neq f(\pi_1)$, then C is a conic set with apex $x^* = \frac{-b}{a\|\pi\|_2^2}\pi$. Furthermore, $\pi^T x^* = \frac{\pi_0 f(\pi_1)-\pi_1 f(\pi_0)}{f(\pi_1)-f(\pi_0)} = \pi_1 + \frac{(\pi_0-\pi_1)f(\pi_1)}{f(\pi_1)-f(\pi_0)} = \pi_0 + \frac{(\pi_0-\pi_1)f(\pi_0)}{f(\pi_1)-f(\pi_0)}$. If $f(\pi_0) < f(\pi_1)$, then $\pi^T x^* \geq \pi_1$ and if $f(\pi_0) > f(\pi_1)$ then $\pi^T x^* \leq \pi_0$. Therefore, friends condition (3.3) follows from Proposition 4. The result then follows from Proposition 1. □

As a direct corollary of Proposition 6, we obtain formulas for elementary split cuts for balls of p-norms.

Corollary 3. *Let F be an elementary split associated with $\pi = e^k$, $r \in \mathbb{R}$ such that $|\pi_0|, |\pi_1| \leq r$,*

$$E_p := \{x \in \mathbb{R}^n \; : \|x\|_p \leq r\},$$

$f(u) := -\left(r^p - |u|^p\right)^{1/p}$, a, b as in Proposition 6, and $\widehat{A} := I - e^k {e^k}^T$. Then $\overline{E_p} = E_p \cap C$, where

$$C = \left\{x \in \mathbb{R}^n : \left\|\widehat{A}x\right\|_p + ax_k + b \leq 0\right\}.$$

Proof. Direct from Proposition 6 by noting that $E_p = \left\{x \in \mathbb{R}^n \; : \left\|\widehat{A}x\right\|_p + f(x_k) \leq 0\right\}$ and $\widehat{A} = P_\pi^\perp$. □

3.2.3 Non-trivial Extensions

In this section we consider two non-trivial extensions/applications of the interpolation technique. The first example considers t-inclusive split cuts for epigraphical sets and illustrates the case when the interpolation coefficients cannot be easily calculated. The second example shows how the technique can be used beyond split sets to construct k-branch split cuts for epigraphical sets. We hope these examples serve as a guide for future applications or extensions of the interpolation technique.

3.2.3.1 t-inclusive Split Cuts for Epigraphical Sets Consider the base set $Q_0 := \{(x,t) \in \mathbb{R}^2 \,:\, x^2 \le t\}$ and the t-inclusive split $x + t \in [0,1]$. The first step to construct the associated split cut $C \subseteq \mathbb{R}^2$ such that $\overline{Q_0} = Q_0 \cap C$ is to find the general form of such a cut. The inclusion of t in the split prevents us from directly using the interpolation arguments for regular splits to construct this general form. However, by extrapolating these arguments to the t-inclusive setting and analyzing the geometry of the problem (e.g. the intersection of Q_0 with $x + t \in \{0,1\}$ corresponds to two ellipses), we may guess that the appropriate interpolation form is

$$C = \left\{(x,t) \in \mathbb{R}^2 \,:\, \sqrt{(ax+b)^2} \le cx + dt + e\right\}, \tag{3.12}$$

for some interpolation coefficients $a, b, c, d, e \in \mathbb{R}$. Unlike the regular split setting, it is not immediately clear what these coefficients should be, but we may try to deduce them by forcing interpolation conditions (3.2). Interpolation condition (3.2a) corresponds to

$$\{(x,t) \in Q_0 \,:\, t = -x\} \;=\; \{(x,t) \in C \,:\, t = -x\} \tag{3.13}$$

$$\{(x,t) \in Q_0 \,:\, t = 1 - x\} \;=\; \{(x,t) \in C \,:\, t = 1 - x\}, \tag{3.14}$$

which induces an infinite number of constraints on the coefficients.[2] We could try to reduce such a set of constraints to find the interpolation coefficients. In particular, the arguments for the regular splits effectively reduce such a set of constraints to two equality constraints. For instance, in the interpolation given in (3.1), the corresponding interpolation conditions

[2]For instance, (3.13) implies $\sqrt{(ax+b)^2} \le (c-d)\,x + e$ for all $(x,-x) \in Q_0$.

analogous to (3.13) and (3.14) reduce to $G(\pi_i) = J(\pi_i)$ for $i \in \{0,1\}$. To obtain a similar reduction, we here take a possibly naive approach that, nonetheless, is successful for several classes of cuts and is flexible enough to be extended to more complicated base and forbidden sets. The idea of this approach is to note that (3.13) and (3.14) can be expressed as

$$\left\{x \in \mathbb{R} : x^2 \leq -x\right\} = \left\{x \in \mathbb{R} : (ax+b)^2 \leq ((c-d)x+e)^2, (c-d)x+e \geq 0\right\} \tag{3.15}$$

$$\left\{x \in \mathbb{R} : x^2 \leq 1-x\right\} = \left\{x \in \mathbb{R} : (ax+b)^2 \leq ((c-d)x+d+e)^2, (c-d)x+d+e \geq 0\right\}. \tag{3.16}$$

A sufficient condition for these constraints is for the quadratic polynomials in both sides of (3.15) and (3.16) to be identical, and for the following condition to hold:

$$\{x \in \mathbb{R} : x^2 \leq -x\} \subseteq \{x \in \mathbb{R} : (c-d)x+e \geq 0\} \tag{3.17}$$

$$\{x \in \mathbb{R} : x^2 \leq 1-x\} \subseteq \{x \in \mathbb{R} : (c-d)x+d+e \geq 0\}. \tag{3.18}$$

Forcing the polynomials to be identical is a simple matter of matching coefficients, which results in the set of polynomial inequalities on a, b, c, d and e given by

$$a^2 - (c-d)^2 = 1, \quad ab - (c-d)e = 1/2,$$
$$ab - (c-d)(d+e) = 1/2, \quad b^2 - e^2 = 0,$$
$$b^2 - (d+e)^2 = -1.$$

The above linear system has four solutions given by:
$(1, \frac{1}{2}, \frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}-1}{2}, \frac{1}{2})$, $(1, \frac{1}{2}, \frac{-\sqrt{5}+1}{2}, \frac{-\sqrt{5}+1}{2}, \frac{-1}{2})$, $(1, \frac{1}{2}, \frac{\sqrt{5}+1}{2}, \frac{\sqrt{5}+1}{2}, \frac{-1}{2})$, and $(1, \frac{1}{2}, \frac{-\sqrt{5}-1}{2}, \frac{-\sqrt{5}-1}{2}, \frac{1}{2})$, of which only the first satisfies the additional conditions (3.17) and (3.18). Note that since $c = d$ in the first solution, checking (3.17) and (3.18) is equivalent to checking $e \geq 0$ and $d + e \geq 0$, which is trivial. Furthermore, this point also satisfies the interpolation condition (3.2b) which in this case, corresponds to

$$\{(x,t) \in Q_0 : x + t \notin (0,1)\} \subseteq \{(x,t) \in C : x + t \notin (0,1)\}. \tag{3.19}$$

Finally, to show that this choice of interpolation coefficients yields the desired split cut, note that C for such coefficients is a conic set with apex $(x^*, t^*) = \left(\frac{-1}{2}, \frac{\sqrt{5}-3}{2\sqrt{5}-2}\right)$ and $x^* + t^* < 0$. Then friends condition (3.3) follows from Proposition 4.

Note that identifying the coefficients of the quadratic polynomials and having (3.17) and (3.18) are sufficient for the interpolation condition (3.2a), but they may not be necessary in general. Hence, there might be other interpolation coefficients for which $\overline{Q_0} = Q_0 \cap C$. Moreover, it is not even clear that (3.12) is the only possible interpolation form for the associated split cut. However, if the described procedure is successful, we need not worry about alternative characterizations, since they will all yield $\overline{Q_0}$ when intersected with Q_0. There is of course no guarantee that the above procedure for finding a representation of C will always succeed. However, as we illustrate in Section 3.3, the procedure is successful in constructing rather complicated cuts for conic quadratic sets.

3.2.3.2 k-branch Split Cuts for Epigraphical Sets We now illustrate how Proposition 1 can be used for sets other than splits by constructing certain k-branch split cuts for separable functions. The following proposition is a direct, but technical, generalization of Proposition 3, which explains our reason to postpone its introduction to this stage of the paper.

Proposition 7. *Let $g : \mathbb{R} \to \mathbb{R}$ and $f_i : \mathbb{R} \to \mathbb{R}$ for each $i \in [k]$ be closed convex functions. Furthermore, let F be a k-branch split such that $\pi_i \perp \pi_j$ for every $i \neq j$. Finally, let $P_\Pi^\perp := I - \sum_{i=1}^k \frac{\pi_i \pi_i^T}{\|\pi_i\|_2^2}$,*

$$B_{g,f} := \left\{ (x,t) \in \mathbb{R}^{n+1} \,:\, g\left(P_\Pi^\perp x\right) + \sum_{i=1}^k f_i\left(\pi_i^T x\right) \le t \right\},$$

$a_i := \frac{f_i(\pi_1^i) - f_i(\pi_0^i)}{\pi_1^i - \pi_0^i}$, $b_i := \frac{\pi_1^i f_i(\pi_0^i) - \pi_0^i f_i(\pi_1^i)}{\pi_1^i - \pi_0^i}$ for all $i \in [k]$, and for every $\mathcal{I} \subseteq [k]$ let

$$h_\mathcal{I}(x) := g\left(P_\Pi^\perp x\right) + \sum_{i \in [k] \setminus \mathcal{I}} f_i\left(\pi_i^T x\right) + \sum_{i \in \mathcal{I}} a_i \pi_i^T x + b_i.$$

Then $\overline{B_{g,f}} = B_{g,f} \cap C$, where $C = \left\{ (x,t) \in \mathbb{R}^{n+1} \,:\, \max_{\mathcal{I} \subseteq [k]} h_\mathcal{I}(x) \le t \right\}$.

Proof. Interpolation condition (3.2) holds by the definition of a_i and b_i and convexity of f_i. Now let $(\bar{x}, \bar{t}) \in C \cap \operatorname{int}(F)$. To construct the friends of $(\bar{x}, \bar{t})$ we proceed as follows.

Let $\mathcal{I} \subseteq [k]$ be such that for all $i \in \mathcal{I}$ we have $\pi_i^T \bar{x} \in (\pi_0^i, \pi_1^i)$, and for all $i \in [k] \setminus \mathcal{I}$ we have $\pi_i^T \bar{x} \notin (\pi_0^i, \pi_1^i)$. For each $s \in \{0,1\}^{\mathcal{I}}$, let

$$x^s = P_\Pi^\perp \bar{x} + \sum_{i \in [k] \setminus \mathcal{I}} \frac{\pi_i^T \bar{x}}{\|\pi_i\|_2^2} \pi_i + \sum_{i \in \mathcal{I}} \frac{s_i \pi_0^i + (1 - s_i)\pi_1^i}{\|\pi_i\|_2^2} \pi_i, \quad t^s = \bar{t} + \sum_{i \in \mathcal{I}} a_i \left(s_i \pi_0^i + (1 - s_i)\pi_1^i - \pi_i^T \bar{x} \right), \tag{3.20}$$

and $\lambda_s = \prod_{i \in \mathcal{I}} \left(s_i \frac{\pi_1^i - \pi_i^T \bar{x}}{\pi_1^i - \pi_0^i} + (1 - s_i) \frac{\pi_i^T \bar{x} - \pi_0^i}{\pi_1^i - \pi_0^i} \right)$.
Note that $(\bar{x}, \bar{t}) = \sum_{s \in \{0,1\}^{\mathcal{I}}} \lambda_s (x^s, t^s)$, $\sum_{s \in \{0,1\}^{\mathcal{I}}} \lambda_s = 1$, and $\lambda_s \geq 0$ for all $s \in \{0,1\}^{\mathcal{I}}$. Furthermore, by construction and the assumption on $\mathcal{I}$, we have that $x^s \in \operatorname{bd}(F)$ and $(x^s, t^s) \in \operatorname{epi}(h_{\mathcal{I}})$ for all $s \in \{0,1\}^{\mathcal{I}}$. The result then follows from Proposition 1 by noting that for all $s \in \{0,1\}^{\mathcal{I}}$, we have $\max_{\mathcal{J} \subseteq [k]} h_{\mathcal{J}}(x^s) = h_{\mathcal{I}}(x^s)$. □

3.3 INTERSECTION CUTS FOR CONIC QUADRATIC SETS

In this section we consider intersection cuts for conic quadratic sets of the form $\mathcal{C} := \{x \in \mathbb{R}^n : Ax - d \in L^m\}$ where $A \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$, and L^m is the m-dimensional Lorentz cone. Note that $\mathcal{C}$ can be written as

$$\mathcal{C} = \left\{ x \in \mathbb{R}^n \,:\, \|A_0 x - d_0\|_2 \leq a_m^T x - d_m \right\}, \tag{3.21}$$

where (A_0, d_0) is obtained from (A, d) by deleting the m-th row, and (a_m, d_m) is the m-th row of (A, d). Using (3.21), one can rewrite $\mathcal{C}$ as

$$\mathcal{Q} := \left\{ x \in \mathbb{R}^n \,:\, x^T Q x - 2h^T x + \rho \leq 0, \quad a_m^T x - d_m \geq 0 \right\},$$

where $Q = A_0^T A_0 - a_m a_m^T$, $h = A_0^T d_0 - a_m d_m$, and $\rho = d_0^T d_0 - d_m^2$. Also note that $Q \in \mathbb{R}^{n \times n}$ is symmetric with at most one negative eigenvalue. Using known classifications of sets described by a quadratic inequality with at most one negative eigenvalue (e.g. see Table 2.1 and the reasoning after the proof of Lemma 2.1 in [15]), we have that all conic quadratic sets of the form $\mathcal{C}$ correspond to the following list:

1. A full dimensional paraboloid,
2. a full dimensional ellipsoid (or a single point),
3. a full dimensional second-order cone,
4. one side of a full dimensional hyperboloid of two sheets,
5. a cylinder generated by a lower-dimensional version of one of the previous sets, or
6. an invertible affine transformation of one of the previous sets.

We first consider split cuts for conic quadratic sets with simple structures that can be obtained as direct corollaries of Propositions 3, 5, and 6. We then consider t-inclusive and k-branch split cuts for conic quadratic sets that require ad-hoc proofs based on Proposition 1. As expected, we see that split cut formulas are significantly simpler than those for t-inclusive and k-branch split cuts. However, in either case, it is crucial to exploit the symmetry of the Euclidean norm through the following standard lemma.

Lemma 2. *For $v \in \mathbb{R}^n$, $\|x\|_2^2 = \|P_v x\|_2^2 + \|P_v^\perp x\|_2^2$.*

To give formulas for split cuts for all the sets 1–6, it suffices to consider cases 1–4. With these, we can construct split cut formulas for cylinders using the following lemma.

Lemma 3. *Let $B \subseteq \mathbb{R}^n$ be a closed convex set of the form $B_0 + L$ where L is a linear subspace, and let $F \subseteq \mathbb{R}^n$ be a split. If $\pi \in L^\perp$ and $\operatorname{conv}(B_0 \setminus \operatorname{int}(F)) = B_0 \cap C$, then $\operatorname{conv}(B \setminus \operatorname{int}(F)) = (B_0 \cap C) + L$. If $\pi \notin L^\perp$, then $\operatorname{conv}(B \setminus \operatorname{int}(F)) = B$.*

Proof. We first prove the second case $\pi \notin L^\perp$. The left to right containment follows from $B \setminus \operatorname{int}(F) \subseteq B$ and convexity of B. To show the right to left containment, let $\bar{x} \in B$ such that $\pi^T \bar{x} \in (\pi_0, \pi_1)$ and $u \in L$. Note that $\pi \notin L^\perp$ implies $\pi^T u \neq 0$. Let $x^i := \bar{x} + \lambda_i u$ for $i \in \{0,1\}$, where $\lambda_i = \frac{\pi_i - \pi^T \bar{x}}{\pi^T u}$, and let $\beta \in (0,1)$ be such that $\pi^T \bar{x} = \beta \pi_0 + (1-\beta)\pi_1$. Because $u \in L$ and since $\pi^T x^i = \pi_i$, we have $x^i \in B \setminus \operatorname{int}(F)$ for $i \in \{0,1\}$. The results then follows by noting that $\bar{x} = \beta x^0 + (1-\beta) x^1$.

We prove the first case by showing that

$$\begin{aligned} \operatorname{conv}(B \setminus \operatorname{int}(F)) &= \operatorname{conv}((B_0 + L) \setminus \operatorname{int}(F)) && (3.22) \\ &= \operatorname{conv}(B_0 \setminus \operatorname{int}(F)) + L && (3.23) \\ &= (B_0 \cap C) + L. && (3.24) \end{aligned}$$

Note that (3.22) and (3.24) follow from the assumptions. To show the left to right containment in (3.23), let $\bar{x} \in \operatorname{conv}((B_0 + L) \setminus \operatorname{int}(F))$. There exist $y^i \in B_0$, $u^i \in L$ for $i \in \{0, 1\}$, and $\beta \in [0, 1]$ such that for $x^i := y^i + u^i$, we have $x^i \notin \operatorname{int}(F)$ and $\bar{x} = \beta x^0 + (1 - \beta) x^1$. Note that $\pi \in L^\perp$ and $x^i \notin \operatorname{int}(F)$ imply $y^i \notin \operatorname{int}(F)$ for $i \in \{0, 1\}$. The result then follows from noting that $\beta y^0 + (1 - \beta) y^1 \in \operatorname{conv}(B_0 \setminus \operatorname{int}(F))$ and $\beta u^0 + (1 - \beta) u^1 \in L$.

To show the right to left containment in (3.23), let $\bar{x} \in \operatorname{conv}(B_0 \setminus \operatorname{int}(F)) + L$. There exist $u \in L$, $y^i \in B_0 \setminus \operatorname{int}(F)$ for $i \in \{0, 1\}$, and $\beta \in [0, 1]$ such that $\bar{x} = \beta y^0 + (1 - \beta) y^1 + u$. If $\beta \in \{0, 1\}$, the result follows by noting that $\pi \in L^\perp$ and $y^0, y^1 \notin \operatorname{int}(F)$ imply $\bar{x} \notin \operatorname{int}(F)$. Assume $\beta \in (0, 1)$ and let $x^0 := y^0 + \frac{u}{2\beta}$ and $x^1 := y^1 + \frac{u}{2(1-\beta)}$. The result then follows by noting that $x^i \in B_0 + L \setminus \operatorname{int}(F)$ for $i \in \{0, 1\}$ and $\bar{x} = \beta x^0 + (1 - \beta) x^1$. □

Finaly, we can construct split cut formulas for affine transformations by using the following straightforward lemma.

Lemma 4. *Let $B \subseteq \mathbb{R}^n$ be a closed convex set, $F \subseteq \mathbb{R}^n$ be a split, and $M : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible affine mapping. If $\operatorname{conv}(B \setminus \operatorname{int}(F)) = B \cap C$ for a closed convex set $C \subseteq \mathbb{R}^n$, then*

$$\operatorname{conv}(M(B) \setminus \operatorname{int}(M(F))) = M(B) \cap M(C).$$

We note that classification 1–6 is not strictly necessary for constructing split cuts for quadratic sets. In particular, an algorithm introduced in [94] can be used to obtain an SDP representation of split cuts for any quadratic set (convex or not) without a priori classifying its specific geometry as in 1–6. However, the procedure in [94] requires the execution of a numerical algorithm to construct split cuts and does not provide closed form expressions of the cuts. Furthermore, such an algorithm requires elaborate algebraic tools specific to quadratic sets that go far beyond a basic property such as that described by Lemma 2. Hence, the objective of the following subsection is not to present the shortest possible constructions of all quadratic split cuts, but to (i) present simple proofs tailored to the specific geometries in classification 1–6 and (ii) present a case study on the power and limitations of the general interpolation approach to split cuts.

3.3.1 Split Cuts for Conic Quadratic Sets

Split cuts can be obtained for ellipsoids when interpreted as lower level sets of quadratic or conic functions (i.e., based on the Euclidean norm). Similarly, split cuts can also be characterized for paraboloids and cones that, when interpreted as epigraphs of quadratic or conic functions, are such that t is unaffected by the split disjunctions. We note that the ellipsoid case has already been proven on [14, 37], and that the conic case generalizes Proposition 2 in [9] which considers elementary disjunctions for the standard three dimensional second-order cone. Through the rest of the section, we let $A \in \mathbb{R}^{n\times n}$ be an invertible matrix and $c \in \mathbb{R}^n$.

Corollary 4 (Split cuts for paraboloids)**.** *Let F be a split, $Q := \{(x,t) \in \mathbb{R}^{n+1} : \|A\,(x-c)\|_2^2 \leq t\}$, $a = \frac{\pi_0+\pi_1-2\pi^T c}{\|A^{-T}\pi\|_2^2}$, $b = -\frac{\left(\pi_1-\pi^T c\right)\left(\pi_0-\pi^T c\right)}{\|A^{-T}\pi\|_2^2}$, and $\widehat{A} = P^{\perp}_{A^{-T}\pi} A$. Then $\overline{Q} = Q \cap C$, where*

$$C = \left\{(x,t) \in \mathbb{R}^{n+1} : \left\|\widehat{A}\,(x-c)\right\|_2^2 + a\pi^T\,(x-c) + b \leq t\right\}.$$

Proof. Note that for the affine mappings M, M^{-1} given by $M(x,t) := (A\,(x-c)\,,t)$ and $M^{-1}(x,t) := (A^{-1}x + c, t)$, we have $Q = M^{-1}\,(Q_0)$ and $Q_0 = M\,(Q)$, where $Q_0 := \left\{(x,t) \in \mathbb{R}^{n+1} : \|x\|_2^2 \leq t\right\}$. Using Lemma 4, we prove the corollary by finding a closed form expression for $\overline{Q_0}$ where the forbidden set is the split $M\,(F)$ associated with $\tilde{\pi} = A^{-T}\pi$, $\tilde{\pi}_0 = \pi_0 - \pi^T c$, and $\tilde{\pi}_1 = \pi_1 - \pi^T c$. By Lemma 2, we have $Q_0 = \{(x,t) \in \mathbb{R}^{n+1} : \|P^{\perp}_{\tilde{\pi}} x\|_2^2 + \frac{(\tilde{\pi}^T x)^2}{\|\tilde{\pi}\|_2^2} \leq t\}$. The result then follows from Proposition 3. □

Corollary 5 (Split cuts for cones)**.** *Let F be a split, $K := \{(x,t) \in \mathbb{R}^{n+1} : \|A\,(x-c)\|_2 \leq t\}$, $a = \frac{\pi_1+\pi_0-2\pi^T c}{\pi_1-\pi_0}$, $b = \frac{-2\left(\pi_1-\pi^T c\right)\left(\pi_0-\pi^T c\right)}{\pi_1-\pi_0}$, $\widehat{A} = \left(P^{\perp}_{A^{-T}\pi} + aP_{A^{-T}\pi}\right) A$, $\widehat{c} = \left(b/\left\|A^{-T}\pi\right\|_2^2\right) A^{-T}\pi$. If $\pi^T c \notin (\pi_0, \pi_1)$, then $\overline{K} = K$. Otherwise, $\overline{K} = K \cap C$, where*

$$C = \left\{(x,t) \in \mathbb{R}^{n+1} : \left\|\widehat{A}\,(x-c) + \widehat{c}\right\|_2 \leq t\right\}. \tag{3.25}$$

Proof. Note that for the affine mappings M, M^{-1} defined in the proof of Corollary 4 we have $K = M^{-1}\,(K_0)$ and $K_0 = M\,(K)$, where $K_0 := \{(x,t) \in \mathbb{R}^{n+1} : \|x\|_2 \leq t\}$. Using Lemma 4, we prove the corollary by finding a closed form expression for $\overline{K_0}$ where the forbidden set is the split $M\,(F)$ defined in the proof of Corollary 4. By Lemma 2, we have $K_0 = \left\{(x,t) \in \mathbb{R}^{n+1} : \left(\|P^{\perp}_{\tilde{\pi}} x\|_2^2 + \frac{(\tilde{\pi}^T x)^2}{\|\tilde{\pi}\|_2^2}\right)^{1/2} \leq t\right\}$. The result then follows from Corollary 1. □

A particularly interesting application of Corollaries 4 and 13 is the Closest Vector Problem [71], which can alternatively be written as

$$\min\left\{\|A(x-c)\|_2^2 \,:\, x\in\mathbb{Z}^n\right\} \text{ or } \min\left\{\|A(x-c)\|_2 \,:\, x\in\mathbb{Z}^n\right\}.$$

In turn, these problems can be reformulated as

$$\min\{t \,:\, (x,t)\in Q,\, x\in\mathbb{Z}^n\} \quad \text{and} \quad \min\{t \,:\, (x,t)\in K,\, x\in\mathbb{Z}^n\},$$

respectively. We can then use Corollaries 4 and 13 with lattice free splits to construct split cuts that could improve the solution speed of these problems. We are currently studying the effectiveness of such cuts.

We can also obtain as a corollary the following result from [14, 37].

Corollary 6 (Split cuts for ellipsoids). *Let F be a split, $r\in\mathbb{R}_+$,*

$$E := \{x\in\mathbb{R}^n : \|A(x-c)\|_2 \le r\},$$

$f(u) := -\sqrt{r^2 - \frac{u^2}{\|A^{-T}\pi\|_2^2}}$, $a = \frac{f(\pi_0-\pi^Tc)-f(\pi_1-\pi^Tc)}{\pi_1-\pi_0}$, *and* $b = \frac{(\pi_1-\pi^Tc)f(\pi_0-\pi^Tc)-(\pi_0-\pi^Tc)f(\pi_1-\pi^Tc)}{\pi_1-\pi_0}$.

If $\pi^Tc - r\|A^{-T}\pi\|_2 \le \pi_0 < \pi_1 \le \pi^Tc + r\|A^{-T}\pi\|_2$, then $\overline{E} = E\cap C$, where

$$C = \left\{x\in\mathbb{R}^n : \|P^{\perp}_{A^{-T}\pi}A(x-c)\|_2 \le a\pi^T(x-c)-b\right\}, \tag{3.26}$$

if $\pi_0 < \pi^Tc - r\|A^{-T}\pi\|_2 < \pi_1 \le \pi^Tc + r\|A^{-T}\pi\|_2$, then

$$\overline{E} = \left\{x\in E : \pi^Tx \ge \pi_1\right\}, \tag{3.27}$$

if $\pi^Tc - r\|A^{-T}\pi\|_2 \le \pi_0 < \pi^Tc + r\|A^{-T}\pi\|_2 < \pi_1$, then

$$\overline{E} = \left\{x\in E : \pi^Tx \le \pi_0\right\}, \tag{3.28}$$

if $\pi^Tc - r\|A^{-T}\pi\|_2 \ge \pi_1$ or $\pi_0 \ge \pi^Tc + r\|A^{-T}\pi\|_2$, then $\overline{E} = E$, and otherwise, $\overline{E} = \emptyset$.

Proof. All the cases except the first one can be shown by studying when the ellipsoid is partially or completely contained in one side of the disjunction, or when it is strictly contained between the disjunction.

We now prove the first case. Note that for the affine mappings M, M^{-1} given by $M(x) := A(x-c)$ and $M^{-1}(x) := A^{-1}x + c$, we have $E = M^{-1}(E_0)$ and $E_0 = M(E)$, where $E_0 := \{(x,t) \in \mathbb{R}^{n+1} : \|x\|_2 \leq r\}$. Using Lemma 4, we prove the corollary by finding a closed form expression for $\overline{E_0}$ where the forbidden set is the split $M(F)$ associated with $\tilde{\pi} = A^{-T}\pi$, $\tilde{\pi}_0 = \pi_0 - \pi^T c$, and $\tilde{\pi}_1 = \pi_1 - \pi^T c$. By Lemma 2, we have $E_0 = \left\{x \in \mathbb{R}^n : \|P^{\perp}_{\tilde{\pi}} x\|_2 - \sqrt{r^2 - \frac{(\tilde{\pi}^T x)^2}{\|\tilde{\pi}\|_2^2}} \leq 0\right\}$. The result then follows from Proposition 6. □

We note that Corollary 6 shows there are two types of split cuts for E. In (3.26), we obtain a nonlinear split cut that we would expect from Proposition 6, while in (3.27)–(3.28) we obtain simple linear split cuts. These linear inequalities are actually Chvátal-Gomory (CG) cuts for E [29, 35, 36, 43, 47], but they are still sufficient to describe $\overline{E}$ together with the original constraint. We hence follow the same MILP convention used in [37] and still consider them split cuts. Note that we can also consider "CG split cuts" in Proposition 6 if we include additional structure on the functions such as g being non-negative. Similarly, we can also do the case analysis for CG cuts in Corollary 3.

Proposition 8 (Split cuts for hyperboloids)**.** *Let F be a split, $l \in \mathbb{R} \setminus \{0\}$,*

$$H := \left\{(x,t) \in \mathbb{R}^{n+1} : \sqrt{\|x\|_2^2 + l^2} \leq t\right\},$$

$a = \frac{\sqrt{l^2\|\pi\|_2^2 + \pi_1^2} - \sqrt{l^2\|\pi\|_2^2 + \pi_0^2}}{\pi_1 - \pi_0}$*, and* $b = \frac{\pi_1\sqrt{l^2\|\pi\|_2^2 + \pi_0^2} - \pi_0\sqrt{l^2\|\pi\|_2^2 + \pi_1^2}}{\pi_1 - \pi_0}$*. Then $\overline{H} = H \cap C$, where*

$$C = \left\{(x,t) \in \mathbb{R}^{n+1} : \left\|P^{\perp}_{\pi} x + \frac{a\pi^T x + b}{\|\pi\|_2^2}\pi\right\|_2 \leq t\right\}.$$

Proof. Direct from Proposition 5 by noting that

$$H = \{(x,t) \in \mathbb{R}^{n+1} : \sqrt{\|P^{\perp}_{\pi} x\|_2^2 + \frac{(\pi^T x)^2}{\|\pi\|_2^2} + l^2 \leq t}\}.$$

□

3.3.2 t-inclusive Split Cuts for Conic Quadratic Sets

The split cut formulas in this section are significantly more complicated. For this reason, we only present them for standard sets (i.e., with $A = I$ and $c = 0$). Formulas for the general case may be obtained by combining the formulas for the standard case with Lemma 4.

