A geometric view of SDP exactness in QCQPs and its applications

Alex L. Wang[1] and Fatma Kılınç-Karzan[1]

[1]Carnegie Mellon University, Pittsburgh, PA, 15213, USA.

May 28, 2021

Abstract

Quadratically constrained quadratic programs (QCQPs) are a highly expressive class of nonconvex optimization problems. While QCQPs are NP-hard in general, they admit a natural convex relaxation via the standard (Shor) semidefinite program (SDP) relaxation. Towards understanding when this relaxation is exact, we study general *quadratically constrained sets* and their projected SDP relaxations. We present and analyze a *systematic* rounding procedure defined in the original space that can be used to decide *convex hull exactness*, i.e., when the convex hull of the possibly nonconvex quadratically constrained set coincides with its projected SDP relaxation. When viewed as a proof technique, this rounding procedure gives rise to a sufficient condition for convex hull exactness. One of our main contributions shows that under a technical assumption, this sufficient condition is also necessary. The motivation and analysis of this procedure stem from a geometric treatment of the projected SDP relaxation. Specifically, we give an explicit description of the set of rounding directions at a given point in the projected SDP relaxation using only constraints in the original space. We close with a number of example applications.

1 Introduction

Quadratically constrained quadratic programs (QCQPs) are a fundamental class of nonconvex optimization problems. QCQPs arise naturally in application areas in operations research, engineering, computer science, and machine learning; see [34] for additional applications. Specifically, any polynomial optimization problem or $\{0,1\}$ integer program may be reformulated as a QCQP.

Although QCQPs are NP-hard to solve in general, they admit a natural tractable convex relaxation known as the standard semidefinite program (SDP) relaxation [30]. This relaxation is also referred to as the Shor SDP relaxation.

A number of exciting results in application areas such as phase retrieval [12] and clustering [1, 23, 29] have shown that under various random data models, the QCQP formulation of the corresponding problem has an exact SDP relaxation.

Inspired by such results, an interesting line of recent research has sought to understand deterministic conditions under which the SDP relaxations of more general QCQPs are exact for different definitions of exactness. Burer and Ye [11] study *objective value exactness*—the condition that the optimal value of the QCQP and the optimal value of its SDP relaxation coincide—and give sufficient conditions under which diagonal QCQPs (those QCQPs with diagonal quadratic forms) have this property. Wang and Kılınç-Karzan [32, 34] continue this line of work by developing a general framework for deriving sufficient conditions for both objective value exactness and *convex hull exactness*—the condition that the convex hull of the QCQP epigraph coincides with the (projected) SDP epigraph—for QCQPs with a polyhedral set of convex Lagrange multipliers Γ (see Section 2). (See [2, 34] and references therein for additional work in this direction.)

Beyond being a natural sufficient condition for objective value exactness, convex hull exactness has its own far-reaching applications and motivation. Such results find use for example in deriving strong relaxations of certain critical substructures in nonconvex problems. Specifically, the convexification of commonly occurring substructures in complex nonconvex problems has been critical in advancing the state-of-the-art computational approaches for general nonlinear nonconvex programs and mixed integer linear programs [14, 31].

As a consequence of their framework, Wang and Kılınç-Karzan [34] show that both objective value exactness and convex hull exactness hold for (vectorized reformulations of) quadratic matrix programs (QMPs) whenever the number of constraints is small enough and the set of convex Lagrange multiplier is polyhedral. A QMP is an optimization problem over a matrix variable $X \in \mathbb{R}^{n\times k}$ where the objective function and constraints are each of the form

$$\operatorname{tr}(X^\top AX) + 2\operatorname{tr}(B^\top X) + c$$

for $A \in \mathbb{S}^n$, $B \in \mathbb{R}^{n\times k}$ and $c \in \mathbb{R}$. QMPs and their SDP relaxations were first studied by Beck [7], Beck et al. [8] who gave sufficient conditions for objective value exactness that notably *did not* rely on a polyhedrality assumption. This class of problems has been used, for example, to model robust least squares problems, the orthogonal Procrustes problem [7], and sphere packing [8].

While the framework presented by Wang and Kılınç-Karzan [34] is general enough to cover and extend many existing results on objective value and convex hull exactness [10, 11, 16, 19, 22, 24, 33, 38], it is still quite limited. In particular, the assumption that the set of convex Lagrange multipliers Γ is polyhedral is rarely satisfied outside of diagonal QCQPs and precludes the results in [34] from being applicable to a wider range of interesting QCQPs. For example, the framework in [34] fails to address sets with general quadratic matrix constraints as discussed above or sets with complementarity constraints modeling big-M restrictions (which are critical in applications like sparse regression). Further highlighting the limitations of the framework in [34], convex hull descriptions for various sets with complementarity constraints are routinely derived via the "perspective reformulation/relaxation trick" [3, 13, 15, 17, 18, 35, 36] which is known to be a relaxed form of the SDP relaxation.

In this paper, we vastly generalize the framework first introduced in [32, 34] by eliminating its reliance on the polyhedrality assumption. Specifically, let $\mathcal{S}$ be a subset of $\mathbb{R}^n$ defined by quadratic equality and inequality constraints and let $\overline{\mathcal{S}}$ denote the (projected) SDP relaxation of $\mathcal{S}$. For example, taking $\mathcal{S}$ to be the epigraph of a QCQP, we have that $\overline{\mathcal{S}}$ is the epigraph of its SDP relaxation. The main contribution of this paper is a sufficient condition for *convex hull exactness*, i.e., the property $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$, that is (i) applicable with no assumption on the set of convex Lagrange multipliers and (ii) both necessary and sufficient under a minor technical assumption (Assumption 2).

Our sufficient condition for convex hull exactness is based on a simple convex decomposition procedure. Whenever convex hull exactness holds, a natural approach to proving this result is to design a convex decomposition procedure (possibly tailored to $\mathcal{S}$ and $\overline{\mathcal{S}}$) which decomposes an arbitrary $x \in \overline{\mathcal{S}}$ as a convex combination of points in $\mathcal{S}$. Indeed, this approach (along with other similar approaches) has been explored multiple times in this line of work [6, 10, 19, 25, 26, 32, 33]. Towards this goal, consider the natural Carathéodory procedure: 1) If $x \in \mathcal{S}$, output x. 2) Else, if $x \notin \mathcal{S}$ and x is an extreme point of $\overline{\mathcal{S}}$, output "Fail: $x \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$." 3) Else, x is not an extreme point of $\overline{\mathcal{S}}$ and there exists a nonzero $y \in \mathbb{R}^n$ such that $[x \pm \epsilon y] \subseteq \overline{\mathcal{S}}$ for all $\epsilon > 0$ small enough (we will refer to the set of all such y's as the set of *rounding directions* at x and denote it $\mathcal{R}(x)$). Finally, pick $\alpha^+, \alpha^- > 0$ so that $x + \alpha^+ y$ and $x - \alpha^- y$ also belong to $\overline{\mathcal{S}}$ and recurse on these two new points. We will see in Section 3 that this procedure (for carefully chosen α^+ and α^-) will necessarily terminate and either correctly decompose x or correctly output an $x' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$. In particular, the task of proving $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$ can be reduced (via the above procedure) to the *geometric* question of understanding when $\mathcal{R}(x)$ is nontrivial, i.e., $\mathcal{R}(x) \neq \{0\}$.

In prior work, establishing that $\mathcal{R}(x)$ is nontrivial for all $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$ was handled in an *ad hoc* manner, requiring insights and technical arguments *specific* to the sets $\mathcal{S}$ and $\overline{\mathcal{S}}$ considered in the particular application. In contrast to prior work, in this paper we will present a *general* decomposition procedure which will be applicable to arbitrary sets $\mathcal{S}$ and $\overline{\mathcal{S}}$ under a minor technical assumption. Specifically, one of the main contributions of this paper is an explicit description of $\mathcal{R}(x)$ involving only linear constraints in the original space and terms depending on the set of Lagrange multipliers Γ; in particular, our description will *not* involve a lifted semidefinite constraint. Notably, the procedure and analysis we present in this paper will replace the critical assumption of [34] that Γ is polyhedral with the weaker assumption that the polar cone of Γ is *facially exposed* (see Assumption 2 and Section 2 for a discussion of this assumption). In contrast to previous work on the geometry of semidefinite programs [21, 26, 28], our analysis is done in the original (projected) space $\mathbb{R}^n$ (as opposed to the lifted space where the SDP relaxation lies). This new characterization of $\mathcal{R}(x)$

allows us to systematically treat QCQPs where the set of convex Lagrange multipliers Γ is simple enough analytically to work with. In particular, our theory applies to quadratic matrix programming, the NP-hard partition problem and a well-studied set involving convex quadratics, binary variables and big-M relations that is frequently used in sparse regression problems.

1.1 Overview and outline of the paper

A summary of our contributions, along with an outline of the remainder of the paper, is as follows:

1. In Section 2, we formally define our setup and assumptions and recall basics of the framework introduced in [34]. We then define and examine a number of faces of the cone of convex Lagrange multipliers Γ and its polar cone Γ° that play key roles in our analysis. Finally, we discuss Assumption 2, namely the assumption that Γ° is facially exposed. This sets up the geometric viewpoint of convex hull exactness that we will employ in this paper.

2. In Section 3, we present and analyze our rounding procedure, Procedure 1. In contrast to the rounding procedure implicit in [34, Lemma 7], Procedure 1 does not make any assumption on the geometry of Γ. In fact, our procedure is well-defined whenever a standard definiteness assumption (Assumption 1) is made, notably even when Γ° is not facially exposed. We will then show in Theorem 1 that when Γ° is facially exposed (Assumption 2), this procedure can only fail to find a convex decomposition of the input point if $\operatorname{conv}(\mathcal{S}) \neq \overline{\mathcal{S}}$. When viewed as a *proof technique*, Procedure 1 gives rise to a sufficient condition for convex hull exactness which is also necessary under Assumption 2. We summarize the sufficient condition and its necessity under Assumption 2 in Theorem 2.

3. In Section 4, we present applications of our framework to sets arising in quadratic matrix programming [7, 8], the partition problem [21], and a simple set that is important in mixed binary programming [17]. We conclude by revisiting the setting of polyhedral Γ:

   In Section 4.1, we show that the SDP relaxation of a quadratic matrix program satisfies convex hull exactness whenever the number of constraints is small (when compared to the rank of the matrix variable). This recovers a result first presented in [34]. In contrast to the *ad hoc* proof given in [34], the proof we present in Section 4.1 follows the outline of our general framework.

   In Section 4.2, we consider the NP-hard partition problem and its SDP relaxation. Using our framework, we give an explicit description of the optimal value and epigraph of the SDP relaxation. Consequently, we recover a result due to Laurent and Poljak [21] stating that deciding whether objective value exactness holds for the partition QCQP is NP-hard. In contrast, we show that convex hull exactness never holds for the partition QCQP (as long as there are at least two nonzero weights). This then implies that deciding whether convex hull exactness holds for the partition QCQP is trivial.

   In Section 4.3, we apply our framework to show that convex hull exactness holds for a well-studied set involving convex quadratics, binary variables and big-M relations. This convex hull characterization is well-known in the literature and is often shown as a consequence of the perspective formulation trick due to Ceria and Soares [13] (see also [15, 17, 18]).

   Finally, in Section 4.4 we revisit the setting of polyhedral Γ and show that a variant of the sufficient condition for convex hull exactness presented in [34, Theorem 1] is also necessary. To the best of our knowledge, this is the first necessary and sufficient condition for convex hull exactness even in the setting of diagonal QCQPs.

1.2 Notation

Let $\mathbb{R}_+$ denote the nonnegative reals. For nonnegative integers $m \leq n$, define $[n] \coloneqq \{1, \dots, n\}$ and $[m, n] \coloneqq \{m, m+1, \dots, n-1, n\}$. Let $\mathbb{S}^n$ denote the set of real symmetric $n \times n$ matrices. We write $A \succeq 0$ (respectively, $A \succ 0$) if A is positive semidefinite (respectively, positive definite). Given $A \in \mathbb{S}^n$, let $\operatorname{range}(A)$ and $\ker(A)$ denote the range and kernel of A respectively. Given $A \in \mathbb{S}^n$ and a vector space $V \subseteq \mathbb{R}^n$, let $A\downharpoonright_V$ denote the restriction of A to V. For $a \in \mathbb{R}^n$, let $\operatorname{Diag}(a)$ denote the diagonal matrix $A \in \mathbb{S}^n$ with diagonal entries $A_{i,i} = a_i$ for all $i \in [n]$. For $a \in \mathbb{R}^n$, let $\operatorname{supp}(a) \subseteq [n]$ and $\operatorname{zeros}(a) \subseteq [n]$ denote the support and zeros of a. For a set $\mathcal{D} \subseteq \mathbb{R}^n$, let $\mathcal{D}^\circ$, $\operatorname{int}(\mathcal{D})$, $\operatorname{rint}(\mathcal{D})$, $\operatorname{rbd}(\mathcal{D})$, $\operatorname{conv}(\mathcal{D})$, $\operatorname{span}(\mathcal{D})$, $\operatorname{aff}(\mathcal{D})$, $\dim(\mathcal{D})$, $\operatorname{aff}\dim(\mathcal{D})$ and $\mathcal{D}^\perp$ denote the polar, interior, relative interior, relative boundary, convex hull, lineal hull, affine hull, dimension, affine dimension, and orthogonal complement of $\mathcal{D}$, respectively. We will write $\mathcal{F} \trianglelefteq \mathcal{D}$ to denote the fact that $\mathcal{F}$ is a face of $\mathcal{D}$. Let $\mathbf{S}^{k-1}$ denote the $k-1$ sphere.