Proposition 9. (t-inclusive split cuts for paraboloids) *Let F be a t-inclusive split and*

$$Q_0 := \{(x,t) \in \mathbb{R}^{n+1} : \|x\|_2^2 \leq t\}.$$

If $\hat{\pi} > 0$ and $\pi_1 \leq \frac{-\|\pi\|_2^2}{4\hat{\pi}}$, or if $\hat{\pi} < 0$ and $\frac{-\|\pi\|_2^2}{4\hat{\pi}} \leq \pi_0$, then

$$\overline{Q_0} = Q_0,$$

if $\hat{\pi} > 0$ and $\pi_0 < \frac{-\|\pi\|_2^2}{4\hat{\pi}} < \pi_1$, then

$$\overline{Q_0} = \left\{(x,t) \in Q_0 : \pi^T x + \hat{\pi} t \geq \pi_1\right\},$$

if $\hat{\pi} < 0$ and $\pi_0 < \frac{-\|\pi\|_2^2}{4\hat{\pi}} < \pi_1$, then

$$\overline{Q_0} = \left\{(x,t) \in Q_0 : \pi^T x + \hat{\pi} t \leq \pi_0\right\},$$

and if $\hat{\pi} > 0$ and $\frac{-\|\pi\|_2^2}{4\hat{\pi}} \leq \pi_0$, or if $\hat{\pi} < 0$ and $\pi_1 \leq \frac{-\|\pi\|_2^2}{4\hat{\pi}}$, then $\overline{Q_0} = Q_0 \cap C$, where

$$C = \left\{(x,t) \in \mathbb{R}^{n+1} : \left\| P_\pi^\perp x + \frac{\pi^T x + b}{\|\pi\|_2^2}\pi \right\|_2 \leq c\pi^T x + dt + e \right\},$$

for

$$b = \frac{\|\pi\|_2^2}{2\hat{\pi}}, \quad c = \frac{f}{\sqrt{2}\left(\pi_1 - \pi_0\right)\hat{\pi}}, \quad d = c\hat{\pi}, \quad e = \frac{\|\pi\|_2^2 + \sqrt{\|\pi\|_2^2 + 4\pi_0\hat{\pi}}\sqrt{\|\pi\|_2^2 + 4\pi_1\hat{\pi}}}{4\sqrt{2}\left(\pi_1 - \pi_0\right)\hat{\pi}^2} f,$$
$$f = \sqrt{\|\pi\|_2^2 + 2\left(\pi_0 + \pi_1\right)\hat{\pi} - \sqrt{\|\pi\|_2^2 + 4\pi_0\hat{\pi}}\sqrt{\|\pi\|_2^2 + 4\pi_1\hat{\pi}}},$$

where we use the convention $0/0 := 0$ for the case $\|\pi\|_2 = 0$.

Proof. We first prove the last case using Proposition 1. Using Lemma 2, we have

$$C = \left\{ (x,t) \in \mathbb{R}^{n+1} \,:\, \left\| P_\pi^\perp x \right\|_2^2 \leq (c\pi^T x + dt + e)^2 - \frac{(\pi^T x + b)^2}{\|\pi\|_2^2}, \quad c\pi^T x + dt + e \geq 0 \right\}. \tag{3.29}$$

Now consider the following two cases.

Case 1. Assume that $\|\pi\|_2 \neq 0$. To prove the right to left containment in (3.2a), let $(\bar{x}, \bar{t}) \in C \cap \mathrm{bd}\,(F)$. We need to show that

$$(c\pi^T \bar{x} + d\bar{t} + e)^2 - \frac{(\pi^T \bar{x} + b)^2}{\|\pi\|_2^2} = \bar{t} - \frac{(\pi^T \bar{x})^2}{\|\pi\|_2^2}. \tag{3.30}$$

Replacing $\bar{t}$ with $(\pi_i - \pi^T \bar{x})/\hat{\pi}$ for $i \in \{0, 1\}$, one can check that (3.30) follows from the definition of b, c, d, and e. To prove the left to right containment in (3.2a), let $(\bar{x}, \bar{t}) \in Q_0 \cap \mathrm{bd}\,(F)$. We only need to show that $c\pi^T \bar{x} + d\bar{t} + e \geq 0$. Since $d = c\hat{\pi}$, we can equivalently show that $c(\pi^T \bar{x} + \hat{\pi}\bar{t}) \geq -e$, which after a few simplifications, can be written as

$$\hat{\pi}(\pi^T \bar{x} + \hat{\pi}\bar{t}) \geq -\left(\|\pi\|_2^2 + \sqrt{\|\pi\|_2^2 + 4\pi_0\hat{\pi}}\sqrt{\|\pi\|_2^2 + 4\pi_1\hat{\pi}} \right)/4. \tag{3.31}$$

(3.31) follows from noting that $\min\{\hat{\pi}(\pi^T x + \hat{\pi}t) \,:\, (x,t) \in Q_0\} = -\frac{\|\pi\|_2^2}{4}$.

To show (3.2b), let $(\bar{x}, \bar{t}) \in Q_0 \setminus \mathrm{int}\,(F)$. Proving $c\pi^T \bar{x} + d\bar{t} + e \geq 0$ is similar as before. We only need to show that $(\bar{x}, \bar{t})$ satisfies the quadratic inequality in (5.2), which we prove by showing that

$$\left((c\pi^T \bar{x} + d\bar{t} + e)^2 - \frac{(\pi^T \bar{x} + b)^2}{\|\pi\|_2^2} \right) - \left(\bar{t} - \frac{(\pi^T \bar{x})^2}{\|\pi\|_2^2} \right) \geq 0. \tag{3.32}$$

One can check that proving (3.32) is equivalent to showing that

$$\frac{f^2(\pi^T \bar{x} + \hat{\pi}\bar{t} - \pi_0)(\pi^T \bar{x} + \hat{\pi}\bar{t} - \pi_1)}{2\,(\pi_1 - \pi_0)^2\,\hat{\pi}^2} \geq 0,$$

which follows from $\pi^T \bar{x} + \hat{\pi}\bar{t} \notin (\pi_0, \pi_1)$. Note that C is a conic set with apex $(x^*, t^*) = (\frac{-b}{\|\pi\|_2^2}\pi, \frac{bc-e}{d})$. Furthermore,

$$(\pi, \hat{\pi})^T\,(x^*, t^*) = -e/c = \frac{-\|\pi\|_2^2}{4\hat{\pi}} - \frac{\sqrt{\|\pi\|_2^2 + 4\pi_0\hat{\pi}}\sqrt{\|\pi\|_2^2 + 4\pi_1\hat{\pi}}}{4\hat{\pi}}.$$

Hence, if $\hat{\pi} < 0$, then $(\pi, \hat{\pi})^T (x^*, t^*) \geq \frac{-\|\pi\|_2^2}{4\hat{\pi}} \geq \pi_1$ and if $\hat{\pi} > 0$, then $(\pi, \hat{\pi})^T (x^*, t^*) \leq \frac{-\|\pi\|_2^2}{4\hat{\pi}} \leq \pi_0$. Friends condition (3.3) then follows from Proposition 4.

Case 2. If $\|\pi\|_2 = 0$, C is simplified to $C = \left\{(x,t) \in \mathbb{R}^{n+1} \,:\, \|x\|_2^2 \leq (dt+e)^2, \quad dt + e \geq 0\right\}$.

Interpolation condition (3.2a) follows from noting that $(d\bar{t} + b)^2 = \bar{t}$. Non-negativity of d, e, and t also imply $d\bar{t} + e \geq 0$. Proving (3.2b) is equivalent to showing that

$$\frac{f^2 \left(\hat{\pi}\bar{t} - \pi_0\right)\left(\hat{\pi}\bar{t} - \pi_1\right)}{2\left(\pi_1 - \pi_0\right)^2 \hat{\pi}^2} \geq 0,$$

which follows from $\hat{\pi}\bar{t} \notin (\pi_0, \pi_1)$. Note that C is a conic set with apex $(x^*, t^*) = \left(0, \frac{-e}{d}\right)$. Furthermore, $(\pi, \hat{\pi})^T (x^*, t^*) = -e/c$. As shown in Case 1, we have $(\pi, \hat{\pi})^T (x^*, t^*) \notin (\pi_0, \pi_1)$. Friends condition (3.3) then follows from Proposition 4.

To prove the other cases, let $S_0 := \{(x,t) \in Q_0 \,:\, \pi^T x + \hat{\pi} t \leq \pi_0\}$ and $S_1 := \{(x,t) \in Q_0 \,:\, \pi^T x + \hat{\pi} t \geq \pi_1\}$. Consider the first case where $\hat{\pi} > 0$ and $\pi_1 \leq \frac{-\|\pi\|_2^2}{4\hat{\pi}}$. We prove the result by showing that $S_0 = \emptyset$ and $S_1 = Q_0$. If $\|\pi\|_2 = 0$, the result follows from non-negativity of t. Now assume that $\|\pi\|_2 \neq 0$. We prove $S_0 = \emptyset$ by showing that $(\pi^T x)^2 / \|\pi\|_2^2 > (\pi_0 - \pi^T x)/\hat{\pi}$ for $x \in \mathbb{R}^n$. This follows from noting that for $y \in \mathbb{R}$, the quadratic equation $\frac{y^2}{\|\pi\|_2^2} = \frac{\pi_0 - y}{\hat{\pi}}$ does not have any solution. To prove $S_1 = Q_0$, we show that $\pi^T x + \hat{\pi} t \geq \pi_1$ is a valid inequality for Q_0. This comes from the fact that the quadratic equation $\frac{y^2}{\|\pi\|_2^2} = \frac{\pi_1 - y}{\hat{\pi}}$ has at most a single solution and we thus have $(\pi_1 - \pi^T x)/\hat{\pi} \leq (\pi^T x)^2 / \|\pi\|_2^2 \leq t$ for $x \in \mathbb{R}^n$. The proof for the case $\hat{\pi} < 0$ and $\frac{-\|\pi\|_2^2}{4\hat{\pi}} \leq \pi_0$ is analogous and follows by noting that $S_0 = Q_0$ and $S_1 = \emptyset$.

We prove the second case where $\hat{\pi} > 0$ and $\pi_0 < \frac{-\|\pi\|_2^2}{4\hat{\pi}} < \pi_1$ by showing that $S_0 = \emptyset, S_1 \subsetneq Q_0$, and $S_1 \neq \emptyset$. Proving $S_0 = \emptyset$ is analogous to the previous case. We have $S_1 \subsetneq Q_0$ since $(\bar{x}, \bar{t}) = (\frac{-\pi}{2\hat{\pi}}, \frac{\|\pi\|_2^2}{4\hat{\pi}^2}) \in Q_0$, but $(\bar{x}, \bar{t}) \notin S_1$. To prove $S_1 \neq \emptyset$, one can check that for any $\bar{x} \in \mathbb{R}^n$ and $\bar{t} = \text{Max}\{\|\bar{x}\|_2^2, \frac{\pi_1 - \pi^T \bar{x}}{\hat{\pi}}\}$, $(\bar{x}, \bar{t}) \in S_1$. The proof of the third case is analogous and follows by noting that $S_1 = \emptyset, S_0 \subsetneq Q_0$, and $S_0 \neq \emptyset$. $\square$

Proposition 10. (t-inclusive split cuts for cones) *Let F be a t-inclusive split and*

$$K_0 := \{(x,t) \in \mathbb{R}^{n+1} : \|x\|_2 \leq t\}.$$

If $0 \notin (\pi_0, \pi_1)$, then $\overline{K_0} = K_0$. Otherwise, if $0 \in (\pi_0, \pi_1)$ and $\hat{\pi} \leq -\|\pi\|_2$, then

$$\overline{K_0} = \left\{(x,t) \in K_0 : \pi^T x + \hat{\pi} t \leq \pi_0\right\},$$

if $0 \in (\pi_0, \pi_1)$ *and* $\hat{\pi} \geq \|\pi\|_2$, *then*

$$\overline{K_0} = \left\{(x,t) \in K_0 : \pi^T x + \hat{\pi} t \geq \pi_1\right\},$$

and if $0 \in (\pi_0, \pi_1)$ *and* $\hat{\pi} \in (-\|\pi\|_2, \|\pi\|_2)$, *then* $\overline{K_0} = K_0 \cap C$, *where*

$$C = \left\{(x,t) \in \mathbb{R}^{n+1} : \left\| P_\pi^\perp x + \frac{a\pi^T x + b}{\|\pi\|_2^2}\pi \right\|_2 \leq c\pi^T x + dt + e\right\},$$

where

$$\begin{aligned}
a &= \frac{(\pi_0 + \pi_1)\left(\|\pi\|_2^2 - \hat{\pi}^2\right)}{f}, \quad b = -\frac{2\pi_0\pi_1 \|\pi\|_2^2}{f}, \\
c &= -\frac{4\pi_0\pi_1\hat{\pi}}{(\pi_1 - \pi_0) f}, \quad d = \frac{f}{(\pi_1 - \pi_0)\left(\|\pi\|_2^2 - \hat{\pi}^2\right)}, \\
e &= \frac{2\pi_0\pi_1(\pi_0 + \pi_1)\hat{\pi}}{(\pi_1 - \pi_0) f}, \quad f = \sqrt{\left(\|\pi\|_2^2 - \hat{\pi}^2\right)\left(\|\pi\|_2^2 (\pi_1 - \pi_0)^2 - \hat{\pi}^2 (\pi_0 + \pi_1)^2\right)}.
\end{aligned}$$

Proof. We first prove the last case using Proposition 1. Note that $\hat{\pi} \neq 0$ and $\hat{\pi} \in (-\|\pi\|_2, \|\pi\|_2)$ imply $\|\pi\|_2 \neq 0$. Using Lemma 2, we have

$$C = \left\{(x,t) \in \mathbb{R}^{n+1} : \left\|P_\pi^\perp x\right\|_2^2 \leq (c\pi^T x + dt + e)^2 - \frac{(a\pi^T x + b)^2}{\|\pi\|_2^2}, \quad c\pi^T x + dt + e \geq 0\right\}. \tag{3.33}$$

Observe that $d > 0$. Similarly to the proof of Proposition 9, one can show that interpolation condition (3.2) holds by the definition of a, b, c, d, and e. If $|\pi_0| = |\pi_1|$, then $u = (\pi, \frac{-c\|\pi\|_2^2}{d}) \in \text{lin}(C)$ and friends condition (3.3) follows from Proposition 2. If $|\pi_0| \neq |\pi_1|$, then C is a conic set with apex $(x^*, t^*) = (\frac{-b}{a\|\pi\|_2^2}\pi, \frac{bc-ae}{ad})$. Furthermore, $(\pi, \hat{\pi})^T (x^*, t^*) = \frac{2\pi_0\pi_1}{\pi_0+\pi_1}$. If $\pi_0 + \pi_1 < 0$, then we have $\frac{2\pi_0\pi_1}{\pi_0+\pi_1} \geq \pi_1$, and if $\pi_0 + \pi_1 > 0$, then we have $\frac{2\pi_0\pi_1}{\pi_0+\pi_1} \leq \pi_0$. Friends condition (3.3) then follows from Proposition 4.

To prove the first case $0 \notin (\pi_0, \pi_1)$, we only need to show that friends condition (3.3) holds. This follows from Proposition 4 by noting that K_0 is a conic set whose apex is the origin.

To prove the other cases, let $S_0 := \{(x,t) \in K_0 : \pi^T x + \hat{\pi} t \leq \pi_0\}$ and $S_1 := \{(x,t) \in K_0 : \pi^T x + \hat{\pi} t \geq \pi_1\}$. Consider the second case $0 \in (\pi_0, \pi_1)$ and $\hat{\pi} \leq -\|\pi\|_2$. We prove the result by showing that $S_1 = \emptyset$, $S_0 \subsetneq K_0$, and $S_0 \neq \emptyset$. If $\|\pi\|_2 = 0$, the result follows from non-negativity of t. Now assume that $\|\pi\|_2 \neq 0$. We prove $S_1 = \emptyset$ by showing that

$(\pi^T x)^2 / \|\pi\|_2^2 > (\pi_1 - \pi^T x)^2/\hat{\pi}^2$. Note that non-negativity of t, $\hat{\pi} < 0$, and $\pi^T x + \hat{\pi} t \geq \pi_1$ imply $\pi^T x \geq \pi_1 > 0$. One can see that $-\pi^T x < \pi_1 - \pi^T x < \pi^T x$, where the first inequality comes from the fact that $\pi_1 > 0$, and the second inequality follows from $\pi_1 \leq \pi^T x$ and $-\pi^T x < 0$. Thus, $(\pi^T x)^2 > (\pi_1 - \pi^T x)^2$ and the result follows by noting that $\frac{1}{\|\pi\|_2^2} \geq \frac{1}{\hat{\pi}^2}$. We have $S_0 \subsetneq K_0$ since $(\bar{x}, \bar{t}) = (0_n, 0) \in K_0$, but $(\bar{x}, \bar{t}) \notin S_0$. To prove $S_0 \neq \emptyset$, one can check that for any $\bar{x} \in \mathbb{R}^n$ and $\bar{t} = \text{Max}\{\|\bar{x}\|_2, \frac{\pi_0 - \pi^T \bar{x}}{\hat{\pi}}\}$, $(\bar{x}, \bar{t}) \in S_0$. The proof of the third case $0 \in (\pi_0, \pi_1)$ and $\hat{\pi} \geq \|\pi\|_2$ is analogous and follows by noting that $S_0 = \emptyset$, $S_1 \subsetneq K_0$, and $S_1 \neq \emptyset$. $\square$

With regards to the general interpolation forms of the obtained split cuts in Sections 4.1 and 4.2, we note that these fall into two categories. The first category corresponds to the case in which the intersection of the boundary of the split and the base set is bounded such as when the base set is an ellipsoid. In such a case, the obtained split cuts are always an ellipsoidal cylinder or a conic set. The second category corresponds to the case in which the intersection of the boundary of the split and the base set is unbounded. In such a case, the obtained split cut is of the same form as the base set. For instance, split cuts for conic sets or sections of conic sets are conic. A nice illustration of this dichotomy is the case of paraboloids, where t-inclusive splits have bounded intersections and yield conic cuts, while splits that are not t-inclusive have unbounded intersections and yield parabolic cuts.

Finally, we note that the only formulas that we did not explicitly characterize here are t-inclusive split cuts for affine transformations of paraboloids and cones, split cuts for affine transformation of hyperboloids, and t-inclusive split cuts for hyperboloids and their affine transformations. All such formulas can be obtained using Lemma 4, except t-inclusive split cuts for hyperboloids. We can still obtain formulas for t-inclusive split cuts for hyperboloids using the interpolation technique; however, the resulting formulas are significantly more involved and no longer fit the "simple" formulas theme of the paper. However, the analysis so far is still a significant generalization of what is known for split cuts for conic quadratic sets. In fact, the most general alternative that we are aware of is the concurrently developed technique in [6], which considers conic sets of the form $\{x \in \mathbb{R}^n : Ax - d \in L^m\}$ for a full rank matrix A, which we do not require. When A does not have full row rank, it is possible to consider a full row rank submatrix of A and use this relaxation to generate the cuts from [6]. However, as noted in Example 1 of [6], this approach fails to give split cuts for hyperboloids which we can obtain from Proposition 8 and Lemma 4. Nevertheless, one advantage of the approach in [6] is the use of a more systematic procedure to obtain the interpolation coefficients, which can be particularly useful when constructing t-inclusive split cuts. For instance, in Proposition 10 we obtain the interpolation coefficients through the heuristic procedure described in Section 3.2.3.1, which required guessing the interpolation form of the split cut and was not guaranteed to be successful even if this guess was accurate. In contrast, the approach in [6] only assumes that the split cut is a polynomial inequality and calculates the coefficients of the associated polynomial through a systematic use of techniques from algebraic geometry. The conversion of this polynomial inequality to a conic quadratic inequality is an ad-hoc procedure that might be limited to quadratic cones. However, the construction of the initial polynomial inequality seems to have a higher chance of being extended to higher order cones or more general semi-algebraic sets than the approach in Section 3.2.3.1. In contrast, when we consider split disjunctions that are not t-inclusive, the approach from Section 3.2.1.2 has an advantage as it is not restricted to semi-algebraic sets.

3.3.3 k-branch Split Cuts for Conic Quadratic Sets

Similarly to Corollary 4, we can use the following direct generalization of Lemma 2 to get formulas for several families of k-branch split cuts for convex quadratic sets.

Lemma 5. *Let $\{\pi_i\}_{i=1}^k \subseteq \mathbb{R}^n \setminus \{0_n\}$ be such that $\pi_i \perp \pi_j$ for every $i \neq j$ and $P_\Pi^\perp := I - \sum_{i=1}^k \frac{\pi_i \pi_i^T}{\|\pi_i\|_2^2}$. Then for any $v \in \mathbb{R}^n$ we have $\|x\|_2^2 = \left\|P_\Pi^\perp x\right\|_2^2 + \sum_{i=1}^k \frac{\left(\pi_i^T x\right)^2}{\|\pi_i\|_2^2}$.*

The following corollary generalizes the result of Corollary 4 to the case of k-branch split cuts for paraboloids.

Corollary 7 (k-branch split cuts for paraboloids)**.** *Let*

$$Q := \left\{(x,t) \in \mathbb{R}^{n+1} : \|A\left(x-c\right)\|_2^2 \leq t\right\}.$$

Also let F be a k-branch split such that $A^{-T}\pi_i \perp A^{-T}\pi_j$ for every $i \neq j$, $a_i = \frac{\pi_0^i + \pi_1^i - 2\pi_i^T c}{\|A^{-T}\pi_i\|_2^2}$ and $b_i = -\frac{\left(\pi_1^i - \pi_i^T c\right)\left(\pi_0^i - \pi_i^T c\right)}{\|A^{-T}\pi_i\|_2^2}$ for all $i \in [k]$, and for every $\mathcal{I} \subseteq [k]$ let

$$h_{\mathcal{I}}(x) := \left\| \left(A - \sum_{i \in \mathcal{I}} \frac{A^{-T}\pi_i \pi_i^T}{\|A^{-T}\pi_i\|_2^2} \right)(x - c) \right\|_2^2 + \sum_{i \in \mathcal{I}} a_i \pi_i^T (x - c) + b_i.$$

Then $\overline{Q} = Q \cap C$, where $C = \left\{(x,t) \in \mathbb{R}^{n+1} \,:\, \max_{\mathcal{I} \subseteq [k]} h_{\mathcal{I}}(x) \leq t\right\}$.

Proof. Note that for the affine mappings M, M^{-1} defined in the proof of Corollary 4 we have $Q = M^{-1}(Q_0)$ and $Q_0 = M(Q)$, where $Q_0 := \left\{(x,t) \in \mathbb{R}^{n+1} \,:\, \|x\|_2^2 \leq t\right\}$. Using Lemma 4, we prove the corollary by finding a closed form expression for $\overline{Q_0}$ where the forbidden set is a k-branch split $M(F)$ associated with $\tilde{\pi}_i = A^{-T}\pi_i$, $\tilde{\pi}_0^i = \pi_0^i - \pi_i^T c$, and $\tilde{\pi}_1^i = \pi_1^i - \pi_i^T c$ for $i \in [k]$. By Lemma 5, we have

$$Q_0 = \left\{ (x,t) \in \mathbb{R}^{n+1} : \left\| P_{\tilde{\Pi}}^{\perp} x \right\|_2^2 + \sum_{i=1}^{k} \frac{\left(\tilde{\pi}_i^T x\right)^2}{\|\tilde{\pi}_i\|_2^2} \leq t \right\}.$$

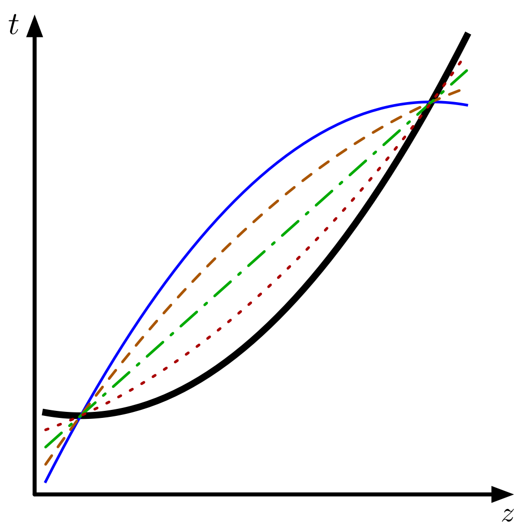
The result then follows from Proposition 7. □

3.4 GENERAL INTERSECTION CUTS THROUGH AGGREGATION

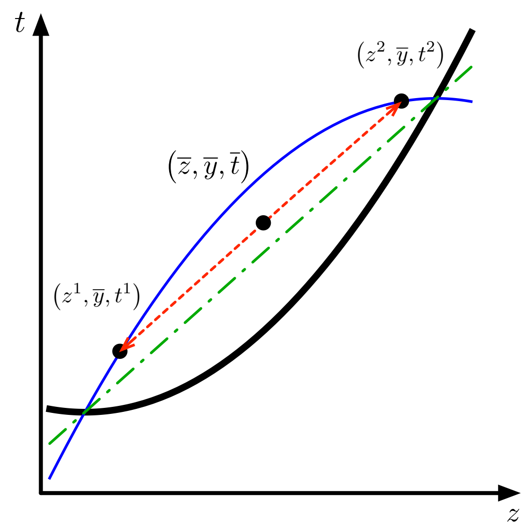
In this section we consider the case in which the base sets are either epigraphs or lower level sets of convex functions and the forbidden sets are hypographs or upper level sets of concave functions. Our cut construction approach in this case is based on a simple aggregation technique, which again can be more naturally explained for epigraphs of specially structured functions. Following the structure of Section 3.2, we also begin by studying the epigraphical sets and then consider the case of non-epigraphical sets. We end this section by illustrating the power and limitations of the aggregation approach by considering intersection cuts for quadratic sets.

3.4.1 Intersection Cuts for Epigraphs

Let $G, J : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a convex and a concave function given by $G(z, y) = z^2 + 2y^2$ and $J(z, y) = -(z-1)^2 + 1 - y^2$, and let $B = \text{epi}\,(G)$ and $F = \text{hyp}(J)$. For $\lambda \in [0, 1]$, let $W_\lambda(z, y) = (1-\lambda)G + \lambda J$. As illustrated in Figure 4(a), for any $\lambda \in [0, 1]$, we have that $W_\lambda(z, y) \le t$ is a binding valid cut for $\overline{B}$. In Figure 4(a), the graph of G is given by the thick black curve, graph of J by the thin blue curve, and valid aggregation cuts W_λ for $\lambda \in \{1/4, 1/2, 3/4\}$ by the red dotted, green dash-dotted, and brown dashed curves, respectively. Figure 4(a) illustrates that, depending on the choice of λ, the inequality could be non-convex, or it could be convex but not sufficient. It is clear from the figure that, in this case, the correct choice of λ is $1/2 = \arg\max\{\lambda \in [0, 1] : W_\lambda \text{ is convex}\}$, which yields the strongest convex cut from this class. Furthermore, as illustrated in Figure 4(b), we have that for any $(\bar{z}, \bar{y}, \bar{t}) \in \text{epi}(W_{1/2}) \cap \text{int}\,(F)$, we can find friends in $\text{epi}\,(W_{1/2}) \cap \text{bd}\,(F)$ by following the slope of $W_{1/2}$ similar to what we did in Section 3.2.1.1 for split cuts of separable functions. We can then show that $\overline{B} = B \cap \text{epi}\,(W_{1/2})$. A similar construction can also be obtained if we instead study $\text{conv}\,(\{(z, y, t) \in \text{epi}(G) : J(z, y) \le 0\})$.



(a) Various aggregations of G and J.



(b) Friends construction by following slope of $W_{1/2}$.

Figure 4: Cuts From Aggregation.

W_λ and the convexity requirement on it are the basis of many techniques such as Lagrangian/SDP relaxations of quadratic programming problems [45, 77, 79, 80], the QCR method for integer quadratic programming [20, 21], and an algorithm for constructing projected SDP representations of the convex hull of quadratic constraints introduced in [94]. It is hence not surprising that the approach works in the quadratic case. However, as shown in [94], even in the quadratic case the approach can fail to yield convex constraints or closed form expressions. Furthermore, for general functions, W_λ can easily be non-convex for every λ. Fortunately, as the following proposition shows, the aggregation approach can yield closed form expressions for general intersection cuts for problems with special structures.

Proposition 11. *Let $g_i : \mathbb{R} \to \mathbb{R}$ be convex functions for each $i \in [n]$, $m, h \in \mathbb{R}^n$, $r, q \in \mathbb{R}$, and $\gamma \in \mathbb{R}_+$. Furthermore, let $\{a_i\}_{i=1}^n \subseteq \mathbb{R}^n$ be such that $a_n \neq 0_n$ and $a_i \perp a_j$ for every $i \neq j$, and $\{\alpha_i\}_{i=1}^n \subseteq \mathbb{R}_+$ be such that $0 \neq \alpha_n \geq \alpha_i$ for all i. Let*

$$G(x) = \sum_{i=1}^n g_i\left(a_i^T x\right) + m^T x + r, \quad J(x) = -\sum_{i=1}^n \alpha_i g_i\left(a_i^T x\right) - h^T x - q,$$

$B := \operatorname{epi}(G)$, and $F := \{(x,t) \in \mathbb{R}^{n+1} \,:\, \gamma t \leq J(x)\}$. If $(1 + \gamma/\alpha_n) > 0$ and

$$\lim_{|s|\to\infty} -\alpha_n g_n\left(s a_n^T a_n\right) - s\left(h^T a_n + \gamma \frac{(m - h/\alpha_n)^T a_n}{1 + \gamma/\alpha_n}\right) = -\infty, \tag{3.34}$$

then

$$\overline{B} = \operatorname{conv}\left(\{(x,t) \in \operatorname{epi}(G) \,:\, J(x) \leq \gamma t\}\right) = \operatorname{epi}(G) \cap \operatorname{epi}(W), \tag{3.35}$$

where $W(x) := \frac{G(x) + (1/\alpha_n)J(x)}{1+\gamma/\alpha_n} = \frac{\sum_{i=1}^{n-1}(1-\alpha_i/\alpha_n) g_i\left(a_i^T x\right) + (m - h/\alpha_n)^T x + (r - q/\alpha_n)}{(1+\gamma/\alpha_n)}$.