2 Preliminaries

2.1 Setup

We will consider subsets of $\mathbb{R}^n$ defined by m-many quadratic constraints: Given $m_I, m_E \in \mathbb{N}$ denoting the number of inequality and equality constraints respectively, define $m \coloneqq m_I + m_E$. Let n_1, $n_2 \in \mathbb{N}$ and define $n \coloneqq n_1 + n_2$. Here, n_2 will denote the number of variables in which the quadratic functions are known *a priori* to vary only linearly.

More formally, decomposing $x \in \mathbb{R}^n$ as $(x^{(1)}, x^{(2)}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, we will consider sets of the form

$$\mathcal{S} \coloneqq \left\{ x \in \mathbb{R}^n : \begin{array}{l} q_i(x) \leq 0,\ \forall i \in [m_I] \\ q_i(x) = 0,\ \forall i \in [m_I + 1, m] \end{array} \right\}$$

where $q_i(x) \coloneqq x^{(1)\top} A_i x^{(1)} + 2 b_i^\top x + c_i$ for some $A_i \in \mathbb{S}^{n_1}$, $b_i \in \mathbb{R}^n$, and $c_i \in \mathbb{R}$ for all $i \in [m]$.

Remark 1. This decomposition of x allows us to specially handle variables that participate only linearly in the constraints. Such variables arise, for example, in epigraph constraints. □

Let $q : \mathbb{R}^n \to \mathbb{R}^m$ denote the vector-valued function with $q(x)_i = q_i(x)$ for all $i \in [m]$. We will work extensively with the aggregated quadratic function $x \mapsto \langle \gamma, q(x) \rangle$ based on the aggregation weights $\gamma \in \mathbb{R}^m$. For notational convenience, define $A(\gamma) \coloneqq \sum_{i \in [m]} \gamma_i A_i$. Similarly define $b(\gamma)$ and $c(\gamma)$ and note that

$$\langle \gamma, q(x) \rangle = x^{(1)\top} A(\gamma) x^{(1)} + 2 b(\gamma)^\top x + c(\gamma).$$

2.2 The convex Lagrange multipliers and the projected SDP relaxation

Recall the following definition from [34].

Definition 1. The set of *convex Lagrange multipliers* associated with $\mathcal{S}$ is

$$\Gamma \coloneqq \left\{ \gamma \in \mathbb{R}^m : \begin{array}{l} A(\gamma) \succeq 0 \\ \gamma_i \geq 0,\ \forall i \in [m_I] \end{array} \right\}. \qquad \square$$

Note that Γ is a closed convex cone. The following assumption can be interpreted as requiring that the dual of the SDP relaxation of $\mathcal{S}$ be strictly feasible. This assumption is standard in the literature [7, 9, 11, 34, 37].

Assumption 1. There exists $\gamma^* \in \Gamma$ such that $A(\gamma^*) \succ 0$ and $\gamma_i > 0$ for all $i \in [m_I]$. □

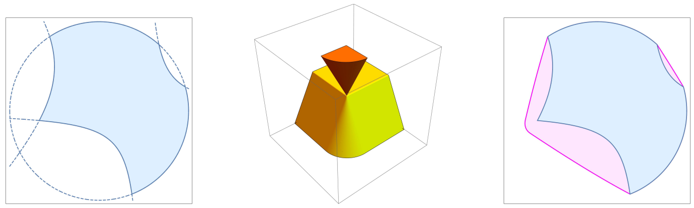
Remark 2. For the sake of simplicity, in a few of our results we will rely on Assumption 1 even though the weaker assumption that Γ is simply nonempty would suffice. □

Definition 2. The *(projected) SDP relaxation* of $\mathcal{S}$ is

$$\overline{\mathcal{S}} \coloneqq \left\{ x : q(x) \in \Gamma^\circ \right\}. \qquad (1)$$

Here, Γ° denotes the polar cone of Γ. □

The following lemma states that under Assumption 1, $\overline{\mathcal{S}}$ is equivalent to the usual (projected) SDP relaxation of $\mathcal{S}$ thus justifying its name.



(a) $\mathcal{S}$ (b) Γ and Γ° (c) $\mathcal{S}$ and $\overline{\mathcal{S}}$

Figure 1: The sets $\mathcal{S}$, $\overline{\mathcal{S}}$, Γ, and Γ° from Example 1 are shown in blue, magenta, orange, and yellow respectively. It is clear from the last picture that $\text{conv}(\mathcal{S}) \neq \overline{\mathcal{S}}$ in this example.

Lemma 1. *Suppose Assumption 1 holds. Then*

$$\overline{\mathcal{S}} = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \exists\, X \succeq x^{(1)} x^{(1)\top} : \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \leq 0,\ \forall i \in [m_I] \\ \langle A_i, X\rangle + 2b_i^\top x + c_i = 0,\ \forall i \in [m_I + 1, m] \end{array} \right\}. \tag{2}$$

Remark 3. In comparison with the right-hand-side of (2), the expression (1) for $\overline{\mathcal{S}}$ makes explicit the role played by Γ°. In particular, the definition of $\overline{\mathcal{S}}$ given in (1) lends itself to a clean analysis whenever Γ° is sufficiently simple. ☐

Let us consider a concrete example to help materialize these definitions.

Example 1. Suppose $n_1 = 2$, $n_2 = 0$, $m_I = 3$, and $m_E = 0$, and consider

$$\mathcal{S} := \left\{ x \in \mathbb{R}^2 : \begin{array}{l} q_1(x) \leq 0 \\ q_2(x) \leq 0 \\ q_3(x) \leq 0 \end{array} \right\},$$

where $q_1(x) := 2x_1x_2 - x_2 - 1/4$, $q_2(x) := x_1^2 - x_2^2 - x_1 + x_2 - 1$, and $q_3(x) := x_1^2 + x_2^2 - 1$. Through a straightforward calculation (given in Appendix A.2), we obtain

$$\Gamma = \left\{ \gamma \in \mathbb{R}^3 : \begin{array}{l} \gamma_3 \geq \sqrt{\gamma_1^2 + \gamma_2^2} \\ \gamma \geq 0 \end{array} \right\}, \qquad \Gamma^\circ = \left\{ \ell \in \mathbb{R}^3 : -\ell_3 \geq \sqrt{(\ell_1)_+^2 + (\ell_2)_+^2} \right\},$$

$$\overline{\mathcal{S}} = \left\{ x \in \mathbb{R}^2 : -q_3(x) \geq \sqrt{q_1(x)_+^2 + q_2(x)_+^2} \right\}.$$

See Figure 1 for the plots of the sets corresponding to $\mathcal{S}$, Γ, Γ°, and $\overline{\mathcal{S}}$. ☐

2.3 Faces of Γ and Γ°

This section defines key faces of Γ and Γ° that will play important roles in our analysis. We will additionally recall a number of elementary properties of convex cones and their faces specialized to our setting. See [4, 5, 26] for a more in-depth treatment of general convex cones and their faces.