Proof. The first equality in (3.35) is direct. For the second equality, we proceed as follows. W is a non-negative linear combination of G and J that is also a convex function from which it is easy to see that the left to right containment holds.

To show the right to left containment, let $\left(\overline{x}, \overline{t}\right) \in \operatorname{epi}(G) \cap \operatorname{epi}(W)$ be such that $J\left(\overline{x}\right) > \gamma\overline{t}$. Let $k = \frac{(m-h/\alpha_n)^T a_n}{1+\gamma/\alpha_n}$. Because of (3.34), there exists $s_1 > 0$ and $s_2 < 0$, for which $(x^i, t^i) = \left(\overline{x} + s_i a_n, \overline{t} + s_i k\right)$ for $i = 1, 2$ are such that $J(x^i) = \gamma t^i$. Furthermore, by design, $(x^i, t^i) \in \operatorname{epi}(W)$ for $i = 1, 2$ which implies $G(x^i) + J(x^i)/\alpha_n \leq (1 + \gamma/\alpha_n)\, t^i$ and hence $G(x^i) \leq t^i$. The result then follows by noting that $\left(\overline{x}, \overline{t}\right) \in \operatorname{conv}\left(\{(x^1, t^1), (x^2, t^2)\}\right)$. □

3.4.2 Intersection Cuts for Level Sets

We can extend the aggregation approach to certain non-epigraphical sets through the following proposition whose proof is a direct analog to that of Proposition 11.

Proposition 12. *Let $G(x)$ be as defined in Proposition 11 and*

$$J(x) = -\sum_{i=1}^{n} \alpha_i g_i \left(a_i^T x\right) - \alpha_n m^T x - q,$$

where $q \in \mathbb{R}$. Also let $B := \{x \in \mathbb{R}^n : G(x) \leq 0\}$, and $F := \{x \in \mathbb{R}^n : J(x) \geq 0\}$. If

$$\lim_{|s| \to \infty} -\alpha_n g_n \left(s a_n^T a_n\right) - s\alpha_n m^T a_n = -\infty, \tag{3.36}$$

then

$$\overline{B} = \operatorname{conv}\left(\{x \in \mathbb{R}^n : G(x) \leq 0, \quad J(x) \leq 0\}\right) = \{x \in \mathbb{R}^n : G(x) \leq 0, \quad W(x) \leq 0\}, \tag{3.37}$$

where $W(x) := G(x) + (1/\alpha_n)J(x) = \sum_{i=1}^{n-1} \left(1 - \alpha_i/\alpha_n\right) g_i(a_i^T x) + (r - q/\alpha_n)$.

The special structure in both of these propositions is extremely simple, but thanks to the symmetry of the quadratic constraints, they can be used to get formulas for several quadratic intersection cuts.

3.4.3 Intersection Cuts for Quadratic Sets

Through the rest of the section, we let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix and $c \in \mathbb{R}^n$.

Corollary 8. *Let $D \in \mathbb{R}^{n\times n}$, $d \in \mathbb{R}^n$, $q \in \mathbb{R}$, $\gamma \in \mathbb{R}_+$, $Q := \{(x,t) \in \mathbb{R}^{n+1} : \|A(x-c)\|_2^2 \leq t\}$, and $F := \left\{(x,t) \in \mathbb{R}^{n+1} : \gamma t + q \leq -\|D(x-d)\|_2^2\right\}$. Then*

$$\overline{Q} = \left\{(x,t) \in \mathbb{R}^{n+1} : \|A(x-c)\|_2^2 \leq t, \quad x^T N x + a^T x + f \leq (\alpha_n + \gamma)t\right\}, \tag{3.38}$$

for $N = A^T R A$, $a = -2A^T e - 2A^T R A c$, $f = c^T A^T R A c + 2\left(A^T e\right)^T c - w - q$, $R = \sum_{i=1}^{n-1} (\alpha_n - \alpha_i) v_i v_i^T$, $e = \sum_{i=1}^{n} \alpha_i v_i^T A(c-d) v_i$, $w = \sum_{i=1}^{n} \alpha_i \left(v_i^T A(c-d)\right)^2$, where $(v_i)_{i=1}^n \subseteq \mathbb{R}^n$ and $(\alpha_i)_{i=1}^n \subseteq \mathbb{R}$ correspond to an eigenvalue decomposition of $A^{-T}D^T D A^{-1}$ so that

$$A^{-T} D^T D A^{-1} = \sum_{i=1}^{n} \alpha_i v_i v_i^T,$$

$\|v_i\|_2 = 1$ for all $i \in [n]$, $v_i^T v_j = 0$ for all $i \neq j$, and $\alpha_n \geq \alpha_i$ for all $i \in [n]$.

Proof. Let $y = A(x - c)$ and $T := Q \setminus \operatorname{int}(F)$. Using orthonormality of the vectors v_i, T can be written on the y variables as $T = \{(y,t) \in \mathbb{R}^{n+1} : \sum_{i=1}^{n} (v_i^T y)^2 \leq t,\ -\sum_{i=1}^{n} \alpha_i (v_i^T y)^2 - 2e^T y - w - q \leq \gamma t\}$. The result then follows by using Proposition 11. □

An interesting case of Corollary 8 arises when $\gamma = 0$. In this case, the base set B corresponds to a paraboloid and the forbidden set F corresponds to an ellipsoidal cylinder. In such a case, the minimization of t over $(x,t) \in B \setminus \operatorname{int}(F)$ is equivalent to the minimization of a convex quadratic function outside an ellipsoid, which corresponds to the simplest indefinite version of the well known trust region problem. While this is a non-convex optimization problem, it can be solved in polynomial time through Lagrangian/SDP approaches [79]. It is known that optimal dual multipliers of an SDP relaxation of a non-convex quadratic programming problem such as the trust region problem can be used to construct a finite convex quadratic optimization problem with the same optimal value as the original non-convex problem (e.g. [46]). Furthermore, the complete feasible region induced by an SDP relaxation on the original space (in this case (x,t)) can be characterized by an infinite number of convex quadratic constraints [60]. This characterization has recently been simplified for the feasible region of the trust region problem in [18, 19]. This work gives a semi-infinite characterization of T for $\gamma = 0$ composed by the convex quadratic constraint $\|A\,(x-c)\|_2^2 \leq t$ plus an infinite number of linear inequalities that can be separated in polynomial time. Corollary 8 shows that these linear inequalities can be subsumed by a single convex quadratic constraint, which gives another explanation for their polynomial time separability[3]. We note that the techniques in [18, 19] are also adapted to other non-convex optimization problems (both quadratic and non-quadratic). Hence, combining Corollary 8 with these techniques could yield valid convex quadratic inequalities for more general non-convex problems.

[3]After our original submission, it was brought to our attention that reduction of the infinite number of inequalities to a single quadratic inequality can also be directly deduced from the formulas for such linear inequalities given in [18, 19].

Another interesting application of Corollary 8 for the case $\gamma = 0$ is the Shortest Vector Problem (SVP) [71] of the form $\min \left\{ \|Ax\|_2^2 \,:\, x \in \mathbb{Z}^n \setminus \{0_n\} \right\}$. Similar to the Closest Vector Problems (CVP) studied in Section 3.3.1, we can transform this problem to $\min_{(x,t)\in Y\cap(\mathbb{Z}^n\times\mathbb{R})} t$ for

$$Y = \left\{ (x,t) \in \mathbb{R}^{n+1} \,:\, \|Ax\|_2^2 \leq t,\, x \neq 0_n \right\},$$

so that we can strengthen the problem by generating valid inequalities for Y. Unfortunately, as the following simple lemma shows, split cuts will not add any strength.

Lemma 6. *Let $Y_0 := Y \cup \{(0_n, 0)\}$ and F be a split. For any $A \in \mathbb{R}^{n\times n}$,*

$$t^* = \min \left\{ t \,:\, (x,t) \in \cap_{(\pi,\pi_0)\in\mathbb{Z}^n\times\mathbb{Z}} \overline{Y_0} \right\} = 0.$$

Proof. Note that for all integer splits $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$, $(\bar{x}, \bar{t}) = (0_n, 0)$ belongs to one side of the disjunction. Thus, we have $t^* \leq 0$ and the result follows from non-negativity of the norm. □

However, we can easily construct *near* lattice free ellipsoids centered at 0_n that do not contain any point from $\mathbb{Z}^n \setminus \{0_n\}$ in their interior, and use them to get some bound improvement. For instance, in the trivial case of $A = I$, Corollary 8 applied to the single *near* lattice free ellipsoid given by the unit ball $\{x \in \mathbb{R}^n \,:\, \|x\|_2 \leq 1\}$ yields a cut that provides the optimal value $t^* = 1$. Similar ellipsoids could be used to generate strong valid convex quadratic inequalities for non-trivial cases to significantly speed up the solution of SVP problems. Studying the effectiveness of these cuts is left for future research.

We end this section with a brief discussion about the strength and possible extensions of the aggregation technique. For this, we begin by presenting the following corollary of Proposition 12 whose proof is analogous to that of Corollary 8.

Corollary 9. *Let $D \in \mathbb{R}^{n\times n}$, $r_1, r_2 \in \mathbb{R}_+$, $E^2 := \{x \in \mathbb{R}^n : \|A\,(x-c)\,\|_2^2 \leq r_1\}$, and*

$$F := \left\{ x \in \mathbb{R}^n \,:\, \|D\,(x-c)\|_2^2 \leq r_2 \right\}.$$

Then

$$\overline{E^2} = \left\{ x \in \mathbb{R}^n : \|A\,(x-c)\|_2^2 \leq r_1, \quad x^T N x + a^T x + f \leq 0 \right\}, \tag{3.39}$$

for $a = -2A^T RAc$, $f = c^T A^T RAc + r_2/\alpha_n - r_1$, $R = \sum_{i=1}^{n-1} (1 - \alpha_i/\alpha_n) v_i v_i^T$, *where* N *is defined in Corollary 8, and* $(v_i)_{i=1}^n \subseteq \mathbb{R}^n$ *and* $(\alpha_i)_{i=1}^n \subseteq \mathbb{R}$ *correspond to an eigenvalue decomposition of* $A^{-T} D^T D A^{-1}$ *given in Corollary 8.*

Corollary 9 shows how to construct the convex hull of the set obtained by removing an ellipsoid or an ellipsoidal cylinder from an ellipsoid. However, this construction only works if the ellipsoids have a common center c. The following example shows how the construction can fail for non-common centers. In addition, the following example shows that the aggregation technique does not subsume the interpolation technique and sheds some light into the relationship between Corollaries 8 and 9 and SDP relaxations for quadratic programming.

Example 1. *Let* $B = \{(z, y) \in \mathbb{R}^2 : z^2 + y^2 \leq 4\}$ *and* F *be a split associated with the split disjunction* $z \leq 0 \vee z \geq 1$*. From Corollary 6, we have that*

$$\begin{aligned} \overline{B} &:= \operatorname{conv}(\{(z, y) \in B : z \leq 0\} \cup \{(z, y) \in B : z \geq 1\}) \\ &= \{(z, y) \in B : |y| \leq (\sqrt{3} - 2)z + 2\}. \end{aligned}$$

Now let $G(z, y) = z^2 + y^2 - 4$ *and* $J(z, y) = -(z - 1/2)^2 + 1/4$*. Since split disjunction* $z \leq 0 \vee z \geq 1$ *is equivalent to* $J(z, y) \leq 0$*, we have* $\overline{B} = \operatorname{conv}(S)$*, where*

$$S = \left\{(z, y) \in \mathbb{R}^2 : G(z, y) \leq 0, \quad J(z, y) \leq 0\right\}. \tag{3.40}$$

Now consider $W_\lambda = (1 - \lambda)G + \lambda J$ *for* $\lambda \in [0, 1]$*. One can check that the split cut* $|y| \leq (\sqrt{3} - 2)z + 2$ *obtained through Corollary 6, can be equivalently written as*

$$y^2 - \left((\sqrt{3} - 2)z + 2\right)^2 \leq 0 \tag{3.41a}$$

$$(\sqrt{3} - 2)z + 2 \geq 0. \tag{3.41b}$$

In turn, (3.41a) is equivalent to $W_{\lambda^*} \leq 0$ *for* $\lambda^* = \frac{4}{33}\left(6 - \sqrt{3}\right)$ *because* $W_{\lambda^*}/\left(\frac{1}{33}\left(9 + 4\sqrt{3}\right)\right) = y^2 - \left((\sqrt{3} - 2)z + 2\right)^2$*. By noting that (3.41b) holds for* B*, we conclude that*

$$\overline{B} = \{(z, y) \in B : W_{\lambda^*}(z, y) \leq 0\}. \tag{3.42}$$

Unfortunately, W_{λ^*} is not a convex function, so it does not fit in the aggregation framework described in this section. In particular, W_{λ^*} is an indefinite quadratic function so it cannot be obtained from an SDP relaxation of S. Indeed, we can show that the SDP relaxation of S strictly contains $\overline{B}$. Finally, while we can obtain W_{λ^*} through a procedure described in [94], this procedure requires the execution of a numerical algorithm and does not give closed form expressions such as those provided by Corollary 6.

3.5 ACKNOWLEDGMENTS

The content of this chapter is a reprint from the published article [70] with the kind permission to reprint from Springer ("Springer") provided by Copyright Clearance Center ("CCC").

4.0 SPLIT CUTS AND EXTENDED FORMULATIONS FOR MIXED INTEGER CONIC QUADRATIC PROGRAMMING

Split cuts, Gomory Mixed Integer (GMI) cuts, and Mixed Integer Rounding (MIR) cuts are some of the most effective valid inequalities for Mixed Integer Linear Programming (MILP). While they are known to be equivalent, each of them provide different advantages and insights. In particular, the split cuts construction shows that they are a particular case of disjunctive cuts [11] and hence have a straightforward extension to Mixed Integer Nonlinear Programming (MINLP). There has been significant work on the computational use of split cuts in MINLP [25, 28, 44, 59, 87] and a recent surge of theoretical developments [5, 9, 14, 15, 37, 57, 70, 81]. In particular, several formulas for split cuts for Mixed Integer Conic Quadratic Programming (MICQP) have been recently developed [5, 14, 15, 37, 70]. While the resulting cuts are strong nonlinear inequalities, adding these cuts to the continuous relaxation of the MICQP can significantly increase its solution time, which could negate the effectiveness of the cuts. One potential solution is to use linearizations of these cuts [25, 59], but in such a case, there is a strong trade-off between their strength and the computational burden of generating them. An alternative approach was introduced by Atamtürk and Narayanan [9] who use the polyhedral portion of a nonlinear *extended formulation* (i.e., a formulation with auxiliary variables) to construct an inexpensive, but potentially strong, linear cut they denote the Conic MIR (CMIR). In this chapter we attempt to broaden our understanding of split cuts for MINLP by providing a precise link between the CMIRs and split cuts for quadratic sets. In particular, this link provides a possible solution to the trade-off between the strength and computational burden resulting from adding the cuts to the relaxation.

In this chapter we study split cuts and extended formulations for MICQP and their relation to Conic Mixed Integer Rounding (CMIR) cuts [9]. Our first contribution is to show that the CMIR is a linear split cut for the polyhedral portion of the nonlinear extended formulation from [9]. Through this equivalence, we can extend the most general version of the CMIRs to the case of variables with unrestricted signs which was not previously possible. Our second contribution is to give a precise relation between the CMIR and nonlinear split cuts for quadratic sets. In particular, we show that, since the CMIR construction does not consider any quadratic information, a single CMIR can be weaker than a single nonlinear split cut. However, we also show that when families of split cuts and CMIRs are considered, CMIRs can provide a significant advantage over nonlinear split cuts by exploiting their common extended formulation. To the best of our knowledge, this is the first illustration of how the power of an extended formulation can improve the strength of a cutting plane procedure in MINLP.

The rest of this chapter is structured as follows. In Section 4.1 we introduce some notation and describe previous results on CMIRs and split cuts for MINLP. In Section 4.2 we establish the equivalency between CMIRs and linear split cuts for an extended formulation. Finally, in Section 4.3 we compare the strength of nonlinear split cuts and CMIRs.

4.1 NOTATION AND PREVIOUS WORK

We use the notation introduced in Chapter 2. Moreover, for notational convenience, we define split cuts while identifying a single set of integer variables $x \in \mathbb{Z}^n$ and three sets of continuous variables $y \in \mathbb{R}^p$, $t \in \mathbb{R}^m$, and $t_0 \in \mathbb{R}$.

Definition 4. *Let $K \subseteq \mathbb{R}^{n+p+m+1}$ be a closed convex set and $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$. A* split cut *for K is any valid inequality for*

$$K^{\pi,\pi_0} := \operatorname{conv}\left(\{(x,y,t,t_0) \in K : \pi^T x \leq \pi_0\} \cup \{(x,y,t,t_0) \in K : \pi^T x \geq \pi_0 + 1\}\right)$$

for some $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$. If $\pi = e^i$ for some $i \in [n]$, we refer to (π, π_0) as an elementary disjunction *and to the obtained cuts as* elementary split cuts.

Because $K^{\pi,\pi_0} \supseteq \text{conv}(K \cap (\mathbb{Z}^n \times \mathbb{R}^{p+m+1}))$, split cuts are valid inequalities for $K \cap (\mathbb{Z}^n \times \mathbb{R}^{p+m+1})$. For MILP, where K is a rational polyhedron, K^{π,π_0} is also a polyhedron and we only need linear split cuts. In contrast, if K is a general closed convex set, K^{π,π_0} is only closed and convex [37]. However, for special classes of K, we can characterize the nonlinear split cuts that need to be added to K to obtain K^{π,π_0} [5, 14, 15, 37, 57, 70]. For instance, the following proposition from [70] characterizes split cuts for conic quadratic sets of the form

$$C := \left\{(x, t_0) \in \mathbb{R}^{n+1} : \|B\,(x - c)\|_2 \leq t_0\right\}, \tag{4.1}$$

where C is in fact an affine transformation of the Quadratic cone $\{(x, t_0) \in \mathbb{R}^{n+1} : \|x\|_2 \leq t_0\}$.

Proposition 13. *Let $B \in \mathbb{R}^{n\times n}$ be an invertible matrix, $c \in \mathbb{R}^n$, $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$, and C be as defined in (4.1). If $\pi^T c \notin (\pi_0, \pi_0 + 1)$, then $C^{\pi,\pi_0} = C$. Otherwise, there exist $\bar{B} \in \mathbb{R}^{n\times n}$ and $\bar{c} \in \mathbb{R}^n$ such that*

$$C^{\pi,\pi_0} = \left\{(x, t_0) \in C \,:\, \left\|\bar{B}\,(x - c) + \bar{c}\,\right\|_2 \leq t_0\right\}.$$

Proposition 13 shows that the single split cut for C is $\left\|\bar{B}\,(x - c) + \bar{c}\,\right\|_2 \leq t_0$ which is of the same class as the inequality describing C. However, this inequality can be too expensive computationally and it can be preferable to add linear cuts instead. One way to achieve this is to add a finite number of linearizations of the nonlinear cuts. Such linearizations can be algorithmically obtained even in the absence of nonlinear cut formulas. Two examples of this are the algorithms introduced in [25, 59] to generate disjunctive inequalities for convex MINLPs.

A completely different linearization scheme was introduced by Atamtürk and Narayanan [9] for the general conic quadratic set given by

$$M_+ := \left\{(x, y, t_0) \in \mathbb{R}^{n+p+1} \,:\, \|Ax + Gy - b\|_2 \leq t_0, \quad x \geq 0, \quad y \geq 0\right\},$$

for rational matrices and vectors $A \in \mathbb{Q}^{m\times n}$, $G \in \mathbb{Q}^{m\times p}$, and $b \in \mathbb{Q}^m$. Instead of considering valid inequalities for $\text{conv}(M_+ \cap (\mathbb{Z}^n \times \mathbb{R}^{p+1}))$ directly, using auxiliary variables $t \in \mathbb{R}^m$, they first introduce the nonlinear extended formulation of M_+ given by

$$|Ax + Gy - b| \leq t, \quad x \geq 0, \quad y \geq 0, \quad \|t\|_2 \leq t_0, \tag{4.2}$$

so that, if $P_+ := \{(x,y,t) \in \mathbb{R}^{n+p+m} \,:\, |Ax + Gy - b| \le t, \quad x \ge 0, \quad y \ge 0\}$ and $\mathrm{Proj}_{(x,y,t_0)}$ is the projection onto the (x, y, t_0) space, then $M_+ = \mathrm{Proj}_{(x,y,t_0)}(\{(x,y,t,t_0) \in \mathbb{R}^{n+p+m+1} \,:\, \|t\|_2 \le t_0, \quad (x,y,t) \in P_+\})$. They then exploit the fact that P_+ is a polyhedron to generate a class of valid inequalities they denote the *Conic MIR* (CMIR). The first version of the CMIR is a simple but strong cut for a four variable and one constraint version of P_+.

Proposition 14 (Simple CMIR). *Let $b_0 \in \mathbb{R}$, $f = b_0 - \lfloor b_0 \rfloor$,*

$$S_0 := \left\{(x,y,t_0) \in \mathbb{R}^4 \,:\, |x + y_1 - y_2 - b_0| \le t_0, \quad y_1, y_2 \ge 0\right\},$$

and let the simple CMIR be the inequality given by

$$(1-2f)(x - \lfloor b_0 \rfloor) + f \le t_0 + y_1 + y_2. \tag{4.3}$$

The simple CMIR is valid for $\mathrm{conv}(S_0 \cap (\mathbb{Z} \times \mathbb{R}^2_+ \times \mathbb{R}_+))$ *and furthermore*

$$\mathrm{conv}(S_0 \cap (\mathbb{Z} \times \mathbb{R}^2_+ \times \mathbb{R}_+)) = \{(x,y,t_0) \in S_0 \,:\, (4.3)\}.$$

The simple CMIR is a linear inequality, but Atamtürk and Narayanan show that it can induce nonlinear inequalities in the (x, t_0) space through (4.2).

Lemma 7 (Nonlinear CMIR). *Let $T_0 := \left\{(x,y,t_0) \in \mathbb{R}^3 \,:\, \sqrt{(x-b_1)^2 + y^2} \le t_0\right\}$, $P_0 := \left\{(x,y,t) \in \mathbb{R}^4 : |x - b_1| \le t_1,\ |y| \le t_2\right\}$, $b_1 \in \mathbb{R}$, and $f = b_1 - \lfloor b_1 \rfloor$. Then the simple CMIR for $|x - b_1| \le t_1$ is given by*

$$(1-2f)(x - \lfloor b_1 \rfloor) + f \le t_1, \tag{4.4}$$

$\mathrm{conv}(T_0 \cap (\mathbb{Z} \times \mathbb{R}^2)) = T_0^{e^1, \lfloor b_1 \rfloor}$, *and*

$$\begin{aligned} T_0^{e^1, \lfloor b_1 \rfloor} &= \left\{(x,y,t_0) \in T_0 \,:\, \sqrt{((1-2f)(x - \lfloor b_1 \rfloor) + f)^2 + y^2} \le t_0\right\} \\ &= \mathrm{Proj}_{(x,y,t_0)}\left(\left\{(x,y,t,t_0) \in \mathbb{R}^5 \,:\, (x,y,t) \in P_0, \quad \|t\|_2 \le t_0, \quad (4.4)\right\}\right). \end{aligned}$$

Atamtürk and Narayanan follows the traditional linear MIR procedure [74, 75] to get CMIRs for M_+ and develop a super-additive version of the CMIR. Their most general version results in the following family of cuts.

Theorem 1 (Super-additive CMIR). *Let $a, v \in \mathbb{R}^n, g, w \in \mathbb{R}^p, h, u \in \mathbb{R}^m$, $S_+ := \{(x, y, t) \in \mathbb{R}^{n+p+m} : \left|a^T x + g^T y + h^T t - b_0\right| \leq u^T t + v^T x + w^T y, \quad x, y, t \geq 0\}$ be a relaxation of P_+ and let $\varphi_f(a) = -a + 2(1-f)\left(\lfloor a \rfloor + \frac{(a - \lfloor a \rfloor - f)^+}{1-f}\right)$. Then for any $\alpha \neq 0$ and $f_\alpha = b_0/\alpha - \lfloor b_0/\alpha \rfloor$, a valid cut for S_+ and P_+ is*

$$\sum\nolimits_{j=1}^{n} \varphi_{f_\alpha}(a_j/\alpha)x_j - \varphi_{f_\alpha}(b_0/\alpha) \leq \left((u + |h|)^T t + (w + |g|)^T y + v^T x\right)/|\alpha|. \tag{4.5}$$

We let a super-additive CMIR be any cut of this form obtained for some relaxation S_+, which can be constructed through various aggregation procedures. Finally, with regards to its relation to the traditional linear MIR, Atamtürk and Narayanan use the aggregation to show that every MIR is a CMIR. In Section 4.2 we show that these two cuts are in fact equivalent.

4.2 CONIC MIR AND LINEAR SPLIT CUTS

We now show that CMIRs are equivalent to linear split cuts for P_+, which are in turn equivalent to traditional linear MIRs for P_+. Through this equivalence, we extend all CMIRs to the case of variables with unrestricted signs and show that such extension follows naturally from the simple CMIR. To show the equivalence between linear split cuts and super-additive conic MIRs, we need the following well-known characterization of split cuts for a polyhedron T (e.g. [89]).

Proposition 15. *Let $T := \{(x, y) \in \mathbb{Z}^n \times \mathbb{R}^p : Cx + Dy \leq d\}$ for $C \in \mathbb{R}^{m \times n}, D \in \mathbb{R}^{m \times p}, d \in \mathbb{R}^m$, and let $\mu \in \mathbb{R}^m$ be such that $C^T \mu = \pi \in \mathbb{Z}^n$ and $D^T \mu = 0 \in \mathbb{R}^p$. Also let $f = \mu^T d - \lfloor \mu^T d \rfloor$. Then every split cut for T is of the form*

$$|\mu|^T (Cx + Dy - d) + (1 - 2f)\left(\pi^T x - \left\lfloor \mu^T d \right\rfloor\right) + f \leq 0.$$

Using this proposition, we show that every linear split cut for P_+ can be obtained from the simple CMIR and that every CMIR is a split cut.

Theorem 2. *Every non-dominated split cut for P_+ is of the form*

$$(1-2f)\left(\pi^T x - \lfloor \mu^T b \rfloor\right) + f \leq |\mu|^T t + |\lambda|^T x + |\gamma|^T y, \tag{4.6}$$

for some $\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^n$, $\gamma \in \mathbb{R}^p$, and $\pi \in \mathbb{Z}^n$ such that $A^T\mu - \lambda = \pi$, $G^T\mu - \gamma = 0$, and $f = \mu^T b - \lfloor \mu^T b \rfloor$. Furthermore, every super-additive CMIR for P_+ is equivalent or dominated by a split cut of this form.

Proof. We first prove formula (4.6) using Proposition 15. We have $P_+ = \{(x, y, t) \in \mathbb{R}^{n+p+m} : \hat{C}x + \hat{D}\begin{bmatrix} y \\ t \end{bmatrix} \leq \hat{d}\}$, where

$$\hat{C} = \begin{bmatrix} A \\ -A \\ -I \\ 0 \end{bmatrix}, \quad \hat{D} = \begin{bmatrix} G & -I \\ -G & -I \\ 0 & 0 \\ -I & 0 \end{bmatrix}, \text{ and } \hat{d} = \begin{bmatrix} b \\ -b \\ 0 \\ 0 \end{bmatrix}.$$

Let $\hat{\mu} = (\mu_1^T, \mu_2^T, \lambda^T, \gamma^T)^T$, where $\mu_1, \mu_2 \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^n$, and $\gamma \in \mathbb{R}^p$. $\hat{D}^T\hat{\mu} = 0$ implies $G^T(\mu_1 - \mu_2) - \gamma = 0$ and $\mu_1 = -\mu_2$. Furthermore, $\hat{C}^T\hat{\mu} = \pi$ implies $A^T(\mu_1 - \mu_2) - \lambda = \pi$. Let $\mu := \mu_1 - \mu_2$ and the result then follows from Proposition 15.