Recall the following definitions.

Definition 3. Given a nonempty face $\mathcal{G} \trianglelefteq \Gamma^\circ$ and $q \in \text{rint}(\mathcal{G})$, the *conjugate face of* $\mathcal{G}$ is $\mathcal{G}^\triangle := \Gamma \cap \mathcal{G}^\perp = \Gamma \cap q^\perp$. Similarly define the conjugate face of $\mathcal{F}$ for a nonempty face $\mathcal{F} \trianglelefteq \Gamma$. ☐

Definition 4. For nonempty $\mathcal{G} \trianglelefteq \Gamma^\circ$, we say that $\mathcal{G}$ is an *exposed face* of Γ° if there exists $\gamma \in \Gamma$ such that $\mathcal{G} = \Gamma^\circ \cap \gamma^\perp$. ☐

Our rounding procedure will seek to understand how a given point $x \in \overline{\mathcal{S}}$ interacts with Γ° (cf. Definition 2). To this end, we define the following faces of Γ and Γ° associated to $x \in \overline{\mathcal{S}}$.

Definition 5. Given $x \in \overline{\mathcal{S}}$, let $\mathcal{G}(x) \trianglelefteq \Gamma^\circ$ denote the minimal face of Γ° containing $q(x)$. $\square$

Fact 1. *Given $x \in \overline{\mathcal{S}}$, we have that $q(x) \in \operatorname{rint}(\mathcal{G}(x))$.*

Definition 6. Given $x \in \overline{\mathcal{S}}$, let $\mathcal{F}(x) \coloneqq \mathcal{G}(x)^\triangle$. That is, $\mathcal{F}(x) \coloneqq \Gamma \cap q(x)^\perp$. $\square$

We will use the following assumption on Γ° to establish the necessity of the sufficient conditions for convex hull exactness that we present in this paper.

Assumption 2. Γ° is facially exposed, i.e., every nonempty face $\mathcal{G} \trianglelefteq \Gamma^\circ$ is exposed. $\square$

This assumption holds for any cone isomorphic to a slice of the nonnegative orthant, the second-order cone, or the positive semidefinite cone. See [27] for a longer discussion of this assumption and its connections to the *nice cones*. In general, all *nice cones* are facially exposed.

The main property we will make use of for exposed faces $\mathcal{G} \trianglelefteq \Gamma^\circ$ is the following:

Fact 2. *A nonempty face $\mathcal{G} \trianglelefteq \Gamma^\circ$ is exposed if and only if $\mathcal{G} = (\mathcal{G}^\triangle)^\triangle$.*

3 A complete rounding scheme

In this section, we present a simple recursive procedure (Procedure 1) which, given $x \in \overline{\mathcal{S}}$, will either output a convex decomposition of x as elements from $\mathcal{S}$ or claim to find a point $x' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$. In particular, a sufficient condition for *convex hull exactness*, i.e., $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$, is the condition that Procedure 1 always returns such a convex decomposition. Our main contribution, Theorems 1 and 2 show that under a technical assumption, this sufficient condition is in fact also necessary.

3.1 The set of rounding directions

We use the following definition in the description and analysis of our procedure.

Definition 7. Given $x \in \overline{\mathcal{S}}$, the set of *rounding directions* at x is

$$\mathcal{R}(x) \coloneqq \{y \in \mathbb{R}^n : q(x + \alpha y) \in \operatorname{span}(\mathcal{G}(x)),\ \forall \alpha \in \mathbb{R}\}. \qquad \square$$

Remark 4. Recalling that $q(x) \in \operatorname{rint}(\mathcal{G}(x))$, we deduce that for all $y \in \mathcal{R}(x)$ and all α small enough, $q(x + \alpha y) \in \mathcal{G}(x) \subseteq \Gamma^\circ$. In particular, for all $y \in \mathcal{R}(x)$ and all $\epsilon > 0$ small enough, $[x \pm \epsilon y]$ is contained in $\overline{\mathcal{S}}$. We will see in Corollary 3 that under Assumption 2, $\mathcal{R}(x)$ is given *exactly* by the set of such vectors. Thus, the name *rounding directions* is justified. $\square$

Definition 8. Given $x \in \overline{\mathcal{S}}$, we will say that $\mathcal{R}(x)$ is *trivial* if $\mathcal{R}(x) = \{0\}$; else it is *nontrivial.* We will refer to a vector $y \in \mathcal{R}(x)$ such that $y^{(1)} \neq 0$ as a *bounded rounding direction.* $\square$

The following lemma gives explicit descriptions of Γ°, $\mathcal{G}(x)$ and $\mathcal{R}(x)$ under Assumption 2. Of note, the explicit description of $\mathcal{R}(x)$ below involves only *linear* constraints. Consequently, the *a priori* possibly nonconvex set $\mathcal{R}(x)$ in fact forms a linear subspace under Assumption 2.

Lemma 2. *Suppose Assumption 1 holds. Then,*

$$\Gamma^\circ = \left\{(\langle A_i, \xi\rangle)_i : \xi \in \mathbb{S}^{n_1}_-\right\} + \left(\mathbb{R}^{m_I}_- \times \{0\}\right).$$

If additionally Assumption 2 holds, then for any $x \in \overline{\mathcal{S}}$ and $f \in \operatorname{rint}(\mathcal{F}(x))$,

$$\mathcal{G}(x) = \left\{ (\langle A_i, \xi \rangle)_i : \begin{array}{l} \xi \in \mathbb{S}^{n_1}_-, \\ \operatorname{range}(\xi) \subseteq \ker(A(f)) \end{array} \right\} + \left\{ \eta \times 0 : \begin{array}{l} \eta \in \mathbb{R}^{m_I}_-, \\ \operatorname{supp}(\eta) \cap \operatorname{supp}(f) = \varnothing \end{array} \right\},$$

$$\mathcal{G}(x)^\perp = \left\{ \gamma : \begin{array}{l} A(\gamma) \restriction_{\ker(A(f))} = 0 \\ \gamma_i = 0, \ \forall i \in [m_I] \cap \operatorname{zeros}(f) \end{array} \right\}, \text{ and}$$

$$\mathcal{R}(x) = \left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \left\langle A(\gamma) x^{(1)}, y^{(1)} \right\rangle + \langle b(\gamma), y \rangle = 0, \ \forall \gamma \in \mathcal{G}(x)^\perp \end{array} \right\}.$$

3.2 The rounding procedure

We are now ready to present our rounding procedure.

Procedure 1. *Suppose Assumption 1 holds and let $x \in \overline{\mathcal{S}}$.*

atemsep=0pt, parsep=0pt, topsep=0pt If $x \in \mathcal{S}$, output x.

btemsep=0pt, pbrsep=0pt, topsep=0pt Else if $\mathcal{R}(x)$ contains a bounded rounding direction y:

(a) Define α^+ and α^- (see Definition 9 below).

(b) Recursively run this procedure on $x + \alpha^+ y$ and $x - \alpha^- y$ and return the appropriate convex combination of the decompositions.

ctemsep=0pt, pcrsep=0pt, topsep=0pt Else, output "Failed at $x \in \overline{\mathcal{S}}$" and stop.

Definition 9. Suppose Assumption 1 holds. Let $x \in \overline{\mathcal{S}}$ and consider a bounded rounding direction $y \in \mathcal{R}(x)$. Define

$$\alpha^+ \coloneqq \max \left\{ \alpha \in \mathbb{R} : x + \alpha y \in \overline{\mathcal{S}} \right\}, \quad \text{and} \quad \alpha^- \coloneqq \max \left\{ \alpha \in \mathbb{R} : x - \alpha y \in \overline{\mathcal{S}} \right\}. \quad \square$$

Lemma 8 in Appendix B.2 shows that α^+ and α^- are in fact well-defined.

Clearly, if this procedure terminates and outputs a convex decomposition, then it does so correctly. Appendix B.3 establishes that $\dim(\mathcal{G}(x))$ decreases with each recursive call so that this procedure necessarily terminates after a finite number of iterations (specifically, after at most m-many recursions). Appendix B.4 will show that under Assumption 2, this procedure fails at x only if $x \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$. Consequently, the following result will follow as a corollary to Appendices B.3 and B.4.

Theorem 1. *Suppose Assumption 1 holds. Then, given $x \in \overline{\mathcal{S}}$, Procedure 1 will terminate in a finite number of steps and either output x as a convex combination of points in $\mathcal{S}$ or fail at some (possibly different) point $x' \in \overline{\mathcal{S}}$. If Procedure 1 fails and Assumption 2 holds, then $x' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$.*

Viewing Procedure 1 as a proof technique, we get the following condition for convex hull exactness.

Theorem 2. *Suppose Assumption 1 holds. The condition:*
"For all $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$, the set $\mathcal{R}(x)$ contains a bounded rounding direction."
is a sufficient condition for $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$. If Assumption 2 additionally holds, it is also necessary.

We emphasize that when Assumption 2 does not hold, Procedure 1 may fail at some point $x \in \overline{\mathcal{S}}$ even if $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$. This happens when $\mathcal{R}(x)$ does not contain a bounded rounding direction.

Appendix B.5 considers a minor modification of Procedure 1 for linear optimization problems over quadratically constrained sets. Given a point $x \in \overline{\mathcal{S}}$ and linear objective $w \in \mathbb{R}^n$, Procedure 2 either outputs $x' \in \mathcal{S}$ with $w^\top x' \leq w^\top x$ or the message "Failed at $x'' \in \overline{\mathcal{S}}$." Similar to Procedure 1, if Assumption 2 holds, then Procedure 2 fails at $x'' \in \overline{\mathcal{S}}$ only if $x'' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$. In contrast to Procedure 1 which makes two recursive calls at each step (to a recursive depth of at most m), Procedure 2 makes a single recursive call in each step. Consequently, Procedure 2 will terminate after computing $\mathcal{R}(x)$ for at most linearly many (in m) choices of x. See Appendix B.5 for details.

Remark 5. Procedures 1 and 2 offer proof techniques that are well-suited for instances where the cones Γ and Γ° are simple enough to complete proofs with. Specifically, these procedures require exact computation and access to geometric objects (e.g., $\mathcal{R}(x)$ and $\mathcal{G}(x)^\perp$) that may be difficult to compute for general instances or via software. ☐

4 Example applications

4.1 Quadratic matrix programming

In this subsection, we consider the special case where each A_i is of the form $I_k \otimes \mathbb{A}_i$ for some $\mathbb{A}_i \in \mathbb{S}^r$ where $r = n_1/k$. This structure arises naturally in vectorized formulations of quadratic matrix programs (QMPs) [7]; see also [34]. Specifically, a QMP is an optimization problem in the variables $(X, x^{(2)}) \in \mathbb{R}^{r\times k} \times \mathbb{R}^{n_2}$, where every constraint and objective function is of the form

$$\operatorname{tr}(X^\top \mathbb{A} X) + 2\operatorname{tr}(B^\top X) + 2\left\langle b^{(2)}, x^{(2)}\right\rangle + c,$$

where $\mathbb{A} \in \mathbb{S}^r$, $B \in \mathbb{R}^{r\times k}$, and $c \in \mathbb{R}$. Letting $x^{(1)} \in \mathbb{R}^{n_1}$ (resp. $b^{(1)} \in \mathbb{R}^{n_1}$) denote the vector formed by stacking the columns of X (resp. B) on top of each other, we can rewrite the above expression as

$$x^{(1)\top}(I_k \otimes \mathbb{A})x^{(1)} + 2\langle b, x\rangle + c.$$

The following lemma establishes that if the number of constraints defining $\mathcal{S}$ is small compared to k, then $\mathcal{R}(x)$ contains a bounded rounding direction for every $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$; see Appendix C.1 for its proof. Theorem 2 then implies that $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$.

Lemma 3. *Suppose Assumption 1 holds. Let $A_i = I_k \otimes \mathbb{A}_i$ for each $i \in [m]$ and assume $k \geq m+1$. Then, for all $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$, the set $\mathcal{R}(x)$ contains a bounded rounding direction. In particular, $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$.*

4.2 The partition problem

We next consider the partition QCQP and its SDP relaxation. The results in this section again highlight that questions of SDP exactness are often easy to answer when the set of convex Lagrange multipliers is "sufficiently simple."

Recall the partition QCQP: Given $a \in \mathbb{R}^n$, we want to minimize

$$\mathrm{Opt}_{\mathrm{QCQP}} \coloneqq \min_{x\in\mathbb{R}^n} \left\{(a^\top x)^2 : \, x_i^2 = 1, \, \forall i \in [n]\right\}.$$

Note that $\mathrm{Opt}_{\mathrm{QCQP}} = 0$ if and only if the vector a can be *partitioned* into two sets of equal weight. Thus deciding whether $\mathrm{Opt}_{\mathrm{QCQP}} = 0$ is NP-hard [20]. We will work with the epigraph of this QCQP:

$$\mathcal{S} \coloneqq \left\{ (x,t) \in \mathbb{R}^n \times \mathbb{R} : \begin{array}{l} (a^\top x)^2 - t \leq 0 \\ x_i^2 - 1 = 0, \, \forall i \in [n] \end{array} \right\}.$$

In a slight departure from the notation of previous sections, we will explicitly refer to the epigraph variable as t instead of x_{n+1}. In particular, we will use the notation $q(x,t)$ to denote the vector in $\mathbb{R}^{n+1}$ where $q(x,t)_1 = (a^\top x)^2 - t$ and $q(x,t)_i = x_{i-1}^2 - 1$ for $i \in [2, n+1]$.