Let

$$C = \begin{bmatrix} \frac{a^T}{2\alpha} - \frac{v^T}{2|\alpha|} \\ -\frac{a^T}{2\alpha} - \frac{v^T}{2|\alpha|} \\ -I \\ 0 \\ 0 \end{bmatrix}, \quad D = \begin{bmatrix} \frac{g^T}{2\alpha} - \frac{w^T}{2|\alpha|} & \frac{h^T}{2\alpha} - \frac{u^T}{2|\alpha|} \\ -\frac{g^T}{2\alpha} - \frac{w^T}{2|\alpha|} & -\frac{h^T}{2\alpha} - \frac{u^T}{2|\alpha|} \\ 0 & 0 \\ -I & 0 \\ 0 & -I \end{bmatrix} \text{ and } d = \begin{bmatrix} \frac{b_0}{2\alpha} \\ -\frac{b_0}{2\alpha} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

so that $S_+ = \left\{(x, y, t) \in \mathbb{R}^{n+p+m} : Cx + D\begin{bmatrix} y \\ t \end{bmatrix} \leq d\right\}$ is a relaxation of P_+. Now let $f_\alpha = b_0/\alpha - \lfloor b_0/\alpha \rfloor$, $\mu = (1, -1, \lambda^T, g^T/\alpha, h^T/\alpha)^T$ where $\lambda_j = a_j/\alpha - \lfloor a_j/\alpha \rfloor$ if $a_j/\alpha - \lfloor a_j/\alpha \rfloor < f_\alpha$, and $\lambda_j = -(1 - a_j/\alpha + \lfloor a_j/\alpha \rfloor)$ if $a_j/\alpha - \lfloor a_j/\alpha \rfloor \geq f_\alpha$. Then, by Proposition 15, we obtain the split cut for S_+ given by

$$\sum_{j\in[n]:a_j/\alpha-\lfloor a_j/\alpha\rfloor<f_\alpha} \Big(-\frac{a_j}{\alpha} + 2(1-f_\alpha)\lfloor a_j/\alpha\rfloor\Big)\, x_j \\ + \sum_{j\in[n]:a_j/\alpha-\lfloor a_j/\alpha\rfloor\geq f_\alpha} \Big(-\frac{a_j}{\alpha} + 2(1-f_\alpha)\lfloor a_j/\alpha\rfloor + 2\left(a_j/\alpha - \lfloor a_j/\alpha\rfloor - f_\alpha\right)\Big)\, x_j \\ -\sum_{j=1}^{n} \frac{v_j}{|\alpha|} x_j - \sum_{j=1}^{p} \frac{w_j + |g_j|}{|\alpha|} y_j - \sum_{j=1}^{p} \frac{u_j + |h_j|}{|\alpha|} t_j \leq 2(1-f_\alpha)\lfloor b_0/\alpha\rfloor - b_0/\alpha.$$

The cut above is precisely super-additive CMIR (4.5). The result follows by noting that since $P_+ \subseteq S_+$, then any split cut for S_+ is also a split cut for P_+. □

From Theorem 2, we have that a natural extension of the super-additive CMIR to the case of variables with unrestricted signs is to consider split cuts. While we can also consider cases with partial non-negativity requirements, because of space limitations, we here focus on the set with no non-negativity constraints given by $M := \{(x, y, t_0) \in \mathbb{R}^{n+p+1} : \|Ax + Gy - b\|_2 \leq t_0\}$. As before, we let the polyhedral portion of the extended formulation of M be

$$P := \left\{(x, y, t) \in \mathbb{R}^{n+p+m} \,:\, |Ax + Gy - b| \leq t\right\}.$$

We can extend the CMIR to this setting through the following theorem.

Theorem 3. *Every non-dominated split cut for P is of the form*

$$(1-2f)\left(\pi^T x - \lfloor \mu^T b\rfloor\right) + f \leq |\mu|^T t, \tag{4.7}$$

for some $\mu \in \mathbb{R}^m$ such that $A^T\mu = \pi \in \mathbb{Z}^n$, $G^T\mu = 0$, and $f = \mu^T b - \lfloor \mu^T b\rfloor$.

Proof. Follows from Proposition 15. □

From (4.6) and (4.7), we can see that all split cuts for P_+ and P can be obtained from the simple CMIR (4.3) and some simple aggregation procedures.

4.3 COMPARISON BETWEEN CUTS

Through Lemma 7, Atamtürk and Narayanan show that using an extended formulation analog to (4.2), the effect of the simple CMIR on the (x, y, t_0) space is equivalent to that of a conic split cut from Proposition 13. We now study to what extent this holds for more general settings. We first study containment relations between the sets obtained by adding nonlinear split cuts and CMIRs to some specific regions bounded by a single conic quadratic inequality. To consider more general sets, we then compare the strength of the bounds generated by the two classes of cuts on some quadratic integer programming problems. In both cases, it will be convenient to use the following direct corollary that specializes Theorem 3 to the polyhedral portion of the analog of extended formulation (4.2) for $C := \{(x, t_0) \in \mathbb{R}^{n+1} : \|B(x - c)\|_2 \leq t_0\}$, which is of the form

$$L := \left\{(x, t) \in \mathbb{R}^{2n} : |B(x - c)| \leq t\right\}. \tag{4.8}$$

Corollary 10. *Let $B \in \mathbb{R}^{n\times n}$ be an invertible matrix, $c \in \mathbb{R}^n$, $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$, and L be as defined in (4.8). If $\pi^T c \notin (\pi_0, \pi_0 + 1)$, then $L^{\pi,\pi_0} = L$. Otherwise*

$$L^{\pi,\pi_0} = \left\{(x, t) \in L : (1 - 2f)\left(\pi^T x - \lfloor \pi^T c \rfloor\right) + f \leq |\mu|^T t\right\},$$

where $\mu \in \mathbb{R}^n$ is the unique solution to $B^T\mu = \pi \in \mathbb{Z}^n$ and $f = \pi^T c - \lfloor \pi^T c \rfloor$.

4.3.1 Containment Relations

Because $C = \mathrm{Proj}_{(x,t_0)}(\{(x, t, t_0) \in \mathbb{R}^{2n+1} : (x, t) \in L, \|t\|_2 \leq t_0\})$, it is natural to compare the strength of the CMIRs (i.e., linear split cuts) for L and the nonlinear split cuts for C from Proposition 13. As discussed, Lemma 7 shows that these cuts can sometimes be equivalent. However, the following proposition shows that this is true only for very specific structures and that a single nonlinear split cut for C is at least as strong as (and many times stronger than) the CMIR associated to the same disjunction. There, we let

$$MIR^{\pi,\pi_0} := \left\{(x, t, t_0) \in \mathbb{R}^{2n+1} : (x, t) \in L^{\pi,\pi_0}, \quad \|t\|_2 \leq t_0\right\}. \tag{4.9}$$

Proposition 16. *Let $B \in \mathbb{R}^{n \times n}$ be invertible, $c \in \mathbb{R}^n$, $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$, and MIR^{π,π_0} as defined in (4.9). Then $C^{\pi,\pi_0} \subseteq \operatorname{Proj}_{(x,t_0)}(MIR^{\pi,\pi_0})$. The containment holds as equality if $B = I$ and $\pi = e^i$ for some $i \in [n]$, but can otherwise be strict even for $n = 2$.*

Proof. We begin with proving the containment. Let $(\bar{x}, \bar{t}_0) \in C^{\pi,\pi_0}$. There exist $\alpha \in [0,1]$, $(x^0, t_0^0) \in C$, and $(x^1, t_0^1) \in C$ such that $(\bar{x}, \bar{t}_0) = \alpha(x^0, t_0^0) + (1-\alpha)(x^1, t_0^1)$, $\pi^T x^0 \leq \pi_0$, and $\pi^T x^1 \geq \pi_0 + 1$. Let $t^0 := |B(x^0 - c)|$, $t^1 := |B(x^1 - c)|$, and $\bar{t} := \alpha t^0 + (1-\alpha) t^1$. Then $(\bar{x}, \bar{t}_0) = \operatorname{Proj}_{(x,t_0)}((\bar{x}, \bar{t}, \bar{t}_0))$ and $(\bar{x}, \bar{t}) \in L^{\pi,\pi_0}$. It then only remains to show that $\|\bar{t}\|_2 \leq \bar{t}_0$, which follows from

$$\begin{aligned} \|\bar{t}\|_2 &= \left\|\alpha t^0 + (1-\alpha) t^1\right\|_2 \leq \alpha \left\|t^0\right\|_2 + (1-\alpha)\left\|t^1\right\|_2 \\ &= \alpha \left\|B(x^0 - c)\right\|_2 + (1-\alpha)\left\|B(x^1 - c)\right\|_2 \leq \alpha t_0^0 + (1-\alpha) t_0^1 = \bar{t}_0. \end{aligned} \tag{4.10}$$

Now we show that the containment holds as equality for $B = I$ and $\pi = e^i$ for some $i \in [n]$. Using Corollary 10, we have

$$MIR^{\pi,\pi_0} = \left\{(x, t, t_0) \in \mathbb{R}^{2n+1} \ : \ |x - c| \leq t, \ \|t\|_2 \leq t_0, \ (1 - 2f_i)(x_i - \lfloor c_i \rfloor) + f_i \leq t_i\right\},$$

where $f_i = c_i - \lfloor c_i \rfloor$. Furthermore, one can check that MIR^{π,π_0} does not change by replacing $(1-2f_i)(x_i - \lfloor c_i \rfloor) + f_i \leq t_i$ with $|(1-2f_i)(x_i - \lfloor c_i \rfloor) + f_i| \leq t_i$. Thus, $\operatorname{Proj}_{(x,t_0)}(MIR^{\pi,\pi_0})$ is defined by the original constraint $\|x - c\|_2 \leq t_0$ and

$$\sqrt{\sum_{j \in [n] : j \neq i} (x_j - c_j)^2 + ((1 - 2f_i)(x_i - \lfloor c_i \rfloor) + f_i)^2} \leq t_0. \tag{4.11}$$

Also using Corollary 5 in [70], the split cut associated to C^{π,π_0} is

$$\sqrt{\sum_{j \in [n] : j \neq i} (x_j - c_j)^2 + (a(x_i - c_i) + b)^2} \leq t_0, \tag{4.12}$$

where $a = \lfloor c_i \rfloor + \lfloor c_i \rfloor + 1 - 2c_i = 1 - 2f_i$ and $b = -2(\lfloor c_i \rfloor - c_i)(\lfloor c_i \rfloor + 1 - c_i) = 2f_i(1 - f_i)$. The result then follows by noting that (4.11) and (4.12) are equivalent.

Finally, we show that the containment is strict for $n = 2$, $B = I$, $c = (1/4, 0)^T$, $\pi = (1,1)^T$, and $\pi_0 = 0$. Again using Corollary 5 in [70], one can check that after a few simplifications, the corresponding split cut is

$$\sqrt{(3x_1 - x_2)^2 + (3x_2 - x_1 + 1)^2} \leq 4t_0, \tag{4.13}$$

and using Corollary 10, the corresponding CMIR cut is $x_1 + x_2 + 1/2 \leq 2(t_1 + t_2)$. Let $(\bar{x}, \bar{t}, \bar{t}_0) = (-0.082, 0.922, 0.337, 0.928, 1)$. We have that $(\bar{x}, \bar{t}, \bar{t}_0) \in MIR^{\pi,\pi_0}$, but $(\bar{x}, \bar{t}_0)$ violates the split cut (4.13). $\square$

While a single CMIR can be weaker than the corresponding nonlinear split cut, a family of CMIRs sharing the same extended formulation can be significantly stronger than the associated family of nonlinear split cuts. This can be illustrated by considering split cuts for C (see Proposition 19 for a result along this line). However, the behavior is more dramatic for an ellipsoid given by

$$E := \{x \in \mathbb{R}^n \, : \, \|B(x - c)\|_2 \leq r\},$$

where $B \in \mathbb{R}^{n \times n}$ is an invertible matrix, $c \in \mathbb{R}^n$, and $r \in \mathbb{R}_+$. As formalized in the following straightforward lemma, an ellipsoid can be described as projections of linear sections of either the cone C defined in (4.1), a paraboloid Q of the form

$$Q := \left\{(x, s_0) \in \mathbb{R}^{n+1} : \|B(x - c)\|_2^2 \leq s_0\right\},$$

and the extended formulation associated to the CMIR, which provides a way of comparing the strength of several cuts.

Lemma 8. *Let $B \in \mathbb{R}^{n \times n}$ be an invertible matrix, $c \in \mathbb{R}^n$, and $r \in \mathbb{R}_+$. Then*

$$\begin{aligned} E &= \operatorname{Proj}_x\left(\{(x, t_0) \in C \, : \, t_0 = r\}\right) = \operatorname{Proj}_x(\{(x, s_0) \in Q \, : \, s_0 = r^2\}) \\ &= \operatorname{Proj}_x\left\{(x, t, t_0) \in \mathbb{R}^{2n+1} \, : \, (x, t) \in L, \quad \|t\|_2 \leq t_0, \quad t_0 = r\right\}. \end{aligned}$$

The CMIR and nonlinear split cuts for C and Q (characterized in [70]) can be used to induce valid inequalities for $E \cap \mathbb{Z}^n$ through the same linear section of Lemma 8. However, the construction of these cuts does not exploit the structure induced by the section and they hence cannot be expected to always achieve the full strength of the nonlinear split cuts for E studied in [14, 37, 70]. The following proposition shows that this is indeed the case and that the cut with the weakest effect on E is the CMIR.

Proposition 17. *Let $B \in \mathbb{R}^{n\times n}$ be an invertible matrix, $c \in \mathbb{R}^n$, $(\pi, \pi_0) \in \mathbb{Z}^n \times \mathbb{Z}$, and $r \in \mathbb{R}_+$. Then*

$$\begin{aligned} E^{\pi,\pi_0} &\subseteq \operatorname{Proj}_x\left(\left\{(x, s_0) \in Q^{\pi,\pi_0} \,:\, s_0 = r^2\right\}\right) \\ &\subseteq \operatorname{Proj}_x(\{(x, t_0) \in C^{\pi,\pi_0} \,:\, t_0 = r\}) \\ &\subseteq \operatorname{Proj}_x\left(\{(x, t, t_0) \in MIR^{\pi,\pi_0} \,:\, t_0 = r\}\right). \end{aligned} \tag{4.14}$$

All containments can be simultaneously strict even for $n = 2$.

Proof. The last containment follows from Proposition 16 by noting that

$$C^{\pi,\pi_0} \subseteq \operatorname{Proj}_{(x,t_0)}\left(MIR^{\pi,\pi_0}\right).$$

We now prove the first containment. If $\bar{x} \in E^{\pi,\pi_0}$, then there exist x^1, x^2 such that $\bar{x} = \alpha x^1 + (1-\alpha)x^2$ for some $\alpha \in [0, 1]$,

$$\left\|B(x^1 - c)\right\|_2 \le r,\ \pi x^1 \le \pi_0 \quad \text{and} \quad \left\|B(x^2 - c)\right\|_2 \le r,\ \pi x^2 \ge \pi_0 + 1,$$

which implies that (x^1, s^*) and (x^2, s^*) - where $s^* = r^2$ - satisfy, respectively

$$\left\|B(x^1 - c)\right\|_2^2 \le s^*,\ \pi x^1 \le \pi_0 \quad \text{and} \quad \left\|B(x^2 - c)\right\|_2^2 \le s^*,\ \pi x^2 \ge \pi_0 + 1.$$

Therefore, $\alpha(x^1, s^*) + (1-\alpha)(x^2, s^*) = (\bar{x}, s^*) = (\bar{x}, r^2)$ belongs to $\{(x, s_0) \in Q^{\pi,\pi_0} \,:\, s_0 = r^2\}$, and thus $\bar{x}$ belongs to the projection of this set on the x-space.

The fact that the second set is contained in the third set can be proved as follows. If $\bar{x}$ belongs to the second set, then $(\bar{x}, r^2) \in \{(x, s_0) \in Q^{\pi,\pi_0} \,:\, s_0 = r^2\}$ which implies that there exist $(x', s'), (x'', s'')$ such that $(\bar{x}, r^2) = \alpha(x', s') + (1-\alpha)(x'', s'')$ for some $\alpha \in [0, 1]$,

$$\|B\,(x' - c)\|_2^2 \le s',\ \pi x' \le \pi_0 \quad \text{and} \quad \|B\,(x'' - c)\|_2^2 \le s'',\ \pi x'' \ge \pi_0 + 1.$$

We can therefore conclude that $(x', r' = \sqrt{s'})$ and $(x'', r'' = \sqrt{s''})$ satisfy

$$\|B(x'-c)\|_2 \leq r',\ \pi x' \leq \pi_0 \quad \text{and} \quad \|B(x''-c)\|_2 \leq r'',\ \pi x'' \geq \pi_0 + 1.$$

As the function $f(x) = \sqrt{x}$ is a concave function for $x \geq 0$, we have

$$r = f(r^2) = f(\alpha s' + (1-\alpha)s'') \geq \alpha f(s') + (1-\alpha) f(s'') = \alpha r' + (1-\alpha) r''.$$

Now, replacing r' by a larger number r'_+, we still have $\|B(x'-c)\|_2 \leq r'_+$; we can choose r'_+ such that $r = \alpha r'_+ + (1-\alpha) r''$, so that $(\bar{x}, r) \in \{(x, t_0) \in C^{\pi,\pi_0} : t_0 = r\}$.

Finally, we show that all three containments are strict for $n = 2$, $B = I$, $c = (1/4, 0)^T$, $\pi = (1,1)^T$, $\pi_0 = 0$, and $r = 1$. The last strict containment follows by considering the example previously provided in the proof of Proposition 16. Using Corollaries 4 and 6 in [70], one can check that after a few simplifications, the corresponding split cuts associated to E^{π,π_0} and Q^{π,π_0} are given by

$$|x_2 - x_1 + 1/4| \leq \left(\left(\sqrt{23} - \sqrt{31}\right)(x_1 + x_2) + \sqrt{31}\right)/4 \tag{4.15}$$

and

$$(x_2 - x_1 + 1/4)^2 + (x_1 + x_2)/2 + 1/16 \leq 2 s_0, \tag{4.16}$$

respectively. The first two strict containments then follow from noting that $(-0.082, 0.903)$ belongs to $\operatorname{Proj}_x(\{(x, s_0) \in Q^{\pi,\pi_0} : s_0 = 1\})$ but violates the split cut (4.15), and $(-0.082, 0.911)$ belongs to $\operatorname{Proj}_x(\{(x, t_0) \in C^{\pi,\pi_0} : t_0 = 1\})$ but violates the split cut (4.16) for $s_0 = 1$. □

For the effect of a single cut on E, the CMIR is the weakest in Proposition 17. However, several CMIRs combined through a common extended formulation (i.e., with a single set of auxiliary variables $t \in \mathbb{R}^n$) can be significantly stronger than even the associated family of split cuts for E. This effectively sidesteps the three potentially strict containments in (4.14). For instance, the following proposition shows that elementary CMIRs are enough to show emptiness of the convex hull of integer points of an ellipsoid with no lattice points, while this cannot be done even with all the nonlinear split cuts (elementary and non-elementary) of E.

Proposition 18. *Let $n \geq 2$, $r = 1/2$, $B = I$, and $c_i = 1/2$ for all $i \in [n]$ so that $E \cap \mathbb{Z}^n = \emptyset$. Then*

$$\emptyset = \operatorname{Proj}_x \left(\left\{ (x, t, t_0) \in \bigcap_{i=1}^n MIR^{e^i, \lfloor c_i \rfloor} \, : \, t_0 = r \right\} \right) \subsetneq \bigcap_{(\pi,\pi_0) \in \mathbb{Z}^n \times \mathbb{Z}} E^{\pi,\pi_0} = \{c\} \,.$$

Proof. For the first equality, note that

$$\bigcap_{i=1}^n MIR^{e^i, \lfloor c_i \rfloor} = \{ (x, t, t_0) \, : \, \|t\|_2 \leq t_0, \, |x_i - 1/2| \leq t_i, \, 1/2 \leq t_i \; \forall \, i \in [n] \} \,,$$

and every point in this set has $t_0 \geq \sqrt{n}/2 > r = 1/2$ for the assumed $n \geq 2$.

For the last equality, we first prove left to the right containment. This follows by noting that the set obtained by adding all the elementary split cuts is equal to $\{c\}$. In particular, the intersection of E with $x_i = 1$ (or $x_i = 0$) is exactly $(1, \ldots, 1)/2 \pm e_i/2$ and $E^{e_i,0}$ is exactly the convex hull of these two points.

The reverse containment directly follows from Lemma 1 in [34], where they show that as long as a convex set contains points where one component is 0/1, and all other components are 1/2, then $\{c\}$ is contained in the split closure of this set (while Lemma 1 in [34] is stated for polyhedra, the extension to general convex sets is straightforward). The strict containment above then follows automatically. □

4.3.2 Bound Strength

The quadratic integer programming problem that we consider is the Closest Vector Problem (CVP) [26, 68] which aims to find the element in an integer lattice that is closest (with respect to the Euclidean distance) to a given target vector not in the lattice. CVP can be equivalently formulated as

$$\min_x \{ \|B(x - c)\|_2 \, : \, x \in \mathbb{Z}^n \} \tag{4.17}$$

or

$$\min_x \left\{ \|B(x - c)\|_2^2 \, : \, x \in \mathbb{Z}^n \right\} , \tag{4.18}$$

where $B \in \mathbb{R}^{n \times n}$ is an invertible matrix whose columns compose the basis of the lattice and $c \in \mathbb{R}^n$ (the target vector is Bc in this case). As noted in [19, 18, 70], because $\operatorname{conv}(\mathbb{Z}^n) = \mathbb{R}^n$, to effectively use cuts in CVP we need the equivalent reformulations of (6.3) and (4.18) given by

$$\min_{(x,t_0)} \{t_0 \,:\, (x,t_0) \in C, \quad x \in \mathbb{Z}^n\} \tag{4.19}$$

for $C := \{(x,t_0) \in \mathbb{R}^{n+1} : \|B\,(x-c)\|_2 \leq t_0\}$, and

$$\min_{(x,s_0)} \{s_0 \,:\, (x,s_0) \in Q, \quad x \in \mathbb{Z}^n\} \tag{4.20}$$

for $Q := \{(x,s_0) \in \mathbb{R}^{n+1} : \|B\,(x-c)\|_2^2 \leq s_0\}$. We can then strengthen these formulations by adding split cuts for C and Q. However, using techniques similar to the proof of Proposition 17, we can show that adding split cuts for Q to (4.20) is always equal or better than adding split cuts for C to (6.4). For this reason, we only compare the strength of nonlinear split cuts for Q to the strength of the CMIRs. In this context, we consider the extended formulation given by

$$\min_{(x,t,t_0)} \left\{t_0^2 \,:\, |B(x-c)| \leq t, \quad \|t\|_2 \leq t_0, \quad t \in \mathbb{R}^n_+,\ t_0 \in \mathbb{R}_+,\ x \in \mathbb{Z}^n\right\}, \tag{4.21}$$

which can be strengthened by adding CMIR cuts. Similarly to Proposition 16, we can show that a single split cut for Q added to (4.20) is at least as strong as the corresponding CMIR added to (4.21). However, as formalized in the following proposition, there are examples where just elementary CMIR cuts can provide a bound that is arbitrarily better than that obtained by all split cuts for Q.

Proposition 19. *Let $B = I$ and $c_i = 1/2$ for all $i \in [n]$. Then*

$$n/4 \quad = \min_{x} \left\{\|x-c\|_2^2 \,:\, x \in \mathbb{Z}^n\right\} = \min_{x,t,t_0} \left\{t_0^2 \,:\, (x,t,t_0) \in \bigcap\nolimits_{i=1}^{n} MIR^{e^i,\lfloor c_i \rfloor}\right\},$$

while $1/4 \geq \min_{x,s_0} \left\{s_0 \,:\, (x,s_0) \in \bigcap\nolimits_{(\pi,\pi_0)\in\mathbb{Z}^n\times\mathbb{Z}} Q^{\pi,\pi_0}\right\}$.

Proof. The first equality is straightforward. From the proof of Proposition 18, we have that $t_0 \geq \sqrt{n}/2$ for any $(x, t, t_0) \in \bigcap_{i=1}^n MIR^{e^i, \lfloor c_i \rfloor}$, which proves the second equality. To prove the inequality above, we show that $(\bar{x}, \bar{s}_0)$ given by $\bar{s}_0 = 1/4$ and $\bar{x}_i = 1/2$ for all $i \in [n]$ satisfies all the quadratic split cuts. For this, note that using Corollary 4 in [70], the quadratic split cut with x replaced by $\bar{x}$ is given by

$$-\left(\pi_0 + 1 - (1/2)\sum\nolimits_{i=1}^n \pi_i\right)\left(\pi_0 - (1/2)\sum\nolimits_{i=1}^n \pi_i\right)/\|\pi\|_2^2 \leq s_0.$$

Then the only interesting cases are those with $\pi_0 < (1/2)\sum_{i=1}^n \pi_i < \pi_0 + 1$, for which the cut reduces to $(1/4\,\|\pi\|_2^2) \leq s_0$. The strongest of these cuts is $1/4 \leq s_0$ which is satisfied by $(\bar{x}, \bar{s}_0)$. □

Note that the example in Proposition 19 is very specific. In fact, our preliminary computational experiments show that for randomly generated CVP instances, the integrality gaps obtained by adding quadratic split cuts and CMIRs are roughly the same. It seems that using an extended formulation is enough to compensate for the lack of non-polyhedral information in the generation of CMIR cuts; however, it does not provide an advantage in general.

Finally, while CVP provides a simple and clean setting to compare the strength of cuts, no class of cuts seems to provide a computational advantage for solving these problems. We are currently exploring the effectiveness of these cuts on more practical MICQPs.

4.4 ACKNOWLEDGMENTS

The content of this chapter is a reprint from the published article [71] with the kind permission to reprint from Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC").

5.0 CONVEX HULL OF TWO QUADRATIC OR A CONIC QUADRATIC AND A QUADRATIC INEQUALITY

Development of strong valid inequalities or cutting planes such as Split cuts, Gomory Mixed Integer (GMI) cuts, and Mixed Integer Rounding (MIR) cuts is one of the most important breakthroughs in the area of Mixed Integer Linear Programming (MILP). Hence, there has recently been a significant interest in extending the theoretical and computational results from MILP to the realm of Mixed Integer Conic Quadratic Programming (MICQP) [8, 25, 28, 35, 36, 44, 56, 87, 81]. Dadush et al. [37] study the split closure of a strictly convex body and characterize split cuts for ellipsoids. Atamtürk and Narayanan [9] study the extension of MIR cuts to sets defined by a single conic quadratic inequality and introduce conic MIR cuts which are linear inequalities derived from an extended formulation. Modaresi et al. [71] then characterize nonlinear split cuts for similar conic quadratic sets and also establish the relation between the split cuts and conic MIR cuts from [9]. Andersen and Jensen [6] also study similar conic quadratic sets as in [9] and derive nonlinear split cuts using the intersection points of the disjunctions and the conic set. Belotti et al. [15] study the families of quadratic surfaces having fixed intersections with two hyperplanes. Following the results in [15], Belotti et al. [14, 16] characterize disjunctive cuts for conic quadratic sets when the sets defined by the disjunctions are bounded and disjoint, or when the disjunctions are parallel. Modaresi et al. [70] characterize intersections cuts for several classes of nonlinear sets with specific structures, including conic quadratic sets. Bienstock and Michalka [18, 19] derive linear inequalities to characterize the convex hull of convex quadratic functions on the complement of a convex quadratic or polyhedral set and they also study the associated separation problem. Morán et al. [81] consider subadditive inequalities for general Mixed Integer Conic Programming and Kılınç-Karzan [57] studies minimal valid linear inequalities to characterize the convex hull of general conic sets with a disjunctive structure. Following the results in [57], Kılınç-Karzan and Yıldız [58] study the structure of the convex hull of a two-term disjunction applied to the second-order cone. Yıldız and Cornuéjols [96] study disjunctive cuts on cross sections of the second-order cone. Finally, Burer and Kılınç-Karzan [27] characterize the closed convex hull of sets defined as the intersection of a conic quadratic and a quadratic inequality that satisfy certain technical conditions.

In this chapter we study the convex hull of regions defined by two quadratic or by a conic quadratic and a quadratic inequality. The technique we use to characterize the convex hulls is an aggregation technique introduced by Yıldıran [94]. In particular, Yıldıran characterizes the convex hull of sets defined by two quadratic inequalities and obtains a Semidefinite Programming (SDP) representation of the convex hull using Linear Matrix Inequalities (LMI). Yıldıran also proposes a polynomial-time algorithm to calculate the convex hull of two quadratics. In this chapter we show that the SDP representation of the convex hull of two quadratics presented in [94] can be described by two conic quadratic inequalities. We also show that the aggregation technique in [94] can be easily extended to derive valid conic quadratic inequalities for the convex hull of sets defined by a conic quadratic and a quadratic inequality. We also show that under an additional assumption, the derived inequalities are sufficient to characterize the convex hull. Therefore, the aggregation technique proposed in [94] provides a unified framework for generating *lattice-free* cuts for quadratic and conic quadratic sets which is independent of the geometry of the *lattice-free* set (e.g., a set that does not contain any integer point in its interior), as long as the lattice-free set can be described by a single quadratic inequality.

We note that the content of this chapter is a reprint from the article [72] which is submitted for publication. Also, the work in [27] contains similar results to those presented here and our main results have been developed independently. In Section 5.3.6 we compare and discuss these various results.

The rest of this chapter is organized as follows. In Section 5.1 we introduce some notation and provide the existing convex hull results from [94]. In Sections 5.2 and 5.3 we introduce the conic quadratic characterization of the convex hull of quadratic and conic quadratic sets and compare the results in this chapter and those in [27].

5.1 NOTATION, PRELIMINARIES, AND EXISTING CONVEX HUL RESULTS

We use the notation introduced in Chapter 2. In Sections 5.1 and 5.2 we follow the convention in [94] and define all sets using strict inequalities. However, in Section 5.3 all sets are defined by non-strict inequalities. This also allows us to compare our results with those in [27]. To simplify the exposition, we use the same notation for sets described by strict and non-strict inequalities; however, if we need to refer to sets defined by strict inequalities in Section 5.3, we use the interior to avoid any ambiguity.

5.1.1 Preliminaries

In this section we first define the quadratic sets that we study. We then provide some useful definitions and results from [94] that are relevant to our analysis. To save space, we do not provide the proofs of such results and we refer the reader to [94].

Our analysis is based on the work in [94] which studies the convex hull of open sets defined by two strict non-homogeneous quadratic inequalities. In particular, let

$$S := \{x \in \mathbb{R}^n \, : \, q_i < 0, \quad i = 1, 2\}, \tag{5.1}$$

where $q_i, \ i = 1, 2$ are quadratic polynomials of the form

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^T \mathcal{P} \begin{bmatrix} x \\ 1 \end{bmatrix} = x^T Q x + 2b^T x + \gamma, \tag{5.2}$$

where $\mathcal{P} = \begin{bmatrix} Q & b \\ b^T & \gamma \end{bmatrix} \in \mathbb{S}^{n+1}$, $Q \in \mathbb{S}^n, b \in \mathbb{R}^n$, and $\gamma \in \mathbb{R}$.

Note that [94] does not require the quadratic functions to satisfy any specific property. In particular, there is no requirement on the convexity or concavity of the quadratic functions defined in (5.2).

To characterize the convex hull of S, [94] considers the aggregated inequalities derived from the convex combinations of the two quadratics. More specifically, denote the pencil of quadratics induced by the convex combination of the two quadratic inequalities as

$$q_\lambda := (1-\lambda)q_1 + \lambda q_2,$$

where $\lambda \in [0,1]$. Similarly, define the associated symmetric matrix pencil

$$\mathcal{P}_\lambda := (1-\lambda)\mathcal{P}_1 + \lambda \mathcal{P}_2,$$

and

$$Q_\lambda := (1-\lambda)Q_1 + \lambda Q_2.$$

For a given quadratic pencil q_λ, define

$$S_\lambda := \left\{x \in \mathbb{R}^n \, : \, q_\lambda < 0\right\}.$$

The aggregation technique in [94] chooses $\lambda \in [0,1]$ such that the aggregated inequalities give conv (S). The characterization of the sets D and E, which are defined below, are crucial to the aggregation technique. Define

$$D := \{\lambda \in [0,1] \, : \, (1-\lambda)Q_1 + \lambda Q_2 \succeq 0\}$$

and

$$E := \left\{\lambda \in [0,1] \, : \, \pi\left(\mathcal{P}_\lambda\right) = 1\right\}.$$

Note that D is the collection of all $\lambda \in [0,1]$ such that the associated quadratic set S_λ is convex. On the other hand, E is the collection of all $\lambda \in [0,1]$ for which $\mathcal{P}_\lambda$ has exactly one negative eigenvalue. Therefore, S_λ may be non-convex for some $\lambda \in E$. However, as shown in Theorem 5, two specific aggregated inequalities associated with E admit a convex representation and these are enough to characterize conv(S). Throughout the paper, we use the following useful lemma from [94] which characterizes the structure of the set E.

Lemma 9. *If $E \neq \emptyset$, then E is the union of at most two disjoint connected intervals of the form*

$$E = [\lambda_1, \lambda_2] \cup [\lambda_3, \lambda_4],$$

where $\lambda_i, \lambda_{i+1} \in [0,1]$ for $i \in \{1,3\}$ are generalized eigenvalues of the pencil $\mathcal{P}_\lambda$.

If E is a single connected interval, we denote $E = [\lambda_1, \lambda_2]$, for $\lambda_1, \lambda_2 \in [0,1]$. Also note that each connected interval of E may contain only a single point. In such a case, we have $\lambda_i = \lambda_{i+1}$. The following proposition from [94] characterizes the relation between D and E.

Proposition 20. *If $S \neq \emptyset$, then D is a closed subset of E.*

Therefore, if E is composed of two disjoint connected intervals, Lemma 9 implies that $D \subseteq [\lambda_i, \lambda_{i+1}]$ for exactly one $i \in \{1,3\}$.

In what follows, we provide the convex hull results from [94]. In Section 5.1.2 we present the convex hull characterization of the homogeneous version of the quadratic set S defined in (5.1). Section 5.1.3 then presents the convex hull characterization of S.

5.1.2 homogeneous quadratic sets

Consider the homogeneous version of the quadratic function q defined in (5.2) as

$$\tilde{\boldsymbol{q}} = y^T \mathcal{P} y, \tag{5.3}$$

where $y = \begin{bmatrix} x \\ x_0 \end{bmatrix} \in \mathbb{R}^{n+1}$. Also consider the homogeneous version of the quadratic set S defined in (5.1) as

$$\boldsymbol{\mathcal{S}} := \left\{ y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\boldsymbol{i}} < 0, \quad i = 1,2 \right\}. \tag{5.4}$$

Analogously, define the associated quadratic pencil $\tilde{\boldsymbol{q}}_{\boldsymbol{\lambda}}$ as

$$\tilde{\boldsymbol{q}}_{\boldsymbol{\lambda}} := (1-\lambda)\tilde{\boldsymbol{q}}_{\boldsymbol{1}} + \lambda \tilde{\boldsymbol{q}}_{\boldsymbol{2}},$$

where $\lambda \in [0,1]$. Also denote the homogeneous version of the set S_λ as

$$\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} := \left\{ y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\boldsymbol{\lambda}} < 0 \right\}.$$

Throughout the paper, we use the following definitions.

Definition 5. $\boldsymbol{\mathcal{C}} \subseteq \mathbb{R}^{n+1}$ *is a cone if for any* $y \in \boldsymbol{\mathcal{C}}$ *and* $\alpha > 0$*, we have* $\alpha y \in \boldsymbol{\mathcal{C}}$*.*

We note that the above definition of a cone $\boldsymbol{\mathcal{C}}$ does not require $0 \in \boldsymbol{\mathcal{C}}$, and it also allows a non-convex set to be a cone.

Definition 6. *The symmetric reflection of* $\boldsymbol{\mathcal{C}} \subseteq \mathbb{R}^{n+1}$ *with respect to the origin is defined as* $-\boldsymbol{\mathcal{C}} := \{-y \in \mathbb{R}^{n+1} \, : \, y \in \boldsymbol{\mathcal{C}}\}$*.*

Definition 7. $\boldsymbol{\mathcal{C}} \subseteq \mathbb{R}^{n+1}$ *is symmetric if* $-\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{C}}$*.*

Also define a linear hyperplane $\boldsymbol{\mathcal{H}} \subseteq \mathbb{R}^{n+1}$ with the associated normal vector $h \in \mathbb{R}^{n+1}$ as

$$\boldsymbol{\mathcal{H}} := \left\{ y \in \mathbb{R}^{n+1} \, : \, h^T y = 0 \right\}.$$

One can see that $\boldsymbol{\mathcal{S}}$ and $\boldsymbol{\mathcal{S}_\lambda}$ for $\lambda \in [0, 1]$ are open symmetric cones. An important notion that we frequently use throughout the paper is the separation of an open symmetric cone which is given in the following definition.

Definition 8. *Consider an open symmetric non-empty cone* $\boldsymbol{\mathcal{C}} \in \mathbb{R}^{n+1}$*. If there exists a linear hyperplane* $\boldsymbol{\mathcal{H}} \subseteq \mathbb{R}^{n+1}$ *such that* $\boldsymbol{\mathcal{H}} \cap \boldsymbol{\mathcal{C}} = \emptyset$*, we say* $\boldsymbol{\mathcal{C}}$ *admits a separation (i.e.,* $\boldsymbol{\mathcal{H}}$ *is a separator of* $\boldsymbol{\mathcal{C}}$ *or separates* $\boldsymbol{\mathcal{C}}$*).*

Denote the two half spaces induced by the hyperplane $\boldsymbol{\mathcal{H}}$ as

$$\boldsymbol{\mathcal{H}}^+ := \left\{ y \in \mathbb{R}^{n+1} \, : \, h^T y > 0 \right\},$$

and

$$\boldsymbol{\mathcal{H}}^- := \left\{ y \in \mathbb{R}^{n+1} \, : \, h^T y < 0 \right\}.$$

Therefore, a separator $\boldsymbol{\mathcal{H}}$ induces two disjoint slices of the set $\boldsymbol{\mathcal{S}}$ denoted by

$$\boldsymbol{\mathcal{S}}^+ := \boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}} \quad \text{and} \quad \boldsymbol{\mathcal{S}}^- := \boldsymbol{\mathcal{H}}^- \cap \boldsymbol{\mathcal{S}}.$$

One can see that the resulting slices of $\boldsymbol{\mathcal{S}}$ satisfy the following properties: (i) $\boldsymbol{\mathcal{S}}^+ = -\boldsymbol{\mathcal{S}}^-$, (ii) $\boldsymbol{\mathcal{S}}^+ \cap \boldsymbol{\mathcal{S}}^- = \emptyset$, and (iii) $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{S}}^+ \cup \boldsymbol{\mathcal{S}}^-$.

Another important definition that we need is the definition of a semi-convex cone.

Definition 9. *A semi-convex cone (SCC) is the union of two convex cones which are symmetric reflections of each other with respect to the origin.*

An SCC is symmetric by definition. Moreover, an open SCC always admits a unique separation. In other words, regardless of the separator we use to separate an SCC with, the associated disjoint slices will always be the same (i.e., after using any one of the valid hyperplanes for separation, the two pieces of the SCC are uniquely defined). This fact is formalized in the following proposition from [94].

Proposition 21. *Let $\boldsymbol{\mathcal{C}} \subseteq \mathbb{R}^{n+1}$ be an open SCC. Assume that there exists a hyperplane $\boldsymbol{\mathcal{H}}$ which separates $\boldsymbol{\mathcal{C}}$. Then, $\boldsymbol{\mathcal{C}}$ admits a unique separation, the slices of which are the convex connected components of $\boldsymbol{\mathcal{C}}$.*

We also use the following useful proposition from [94].

Proposition 22. *Consider an open symmetric non-empty cone given by*

$$\boldsymbol{\mathcal{C}} := \left\{ y \in \mathbb{R}^{n+1} \, : \, y^T \mathcal{P} y < 0 \right\}.$$

Then the following statements are equivalent:

1. *There exists a linear hyperplane which separates $\boldsymbol{\mathcal{C}}$,*
2. *$\pi\left(\mathcal{P}\right) = 1$, and*
3. *$\boldsymbol{\mathcal{C}}$ is an SCC.*

Remark 1. *Note that when $\pi\left(\mathcal{P}\right) = 1$, one can do the spectral decomposition of $\mathcal{P}$ as*

$$\mathcal{P} = VV^T - nn^T,$$

for $n \in \mathbb{R}^{n+1}$ and $V \in \mathbb{R}^{(n+1) \times \nu}$, where ν represents the number of positive eigenvalues of $\mathcal{P}$. One can check that

$$\boldsymbol{\mathcal{H}_n} := \left\{ y \in \mathbb{R}^{n+1} \, : \, n^T y = 0 \right\}$$

separates $\boldsymbol{\mathcal{C}}$ and we call $\boldsymbol{\mathcal{H}_n}$ a natural *separator of $\boldsymbol{\mathcal{C}}$.*

The following Theorem from [94] characterizes the convex hull of any set of the form $\boldsymbol{\mathcal{S}}$ defined by two homogeneous quadratic inequalities.

Theorem 4. *Consider the non-empty open set $\boldsymbol{\mathcal{S}}$ defined in (5.4) and let $\boldsymbol{\mathcal{H}}$ be a separator of $\boldsymbol{\mathcal{S}}$. Then $E \neq \emptyset$ and exactly one of the connected components $[\lambda_i, \lambda_{i+1}]$ of E is such that*

$$\boldsymbol{\mathcal{H}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}} = \emptyset.$$

For such λ_i and λ_{i+1} we have that $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}}$ is an SCC,

$$\operatorname{conv}\left(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}\right) = \boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}}$$

and there exists $\boldsymbol{\mathcal{H}_s}$ which separates both $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}}$ and $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}}$ such that

$$\operatorname{conv}\left(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}\right) = \left(\boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}}\right) \cap \left(\boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}}\right).$$

5.1.3 Quadratic sets

Using the results from Theorem 4, the following theorem from [94] characterizes the convex hull of any set of the form S defined by two strict quadratic inequalities.

Theorem 5. *Consider the non-empty open set S defined in (5.1). If $D = \emptyset$, then* $\operatorname{conv}(S) = \mathbb{R}^n$. *Otherwise, let $i \in \{1, 3\}$ be such that $[\lambda_i, \lambda_{i+1}]$ is the unique connected component of E such that $D \subseteq [\lambda_i, \lambda_{i+1}]$. For such λ_i and λ_{i+1} we have*

$$\operatorname{conv}(S) = S_{\lambda_i} \cap S_{\lambda_{i+1}}.$$

5.2 CONIC QUADRATIC CHARACTERIZATION OF CONVEX HULLS

In this section we first show that the convex hull characterizations presented in Section 5.1 can be described by two strict conic quadratic inequalities. Using results from Theorem 4, we then derive strict conic quadratic inequalities which provide a relaxation for the convex hull of sets defined as the intersection of a strict conic quadratic and a quadratic inequality. We also show that such valid inequalities characterize the convex hull exactly under an additional assumption.

5.2.1 Conic quadratic representation of convex hulls

In what follows, we show that each side of $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}_i}$ and $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}_{i+1}}$ can be described by a single conic quadratic inequality, where $[\lambda_i, \lambda_{i+1}]$ for $i \in \{1, 3\}$ is one of the connected components of E.

Proposition 23. *Let $\lambda \in [0,1]$ be such that $\pi(\mathcal{P}_\lambda) = 1$ and let $\boldsymbol{\mathcal{H}}$ be a separator of $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}}$. Then $\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}}$ can be described by a single strict conic quadratic inequality.*

Proof. We have

$$\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} = \left\{ y \in \mathbb{R}^{n+1} \,:\, y^T \mathcal{P}_\lambda y < 0 \right\}.$$

Since $\pi(\mathcal{P}_\lambda) = 1$, using Proposition 22, one can see that $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}}$ is an SCC. Thus, using Remark 1, one can decompose $\mathcal{P}_\lambda$ as $\mathcal{P}_\lambda = VV^T - nn^T$ for the appropriately chosen matrix and vector V and n. Therefore, we have

$$\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} = \left\{ y \in \mathbb{R}^{n+1} \,:\, \left\| V^T y \right\|_2^2 < \left(n^T y \right)^2 \right\}. \tag{5.5}$$

Let $\boldsymbol{\mathcal{H}}_{\boldsymbol{n}}$ be the natural separator of $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}}$. Using Proposition 21, we have that $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}}$ admits a unique separation, that is,

$$\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} = \boldsymbol{\mathcal{H}}_{\boldsymbol{n}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} \quad \text{or} \quad \boldsymbol{\mathcal{H}}^- \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} = \boldsymbol{\mathcal{H}}_{\boldsymbol{n}}^- \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}}. \tag{5.6}$$

Therefore, from (5.5) and (5.6) we get

$$\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}} = \left\{ y \in \mathbb{R}^{n+1} \,:\, \left\| V^T y \right\|_2 < s\left(n^T y \right) \right\},$$

for some $s \in \{-1, 1\}$. □

A similar argument to the proof of Proposition 23 can be used to show that $\operatorname{conv}(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}})$ given in Theorem 4 can be written as

$$\operatorname{conv}\left(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}} \right) = \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_i} \cap \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_{i+1}},$$

where

$$\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_i} = \boldsymbol{\mathcal{H}}_{\boldsymbol{i}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}_i} \quad \text{and} \quad \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_{i+1}} = \boldsymbol{\mathcal{H}}_{\boldsymbol{i+1}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}_{i+1}}, \tag{5.7}$$

$\boldsymbol{\mathcal{H}}_i^+ \in \left\{\boldsymbol{\mathcal{H}}_{n_i}^+, \boldsymbol{\mathcal{H}}_{n_i}^-\right\}$ and $\boldsymbol{\mathcal{H}}_{i+1}^+ \in \left\{\boldsymbol{\mathcal{H}}_{u_{i+1}}^+, \boldsymbol{\mathcal{H}}_{u_{i+1}}^-\right\}$, and where $\boldsymbol{\mathcal{H}}_{u_i}$ and $\boldsymbol{\mathcal{H}}_{u_{i+1}}$ are natural separators of $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}_i}$ and $\boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda}_{i+1}}$, respectively. In particular, each of the sets $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_i}$ and $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_{i+1}}$ is described by a single strict conic quadratic inequality.

Similarly, conv (S) given in Theorem 5 can be expressed as

$$\text{conv}\,(S) = K_{\lambda_i} \cap K_{\lambda_{i+1}},$$

where

$$K_{\lambda_i} = \left\{ x \in \mathbb{R}^n \,:\, \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_i} \right\} \quad \text{and} \quad K_{\lambda_{i+1}} = \left\{ x \in \mathbb{R}^n \,:\, \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_{i+1}} \right\}, \tag{5.8}$$

for $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_i}$ and $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_{i+1}}$ defined in (5.7). In particular, K_{λ_i} and $K_{\lambda_{i+1}}$ can be described by a single strict conic quadratic inequality. An alternate way of obtaining such conic quadratic inequalities is to apply Schur's Lemma to a homogeneous version of the SDP representation of S_{λ_i} and $S_{\lambda_{i+1}}$ given in Proposition A1 in [94].

5.2.2 Conic quadratic sets

In this section we aim to characterize the convex hull of sets defined by a strict conic quadratic and a strict quadratic inequality.

Using Theorem 4, we first derive valid conic quadratic inequalities for the convex hull of any set defined by a strict conic quadratic and a strict quadratic inequality. We then show that such valid inequalities characterize the convex hull exactly under an additional assumption.

We study open sets of the form

$$C := \left\{ x \in \mathbb{R}^n \,:\, L_1 < 0, \quad q_2 < 0 \right\}, \tag{5.9}$$

where $L_1 < 0$ is a strict conic quadratic inequality of the form

$$\left\| A_1 x - d_1 \right\|_2 < a_1^T x - a_0,$$

where $A_1 \in \mathbb{R}^{n \times n}$, $d_1, a_1 \in \mathbb{R}^n$, $a_0 \in \mathbb{R}$, and $q_2 < 0$ is a strict quadratic inequality of the form

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^T \mathcal{P}_2 \begin{bmatrix} x \\ 1 \end{bmatrix} = x^T Q_2 x + 2b_2^T x + \gamma_2 < 0,$$

where $\mathcal{P}_2 = \begin{bmatrix} Q_2 & b_2 \\ b_2^T & \gamma_2 \end{bmatrix} \in \mathbb{S}^{n+1}$, $Q_2 \in \mathbb{S}^n, b_2 \in \mathbb{R}^n$, and $\gamma_2 \in \mathbb{R}$.

Our goal is to derive strong valid inequalities for $\operatorname{conv}(C)$ and characterize the convex hull exactly when possible. Since we will use results from Theorem 4, we also need to consider the homogeneous version of the set C. Therefore, we define

$$\boldsymbol{\mathcal{C}} := \left\{ y \in \mathbb{R}^{n+1} \, : \, \boldsymbol{\mathcal{L}_1} < 0, \quad \boldsymbol{\tilde{q}_2} < 0 \right\}, \tag{5.10}$$

where $\boldsymbol{\mathcal{L}_1} < 0$ is a strict homogeneous conic quadratic inequality of the form

$$\|A_1 x - d_1 x_0\|_2 < a_1^T x - a_0 x_0,$$

and $\boldsymbol{\tilde{q}_2}$ is a strict homogeneous quadratic function as defined in (5.3). By squaring both sides of the strict conic quadratic inequality $\boldsymbol{\mathcal{L}_1} < 0$, we define

$$\boldsymbol{\mathcal{S}}\left(\boldsymbol{C}\right) := \left\{ y \in \mathbb{R}^{n+1} \, : \, \boldsymbol{\tilde{q}_1} < 0, \quad \boldsymbol{\tilde{q}_2} < 0 \right\}, \tag{5.11}$$

where $\boldsymbol{\tilde{q}_1} = y^T \mathcal{P}_1 y$ such that $Q_1 = A_1^T A_1 - a_1 a_1^T$, $b_1 = -A_1^T d_1 + a_0 a_1$, and $\gamma_1 = d_1^T d_1 - a_0^2$. We also define the hyperplane

$$\boldsymbol{\mathcal{H}_0} := \left\{ y \in \mathbb{R}^{n+1} \, : \, (a_1, -a_0)^T \, y = 0 \right\}. \tag{5.12}$$

One can see that $\boldsymbol{\mathcal{H}_0}$ is a separator for $\boldsymbol{\mathcal{S}}\left(\boldsymbol{C}\right)$,

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{H}_0^+} \cap \boldsymbol{\mathcal{S}}\left(\boldsymbol{C}\right), \tag{5.13}$$

and

$$C = \boldsymbol{\mathcal{H}_0^+} \cap \boldsymbol{\mathcal{S}}\left(\boldsymbol{C}\right) \cap \boldsymbol{\mathcal{E}^1}, \tag{5.14}$$

where $\boldsymbol{\mathcal{E}^1} := \{(x, x_0) \in \mathbb{R}^{n+1} \, : \, x_0 = 1\}$. In Proposition 24, we use (5.13) and (5.14) together with Theorem 4 to characterize $\operatorname{conv}(C)$. We note that the proof of Proposition 24 is a direct adaptation of the proof of Theorem 1 in [94].

Proposition 24. *Consider the non-empty open set C defined in (5.9). Then exactly one of the connected components $[\lambda_i, \lambda_{i+1}]$ of E is such that*

$$\boldsymbol{\mathcal{H}_0} \cap \boldsymbol{\mathcal{S}_{\lambda_i}} \cap \boldsymbol{\mathcal{S}_{\lambda_{i+1}}} = \emptyset, \tag{5.15}$$

where $\boldsymbol{\mathcal{H}_0}$ is defined in (5.12). For such λ_i and λ_{i+1} we have that

$$\text{conv}\,(C) \subseteq K_{\lambda_i} \cap K_{\lambda_{i+1}}, \tag{5.16}$$

where K_{λ_i} and $K_{\lambda_{i+1}}$ are defined in (5.8). Furthermore, if $\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{E}}^+$ for $\boldsymbol{\mathcal{E}} := \{(x, x_0) \in \mathbb{R}^{n+1} \,:\, x_0 = 0\}$, then (5.16) holds as equality.

Proof. Consider $\boldsymbol{\mathcal{C}}$, $\boldsymbol{\mathcal{S}}\,(\boldsymbol{C})$, and $\boldsymbol{\mathcal{H}_0}$ as defined in (5.10), (5.11), and (5.12), respectively. One can see that (5.15) directly follows from Theorem 4. To prove the containment in (5.16), recall from (5.13) and (5.14) that

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{H}_0^+} \cap \boldsymbol{\mathcal{S}}\,(\boldsymbol{C})$$

and

$$C = \boldsymbol{\mathcal{H}_0^+} \cap \boldsymbol{\mathcal{S}}\,(\boldsymbol{C}) \cap \boldsymbol{\mathcal{E}^1},$$

where $\boldsymbol{\mathcal{E}^1} := \{(x, x_0) \in \mathbb{R}^{n+1} \,:\, x_0 = 1\}$. Therefore, $\text{conv}\,(C)$ can be expressed as

$$\begin{aligned}
\text{conv}\,(C) &= \left\{ x \in \mathbb{R}^n \,:\, \begin{bmatrix} x \\ 1 \end{bmatrix} = \sum_{j=1}^{n+1} \theta_j \begin{bmatrix} z_j \\ 1 \end{bmatrix},\ \sum_{j=1}^{n+1} \theta_j = 1,\ \theta_j \geq 0,\ \begin{bmatrix} z_j \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{C}},\ j \in [n+1] \right\} \\
&\subseteq \left\{ x \in \mathbb{R}^n \,:\, \begin{bmatrix} x \\ 1 \end{bmatrix} = \sum_{j=1}^{n+1} \theta_j \tilde{z}_j,\ \sum_{j=1}^{n+1} \theta_j = 1,\ \theta_j \geq 0,\ \tilde{z}_j \in \boldsymbol{\mathcal{C}},\ j \in [n+1] \right\} &(5.17) \\
&= \left\{ x \in \mathbb{R}^n \,:\, \begin{bmatrix} x \\ 1 \end{bmatrix} \in \text{conv}\,(\boldsymbol{\mathcal{C}}) \right\}, \\
&= \left\{ x \in \mathbb{R}^n \,:\, \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{K}_{\lambda_i}} \cap \boldsymbol{\mathcal{K}_{\lambda_{i+1}}} \right\} = K_{\lambda_i} \cap K_{\lambda_{i+1}}, &(5.18)
\end{aligned}$$

where the first equality holds by Carathéodory's Theorem, the first equality in (5.18) follows from Theorem 4, and where $i \in \{1, 3\}$ is an appropriate index evident from Theorem 4. The reverse containment in (5.17) trivially holds when $\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{E}}^+$. □

5.3 CONIC QUADRATIC CHARACTERIZATION OF CLOSED CONVEX HULLS

In this section we study conic and quadratic sets defined by non-strict inequalities instead of strict inequalities. In particular, whenever we refer to a previously defined set, such as $\boldsymbol{\mathcal{S}}$ or $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda}_i}$, we redefine such a set by replacing strict inequalities with non-strict inequalities. Working with non-strict inequalities requires the study of closed convex hulls instead of convex hulls. However, under certain topological assumptions, the strict inequality results directly imply non-strict analogs. One such assumption is condition (5.19) in the following lemma.

Lemma 10. *Let A and B be two non-empty closed sets such that*

$$A \subseteq \overline{\operatorname{int}(A)} \tag{5.19}$$

and B is convex. If $\operatorname{conv}(\operatorname{int}(A)) \subseteq \operatorname{int}(B)$, then $\overline{\operatorname{conv}}(A) \subseteq B$ and if $\operatorname{conv}(\operatorname{int}(A)) = \operatorname{int}(B)$, then $\overline{\operatorname{conv}}(A) = B$.

Proof. First note that (5.19) implies $A = \overline{\operatorname{int}(A)}$ and hence

$$\overline{\operatorname{conv}}(A) = \overline{\operatorname{conv}}\left(\overline{\operatorname{int}(A)}\right) = \overline{\operatorname{conv}}\left(\operatorname{int}(A)\right) = \overline{\operatorname{conv}}\left(\operatorname{conv}\left(\operatorname{int}(A)\right)\right). \tag{5.20}$$

Furthermore,

$$B = \overline{\operatorname{int}\left(B\right)} = \overline{\operatorname{conv}}\left(\operatorname{int}\left(B\right)\right) \tag{5.21}$$

because B is closed and convex and $\operatorname{int}(B) \neq \emptyset$ (because $\operatorname{int}(A) \subseteq \operatorname{int}(B)$ and because (5.19) and $A \neq \emptyset$ imply $\operatorname{int}(A) \neq \emptyset$). The result then follows from (5.20)–(5.21) by taking the closed convex hull on both sides of the corresponding containment or equality. □

In the following subsections we show how Lemma 10 can be used to adapt the convex hull results from Sections 5.1 and 5.2 to the non-strict setting. We then give several examples that illustrate condition (5.19) and some characteristics of the closed convex hull results. Finally, considering sets defined by non-strict inequalities allows us to compare our results with those in [27].

5.3.1 Homogeneous quadratic sets

We consider homogeneous quadratic sets of the form

$$\boldsymbol{\mathcal{S}} := \left\{ y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\boldsymbol{i}} \leq 0, \quad i = 1, 2 \right\}.$$

In this section with a slight abuse of notation, we say that the hyperplane $\boldsymbol{\mathcal{H}} \subseteq \mathbb{R}^{n+1}$ separates $\boldsymbol{\mathcal{S}}$ when $\boldsymbol{\mathcal{H}}$ is in fact a separator of $\operatorname{int}(\boldsymbol{\mathcal{S}}) = \{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\boldsymbol{i}} < 0, \quad i = 1, 2\}$. The following corollary characterizes the non-strict inequality version of Theorem 4 and follows directly from that theorem and Lemma 10.

Corollary 11. *Let* $\boldsymbol{\mathcal{S}} := \{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\boldsymbol{i}} \leq 0, \quad i = 1, 2\}$ *such that* $\operatorname{int}(\boldsymbol{\mathcal{S}}) \neq \emptyset$*,* $\boldsymbol{\mathcal{H}}$ *be a separator of* $\boldsymbol{\mathcal{S}}$*, and* $i \in \{1, 3\}$ *be such that* $[\lambda_i, \lambda_{i+1}]$ *is the unique connected component of* E *such that*

$$\boldsymbol{\mathcal{H}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}} = \emptyset.$$

If

$$\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}} \subseteq \overline{\operatorname{int}\left(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}\right)}, \tag{5.22}$$

then

$$\overline{\operatorname{conv}}\left(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}}\right) = \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda_i}} \cap \boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda_{i+1}}},$$

where $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda_i}}$ *and* $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda_{i+1}}}$ *are as in* (5.7) *defined with non-strict inequalities.*

5.3.2 Conic quadratic sets

The following corollary characterizes the non-strict inequality version of Proposition 24 and follows directly from that proposition and Lemma 10.

Corollary 12. *Let* $C := \{x \in \mathbb{R}^n \,:\, L_1 \leq 0, \quad q_2 \leq 0\}$ *such that* $\operatorname{int}(C) \neq \emptyset$ *and* $i \in \{1, 3\}$ *be such that* $[\lambda_i, \lambda_{i+1}]$ *is the unique connected component of* E *such that*

$$\boldsymbol{\mathcal{H}}_{\mathbf{0}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_i}} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{\lambda_{i+1}}} = \emptyset,$$

where $\boldsymbol{\mathcal{H}}_{\mathbf{0}}$ *is defined in* (5.12)*. If*

$$C \subseteq \overline{\operatorname{int}(C)}, \tag{5.23}$$

then

$$\overline{\text{conv}}\,(C) \subseteq K_{\lambda_i} \cap K_{\lambda_{i+1}}, \tag{5.24}$$

where K_{λ_i} *and* $K_{\lambda_{i+1}}$ *are as in* (5.8) *defined with non-strict inequalities. Furthermore, if*

$$\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{E}}^+, \tag{5.25}$$

for $\boldsymbol{\mathcal{E}} := \{(x, x_0) \in \mathbb{R}^{n+1} \,:\, x_0 = 0\}$*, then* (5.24) *holds as equality.*

Note that $\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{E}}^+$ provides a sufficient condition under which (5.24) trivially holds as equality; however, equality in (5.24) may still hold even if $\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{E}}^+$ is violated.

5.3.3 Quadratic sets

The following corollary characterizes the non-strict inequality version of Theorem 5 and follows directly from that theorem and Lemma 10.