Let $\overline{\mathcal{S}}$ denote the (projected) SDP relaxation of $\mathcal{S}$ and define $\mathrm{Opt}_{\mathrm{SDP}} \coloneqq \min\left\{t : \, (x,t) \in \overline{\mathcal{S}}\right\}$ to be the optimal value of the SDP relaxation. Lemma 4 below gives an explicit description of Γ and $\overline{\mathcal{S}}$ under a minor assumption.

Assumption 3. $a \in \mathbb{R}^n_{++}$ and $n \geq 2$. ☐

This assumption is essentially without loss of generality: It is straightforward to derive closed form descriptions of $\mathcal{S}$, $\operatorname{conv}(\mathcal{S})$, and $\overline{\mathcal{S}}$ when $n = 1$. Similarly, it is not difficult to relate the sets $\mathcal{S}$, $\operatorname{conv}(\mathcal{S})$, and $\overline{\mathcal{S}}$ defined by an arbitrary $a \in \mathbb{R}^n$ with a corresponding set defined by some $a' \in \mathbb{R}^{n'}_{++}$.

Lemma 4. *Suppose Assumption 3 holds. Then,*

$$\Gamma = \left\{ (\mu, \tau) \in \mathbb{R} \times \mathbb{R}^n : \begin{array}{l} \mu a a^\top + \operatorname{diag}(\tau) \succeq 0 \\ \mu \geq 0 \end{array} \right\}, \text{ and}$$

$$\overline{\mathcal{S}} = \left\{ (x,t) \in [-1,1]^n \times \mathbb{R} : (a^\top x)^2 + \max_{i \in [n]} \left(a_i \sqrt{1 - x_i^2} - \sum_{j \neq i} a_j \sqrt{1 - x_j^2} \right)_+^2 \leq t \right\}.$$

See Appendix C.2 for a proof of this statement.

Recall from [21] that a vector $a \in \mathbb{R}^n_{++}$ is said to be *balanced* if for all $i \in [n]$, $a_i \leq \sum_{j \neq i} a_j$. The following result then follows as a corollary to Lemma 4.

Corollary 1. *Suppose Assumption 3 holds. Then,* $\mathrm{Opt}_{\mathrm{SDP}} = 0$ *if and only if* a *is balanced.*

As a consequence of Corollary 1 (and the NP-hardness of deciding whether $\mathrm{Opt}_{\mathrm{QCQP}} = 0$ for the partition QCQP), we see that it is NP-hard to decide whether objective value exactness holds for the partition QCQP. This recovers a result due to Laurent and Poljak [21].

In contrast to the NP-hardness of checking *objective value exactness* for the partition QCQP, the following theorem shows that checking *convex hull exactness* for the partition QCQP is a trivial task.

Theorem 3. *Suppose Assumption 3 holds. Then,* $\operatorname{conv}(\mathcal{S}) \neq \overline{\mathcal{S}}$.

The proof of Theorem 3 follows from the observation that $\operatorname{conv}(\mathcal{S})$ is polyhedral and that $\overline{\mathcal{S}}$ is not polyhedral. See Appendix C.4 for details.

4.3 Mixed binary programming

We now apply our framework to a well-studied prototypical set involving a convex quadratic function, a binary variable and a big-M relation.

Take $n_1 = 2$, $n_2 = 1$, $m_I = 1$, and $m_E = 2$. Define

$$\mathcal{S} = \left\{ x \in \mathbb{R}^3 : q_1(x) \leq 0,\ q_2(x) = 0,\ q_3(x) = 0 \right\},$$

where $q_1(x) := x_2^2 - x_3$, $q_2(x) := x_1(x_1 - 1)$, and $q_3(x) := \sqrt{2} x_2 (x_1 - 1)$. In particular, $x^{(1)} = (x_1, x_2)$. In words, x_1 is a binary on-off variable, x_2 is a continuous variable which is constrained to be off whenever x_1 is off, and x_3 is the epigraph variable corresponding to x_2^2. Here, q_1 models the epigraph relation between the variables x_2 and x_3, q_2 models the binary requirement on x_1, and q_3 models the scaled big-M/complementarity relation between x_1 and x_2.

It is well-known that $\operatorname{conv}(\mathcal{S})$ is given by the perspective reformulation of $\mathcal{S}$ (see e.g., [17, 18]):

$$\operatorname{conv}(\mathcal{S}) = \left\{ x \in \mathbb{R}^3 : \ x_2^2 - x_3 x_1 \leq 0,\ 0 \leq x_1 \leq 1 \right\}. \tag{3}$$

We give a different proof of this description. We will show that $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$, the projected SDP relaxation, by showing that $\mathcal{R}(x)$ contains a bounded rounding direction for every $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$. Then using an explicit description of Γ°, it is possible to show that $\overline{\mathcal{S}}$ and the right hand side of (3) coincide.

A simple computation shows that in this setting, we have

$$\Gamma = \left\{ \gamma \in \mathbb{R}^3 : \gamma_1 + \gamma_2 \geq \sqrt{(\gamma_1 - \gamma_2)^2 + \left(\sqrt{2}\gamma_3\right)^2} \right\} \text{ and}$$

$$\Gamma^\circ = \left\{ \ell \in \mathbb{R}^3 : -\ell_1 - \ell_2 \geq \sqrt{(\ell_1 - \ell_2)^2 + \left(\sqrt{2}\ell_3\right)^2} \right\}.$$

Hence, Γ and Γ° are both second-order cones and Assumptions 1 and 2 hold.

It remains to show that for all $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$, the set $\mathcal{R}(x)$ contains a bounded rounding direction. To this end, let $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$. Recall that Γ° has three types of nonempty faces: the two trivial faces (the apex and the cone itself) and the one-dimensional proper faces. Thus, there are three cases to consider: (i) $\mathcal{G}(x) = \{0\}$, (ii) $\mathcal{G}(x) = \Gamma^\circ$, and (iii) $\mathcal{G}(x)$ is a one-dimensional face of Γ°.

In case (i), $q(x) = 0$ and $x \in \mathcal{S}$, a contradiction. In case (ii), $\operatorname{span}(\mathcal{G}(x)) = \mathbb{R}^3$ so that $\mathcal{R}(x) = \mathbb{R}^3$ and clearly contains a bounded rounding direction. In the final case, an application of Lemma 2 gives

$$\mathcal{R}(x) = \left\{ \begin{pmatrix} x_3 \\ -x_2 \\ 0 \end{pmatrix}, \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}, \begin{pmatrix} -x_3/2 \\ x_2 \\ -x_1/2 \end{pmatrix}, \begin{pmatrix} x_2(1 - x_1 - x_3) \\ x_1^2 - x_1 + x_1 x_3 + x_3 \\ -x_2 \end{pmatrix} \right\}^\perp . \quad (4)$$

Finally, it is not difficult to verify that $x \in \mathcal{R}(x)$ is a bounded rounding direction.