Corollary 13. *Let* $S := \{x \in \mathbb{R}^n \,:\, q_i \leq 0, \quad i = 1, 2\}$ *such that* $\text{int}\,(S) \neq \emptyset$*. If* $D = \emptyset$*, then* $\overline{\text{conv}}\,(S) = \mathbb{R}^n$*. Otherwise, let* $i \in \{1, 3\}$ *be such that* $[\lambda_i, \lambda_{i+1}]$ *is the unique connected component of* E *such that* $D \subseteq [\lambda_i, \lambda_{i+1}]$*. If*

$$S \subseteq \overline{\text{int}\,(S)}, \tag{5.26}$$

then

$$\overline{\text{conv}}\,(S) = K_{\lambda_i} \cap K_{\lambda_{i+1}},$$

where K_{λ_i} *and* $K_{\lambda_{i+1}}$ *are as in* (5.8) *defined with non-strict inequalities.*

Finally, note that λ_i and λ_{i+1} can be obtained by a polynomial-time algorithm based on the S-Lemma and the calculation of generalized eigenvalues of the pencil $\mathcal{P}_\lambda$ (Algorithm 1 in [94]). Therefore, once topological condition (5.19) is verified, one can obtain the conic quadratic relaxation of $\overline{\text{conv}}\,(\boldsymbol{\mathcal{H}}^+ \cap \boldsymbol{\mathcal{S}})$, $\overline{\text{conv}}\,(S)$, or $\overline{\text{conv}}\,(C)$ in polynomial time. This relaxation is always a characterization for the first two sets and if condition (5.25) is satisfied (which can be checked in polynomial time), it is also a characterization for the last one. A simple version of the algorithm to obtain λ_i and λ_{i+1} calculates all generalized eigenvalues $\{\lambda_i\}_{i=1}^r$ of the pencil $\mathcal{P}_\lambda$, which can be done in polynomial-time. Assuming such generalized eigenvalues are ordered such that $\lambda_i < \lambda_{i+1}$ for all $i \in [r-1]$, we can construct E by evaluating the number of negative eigenvalues of $\mathcal{P}_\lambda$ for all $\lambda = (\lambda_i + \lambda_{i+1})/2$ for $i \in [r-1]$. The remaining check to determine the appropriate connected component of E can also be done in polynomial time.

5.3.4 Verifying the topological condition

In this section we give two lemmas that are useful when checking the topological condition (5.19). The first lemma shows that the condition is automatically satisfied for a wide range of sets and the second lemma gives a sufficient condition that can often be easier to check than the original condition (5.19).

Lemma 11. *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be continuous functions and K be a closed convex set or the complement of an open convex set. Then (5.19) holds for $A = \{(x, x_0) : f(x) \leq x_0, \quad g(x) \leq x_0\}$ and $A = \{(x, x_0) : f(x) \leq x_0, \quad x \in K\}$.*

Proof. The first case follows by noting that for any $(x, x_0) \in A$ and for every $\varepsilon > 0$, we have that $(x, x_0 + \varepsilon) \in \operatorname{int}(A)$. For the second case, note that for any $\bar{x} \in \operatorname{bd}(K)$, there exist $d \in \mathbb{R}^n$ such that $\bar{x} + \varepsilon d \in \operatorname{int}(K)$ for all sufficiently small $\varepsilon > 0$. Furthermore, $(\bar{x}, f(\bar{x})) = \lim_{\varepsilon \to 0} (\bar{x} + \varepsilon d, f(\bar{x} + \varepsilon d))$. Hence, it suffices to show that $(\bar{x} + \varepsilon d, f(\bar{x} + \varepsilon d)) \in \overline{\operatorname{int}(A)}$ for all sufficiently small $\varepsilon > 0$. This follows from noting that $(\bar{x} + \varepsilon d, f(\bar{x} + \varepsilon d) + \delta) \in \operatorname{int}(A)$ for all sufficiently small $\varepsilon > 0$ and for any $\delta > 0$. $\square$

Sets of the form considered by Lemma 11 include a wide range of quadratic sets such as the intersection of a paraboloid with a general quadratic inequality. It also includes trust region problems and hence, together with Corollary 13, this lemma can be used to show that such problems are equivalent to simple convex optimization problems (e.g. [17, Corollary 8] and [27, Section 7.2])

Lemma 12. *If $A = \bigcup_{i=1}^{l} A_i$ and A_i satisfies (5.19) for each $i \in [l]$, then A satisfies (5.19). In particular, if A_i is convex and $\operatorname{int}(A_i) \neq \emptyset$ for each $i \in [l]$, then A satisfies (5.19).*

Proof. The first part follows from $A = \bigcup_{i=1}^{l} A_i \subseteq \bigcup_{i=1}^{l} \overline{\operatorname{int}(A_i)} \subseteq \overline{\operatorname{int}(A)}$. The second follows from the fact that (5.19) is naturally satisfied by convex sets with non-empty interiors. $\square$

Sets considered by Corollaries 11–13 that are unions of convex sets include those constructed from two-term disjunctions such as ones considered in [27, Section 6]. Such sets are the unions of two convex sets defined by a single quadratic or conic quadratic inequality and two linear inequalities. In the next sub-section we show that checking that these two convex sets have non-empty interior is often easy and that when one of the sets has an empty interior, the topological condition (5.19) can be violated.

5.3.5 Illustrative examples

We now illustrate the results in this section through several examples. In particular, we show how the two inequalities in the closed convex hull or relaxation characterization may include one of the original inequalities, one or two new inequalities, or even a redundant inequality.

We begin with three examples for which the description of the closed convex hull only requires one additional inequality (i.e. one of the inequalities associated to $\lambda_i,\ \lambda_{i+1}$ is one of the original inequalities). In the first two examples, Corollaries 12 and 13 are able to prove that adding this additional inequality yields the closed convex hull. However, in the third example, Corollary 12 cannot prove that adding the additional inequality yields the closed convex hull even though it actually does.

Example 2. *Here we consider Example 3 in [73], which is given by*

$$S_1 := \left\{x \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 - x_3 - 4 \leq 0, \quad x_1^2 + x_2^2 - x_3^2 + 1 \leq 0\right\}.$$

To check condition (5.26) *of Corollary 13, first note that* $S_1 = S_1' \cup S_1''$ *for convex sets*

$$S_1' := \left\{x \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 - x_3 - 4 \leq 0, \quad \sqrt{x_1^2 + x_2^2 + 1} \leq x_3\right\}$$

and

$$S_1'' := \left\{x \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 - x_3 - 4 \leq 0, \quad \sqrt{x_1^2 + x_2^2 + 1} \leq -x_3\right\}.$$

Furthermore, both sets have non-empty interiors (e.g. $(0,0,2) \in \operatorname{int}(S_1')$ *and* $(0,0,-2) \in \operatorname{int}(S_1'))$*. Hence, by Lemma 12 condition* (5.26) *is satisfied. We can also check that*

$$E = \left[0, \frac{1}{21}\left(9 - 2\sqrt{15}\right)\right] \cup \left[\frac{1}{21}\left(9 + 2\sqrt{15}\right), 1\right]$$

and $D = \{0\}$ *is contained in the first interval. Then,* $\lambda_i = 0$ *and* $\lambda_{i+1} = \frac{1}{21}\left(9 - 2\sqrt{15}\right)$ *and Corollary 13 yields*

$$\overline{\operatorname{conv}}\,(S_1) = \left\{ x \in \mathbb{R}^2 \;:\; \begin{array}{r} x_1^2 + x_2^2 - x_3 - 4 \leq 0, \\ \sqrt{x_1^2 + x_2^2} \leq \frac{1}{21}\sqrt{9 + 2\sqrt{15}}\left(\left(9 - 2\sqrt{15}\right)x_3 + \sqrt{15} + 6\right) \end{array} \right\}.$$

Because $\lambda_i = 0$ *and* $\lambda_{i+1} \notin \{0, 1\}$*, the first inequality given by Corollary 13 is one of the original inequalities and the second one is a new inequality, which we can check is non-redundant for the description of* $\overline{\operatorname{conv}}\,(S_1)$*.*

Example 3. *Here we consider an example proposed by Burer and Kılınç-Karzan [27], which is given by*

$$C_2 := \left\{ x \in \mathbb{R}^3 \;:\; \sqrt{x_1^2 + x_2^2} \leq x_3, \quad (x_1 + x_3 - 3)(x_3 - 2) \leq 0 \right\}.$$

The homogeneous version of this set is given by

$$\boldsymbol{\mathcal{C}_2} := \boldsymbol{\mathcal{H}_2^+} \cap \boldsymbol{\mathcal{S}_2} = \left\{ (x, x_0) \in \mathbb{R}^4 \;:\; \sqrt{x_1^2 + x_2^2} \leq x_3, \quad x_3^2 + x_1x_3 - 2x_1x_0 - 5x_3x_0 + 6x_0^2 \leq 0 \right\},$$

for $\boldsymbol{\mathcal{H}_2^+} := \{(x, x_0) \in \mathbb{R}^4 \;:\; x_3 \geq 0\}$ *and*

$$\boldsymbol{\mathcal{S}_2} := \left\{ (x, x_0) \in \mathbb{R}^4 \;:\; x_1^2 + x_2^2 - x_3^2 \leq 0, \quad x_3^2 + x_1x_3 - 2x_1x_0 - 5x_3x_0 + 6x_0^2 \leq 0 \right\}.$$

To check condition (5.23) *of Corollary 12, first note that* $C_2 = C_2' \cup C_2''$ *for convex sets*

$$C_2' := \left\{ x \in \mathbb{R}^3 \;:\; \sqrt{x_1^2 + x_2^2} \leq x_3, \quad (x_1 + x_3 - 3) \leq 0, \quad (x_3 - 2) \geq 0 \right\},$$

and

$$C_2'' := \left\{ x \in \mathbb{R}^3 \;:\; \sqrt{x_1^2 + x_2^2} \leq x_3, \quad (x_1 + x_3 - 3) \geq 0, \quad (x_3 - 2) \leq 0 \right\}.$$

Furthermore, both C_2' *and* C_2'' *have non-empty interiors. Hence, by Lemma 12 condition* (5.23) *is satisfied. We can also check that* $E = [0, 8/9] \cup [1, 1]$ *and* $\boldsymbol{\mathcal{H}_2}$ *only separates the set associated to the first interval. Then* $\lambda_i = 0$ *and* $\lambda_{i+1} = 8/9$ *and*

$$\overline{\operatorname{conv}}\,(C_2) \subseteq \left\{ x \in \mathbb{R}^3 \;:\; \sqrt{x_1^2 + x_2^2} \leq x_3^2, \quad \sqrt{(ax + bz + c)^2 + \frac{y^2}{9}} \leq dx + ez + f \right\}, \quad (5.27)$$

where

$$
\begin{aligned}
a &= \frac{1}{132}\left(53+3\sqrt{97}\right)\sqrt{\frac{1}{582}\left(53\sqrt{97}-291\right)},\\
b &= \frac{1}{33}\left(20+3\sqrt{97}\right)\sqrt{\frac{1}{582}\left(53\sqrt{97}-291\right)},\\
c &= \frac{1}{132}\left(-248-24\sqrt{97}\right)\sqrt{\frac{1}{582}\left(53\sqrt{97}-291\right)},\\
d &= \frac{1}{132}\sqrt{\frac{1}{2}+\frac{53}{6\sqrt{97}}}\left(3\sqrt{97}-53\right),\\
e &= \frac{1}{33}\sqrt{\frac{1}{2}+\frac{53}{6\sqrt{97}}}\left(3\sqrt{97}-20\right),
\end{aligned}
$$

and

$$
f = \frac{1}{132}\sqrt{\frac{1}{2}+\frac{53}{6\sqrt{97}}}\left(248-24\sqrt{97}\right).
$$

Finally, to check condition (5.25), first note that $\boldsymbol{\mathcal{C}_2} = \boldsymbol{\mathcal{C}'_2} \cup \boldsymbol{\mathcal{C}''_2}$ *for convex sets*

$$
\boldsymbol{\mathcal{C}'_2} := \left\{(x, x_0) \in \mathbb{R}^4 \,:\, \sqrt{x_1^2 + x_2^2} \le x_3, \quad (x_1 + x_3 - 3x_0) \le 0, \quad (x_3 - 2x_0) \ge 0\right\},
$$

and

$$
\boldsymbol{\mathcal{C}''_2} := \left\{(x, x_0) \in \mathbb{R}^4 \,:\, \sqrt{x_1^2 + x_2^2} \le x_3, \quad (x_1 + x_3 - 3x_0) \ge 0, \quad (x_3 - 2x_0) \le 0\right\}.
$$

The conic inequality of $\boldsymbol{\mathcal{C}'_2}$ *implies* $-x_1 - x_3 \le 0$*, which together with its first linear inequality implies* $x_0 \ge 0$*. Similarly, the conic inequality of* $\boldsymbol{\mathcal{C}''_2}$ *implies* $-x_3 \le 0$*, which together with its second linear inequality implies* $x_0 \ge 0$*. Hence, condition (5.25) holds and Corollary 12 implies that (5.27) holds as equality.*

Example 4. *Here we consider an example similar to those of Section 6.2 in [27], which is given by*

$$C_3 := \left\{ x \in \mathbb{R}^2 \,:\, |x_1| \leq x_2, \quad (2x_1 + x_2 - 1)\,(-2x_1 - x_2 - 1) \leq 0 \right\}.$$

The homogeneous version of this set is given by

$$\boldsymbol{\mathcal{C}_3} := \boldsymbol{\mathcal{H}_3^+} \cap \boldsymbol{\mathcal{S}_3} = \left\{ (x, x_0) \in \mathbb{R}^3 \,:\, |x_1| \leq x_2, \quad (2x_1 + x_2 - x_0)\,(-2x_1 - x_2 - x_0) \leq 0 \right\},$$

where $\boldsymbol{\mathcal{H}_3^+} := \{(x, x_0) \in \mathbb{R}^3 \,:\, x_2 \geq 0\}$ *and*

$$\boldsymbol{\mathcal{S}_3} := \left\{ (x, x_0) \in \mathbb{R}^3 \,:\, x_1^2 \leq x_2^2, \quad (2x_1 + x_2 - x_0)\,(-2x_1 - x_2 - x_0) \leq 0 \right\}.$$

Similarly to Example 3 we can check condition (5.23) of Corollary 12 through Lemma 12 as C_3 *is the union of two convex sets with non-empty interior. We can also check that* $E = [0, 1/4] \cup [1, 1]$ *and* $\boldsymbol{\mathcal{H}_3}$ *only separates the set associated to the first interval. Then* $\lambda_i = 0$ *and* $\lambda_{i+1} = 1/4$ *and*

$$\overline{\text{conv}}\,(C_3) \subseteq \left\{ x \in \mathbb{R}^2 \,:\, |x_1| \leq x_2, \quad 1 - x - 2y \leq 0 \right\}. \tag{5.28}$$

We can check that equality holds in (5.28), but condition (5.25) does not hold so Corollary 12 cannot prove this.

For next pair of examples, we have that neither of the inequalities needed to describe the closed convex hull is one of the original inequalities.

Example 5. *Here we consider the set given by*

$$\boldsymbol{\mathcal{S}_1} := \left\{ (x, x_0) \in \mathbb{R}^3 \,:\, 2x_1^2 - x_2^2 - x_0^2 \leq 0, \quad -x_1^2 + x_2^2 - x_0^2 \leq 0 \right\}.$$

One can see that $\boldsymbol{\mathcal{E}} := \{(x, x_0) \in \mathbb{R}^3 \,:\, x_0 = 0\}$ *separates* $\boldsymbol{\mathcal{S}_1}$*. Let* $\boldsymbol{\mathcal{S}_1^+} := \boldsymbol{\mathcal{E}}^+ \cap \boldsymbol{\mathcal{S}_1}$ *and let* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *be the matrices associated with the quadratic inequalities. Condition (5.22) of Corollary 11 can easily be checked using Lemma 11 or by noting that for every* $(x, x_0) \in \boldsymbol{\mathcal{S}_1^+}$ *and* $\varepsilon > 0$ *we have that* $(x, x_0 + \varepsilon) \in \text{int}\left(\boldsymbol{\mathcal{S}_1^+}\right)$*. We can also check that* $E = [1/2, 2/3]$ *and* *that $\boldsymbol{\mathcal{E}}$ separates the set associated to this interval. Then, $\lambda_i = 1/2$ and $\lambda_{i+1} = 2/3$ and Corollary 11 yields*

$$\overline{\operatorname{conv}}\left(\boldsymbol{\mathcal{S}}_{\mathbf{1}}^{+}\right) = \left\{(x, x_0) \in \mathbb{R}^3 \,:\, |x_1| \leq \sqrt{2}x_0, \quad |x_2| \leq \sqrt{3}x_0\right\}.$$

In contrast to Examples 2–4, because $\lambda_i,\ \lambda_{i+1} \notin \{0,1\}$, neither of the inequalities given by Corollary 11 is one of the original inequalities. We can also check that the two new inequalities given by Corollary 11 are non-redundant for the description of $\overline{\operatorname{conv}}\left(\boldsymbol{\mathcal{S}}_{\mathbf{1}}^{+}\right)$.

Example 6. *Consider the Example 1 in [94] and Example 2 in [73], which is given by*

$$S_5 := \left\{x \in \mathbb{R}^2 \,:\, x_1^2 - x_2^2 + 2x_1 + 2 \leq 0, \quad -x_1^2 + x_2^2 + 2x_1 - 2 \leq 0\right\}.$$

We can check condition (5.26) of Corollary 13 through Lemma 12 by noting that S_5 is the union of two (non-convex) sets that satisfy condition (5.19). Alternatively, we can first note that if $x \in S_5$ satisfies both inequalities of S_5 strictly, then $x \in \operatorname{int}(S_5)$ and the condition is trivially satisfied. Furthermore, if $x \in S_5$ satisfies one of the inequalities strictly, we can trivially perturb x so that it remains in S_5 and satisfies both inequalities strictly. Hence, the only nontrivial check of the condition is for points $x \in S_5$ that satisfy both inequalities of S_5 at equality. We can easily check that only two such points exist and each of them satisfy $(x_1 - \varepsilon, x_2) \in \operatorname{int}(S_5)$ for all sufficiently small $\varepsilon > 0$. We can also check that $E = [0, 1/2 - 1/(2\sqrt{2})] \cup [1/2, 1/2 + 1/(2\sqrt{2})]$ and $D = \{1/2\} \subseteq [1/2, 1/2 + 1/(2\sqrt{2})]$. Then, $\lambda_i = 1/2$ and $\lambda_{i+1} = 1/2 + 1/(2\sqrt{2})$ and Corollary 13 yields

$$\overline{\operatorname{conv}}\left(S_5\right) = \left\{x \in \mathbb{R}^2 \,:\, x_1 \leq 0, \quad |y| \leq \sqrt{2} - x\right\}.$$

Again, because $\lambda_i,\ \lambda_{i+1} \notin \{0,1\}$, neither of the inequalities given by Corollary 13 is one of the original inequalities. We can also check that the two new inequalities given by Corollary 13 are non-redundant for the description of $\overline{\operatorname{conv}}\left(S_5\right)$.

For the following example, we have that $\lambda_i = \lambda_{i+1}$, so Corollary 11 yields a unique inequality.

Example 7. *Here we consider the homogeneous version of the example from Section 4.4 in [27], which is given by*

$$\boldsymbol{\mathcal{C}_6} := \boldsymbol{\mathcal{H}_6^+} \cap \boldsymbol{\mathcal{S}_6} = \left\{(x, x_0) \in \mathbb{R}^3 \,:\, |x_1| \le x_2, \quad x_1(x_2 - x_0) \le 0\right\},$$

where $\boldsymbol{\mathcal{H}_6^+} := \{(x, x_0) \in \mathbb{R}^3 \,:\, x_2 \ge 0\}$ *and* $\boldsymbol{\mathcal{S}_6} := \{(x, x_0) \in \mathbb{R}^3 \,:\, x_1^2 \le x_2^2, \quad x_1x_2 - x_1x_0 \le 0\}$. *Condition* (5.26) *of Corollary 11 can easily be checked through Lemma 12 by noting that* $\boldsymbol{\mathcal{C}_6}$ *is the union of two convex sets with non-empty interior. We may hence use Corollary 11 to construct* $\overline{\text{conv}}\left(\boldsymbol{\mathcal{C}_6}\right)$*. For that note that* $E = [0, 0] \cup [1, 1]$*, and that* $\boldsymbol{\mathcal{H}_6}$ *only separates the set associated to the first interval. Hence,* $\lambda_i = \lambda_{i+1} = 0$ *and*

$$\overline{\text{conv}}\left(\boldsymbol{\mathcal{C}_6}\right) = \boldsymbol{\mathcal{K}_0} = \left\{(x, x_0) \in \mathbb{R}^3 \,:\, |x_1| \le x_2\right\}.$$

Finally, note that we trivially have $\overline{\text{conv}}\left(\boldsymbol{\mathcal{C}_6}\right) \subseteq \{(x, x_0) \in \mathbb{R}^3 \,:\, |x_1| \le x_2\}$*. However, the equality in this containment proven by Corollary 11 is not trivial.*

We end this section by considering an example where topological condition (5.19) fails and discussing one possible way to adapt the results in this paper to such a setting. This example also illustrates how one of the inequalities in the closed convex hull characterization may be redundant.

Example 8. *For any* $\varepsilon \ge 0$*, consider the generalization of the example from Section 4.5 in [27], which is given by*

$$\boldsymbol{\mathcal{C}_7}\left(\varepsilon\right) := \boldsymbol{\mathcal{H}_7^+} \cap \boldsymbol{\mathcal{S}_7}\left(\varepsilon\right) := \left\{(x_1, x_0) \in \mathbb{R}^2 \,:\, |x_1| \le x_0, \quad 2x_1x_0 - (2+\varepsilon)x_1^2 \le 0\right\},$$

where $\boldsymbol{\mathcal{H}_7^+} := \{(x_1, x_0) \in \mathbb{R}^2 \,:\, x_0 \ge 0\}$ *and* $\boldsymbol{\mathcal{S}_7}\left(\varepsilon\right) := \{(x_1, x_0) \in \mathbb{R}^2 \,:\, x_1^2 \le x_0^2, \quad 2x_1x_0 - (2+\varepsilon)x_1^2 \le 0\}$*. If we let* $\mathcal{P}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ *and* $\mathcal{P}_2\left(\varepsilon\right) = \begin{bmatrix} -(2+\varepsilon) & 1 \\ 1 & 0 \end{bmatrix}$ *be the matrices associated to* $\boldsymbol{\mathcal{S}_7}\left(\varepsilon\right)$*, we have that*

$$E = \left[0, \frac{1}{2} - f(\varepsilon)\right] \cup \left[\frac{1}{2} + f(\varepsilon), 1\right],$$

where $f\left(\varepsilon\right) := \frac{1}{2}\sqrt{\frac{\varepsilon}{4+\varepsilon}}$*. If* $\varepsilon > 0$*, then* E *is composed of two intervals and we can check that* $\boldsymbol{\mathcal{H}_7}$ *only separates the sets associated to the first interval. The inequality associated to* $\lambda_i = 0$ *is the conic constraint* $|x_1| \le x_0$ *and the one associated to* $\lambda_{i+1} = \frac{1}{2} - f(\varepsilon)$ *is dominated by* *this conic constraint and is hence redundant. We can also check that condition (5.22) is satisfied and then by Corollary 11, we have*

$$\overline{\operatorname{conv}}\left(\boldsymbol{\mathcal{C}_7}\left(\varepsilon\right)\right) = \left\{(x_1, x_0) \in \mathbb{R}^2 \,:\, |x_1| \leq x_0\right\}. \tag{5.29}$$

In contrast, if $\varepsilon = 0$*, we have that* E *becomes the complete interval* $[0, 1]$ *and we instead obtain* $\lambda_{i+1} = 1$*. We can check that in this case (5.29) still holds, but the inequality associated to* $\lambda_{i+1} = 1$ *implies* $x_1 \leq 0$*, which removes a portion of the closed convex hull and is hence invalid. This aligns with the fact that condition (5.22) is not satisfied for* $\varepsilon = 0$ *and hence Corollary 11 cannot characterize relaxations of* $\overline{\operatorname{conv}}\left(\boldsymbol{\mathcal{C}_5}\left(\varepsilon\right)\right)$*.*

The construction of E *in [94] explicitly considers the possibility of* $E = [\lambda_1, \lambda_2] \cup [\lambda_3, \lambda_4]$ *with* $\lambda_2 = \lambda_3$ *and relates the* λ_i*'s to the rank (and in particular singularity) of the pencil* $\mathcal{P}_\lambda = (1-\lambda)\mathcal{P}_1 + \lambda\mathcal{P}_2$*. However, special treatment of degenerate cases such as* $\varepsilon = 0$ *in this example is not considered in [94], since it is not required for the case of strict inequalities (indeed for the strict inequality version for* $\varepsilon = 0$*, the choice* $\lambda_{i+1} = 1$ *is correct). Recognizing such degenerate cases may allow relaxing the assumption (5.22) in Corollary 11. However, achieving this will likely require adapting the proofs of some of the technical results from [94] or combining them with additional results. For instance, in this example maintaining* $\lambda_{i+1} = \frac{1}{2} - f(\varepsilon)$ *even for* $\varepsilon = 0$ *yields a correct characterization of* $\overline{\operatorname{conv}}\left(\boldsymbol{\mathcal{C}_5}\left(\varepsilon\right)\right)$*, so perhaps some type of perturbation analysis could resolve the issues with the non-compliance with condition (5.22).*

5.3.6 Comparison to the closed convex hull characterizations by Burer and Kılınç-Karzan

The work in [27] studies the closed convex hull characterization of sets defined as the intersection of a conic quadratic and a quadratic inequality similar to those defined in (5.9) and (5.10) given by non-strict inequalities. The work in [27] studies a similar aggregation technique and identifies a set of assumptions that need to be verified in order to get the closed convex hull. Theorem 1 in [27] states the main result of the paper. In this section we do a comparison between the results in [27] and our work and highlight the similarities and differences of the two approaches.

In the language of this paper the first assumption in [27] is:

$$\mathcal{P}_1 \text{ has exactly one negative eigenvalue and } \boldsymbol{\mathcal{H}} \text{ is a separator of } \left\{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\mathbf{1}} \leq 0\right\}. \tag{A1}$$

Assumption (A1) simply formalizes the fact that [27] studies the intersection of a conic quadratic and a general quadratic inequality and hence is not an actual restriction in the context of [27]. Under Assumption (A1), the second assumption of [27] simply requires $\text{int}\,(\boldsymbol{\mathcal{S}}) \neq \emptyset$. This assumption is shared by this paper and we denote it (A2). The third assumption in [27] is a minor technical assumption on the singularity of $\mathcal{P}_1$ and $\mathcal{P}_2$ as follows: either (i) $\mathcal{P}_1$ is nonsingular, (ii) $\mathcal{P}_1$ is singular and $\mathcal{P}_2$ is positive definite on $\text{null}(\mathcal{P}_\lambda)$, or (iii) $\mathcal{P}_1$ is singular and $\mathcal{P}_2$ is negative definite on $\text{null}(\mathcal{P}_\lambda)$. We denote this assumption (A3) and show that this assumption seems to be mildly restrictive. Using Assumption (A3), [27] defines an $s \in [0,1]$ that allows then to describe the closed convex hull using conic quadratic inequalities associated to the pencils $\mathcal{P}_\lambda := (1-\lambda)\mathcal{P}_1 + \lambda\mathcal{P}_2$ at $\lambda = 0$ and $\lambda = s$. In particular, this forces one of the inequalities to be the original conic quadratic inequality, which is a natural choice in the context of [27]. Depending on the details of Assumption (A3), the choice of s is either 0 or the minimum $s \in (0,1]$ such that the pencil $\mathcal{P}_s$ is singular. The last two assumptions of [27] are geometric conditions on the inequalities used to describe the closed convex hull. To state these assumptions, let $\boldsymbol{\mathcal{H}}_{\boldsymbol{n_s}}$ be the natural separator of $\boldsymbol{\mathcal{S}}_{\boldsymbol{s}} := \{y \in \mathbb{R}^{n+1} : \tilde{\boldsymbol{q}}_{\boldsymbol{s}} < 0\}$ and let $\boldsymbol{\mathcal{K}}_{\boldsymbol{s}} := \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^+ \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{s}}$ for $\boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^+ \in \left\{\boldsymbol{\mathcal{H}}_{\boldsymbol{n_s}}^+, \boldsymbol{\mathcal{H}}_{\boldsymbol{n_s}}^-\right\}$ be defined analogously to $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda_i}}$ and $\boldsymbol{\mathcal{K}}_{\boldsymbol{\lambda_{i+1}}}$ in (5.7). With this notation, the homogeneous version of the geometric conditions is

$$s = 1 \quad \text{or} \quad \boldsymbol{\mathcal{K}}_{\boldsymbol{s}} \cap \boldsymbol{\mathcal{H}}_{\boldsymbol{n_s}} \cap \left\{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\mathbf{2}} < 0\right\} \neq \emptyset, \tag{A4}$$

while the non-homogeneous version is

$$\begin{gathered} s = 1 \quad \text{or} \quad \left\{\begin{bmatrix} x \\ 0 \end{bmatrix} \in \mathbb{R}^{n+1} \,:\, \begin{bmatrix} x \\ 0 \end{bmatrix} \in \boldsymbol{\mathcal{K}}_{\boldsymbol{s}} \cap \boldsymbol{\mathcal{H}}_{\boldsymbol{n_s}} \right\} \cap \left\{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\mathbf{2}} < 0\right\} \neq \emptyset \\ \text{or} \\ \left\{\begin{bmatrix} x \\ 0 \end{bmatrix} \in \mathbb{R}^{n+1} \,:\, \begin{bmatrix} x \\ 0 \end{bmatrix} \in \boldsymbol{\mathcal{K}}_{\boldsymbol{s}} \right\} \cap \left\{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\mathbf{1}} \leq 0\right\} \cap \boldsymbol{\mathcal{H}}^+ \subseteq \left\{y \in \mathbb{R}^{n+1} \,:\, \tilde{\boldsymbol{q}}_{\mathbf{2}} \leq 0\right\}. \end{gathered} \tag{A5}$$

With this notation, Theorem 1 in [27] can be written as follows.

Theorem 6. *Let* $\boldsymbol{\mathcal{S}} := \{y \in \mathbb{R}^{n+1} : \tilde{\boldsymbol{q}}_{\boldsymbol{i}} \leq 0, \quad i = 1, 2\}$. *If Assumptions* (A1)–(A3) *hold, then there exists* $s \in [0, 1]$ *such that*

$$\overline{\text{conv}}\left(\boldsymbol{\mathcal{H}}^{+} \cap \boldsymbol{\mathcal{S}}\right) \subseteq \left\{y \in \mathbb{R}^{n+1} : \tilde{\boldsymbol{q}}_{\boldsymbol{1}} \leq 0\right\} \cap \boldsymbol{\mathcal{H}}^{+} \cap \boldsymbol{\mathcal{K}}_{\boldsymbol{s}}, \tag{5.30}$$

where $\boldsymbol{\mathcal{H}}$ *is a separator of* $\boldsymbol{\mathcal{S}}' := \{y \in \mathbb{R}^{n+1} : \tilde{\boldsymbol{q}}_{\boldsymbol{1}} \leq 0\}$. *In such a case, the right hand side of* (5.30) *can be described by two conic quadratic inequalities. If additionally Assumption* (A4) *is satisfied, then* (5.30) *holds at equality. Finally, if Assumptions* (A1)–(A5) *hold, then there exists* $s \in [0, 1]$ *such that*

$$\overline{\text{conv}}\,(C) = \left\{x \in \mathbb{R}^{n} : \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{H}}^{+} \cap \boldsymbol{\mathcal{S}}'\right\} \cap \left\{x \in \mathbb{R}^{n} : \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{K}}_{\boldsymbol{s}}\right\}, \tag{5.31}$$

for $C = \left\{x \in \mathbb{R}^{n} : \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{H}}^{+} \cap \boldsymbol{\mathcal{S}}\right\}$. *In such a case, the right hand side of* (5.31) *can be described by two conic quadratic inequalities.*

We now compare Theorem 6 and the results in this paper using examples from Section 5.3.5. We begin by showing examples where Assumptions (A1) and (A3) restrict the applicability of Theorem 6 as compared to the results in this paper. We then show how condition (5.22) restricts the applicability of Corollary 11 as compared with Theorem 6 and how Assumption (5.25) restricts the applicability of Corollary 12 as compared with Theorem 6. Finally, we comment on the results of Section 7 in [27].

To show how Assumption (A1) can be a tangible restriction when compared with the results in this paper we can use Examples 5 and 6 from Section 5.3.5. For Example 5, we have that Assumption (A1) is violated because neither $\mathcal{P}_1$ nor $\mathcal{P}_2$ have exactly one negative eigenvalue. Hence, Theorem 6 cannot characterize a relaxation for $\overline{\text{conv}}\left(\boldsymbol{\mathcal{S}}_{\boldsymbol{1}}^{+}\right)$. For Example 6, we have that Assumption (A1) is violated, since there is no separator $\boldsymbol{\mathcal{H}}$ of the first homogeneous quadratic inequality which can be used to write S_5 as

$$S_5 = \left\{(x, x_0) \in \mathbb{R}^3 : \begin{bmatrix} x \\ 1 \end{bmatrix} \in \boldsymbol{\mathcal{H}}^{+} \cap \boldsymbol{\mathcal{S}}_{\boldsymbol{3}}\right\},$$

where $\boldsymbol{\mathcal{S}_5}$ is the homogeneous version of S_5. Hence, Theorem 6 cannot characterize a relaxation for $\overline{\text{conv}}\,(S_5)$. We note that considering cases beyond Assumption (A1) was out of the intended scope of [27]. Indeed, one important difference between Theorem 6 and Corollaries 11–13 is that the former only adds one inequality and the later can add two inequalities. Adding one inequality is sufficient for the intended scope of [27], but two inequalities may necessary for other cases such as Examples 5 and 6.

To show how technical Assumption (A3) can be mildly restrictive when compared with the results in this paper we can use Example 7 from Section 5.3.5. Because Assumption (A3) is violated, the only relaxation for $\overline{\text{conv}}\,(\boldsymbol{\mathcal{C}_6})$ that Theorem 6 can characterize is the trivial relaxation $\{(x, x_0) \in \mathbb{R}^3 \,:\, |x_1| \leq x_2\}$. While this relaxation happens to characterize the closed convex hull, Theorem 6 cannot prove this. In contrast, Corollary 11 can prove that the trivial relaxation yields the closed convex hull.

To show how condition (5.25) of Corollary 12 is mildly restrictive as compared with Assumption (A5) of Theorem 6 we can use Example 4 from Section 5.3.5. Corollary 12 can show

$$\overline{\text{conv}}\,(C_3) \subseteq \left\{x \in \mathbb{R}^2 \,:\, |x_1| \leq x_2 \quad 1 - x - 2y \leq 0\right\}, \tag{5.32}$$

but since condition (5.25) is violated, it cannot prove that equality holds in (5.32). In contrast, Theorem 6 can construct the relaxation and prove the equality in (5.32).

To show how Assumption (5.22) from Corollary 11 can be restrictive when compared with Theorem 6 we can use Example 8 from Section 5.3.5 with $\varepsilon = 0$. Since condition (5.22) does not hold, the only relaxation for $\overline{\text{conv}}\,(\boldsymbol{\mathcal{C}_7})$ that Corollary 11 can characterize is the trivial relaxation $\{(x_1, x_0) \in \mathbb{R}^2 \,:\, |x_1| \leq x_0\}$. Theorem 6 can also characterize this relaxation, but in a more systematic way that could provide non-trivial relaxations for other sets for which condition (5.22) fails. Analyzing how Theorem 6 characterizes this relaxation provides a convenient way to compare the technical results related to the selection of s in [27] and λ_i and λ_{i+1} in [94]. For this, let $\mathcal{P}_1$ and $\mathcal{P}_1$ be the matrices defined in Example 8. The value s from Theorem 6 is the minimum $s \in (0, 1]$ such that the pencil $(1 - s)\mathcal{P}_1 + s\mathcal{P}_2$ is singular, which corresponds to $s = \frac{1}{2} - f(\varepsilon)$ for f defined in Example 8. For $\varepsilon > 0$, this s is identical to λ_{i+1} obtained by Corollary 11 which yields the relaxation for Example 8. In contrast, for $\varepsilon = 0$, we have $s = 1/2$ and Theorem 6 yields an inequality that is valid for $\overline{\text{conv}}\,(\boldsymbol{\mathcal{C}_7})$, while $\lambda_{i+1} = 1$ and Corollary 11 yields an invalid inequality. Hence, Theorem 6 seems to be less sensitive to the degeneracy issues caused by the violation of condition (5.22) that we discussed at the end of Example 8. We end the discussion of Example 8 by noting that for all $\varepsilon \geq 0$, we have that $\boldsymbol{\mathcal{K}_s}$ is dominated by the original conic inequality $|x_1| \leq x_0$. This shows that, similarly to the results in this paper, Theorem 6 can also yield a redundant inequality $\boldsymbol{\mathcal{K}_s}$.

We note that for Examples 2 and 3, Theorem 6 yields the same results as Corollaries 12 and 13.

Finally, we consider the sets studied in Section 7 of [27]. This section develops simplifications of Assumptions A1–A5 for intersections of a conic section and a general quadratic constraint. All resulting sets correspond to the intersection of a convex quadratic inequality with a general quadratic inequality. The convex hull of the strict inequality version of all these sets can be characterized without any assumptions by Theorem 5. Similarly, characterizing the closed convex hull of the non-strict inequality versions through Corollary 13 only requires the sets to be contained in the closure of their interiors. Because this last condition is not too restrictive, we can find examples where Corollary 13 can construct the closed convex hull of the intersections of a conic section and a general quadratic constraint, while the simplified assumptions from Section 7 of [27] do not hold. For instance, Example 3 in [73] shows how Corollary 13 yields the closed convex hull of a paraboloid intersected with a non-convex quadratic constraint. This example does not satisfy the simplified assumptions in Section 7 of [27]; however, it satisfies the more general Assumptions A1–A5. Hence there does not seem to be a major difference on the applicability of the two techniques on this class of problems.

6.0 COMPUTATIONAL EXPERIMENTS WITH CUTS AND EXTENDED FORMULATIONS

In this chapter we compare the strength of the introduced linear and nonlinear cuts and extended formulations on two classes of MINLP problems. After conducting preliminary computational experiments, we select the following MINLP problems for our computational tests:

- Closest Vector Problem (CVP) [26, 68]
- Mean-Variance Capital Budgeting (MVCB) problem [9, 13, 67, 92]

These two problems can be formulated as

$$\begin{aligned}
\min \quad & c^T y + f^T x \\
\text{s.t.} \quad & \|Dy + Ex - d\|_2 \leq \rho^T y + w^T x - q \\
& l_y \leq y \leq u_y \\
& l_x \leq x \leq u_x \\
& y \in \mathbb{R}^p, \quad x \in \mathbb{Z}^n,
\end{aligned} \tag{6.1}$$

where $D \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{m\times n}, d \in \mathbb{R}^m, c, \rho, l_y, u_y \in \mathbb{R}^p, f, w, l_x, u_x \in \mathbb{R}^n$, and $q \in \mathbb{R}$. We refer to this formulation as the *original formulation*.

Note that the feasible regions of all these problems share a similar structure; they are composed of a conic quadratic inequality (i.e., the non-polyhedral portion of the formulation) intersected with a set of linear inequalities (i.e., the polyhedral portion of the formulation). We provide detailed formulation of each problem in Sections 6.0.8 and 6.0.9.

The second formulation that we consider is an extended formulation similar to the one presented in Section 4, which is of the form

$$
\begin{aligned}
\min \quad & c^T y + f^T x \\
\text{s.t.} \quad & |Dy + Ex - d| \leq t \\
& t_0 \leq \rho^T y + w^T x - q \\
& \|t\|_2 \leq t_0 \\
& l_y \leq y \leq u_y \\
& l_x \leq x \leq u_x \\
& y \in \mathbb{R}^p, \quad x \in \mathbb{Z}^n, \quad t \in \mathbb{R}^m, \quad t_0 \in \mathbb{R}.
\end{aligned}
\tag{6.2}
$$

We refer to this formulation as the *extended formulation*. We compare the strengths of the nonlinear split cuts (3.25) added to the original formulation (6.1) versus the linear CMIR cuts (4.7) added to the extended formulation (6.2). Note that adding nonlinear cuts to the continuous relaxation of a MINLP could significantly increase its solution time. Hence there will likely be a strong trade-off between the strength provided by the added cuts and their computational cost. We propose to further study such a trade-off on the above set of MINLP problems.

6.0.7 Implementation and Computational Settings

All experiments are performed on an Intel Core(TM)2 Quad PC with 3.0 GHz CPU, 4 GB RAM, and Windows 7 (64-bit) operating system. All models and formulations are implemented using the JuMP modeling language [2, 3, 65] and solved with CPLEX v12.6 [1].

Our base algorithms to solve MICQP problems are CPLEX standard algorithms for solving MICQP. CPLEX implements an NLP-based branch-and-bound algorithm and a standard LP-based branch-and-bound algorithm. Each of these implementations include advanced features such as cutting planes, heuristics, preprocessing, and elaborate branching and node selection strategies. We do not make any changes to CPLEX's default settings. We refer to the NLP-based algorithm as *CPLEXNLP* and to the LP-based algorithm as *CPLEXLP*. Both algorithms are limited to a single thread by appropriately setting CPLEX parameters and to a total run time of 1800 seconds. We present our computational results in four tables. The first table summarizes the percentage integrality gap (the percentage gap between the optimal solution and the continuous relaxation) defined as

$$\frac{|\text{best integer} - \text{ best node}|}{1e-10+|\text{best integer}|},$$

where *best integer* is the best integer solution found so far and *best node* is the relaxation solution at the given node. We report the percentage integrality gaps both at the root node and the terminal node (the final node before the time limit is reached or the instance is solved to optimality). The other three tables summarize solve times, node counts, and the total number of cuts added by the solver (other than the cuts we add to the problems).

We consider adding the nonlinear split cuts (3.25) to the original formulation (6.1) and linear CMIR cuts (4.7) to the extended formulation (6.2). The disjunctions we use to generate cuts are elementary disjunctions $e^i \in \mathbb{Z}^n$, where e^i is the i-th unit vector, and all cuts are added at the root node as constraints. As discussed, we consider two formulations for each problem, the two available MICQP implementations, and the formulations with and without cuts. Therefore, each of the above tables has eight columns for the eight possible combinations of the formulation, MICQP implementation, and cuts.

6.0.8 Closest Vector Problem

Closest Vector Problem (CVP) [26, 68] aims to find the element in an integer lattice that is closest (with respect to the Euclidean distance) to a given target vector not in the lattice. This problem has a wide range of theoretical and practical applications and is hard to even approximate [26, 68]. CVP can be formulated as

$$\min_x \left\{ \|B(x-c)\|_2 \,:\, x \in \mathbb{Z}^n \right\}, \tag{6.3}$$

where $B \in \mathbb{R}^{n\times n}$ is an invertible matrix whose columns compose the basis of the lattice and $c \in \mathbb{R}^n$ (the target vector is Bc in this case). As noted in [18, 70], since $\operatorname{conv}(\mathbb{Z}^n) = \mathbb{R}^n$, to effectively use cuts in CVP we need the equivalent reformulation of (6.3) given by

$$\min_{(x,t_0)} \{t_0 \,:\, \|B(x-c)\|_2 \leq t_0, \quad x \in \mathbb{Z}^n\}. \tag{6.4}$$

Note that the above formulation is the analog of the original formulation (6.1). The only difference between the formulation above and (6.1) is that (6.4) does not include any linear inequalities.

We can then strengthen the above formulation by adding nonlinear split cuts (3.25). The second formulation of CVP that we consider is the extended formulation

$$\min_{(x,t,t_0)} \left\{t_0 \,:\, |B(x-c)| \leq t, \quad \|t\|_2 \leq t_0, \quad t \in \mathbb{R}^n_+,\ t_0 \in \mathbb{R}_+,\ x \in \mathbb{Z}^n\right\}, \tag{6.5}$$

which is the analog of the extended formulation (6.2). The above extended formulation can be strengthened by adding linear CMIR cuts (4.7). As shown in Proposition 16, a single nonlinear split cut added to the original formulation (6.4) is at least as strong as the corresponding linear CMIR cut added to extended formulation (6.5). However, as also formalized in Proposition 19, there are examples where just elementary CMIR cuts can provide a bound that is arbitrarily better than that obtained by all nonlinear split cuts. We further study this by comparing the strengths of the linear CMIRs and nonlinear split cuts for randomly generated CVP instances.

Note that the disadvantage of the CMIRs over the nonlinear split cuts is that CMIRs are derived by only using the set of linear constraints $|B(x-c)| \leq t$. In contrast, nonlinear split cuts are derived using the nonlinear constraint $\|B(x-c)\|_2 \leq t_0$. However, the CMIR bound does use the nonlinear constraint $\|t\|_2 \leq t_0$ after the derivation of the cuts. This suggests that bounds stronger than those obtained by both the CMIR and nonlinear split cuts could be obtained by deriving cuts using the complete feasible region of the extended formulation (6.5). Deriving closed form expressions of such cuts are beyond the current research, but we can use the *LP-based* cut generation techniques (e.g., [25, 56]) to approximate the associated bound.

6.0.8.1 Test Instances We generate 5 randomly generated instances for each $n \in \{10, 20, 30, 40, 50\}$ as described in [26]. In particular, entries of matrix B are generated as uniformly random integers in $\{-3, \ldots, 3\}$ and entries of vector c are chosen uniformly at random in $[-1, 1]$. For simplicity, we only consider the n CMIR and split cuts associated to the elementary disjunctions. However, because B is a general matrix, the cut equivalence of Proposition 16 does not hold and hence this is not necessarily a favorable case for the CMIRs.

6.0.8.2 Results We now present the computation results for CVP instances. Table 1 summarizes the percentage integrality gaps at the root and terminal nodes for randomly generated CVP instances. In Table 1, n denote the problem size, *rep* denote the replication number, R denote the root node, and T denote the terminal node. Note that if an instance is solved to optimality, we report the integrality gap at the terminal node as 0.

Table 1: Gaps for CVP Instances

		Original Formulation								Extended Formulation							
		CPLEXLP				CPLEXNLP				CPLEXLP				CPLEXNLP			
n	rep	No Cuts		With Cuts		No Cuts		With Cuts		No Cuts		With Cuts		No Cuts		With Cuts	
		R	T	R	T	R	T	R	T	R	T	R	T	R	T	R	T
10	1	100	0	82.78	0	100	0	77.15	0	100	0	77.88	0	100	0	77.84	0
10	2	100	0	47.9	0	100	0	44.21	0	100	0	45.7	0	100	0	45.48	0
10	3	100	0	78.38	0	100	0	85.02	0	100	0	77.99	0	100	0	77.87	0
10	4	100	0	55.5	0	100	0	54.52	0	100	0	55.71	0	100	0	55.27	0
10	5	100	0	99.6	0	100	0	99.58	0	100	0	99.58	0	100	0	99.58	0
20	1	100	0	87	40.32	100	0	83.88	0	100	0	85.76	0	100	0	84.78	0
20	2	100	11.13	83.93	35.21	100	0	79.37	0	100	0	80.5	0	100	0	77.42	0
20	3	100	0	91.22	46.74	100	0	86.34	9.84	100	0	92.58	0	100	0	89.16	0
20	4	100	0	89.83	26.97	100	0	86.05	0	100	0	87.03	0	100	0	86.08	0
20	5	100	33.6	98.2	55.75	100	0	97.2	26.02	100	0	97.52	0	100	0	97.36	0
30	1	100	69.44	95.56	86.06	100	8.7	92.35	68.23	100	21.3	93.13	22.66	100	19.18	90.96	25.36
30	2	100	69.7	94.04	82.07	100	0	82.58	61.7	100	13.59	90.29	21.78	100	7.97	85.99	17.42
30	3	100	74.89	94.66	84.87	100	0	89.2	71.03	100	15.57	90.54	18.04	100	15.51	89.53	22.66
30	4	100	69.06	93.87	80.17	100	0	88.01	66.63	100	13.77	88.64	12.9	100	12.96	88.18	18.6
30	5	100	74.78	96.92	88.68	100	19.57	90.73	78.45	100	23.42	94.63	24.72	100	28.43	93.76	33.9
40	1	100	86.74	100	92.17	100	36.73	91.46	82.86	100	48.66	91.99	49.5	100	46.06	89.88	49.88
40	2	100	87.72	96.76	90.17	100	40.95	86.76	83.33	100	43.74	92.34	47.67	100	52.58	87.77	54.84
40	3	100	86.23	96.86	92.79	100	42.48	92.16	87.57	100	51.05	94.88	49.75	100	48.85	91.96	51.79
40	4	100	90.22	99.12	96.45	100	54.97	98.26	96.52	100	53.14	98.76	56.75	100	62.27	98.26	68.33
40	5	100	87.4	98.81	92.89	100	38.13	93.96	85.9	100	54.06	95.81	52.34	100	48.3	93.5	47.6
50	1	100	91.47	98.91	95.03	100	48.33	93.73	88.36	100	62.26	95.44	60.17	100	54.95	94.04	56.87
50	2	100	94.52	99.26	96.53	100	63.8	94.89	94.41	100	67.35	97.3	66.45	100	69.17	95.73	69.83
50	3	100	92.95	98.08	94.92	100	57.7	94.2	90.85	100	66.05	96.01	67.95	100	64.3	93.96	68.93
50	4	100	95.64	99.24	97.62	100	64.06	97.72	96.76	100	72.46	97.86	71.37	100	69.5	97.81	71.57
50	5	100	90.47	98.57	95.17	100	52.66	91.11	90.24	100	63.37	94.05	55.66	100	59.12	92.64	63.32
Average		100	52.24	91	62.82	100	21.12	86.82	51.15	100	26.79	88.48	27.11	100	26.37	86.99	28.84

We first compare the strength of nonlinear split cuts (3.25) and CMIR cuts (4.7). For the CPLEXLP strategy, we observe that in average, nonlinear split cuts and CMIR cuts close 9% and 11.52% of the integrality gap at the root node, respectively. On the other hand, for CPLEXNLP strategy, the average gap closed at the root node using the nonlinear split cuts and CMIR cuts are 13.18% and 13.01%, respectively. Therefore, there is no significant domination between the two classes of cuts in terms of closing the integrality gap at the root node and and hence the lack of non-polyhedral information in the generation of CMIR cuts seems to be mostly compensated by the strength of their common extended formulation. On the other hand, the average integrality gaps at the terminal nodes are significantly smaller for CMIR cuts. This could partially be due to the fact that adding linear CMIR cuts is less computationally expensive than adding nonlinear split cuts and therefore the extended formulation with CMIRs outperforms the original formulation with the nonlinear split cuts. Moreover, this could also be due to the power of the extended formulation to generate CMIRs, which we study next.

We now compare the strength of the original formulation (6.4) (with no cuts) versus the extended formulation (6.5) (with no cuts). Using the extended formulation (6.5) does provide a significant computational advantage by improving the effectiveness of the CPLEXLP strategy as the average integrality gap at the terminal node is significantly smaller for the extended formulation; however, we do not see such an improvement for CPLEXNLP strategy.

Finally, the average integrality gaps at the terminal nodes for both formulations with no cuts are smaller than those for the formulations with cuts. This suggests that while adding cuts improve the formulations by closing the integrality gap at the root nodes, it does not provide much computational advantage for solving CVP instances. We further study this by comparing the effectiveness of the cuts and formulations using solve times.

Table 2 presents the solve times for the randomly generated CVP instances. Note that if an instance is not solved to optimality in the time limit of 1800 seconds, we report the solve time as 1800.

We see that the average solve times for the extended formulation with CMIR cuts are smaller than those of the the original formulation with nonlinear split cuts. Hence, CMIRs seem to provide the best balance between strength, cut generation expense, and compu-

Table 2: Solve Times for CVP Instances

n	rep	Original Formulation				Extended Formulation			
		CPLEXLP		CPLEXNLP		CPLEXLP		CPLEXNLP	
		No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts
10	1	0.31	0.68	0.23	1.95	0.05	0.1	0.37	0.43
10	2	0.11	0.13	0.06	0.61	0.03	0.07	0.09	0.14
10	3	0.57	1.84	0.61	4.96	0.1	0.14	0.95	0.93
10	4	0.4	0.98	0.23	3.64	0.07	0.13	0.37	0.59
10	5	1.22	6.02	2.54	23.33	0.36	0.36	4.03	5.53
20	1	932.54	1800	9.97	902.11	6.79	6.83	18.79	28.08
20	2	1800	1800	13.42	755.78	8.55	9.26	24.9	34.23
20	3	1538.93	1800	16.39	1800	8.89	10.22	31.22	48.17
20	4	474.1	1800	3.93	342.33	3.27	3.23	7.51	12.29
20	5	1800	1800	64.47	1800	19.13	22.02	121.12	148.82
30	1	1800	1800	1800	1800	1800	1800	1800	1800
30	2	1800	1800	1011.67	1800	1800	1800	1800	1800
30	3	1800	1800	1669.87	1800	1800	1800	1800	1800
30	4	1800	1800	1394.62	2166.11	1800	1800	1800	1800
30	5	1800	1800	1800	1800	1800	1800	1800	1800
40	1	1800	1800	1800	1800	1800	1800	1800	1800
40	2	1800	1800	1800	1800	1800	1800	1800	1800
40	3	1800	1800	1800	1800	1800	1800	1800	1800
40	4	1800	1800	1800	1800	1800	1800	1800	1800
40	5	1800	1800	1800	1800	1800	1800	1800	1800
50	1	1800	1800	1800	1800	1800	1800	1800	1800
50	2	1800	1800	1800	1800	1800	1800	1800	1800
50	3	1800	1800	1800	1800	1800	1800	1800	1800
50	4	1800	1800	1800	1800	1800	1800	1800	1800
50	5	1800	1800	1800	1800	1800	1800	1800	1800
Average		1341.93	1440.47	1031.61	1320.09	1081.93	1082.13	1088.4	1091.18

tational burden resulting from adding the cuts to the formulation. However, no class of cuts seems to provide a computational advantage for solving CVP (we found similar results, both in absolute and relative strength among the cut classes, by trying non-elementary disjunctions). Using the extended formulation (6.5) does provide a significant computational advantage by improving the effectiveness of CPLEXLP strategy (see [53] for a detailed discussion of how extended formulations can help such algorithms). However, it is unlikely that this approach will outperform specialized CVP algorithms such as the SDP-inspired branch-and-bound algorithm from [26].

Finally, Table 3 summarizes the node counts for randomly generated CVP instances. Also note that CPLEX did not add any cuts while solving the CVP instances.

Table 3: Node Counts for CVP Instances

n	rep	Original Formulation				Extended Formulation			
		CPLEXLP		CPLEXNLP		CPLEXLP		CPLEXNLP	
		No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts
10	1	2112	1387	189	145	294	289	189	166
10	2	913	278	35	31	158	73	35	47
10	3	3945	3458	539	398	876	603	537	391
10	4	2729	1955	197	259	508	563	197	241
10	5	8578	9762	1942	1998	2937	2287	1942	1980
20	1	437101	65826	4167	3940	29798	25648	4167	4180
20	2	457287	51922	5670	4107	30199	29722	5609	5434
20	3	676760	79419	7511	7839	40565	40918	7532	7857
20	4	357009	90817	1695	1856	15840	11369	1695	1945
20	5	614300	79946	26665	10979	83561	79132	26629	22604
30	1	332732	34000	429800	1272	1682270	1530924	179127	118531
30	2	280680	33428	228347	1173	1701093	1180753	178024	117775
30	3	290297	31500	383767	1276	2791983	1832705	182690	115060
30	4	300114	35366	318779	1251	2151980	1936830	179833	118518
30	5	306756	37335	374596	1199	1909900	1531080	169680	103394
40	1	234365	29228	241799	240	655700	651676	95040	60760
40	2	251785	29696	242200	210	1157534	984761	94652	59677
40	3	263160	29035	244960	230	830653	862747	95659	63671
40	4	242637	31046	193293	260	1230137	909597	76527	46824
40	5	250585	31900	250319	251	642834	761281	99071	60950
50	1	215187	24382	157715	70	434315	512084	56660	37840
50	2	208389	23574	152039	69	438149	447115	55504	35100
50	3	206040	19698	151980	61	453836	461091	58547	36439
50	4	200474	19946	153575	63	470466	416099	55782	34010
50	5	214239	22425	150550	57	409695	771835	56980	37074
Average		254326.96	32693.16	148893.16	1569.36	686611.24	599247.28	67292.32	43618.72

6.0.9 Mean-Variance Capital Budgeting (MVCB)

The second class of problems we study consists of capital budgeting problems with a mean-variance objective [9, 13, 67, 92] of the form

$$\max_{x} \left\{ r^T x - \gamma x^T V x \;:\; c^T x \leq d, \quad x \in \mathbb{Z}^n \right\}, \tag{6.6}$$

where r is the expectation vector and V is the covariance matrix of uncertain return for n projects. There is also a budget constraint $c^T x \leq d$, and $\gamma > 0$ is the investor's risk-averseness parameter.

As described in [9], for $a \in \mathbb{R}^n$ such that $a^T V^{1/2} = \frac{1}{2\gamma} r$, (6.6) can be written as

$$\gamma a^T a - \gamma \min_x \left\{ x^T V x - 2a^T V^{1/2} x + a^T a \, : \, c^T x \leq d, \quad x \in \mathbb{Z}^n \right\},$$

which in turn can be formulated in the conic quadratic form as

$$\gamma a^T a - \gamma \left(\min_{(x,t_0)} \left\{ t_0 \, : \, \left\| V^{1/2} x - a \right\|_2 \leq t_0, \quad c^T x \leq d, \quad x \in \mathbb{Z}^n, \ t_0 \in \mathbb{R}_+ \right\} \right)^2 . \tag{6.7}$$

Therefore, the optimization problem we need to solve for each instance of MVCB is of the form

$$\min_{(x,t_0)} \left\{ t_0 \, : \, \left\| V^{1/2} x - a \right\|_2 \leq t_0, \quad c^T x \leq d, \quad x \in \mathbb{Z}^n, \ t_0 \in \mathbb{R}_+ \right\}. \tag{6.8}$$

In order to obtain the actual optimal objective value, we need to first solve the optimization problem (6.8) and then transform the optimal objective value to the correct one using (6.7). However, since the purpose of our experiments is to compare the strength of nonlinear split cuts and linear CMIR cuts, we only solve the optimization problem (6.7) and do not consider the transformation given by (6.8).

The second formulation of MVCB that we consider is the extended formulation

$$\min_{(x,t,t_0)} \left\{ t_0 \, : \, \left| V^{1/2} x - a \right| \leq t, \quad \|t\|_2 \leq t_0, \quad c^T x \leq d, \quad x \in \mathbb{Z}^n, \ t \in \mathbb{R}^n, \ t_0 \in \mathbb{R}_+ \right\}. \tag{6.9}$$

As before, we compare the effectiveness of nonlinear split and linear CMIR cuts by adding them to the original formulation (6.8) and the extended formulation (6.9), respectively.

6.0.9.1 Test Instances We generate 5 randomly generated instances for each $n \in \{60, 80, 100, 110, 120\}$ and $\gamma \in \{1\}$ similar to the data generation of [9]. In particular, entries of matrix B are chosen uniformly at random in $[-1, 1]$ and entries of the vectors b and c are chosen uniformly at random in $[0, 1]$. Moreover, the budget d is chosen as $d = \sum_{i=1}^{n} |a_i|$. Note that since V is the covariance matrix, we need to make sure all its diagonal entries are non-negative; however, since the purpose of our experiments is to compare the strength of nonlinear split cuts and linear CMIR cuts, we relax this requirement. As in Section 6.0.8, for simplicity, we only consider the n CMIR and split cuts associated to elementary disjunctions.