Remark 6. Here, the motivation for the final step of checking that $x \in \mathcal{R}(x)$ is as follows: One can show that $x_1 \neq 0$ and $x_1 x_3 = x_2^2$, whence the first three vectors in (4) span the 2-dimensional subspace orthogonal to x. In particular, any $y \in \mathcal{R}(x)$ must be a scalar multiple of x. □

4.4 Revisiting the setting of polyhedral Γ

Wang and Kılınç-Karzan [34] give sufficient conditions for convex hull exactness under the assumption that Γ is polyhedral. This assumption holds, for example, when the set of quadratic forms $\{A_1, \dots, A_m\}$ defining $\mathcal{S}$ is simultaneously diagonalizable. Specializing our Lemma 2 and Theorem 2 to this setting, we prove the following *necessary and sufficient* variant of [34, Theorem 1].

Theorem 4. *Suppose Assumption 1 holds and that Γ is polyhedral. Then* $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$ *if and only if*

$$\left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \langle b(\gamma), y \rangle = 0,\ \forall \gamma \in \mathcal{F} \end{array} \right\}$$

contains a bounded rounding direction for every $\mathcal{F} \trianglelefteq \Gamma$ which is exposed by some vector $q(x)$ for $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$. Here, f is any vector in $\operatorname{rint}(\mathcal{F})$.

This theorem follows from the observation that for polyhedral Γ, we have $\operatorname{span}(\mathcal{F}(x)) = \mathcal{G}(x)^\perp$ so that the constraints defining $\mathcal{R}(x)$ only need to be imposed for $\gamma \in \mathcal{F}(x)$ (see Appendix C.5).

Remark 7. The main difference between Theorem 4 and [34, Theorem 1] is that Theorem 4 only considers certain distinguished faces of Γ whereas [34, Theorem 1] imposes a constraint on every *semidefinite* face of Γ. □

Acknowledgments

This research is supported in part by NSF grant CMMI 1454548 and ONR grant N00014-19-1-2321.

References

[1] E. Abbe, A. S. Bandeira, and G. Hall. Exact recovery in the stochastic block model. *IEEE Trans. Inform. Theory*, 62(1):471–487, 2015.

[2] C.J. Argue, F. Kılınç-Karzan, and A. L. Wang. Necessary and sufficient conditions for rank-one generated cones. Technical Report arXiv:2007.07433, 2020.

[3] A. Atamturk, A. Gómez, and S. Han. Sparse and smooth signal estimation: Convexification of l0 formulations. Technical Report 1811.02655, 2018.

[4] G. P. Barker. Faces and duality in convex cones. *Linear and Multilinear Algebra*, 6(3):161–169, 1978.

[5] G. P. Barker. Theory of cones. *Linear Algebra Appl.*, 39:263–291, 1981.

[6] A. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete Comput. Geom.*, 13:189–202, 1995.

[7] A. Beck. Quadratic matrix programming. *SIAM J. Optim.*, 17(4):1224–1238, 2007.

[8] A. Beck, Y. Drori, and M. Teboulle. A new semidefinite programming relaxation scheme for a class of quadratic matrix problems. *Oper. Res. Lett.*, 40(4):298–302, 2012.

[9] A. Ben-Tal and M. Teboulle. Hidden convexity in some nonconvex quadratically constrained quadratic programming. *Math. Program.*, 72:51–63, 1996.

[10] S. Burer and F. Kılınç-Karzan. How to convexify the intersection of a second order cone and a nonconvex quadratic. *Math. Program.*, 162:393–429, 2017.

[11] S. Burer and Y. Ye. Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs. *Math. Program.*, 181:1–17, 2019.

[12] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM Rev.*, 57(2):225–251, 2015.

[13] S. Ceria and J. Soares. Convex programming for disjunctive convex optimization. *Math. Program.*, 86: 595–614, 1999.

[14] M. Conforti, G. Cornuéjols, and G. Zambelli. *Integer Programming*, volume 271 of *Grad. Texts in Math.* Springer, 2014.

[15] H. Dong and J. Linderoth. On valid inequalities for quadratic programming with continuous variables and binary indicators. In *Integer Programming and Combinatorial Optimization (IPCO 2013)*, pages 169–180. Springer, 2013.

[16] A. L. Fradkov and V. A. Yakubovich. The S-procedure and duality relations in nonconvex problems of quadratic programming. *Vestnik Leningrad Univ. Math.*, 6:101–109, 1979.

[17] A. Frangioni and C. Gentile. Perspective cuts for a class of convex 0–1 mixed integer programs. *Math. Program.*, 106:225–236, 2006.

[18] O. Günlük and J. Linderoth. Perspective reformulations of mixed integer nonlinear programs with indicator variables. *Math. Program.*, 124:183–205, 2010.

[19] N. Ho-Nguyen and F. Kılınç-Karzan. A second-order cone based approach for solving the Trust Region Subproblem and its variants. *SIAM J. Optim.*, 27(3):1485–1512, 2017.

[20] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

[21] M. Laurent and S. Poljak. On a positive semidefinite relaxation of the cut polytope. *Linear Algebra Appl.*, 223-224:439–461, 1995.

[22] M. Locatelli. Some results for quadratic problems with one or two quadratic constraints. *Oper. Res. Lett.*, 43(2):126–131, 2015.

[23] D. G. Mixon, S. Villar, and R. Ward. Clustering subgaussian mixtures by semidefinite programming. Technical report, 2016.

[24] S. Modaresi and J. P. Vielma. Convex hull of two quadratic or a conic quadratic and a quadratic inequality. *Math. Program.*, 164:383–409, 2017.

[25] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358, 1998.

[26] G. Pataki. The geometry of semidefinite programming. In *Handbook of semidefinite programming*, pages 29–65. Springer, 2000.

[27] G. Pataki. On the connection of facially exposed and nice cones. *Journal of Mathematical Analysis and Applications*, 400(1):211–221, 2013.

[28] M. Ramana and A. J. Goldman. Some geometric results in semidefinite programming. *J. Global Optim.*, 7(1):33–50, 1995.

[29] N. Rujeerapaiboon, K. Schindler, D. Kuhn, and W. Wiesemann. Size matters: Cardinality-constrained clustering and outlier detection via conic optimization. *SIAM J. Optim.*, 29(2):1211–1239.

[30] N. Z. Shor. Dual quadratic estimates in polynomial and boolean programming. *Ann. Oper. Res.*, 25: 163–168, 1990.

[31] M. Tawarmalani and N. Sahinidis. *Convexification and global optimization in continuous and mixed-integer nonlinear programming: Theory, algorithms, software, and applications*, volume 65 of *Nonconvex Optim. Appl.* Springer US, 2002.

[32] A. L. Wang and F. Kılınç-Karzan. On convex hulls of epigraphs of QCQPs. In *Integer Programming and Combinatorial Optimization (IPCO 2020)*, pages 419–432. Springer, 2020.