6.0.9.2 Results We now present the computation results for MVCB instances. Table 4 summarizes the percentage integrality gaps at the root and terminal nodes for randomly generated MVCB instances. In Table 4, n denote the problem size, *rep* denote the replication number, R denote the root node, and T denote the terminal node. As before, if an instance is solved to optimality, we report the integrality gap at the terminal node as 0.

Table 4: Gaps for MVCB Instances

		Original Formulation								Extended Formulation							
		CPLEXLP				CPLEXNLP				CPLEXLP				CPLEXNLP			
n	rep	No Cuts		With Cuts		No Cuts		With Cuts		No Cuts		With Cuts		No Cuts		With Cuts	
		R	T	R	T	R	T	R	T	R	T	R	T	R	T	R	T
60	1	20.05	12.16	15.51	14.24	0.22	0	0.22	0	17.38	1.18	17.38	1.36	0.22	0	0.22	0
60	2	22.36	0	22.29	0	3.68	0	3.57	0	8.16	0	8.16	0	3.68	0	3.68	0
60	3	42.38	31.5	41.72	37.48	1.25	0	1.25	0	40.47	0	40.47	0	1.25	0	1.25	0
60	4	28.8	0	24.72	0	5.35	0	4.18	0	12.23	0	12.23	0	5.35	0	5.35	0
60	5	26.1	16.72	26.1	19.94	0.46	0	0.46	0	19.5	1.98	22.47	1.68	0.46	0	0.46	0
80	1	31.09	0	26.78	18.97	4.3	0	3.27	1.2	13.52	0	13.52	0	4.3	0	4.3	0
80	2	60.44	38.68	57.36	50.77	1.09	0	1.05	0.71	25.37	0	25.37	0	1.09	0	1.09	0
80	3	35.51	0	32.61	25.8	5.72	0	4.75	3.62	14.3	0	14.3	0	5.72	0	5.72	0
80	4	57.55	30.11	57.81	49.35	1.08	0	1.07	0.8	5.73	0	5.73	0	1.08	0	1.08	0
80	5	78.13	57.24	73.46	71.25	6.8	0	6.05	2.15	57.36	0	57.36	0	43.9	0	43.9	0
100	1	54.43	40.4	53.01	48.03	0.5	0	0.49	0.3	34.58	0.13	34.58	0	0.49	0	0.5	0
100	2	14.08	8.43	12.32	9.18	0.07	0	0.07	0.04	12.74	7.34	12.74	7.14	0.07	0	0.07	0
100	3	65.27	23.18	56.88	28.52	2.88	0	2.77	1.13	9.74	0	9.74	0	2.88	0	2.88	0
100	4	60.07	46.9	59.87	56.2	1.06	0	1.04	0.39	40.93	0	40.93	0	1.06	0	1.06	0
100	5	7.84	4.71	6.44	5.04	0	0	0.02	0	6.58	4.09	6.58	3.87	0.02	0	0	0
110	1	75.04	57.19	74.8	68.61	1.85	0	1.76	0.73	9.62	0	9.62	0	1.85	0	1.85	0
110	2	68.32	54.46	68.43	63.39	1.27	0	1.14	0.88	56.03	0	56.03	0	1.27	0	1.27	0
110	3	1.1	0.8	1.33	0.91	0	0	0	0	1.59	0.9	1.59	0.91	0	0	0	0
110	4	74.24	60.12	73.6	68.11	1.52	0	1.49	1.21	8.41	0	8.41	0	1.52	0	1.52	0
110	5	3.85	1.9	3.05	2.6	0	0	0	0	3.29	1.82	3.55	1.82	0	0	0	0
120	1	46.37	18.93	43.84	40.19	6.02	0	5.58	4.29	17.63	0	17.63	0	6.02	0	6.02	0
120	2	53.4	42.98	56.38	53.13	0.34	0	0.34	0.25	2.86	0	2.86	0	0.34	0	0.34	0
120	3	52.4	46.86	56.86	52.31	0.56	0	0.56	0.36	49.21	4.69	49.21	4.65	0.56	0	0.56	0
120	4	69.44	21.12	34.04	31.49	3.64	0	3.51	2.54	10.38	0	10.38	0	3.64	0	3.64	0
120	5	45.1	30.14	44.98	38.53	0.16	0	0.16	0.11	37.63	2.59	37.63	2.6	0.16	0	0.16	0
Average		43.73	25.78	40.97	34.16	1.99	0	1.79	0.83	20.61	0.99	20.74	0.96	3.48	0	3.48	0

We first compare the strength of nonlinear split cuts (3.25) and CMIR cuts (4.7). We see that in average, nonlinear split cuts close 6.44% and 4.7% of the integrality gap at the root node for CPLEXLP and CPLEXNLP strategies, respectively. On the other hand, CMIR cuts do not seem to close much of the integrality gap at the root node. However, the average integrality gaps at the terminal nodes are smaller for CMIR cuts.

We now compare the strength of the original formulation (6.8) (with no cuts) versus the extended formulation (6.9) (with no cuts). Using the extended formulation (6.9) does provide a significant computational advantage by improving the effectiveness of the CPLEXLP strategy as the average integrality gap at the terminal node is significantly smaller for the extended formulation; however, we do not see such an improvement for CPLEXNLP strategy as all the instances are solved to optimality for both formulations.

Finally, the average integrality gaps at the terminal nodes of the original formulation with no cuts are smaller than that with the nonlinear split cuts. On the other hand, the average integrality gaps at the terminal nodes of the extended formulation with and without CMIR cuts seem to be comparable.

Table 5 presents the solve times for the randomly generated MVCB instances. As before, if an instance is not solved to optimality in the time limit of 1800 seconds, we report the solve time as 1800.

Table 5: Solve Times for MVCB Instances

		Original Formulation				Extended Formulation			
n	rep	CPLEXLP		CPLEXNLP		CPLEXLP		CPLEXNLP	
		No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts
60	1	1800	1800	1.45	1011.21	1800	1800	3.49	11.19
60	2	28.34	1316.62	1.38	1193.32	0.55	0.67	2.91	9.95
60	3	1800	1800	1.37	1026.65	47.17	49.76	3.02	11.66
60	4	71.37	1758.53	2.01	1269.5	0.58	0.73	4.06	10.39
60	5	1800	1800	1.98	1093.09	1800	1800	3.67	9.27
80	1	135.09	1800	4.39	1800	1.04	1.42	10.49	19.49
80	2	1800	1800	3.91	1800	20.32	23.83	11.26	23.43
80	3	878.5	1800	4.57	1800	1.44	1.84	10.76	22.57
80	4	1800	1800	4.05	1800	12.22	18.02	10.53	24.25
80	5	1800	1800	6.13	1800	5.18	6.45	12	21.15
100	1	1800	1800	9.12	1800	1800	1461.55	22.32	45.77
100	2	1800	1800	6.88	1800	1800	1800	14.8	41.04
100	3	1800	1800	7.83	1800	2.72	3.54	20.03	49.41
100	4	1800	1800	9.4	1800	38.56	39.14	19.95	48.2
100	5	1800	1800	7.97	702.24	1800	1800	2.52	31.6
110	1	1800	1800	11.7	1800	28.9	38.93	28.3	71.3
110	2	1800	1800	11.31	1800	98.81	126.66	27.73	64.53
110	3	1800	1800	0.17	288.57	1800	1800	0.34	0.99
110	4	1800	1800	12.21	1800	12.22	14.49	28.43	61.91
110	5	1800	1800	0.33	365.16	1800	1800	0.44	1.23
120	1	1800	1800	16.66	1800	3.37	4.33	44.98	88.28
120	2	1800	1800	14.14	1800	73.06	87.66	42.08	96.45
120	3	1800	1800	17.27	1800	1800	1800	58.85	88.17
120	4	1800	1800	18.44	1800	3.28	4.44	44.52	92.99
120	5	1800	1800	15.76	1800	1800	1800	43.88	92.69
Average		1556.61	1779.53	7.62	1504.55	662.02	651.36	18.85	41.52

We see that the average solve times of the extended formulation with CMIR cuts are significantly smaller than those of the the original formulation with nonlinear split cuts. Hence, CMIRs seem to provide the best balance between strength, cut generation expense, and computational burden resulting from adding the cuts to the formulation. Using the extended formulation (6.5) does provide a significant computational advantage by improving the effectiveness of CPLEXLP strategy. Furthermore, while CMIR cuts improve the average solve time of the extended formulation with CPLEXLP strategy, no class of cuts seems to provide a significant computational advantage for solving MVCB (we found similar results, both in absolute and relative strength among the cut classes, by trying non-elementary disjunctions).

Finally, Tables 6 and 7 present the node counts and the total number of added cuts by CPLEX for randomly generated MVCB instances.

Table 6: Node Counts for MVCB Instances

		Original Formulation				Extended Formulation			
		CPLEXLP		CPLEXNLP		CPLEXLP		CPLEXNLP	
n	rep	No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts
60	1	177592	29148	63	98	241402	269833	65	100
60	2	44607	56608	113	110	199	206	113	112
60	3	176667	18039	112	109	72652	63647	111	104
60	4	110089	78079	107	109	221	218	107	114
60	5	177213	17503	71	103	198996	265206	71	103
80	1	150555	40476	148	27	190	177	141	148
80	2	156944	11186	147	39	15056	14724	137	144
80	3	891973	34322	152	30	393	363	152	152
80	4	213851	14518	143	26	9479	10322	142	140
80	5	148000	13993	142	27	1574	2012	141	141
100	1	134727	9622	174	15	366495	323741	165	163
100	2	152837	18695	137	14	190136	197500	99	127
100	3	184046	21780	182	15	554	559	181	183
100	4	142806	9900	181	14	16494	10861	178	177
100	5	147943	22986	130	5	165627	164677	5	79
110	1	137078	8627	200	10	9970	10422	200	188
110	2	128363	7900	194	11	36840	33710	194	190
110	3	129017	28907	0	0	142934	127729	0	0
110	4	134900	8400	194	9	3500	3168	196	189
110	5	138541	24127	0	0	116179	135774	0	0
120	1	915701	15663	223	5	245	250	221	222
120	2	133201	7550	211	6	27488	25606	190	190
120	3	121553	6700	204	6	219100	282462	203	192
120	4	1004059	16772	226	5	310	279	217	219
120	5	142714	7179	197	7	295400	231199	195	200
Average		239799.08	21147.2	146.04	32	85257.36	86985.8	136.96	143.08

Table 7: Number of Added Cuts by CPLEX for MVCB Instances

n	rep	Original Formulation				Extended Formulation			
		CPLEXLP		CPLEXNLP		CPLEXLP		CPLEXNLP	
		No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts	No Cuts	With Cuts
60	1	1	2	0	0	3	3	0	0
60	2	4	2	0	0	3	3	0	0
60	3	2	0	0	0	4	4	0	0
60	4	4	2	0	0	1	1	0	0
60	5	1	1	0	0	2	2	0	0
80	1	3	3	0	0	4	4	0	0
80	2	2	1	0	0	2	2	0	0
80	3	4	3	0	0	5	5	0	0
80	4	3	0	0	0	5	5	0	0
80	5	2	2	0	0	4	4	0	0
100	1	2	2	0	0	3	3	0	0
100	2	1	0	0	0	3	3	0	0
100	3	2	3	0	0	4	4	0	0
100	4	4	3	0	0	3	3	0	0
100	5	0	0	0	0	0	0	0	0
110	1	1	1	0	0	4	4	0	0
110	2	0	0	0	0	3	3	0	0
110	3	1	0	0	0	1	1	0	0
110	4	2	1	0	0	6	6	0	0
110	5	1	0	0	0	0	1	0	0
120	1	6	2	0	0	7	7	0	0
120	2	2	1	0	0	5	5	1	1
120	3	2	2	0	0	4	4	0	0
120	4	3	1	0	0	4	4	0	0
120	5	1	1	0	0	5	5	0	0
Average		2.16	1.32	0	0	3.4	3.44	0.04	0.04

6.0.10 Concluding Remarks

Our computational experiments on CVP and MVCB instances show that using extended formulations provides a significant computational advantage by improving the effectiveness of CPLEXLP strategy. One can see that the solve times of the extended formulations of these problems (with no cuts) are significantly smaller than those of the original formulations (with no cuts).

On the other hand, while CVP and MVCB problems provide simple and clean settings to compare the strength of cuts, no class of cuts seems to provide a computational advantage for solving these problems. In particular, while adding cuts improves the formulations by reducing the integrality gap at the root node, it does not provide any additional computational advantage, as solve times of the formulations with cuts generally increase.

Finally, we see that there is no significant domination between the two classes of cuts and hence the lack of non-polyhedral information in the generation of CMIR cuts seems to be mostly compensated by the strength of their common extended formulation. Furthermore, the average solve times of the extended formulations with CMIR cuts are significantly smaller than those of the original formulations with nonlinear split cuts. Therefore, CMIRs seem to provide the best balance between strength, cut generation expense, and computational burden resulting from adding the cuts to the relaxation. However, it seems that most of the computational advantage is due to the extended formulation and the CMIR cuts themselves do not provide much additional computational advantage. We expect the CMIRs to have a better performance in problems with few conic quadratic constraints and many linear constraints, where we can generate linear split cuts derived by jointly using the original linear constraints and the polyhedral portion of the CMIR extended formulation. Studying such problems is beyond the current research, but taking advantage of advanced separation techniques (e.g., [39]) can potentially help CMIRs provide additional computational advantage on top of the extended formulation.

7.0 CONCLUSIONS

This dissertation is focused on the development of new cuts and extended formulations for Mixed Integer Nonlinear Programs. We introduce two techniques to give precise characterization of general intersection cuts for several classes of MINLP problems with specific structures. We also study the relation between the introduced cuts and some known classes of cutting planes from MILP. Furthermore, we show how an aggregation technique can be easily extended to characterize the convex hull of sets defined by two quadratic or by a conic quadratic and a quadratic inequality. We also computationally evaluate the performance of the introduced cuts and extended formulations on two classes of MINLP problems.

In Chapter 3 we study the generalization of split, k-branch split, and intersection cuts from MILP to MINLP. We propose two simple techniques to derive general intersection cuts for several classes of MINLP problems with specific structures. In particular, we give simple formulas for split cuts for essentially all convex sets described by a single conic quadratic inequality. We also give simple formulas for k-branch split cuts and some general intersection cuts for a wide variety of convex quadratic sets.

In Chapter 4 we study split cuts and extended formulations for MICQP. In particular, we study the relation between Conic MIR (CMIR) cuts [9] and nonlinear split cuts for a class of MICQP problems. We also study an extended formulation for such a class of MICQP and illustrate how the power of an extended formulation can improve the strength of a cutting plane procedure in MINLP.

In Chapter 5 we consider an aggregation technique introduced by Yıldıran [94] to study the convex hull of regions defined by two quadratic or by a conic quadratic and a quadratic inequality. We show how this aggregation technique can be easily extended to yield valid conic quadratic inequalities for the convex hull of sets defined by two quadratic or by a conic quadratic and a quadratic inequality. We also show that in many cases under additional assumptions, these valid inequalities characterize the convex hull exactly.

In Chapter 6 we computationally evaluate the performance of the introduced linear and nonlinear cuts and extended formulations on two classes of MINLP problems (Closest Vector Problem and Mean-variance Capital Budgeting). We compare the strength of the nonlinear cuts added to the original formulation versus the linear cuts added to an extended formulation.

BIBLIOGRAPHY

[1] Ibm ilog: Cplex high-performance mathematical programming engine. http://www.ibm.com/software/integration/optimization/cplex/.

[2] Julia interface for the cplex optimization software. https://github.com/JuliaOpt/CPLEX.jl.

[3] Jump modeling language for mathematical programming. https://github.com/JuliaOpt/JuMP.jl.

[4] T. Achterberg. SCIP: solving constraint integer programs. *Mathematical Programmign Computation*, 1:1–41, 2009.

[5] K. Andersen, G. Cornuéjols, and Y. Li. Split closure and intersection cuts. *Mathematical Programming*, 102:457–493, 2005.

[6] K. Andersen and A. N. Jensen. Intersection cuts for mixed integer conic quadratic sets. In M. Goemans and J. Correa, editors, *16th international IPCO Conference, Valparaiso*, Lecture Notes in Computer Science, pages 37–48. Springer, 2013.

[7] K. Andersen, Q. Louveaux, and R. Weismantel. An analysis of mixed integer linear sets based on lattice point free convex sets. *Mathematics of Operations Research*, 35:233–256, 2010.

[8] A. Atamtürk and V. Narayanan. Cuts for conic mixed-integer programming. In M. Fischetti and D. P. Williamson, editors, *IPCO*, volume 4513 of *LNCS*, pages 16–29. Springer, 2007.

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

[10] E. Balas. Intersection cuts-a new type of cutting planes for integer programming. *Operations Research*, 19:19–39, 1971.

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

[12] E. Balas and F. Margot. Generalized intersection cuts and a new cut generating paradigm. *Mathematical Programming*, 137:19–35, 2013.

[13] S. Baum, R.C. Carlson, and J.V. Jucker. Some problems in applying the continuous portfolio selection model to the discrete capital budgeting problem. *Journal of Financial and Quantitative Analysis*, 13(02):333–344, 1978.

[14] P. Belotti, J. C. Góez, I. Pólik, T. K. Ralphs, and T. Terlaky. A conic representation of the convex hull of disjunctive sets and conic cuts for integer second order cone optimization. *Optimization Online*, 2012.

[15] P. Belotti, J. C. Góez, I. Pólik, T. K. Ralphs, and T. Terlaky. On families of quadratic surfaces having fixed intersections with two hyperplanes. *Discrete Applied Mathematics*, 161(16):2778–2793, 2013.

[16] P. Belotti, J. C. Góez, I. Pólik, T. K. Ralphs, and T. Terlaky. Disjunctive conic cuts for mixed integer second order cone optimization. 2014.

[17] Aharon Ben-Tal and Marc Teboulle. Hidden convexity in some nonconvex quadratically constrained quadratic programming. *Mathematical Programming*, 72(1):51–63, 1996.

[18] D. Bienstock and A. Michalka. Strong formulations for convex functions over nonconvex sets. *Optimization Online*, 2011.

[19] D. Bienstock and A. Michalka. Cutting-planes for optimization of convex functions over nonconvex sets. *SIAM Journal on Optimization*, 24(2):643–677, 2014.

[20] A. Billionnet, S. Elloumi, and A. Lambert. Extending the QCR method to general mixed-integer programs. *Mathematical programming*, 131:381–401, 2012.

[21] A. Billionnet, S. Elloumi, and M.C. Plateau. Improving the performance of standard solvers for quadratic 0-1 programs by a tight convex reformulation: The QCR method. *Discrete Applied Mathematics*, 157:1185–1197, 2009.

[22] R. Bixby and E. Rothberg. Progress in computational mixed integer programming - a look back from the other side of the tipping point. *Annals of Operations Research*, 149:37–41, 2007.

[23] R.E. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling. Mixed-integer programming: a progress report. In *The sharpest cut: the impact of Manfred Padberg and his work*, chapter 18, pages 309–326. SIAM, Philadelphia, PA, 2004.

[24] G. Blekherman, P.A. Parrilo, and R. Thomas. *Semidefinite Optimization and Convex Algebraic Geometry*. MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2013.

[25] P. Bonami. Lift-and-project cuts for mixed integer convex programs. In *Integer Programming and Combinatoral Optimization*, pages 52–64. Springer, 2011.

[26] C. Buchheim, A. Caprara, and A. Lodi. An effective branch-and-bound algorithm for convex quadratic integer programming. *Mathematical Programming*, 135:369–395, 2012.

[27] S. Burer and F. Kılınç-Karzan. How to convexify the intersection of a second order cone and a nonconvex quadratic. *arXiv preprint arXiv:1406.1031*, 2014.

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

[29] V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4:305–337, 1973.

[30] M. Conforti, G. Cornuéjols, and G. Zambelli. Polyhedral approaches to mixed integer linear programming. *50 Years of Integer Programming 1958-2008*, pages 343–385, 2010.

[31] M. Conforti, G. Cornuéjols, and G. Zambelli. Corner polyhedron and intersection cuts. *Surveys in Operations Research and Management Science*, 16:105–120, 2011.

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

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

[34] G. Cornuéjols and Y. Li. On the rank of mixed 0, 1 polyhedra. *Mathematical Programming*, 91(2):391–397, 2002.

[35] D. Dadush, S. S. Dey, and J. P. Vielma. The Chvátal-Gomory closure of a strictly convex body. *Mathematics of Operations Research*, 36:227–239, 2011.

[36] D. Dadush, S. S. Dey, and J. P. Vielma. On the Chvátal-Gomory closure of a compact convex set. In O. Günlük and G. J. Woeginger, editors, *IPCO*, volume 6655 of *LNCS*, pages 130–142. Springer, 2011.

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

[38] S. Dash, S. S. Dey, and O. Günlük. Two dimensional lattice-free cuts and asymmetric disjunctions for mixed-integer polyhedra. *Mathematical programming*, 135:221–254, 2012.

[39] S. Dash, O. Günlük, and A. Lodi. MIR closures of polyhedral sets. *Mathematical Programming*, 121:33–60, 2010.

[40] S. Dash, O. Günlük, and C. Raack. A note on the MIR closure and basic relaxations of polyhedra. *Operations Research Letters*, 39:198–199, 2011.

[41] S. Dash, O. Günlük, and J. P. Vielma. Computational experiments with cross and crooked cross cuts. *INFORMS Journal on Computing*, 26:780–797, 2014.

[42] A. Del Pia and R. Weismantel. Relaxations of mixed integer sets from lattice-free polyhedra. *4OR: A Quarterly Journal of Operations Research*, 10:1–24, 2012.

[43] S. S. Dey and J. P. Vielma. The chvátal-gomory closure of an ellipsoid is a polyhedron. In *IPCO*, volume 14, pages 327–340. Springer, 2010.

[44] S. Drewes. *Mixed Integer Second Order Cone Programming.* PhD thesis, Technische Universität Darmstadt, 2009.

[45] T. Fujie and M. Kojima. Semidefinite programming relaxation for nonconvex quadratic programs. *Journal of Global Optimization*, 10:367–380, 1997.

[46] M. Giandomenico, A. N. Letchford, F. Rossi, and S. Smriglio. A new approach to the stable set problem based on ellipsoids. In *Integer Programming and Combinatoral Optimization*, pages 223–234. Springer, 2011.

[47] R. E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64:275–278, 1958.

[48] R. E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and its Applications*, 2:451 – 558, 1969.

[49] R. E. Gomory and E. L. Johnson. Some continuous functions related to corner polyhedra. *Mathematical Programming*, 3:23–85, 1972.

[50] J. Gouveia and R. Thomas. Convex hulls of algebraic sets. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 113–138. Springer, 2012.

[51] J. W. Helton and J. Nie. Semidefinite representation of convex sets and convex hulls. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 77–112. Springer, 2012.

[52] D. Henrion. Semidefinite representation of convex hulls of rational varieties. *Acta applicandae mathematicae*, 115:319–327, 2011.

[53] H. Hijazi, P. Bonami, and A. Ouorou. An outer-inner approximation for separable mixed-integer nonlinear programs. *INFORMS Journal on Computing*, 26(1):31–44, 2013.

[54] R. Horst and H. Tuy. *Global optimization: Deterministic approaches.* Springer, 2003.

[55] E. L. Johnson, G. L. Nemhauser, and M. W. P. Savelsbergh. Progress in linear programming-based algorithms for integer programming: An exposition. *INFORMS Journal on Computing*, 12:2–23, 2000.

[56] M. R. Kılınç, J. Linderoth, and J. Luedtke. Effective separation of disjunctive cuts for convex mixed integer nonlinear programs. *Optimization Online*, 2010.

[57] F. Kılınç-Karzan. On minimal valid inequalities for mixed integer conic programs. *arXiv preprint arXiv:1408.6922*, 2014.

[58] F. Kılınç-Karzan and S. Yıldız. Two-term disjunctions on the second-order cone. In *Integer Programming and Combinatorial Optimization*, pages 345–356. Springer, 2014.

[59] M. R. Kılınç, J.T. Linderoth, and J. Luedtke. Effective separation of disjunctive cuts for convex mixed integer nonlinear programs. Technical report, University of Wisconsin-Madison, 2010.

[60] M. Kojima and L. Tunçel. Cones of matrices and successive convex relaxations of nonconvex sets. *SIAM Journal on Optimization*, 10:750–778, 2000.

[61] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11:796–817, 2001.

[62] Y. Li and J-P. P. Richard. Cook, Kannan and Schrijver's example revisited. *Discrete Optimization*, 5:724–734, 2008.

[63] A. Lodi. Mixed integer programming computation. chapter 16, pages 619–645. Springer-Verlag, New York, 2010.

[64] L. Lovász. Geometry of numbers and integer programming. In M. Iri and K. Tanabe, editors, *Mathematical Programming: Recent Developments and Applications*, pages 177–210. Kluwer, 1989.

[65] M. Lubin and I. Dunning. Computing in operations research using julia. *INFORMS Journal on Computing*, 27(2):238–248, 2015.

[66] H. Marchand and L.A. Wolsey. Aggregation and Mixed Integer Rounding to solve MIPs. *Operations Research*, 49:363–371, 2001.

[67] R.D. McBride. Finding the integer efficient frontier for quadratic capital budgeting problems. *Journal of Financial and Quantitative Analysis*, 16(02):247–253, 1981.

[68] D. Micciancio and S. Goldwasser. *Complexity of lattice problems: a cryptographic perspective*, volume 671. Springer Science & Business Media, 2012.

[69] S. Modaresi, M. R. Kılınç, and J. P. Vielma. Split cuts for conic programming. 9th Mixed Integer Programming Workshop (MIP 2012), July 16–19, 2012, Davis, CA, Poster., 2012.

[70] S. Modaresi, M. R. Kılınç, and J. P. Vielma. Intersection cuts for nonlinear integer programming: Convexification techniques for structured sets. *Mathematical Programming*, pages 1–37, 2015.

[71] S. Modaresi, M. R. Kılınç, and J. P. Vielma. Split cuts and extended formulations for mixed integer conic quadratic programming. *Operations Research Letters*, 43(1):10 – 15, 2015.

[72] S. Modaresi and J. P. Vielma. Convex hull of two quadratic or a conic quadratic and a quadratic inequality. Technical report, MIT, 2014.

[73] S. Modaresi and J. P. Vielma. The power of a negative eigenvalue: Aggregation cuts for nonlinear integer programming. 2014 Mixed Integer Programming Workshop, July 21–24, 2014, Columbus, OH, Poster., 2014.

[74] G. L. Nemhauser and L. A. Wolsey. *Integer and combinatorial optimization.* Wiley, 1988.

[75] G. L. Nemhauser and L. A. Wolsey. A recursive procedure to generate all cuts for 0-1 mixed integer programs. *Mathematical Programming*, 46:379–390, 1990.

[76] Y. Nesterov, H. Wolkowicz, and Y. Ye. Nonconvex Quadratic Optimization. In R. Saigal, L. Vandenberghe, and H. Wolkowicz, editors, *Handbook of Semidefinite Programming*, pages 361 – 420. Kluwer Academic Publishers, 2000.

[77] C.L.F. Oustry. SDP relaxations in combinatorial optimization from a Lagrangian viewpoint. *Advances in Convex Analysis and Global Optimization: Honoring the Memory of C. Caratheodory (1873-1950)*, 54:119–134, 2001.

[78] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.

[79] I. Pólik and T. Terlaky. A survey of the S-lemma. *SIAM review*, 49:371–418, 2007.

[80] S. Poljak, F. Rendl, and H. Wolkowicz. A recipe for semidefinite relaxation for (0, 1)-quadratic programming. *Journal of Global Optimization*, 7:51–73, 1995.

[81] D. A. Morán R, S. Dey, and J. P. Vielma. A strong dual for conic mixed-integer programs. *SIAM Journal on Optimization*, 22(3):1136–1150, 2012.

[82] K. Ranestad and B. Sturmfels. The convex hull of a variety. *Notions of Positivity and the Geometry of Polynomials*, pages 331–344, 2011.

[83] K. Ranestad and B. Sturmfels. On the convex hull of a space curve. *Advances in Geometry*, 12:157–178, 2012.

[84] R. Sanyal, F. Sottile, and B. Sturmfels. Orbitopes. *Mathematika*, 57:275–314, 2011.

[85] C. Scheiderer. Convex hulls of curves of genus one. *Advances in Mathematics*, 228:2606 – 2622, 2011.

[86] H.D. Sherali and W.P. Adams. *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer, 1998.

[87] R. A. Stubbs and S. Mehrotra. A branch-and-cut method for 0-1 mixed convex programming. *Mathematical Programming*, 86:515–532, 1999.

[88] M. Tawarmalani and N.V. Sahinidis. *Convexification and global optimization in continuous and mixed-integer nonlinear programming: theory, algorithms, software, and applications*, volume 65. Springer, 2002.

[89] J. P. Vielma. A constructive characterization of the split closure of a mixed integer linear program. *Operations Research Letters*, 35:29–35, 2007.

[90] J. P. Vielma. Mixed integer linear programming formulation techniques. *SIAM Review*, 57(1):3–57, 2015.

[91] J. P. Vielma, I. Dunning, J. Huchette, and M. Lubin. Extended formulations in mixed integer conic quadratic programming. 2015.

[92] H.M. Weingartner. Capital budgeting of interrelated projects: survey and synthesis. *Management Science*, 12(7):485–516, 1966.

[93] L. A. Wolsey. *Integer Programming*. Wiley, 1998.

[94] U. Yıldıran. Convex hull of two quadratic constraints is an LMI set. *IMA Journal of Mathematical Control and Information*, 26:417–450, 2009.

[95] U. Yıldıran and I. E. Kose. LMI representations of the convex hulls of quadratic basic semialgebraic sets. *Journal of Convex Analysis*, 17:535–551, 2010.

[96] S. Yıldız and G. Cornuéjols. Disjunctive cuts for cross-sections of the second-order cone. 2014.