[33] A. L. Wang and F. Kılınç-Karzan. The generalized trust region subproblem: solution complexity and convex hull results. *Math. Program.*, 2020.

[34] A. L. Wang and F. Kılınç-Karzan. On the tightness of SDP relaxations of QCQPs. *Math. Program.*, 2020.

[35] L. Wei, A. Gómez, and S. Küçükyavuz. On the convexification of constrained quadratic optimization problems with indicator variables. In *Integer Programming and Combinatorial Optimization (IPCO 2020)*, pages 433–447. Springer, 2020.

[36] L. Wei, A. Gómez, and S. Küçükyavuz. Ideal formulations for constrained convex optimization problems with indicator variables. Technical Report arXiv:2007.00107, 2020.

[37] Y. Ye and S. Zhang. New results on quadratic minimization. *SIAM J. Optim.*, 14(1):245–267, 2003.

[38] U. Yıldıran. Convex hull of two quadratic constraints is an LMI set. *IMA J. Math. Control Inform.*, 26 (4):417–450, 2009.

A Supplementary material for Section 2

A.1 Proof of Lemma 1

Lemma 1. *Suppose Assumption 1 holds. Then*

$$\overline{\mathcal{S}} = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \exists\, X \succeq x^{(1)}x^{(1)\top} : \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \leq 0,\ \forall i \in [m_I] \\ \langle A_i, X\rangle + 2b_i^\top x + c_i = 0,\ \forall i \in [m_I+1, m] \end{array} \right\}. \tag{2}$$

Proof. Fix $x \in \mathbb{R}^n$. Note that

$$\begin{aligned} \sup_{\gamma\in\Gamma} \langle \gamma, q(x)\rangle &= \sup_{\gamma\in\mathbb{R}^m} \left\{ \langle \gamma, q(x)\rangle : \begin{array}{l} A(\gamma) \succeq 0 \\ \gamma_i \geq 0,\ \forall i \in [m_I] \end{array} \right\} \\ &= \inf_{\xi \in \mathbb{S}^{n_1}} \left\{ 0 : \begin{array}{l} q_i(x) + \langle A_i, \xi\rangle \leq 0,\ \forall i \in [m_I] \\ q_i(x) + \langle A_i, \xi\rangle = 0,\ \forall i \in [m_I+1, m] \\ \xi \succeq 0 \end{array} \right\}, \end{aligned}$$

where the second equation follows from the strong conic duality theorem and Assumption 1. Taking $X := x^{(1)}x^{(1)\top} + \xi$, we conclude that $x \in \overline{\mathcal{S}}$ if and only if x lies in the right hand side of (2). ∎

A.2 Computation for Example 1

We compute

$$\begin{aligned}\Gamma &= \left\{\gamma\in\mathbb{R}^3:\ \begin{array}{l}\begin{pmatrix}\gamma_2+\gamma_3 & \gamma_1\\ \gamma_1 & \gamma_3-\gamma_2\end{pmatrix}\succeq 0\\ \gamma\geq 0\end{array}\right\} = \left\{\gamma\in\mathbb{R}^3:\ \begin{array}{l}\gamma_3^2\geq\gamma_1^2+\gamma_2^2\\ \gamma\geq 0\end{array}\right\}\\
&= \left\{\gamma\in\mathbb{R}^3:\ \begin{array}{l}\gamma_3\geq\sqrt{\gamma_1^2+\gamma_2^2}\\ \gamma\geq 0\end{array}\right\},\\
\Gamma^\circ &= \operatorname{cl}\left(\left\{\gamma\in\mathbb{R}^3:\ \gamma_3\geq\sqrt{\gamma_1^2+\gamma_2^2}\right\}^\circ+\left\{\gamma\in\mathbb{R}^3:\ \gamma\geq 0\right\}^\circ\right)\\
&= \left\{\ell\in\mathbb{R}^3:\ -\ell_3\geq\sqrt{\ell_1^2+\ell_2^2}\right\}+\mathbb{R}^3_-\\
&= \left\{\ell\in\mathbb{R}^3:\ -\ell_3\geq\sqrt{(\ell_1)_+^2+(\ell_2)_+^2}\right\},\ \text{and}\\
\overline{\mathcal{S}} &= \left\{x\in\mathbb{R}^2:\ -q_3(x)\geq\sqrt{q_1(x)_+^2+q_2(x)_+^2}\right\}.\end{aligned}$$

B Supplementary material for Section 3

B.1 Proof of Lemma 2

We prove Lemma 2 by establishing a number of smaller lemmas.

Lemma 5. *Suppose Assumption 1 holds. Then*

$$\Gamma^\circ = \left\{(\langle A_i,\xi\rangle)_i:\ \xi\in\mathbb{S}^{n_1}_-\right\}+\left(\mathbb{R}^{m_I}_-\times\{0\}\right).$$

Proof. Let $\ell\in\mathbb{R}^m$. By Assumption 1 and strong conic duality,

$$\begin{aligned}&\sup_{\gamma\in\mathbb{R}^m}\left\{\langle\ell,\gamma\rangle:\ \begin{array}{l}A(\gamma)\succeq 0\\ \gamma_i\geq 0,\ \forall i\in[m_I]\end{array}\right\}\\
&\qquad=\inf_{\xi\in\mathbb{S}^{n_1},\,\eta\in\mathbb{R}^{m_I}}\left\{0:\ \begin{array}{l}\ell_i=\langle A_i,\xi\rangle+\eta_i,\ \forall i\in[m_I]\\ \ell_i=\langle A_i,\xi\rangle,\ \forall i\in[m_I+1,m]\\ \xi\preceq 0,\\ \eta\leq 0\end{array}\right\}.\end{aligned}$$

■

Lemma 6. *Suppose Assumptions 1 and 2 hold. Let $x\in\overline{\mathcal{S}}$ and let $f\in\operatorname{rint}(\mathcal{F}(x))$. Then*

$$\mathcal{G}(x)=\left\{(\langle A_i,\xi\rangle)_i:\ \begin{array}{l}\xi\in\mathbb{S}^{n_1}_-\\ \operatorname{range}(\xi)\subseteq\ker(A(f))\end{array}\right\}+\left\{\eta\times 0:\ \begin{array}{l}\eta\in\mathbb{R}^{m_I}_-\\ \operatorname{supp}(\eta)\cap\operatorname{supp}(f)=\varnothing\end{array}\right\}.$$

Proof. By Assumption 2, $\mathcal{G}(x)=\mathcal{F}(x)^\triangle=\Gamma^\circ\cap f^\perp$. Then by Lemma 5, we have that

$$\begin{aligned}\mathcal{G}(x)&=\Gamma^\circ\cap f^\perp\\
&=\left(\left\{(\langle A_i,\xi\rangle)_i:\ \xi\in\mathbb{S}^{n_1}_-\right\}\cap f^\perp\right)+\left(\left(\mathbb{R}^{m_I}_-\times\{0\}\right)\cap f^\perp\right)\\
&=\left\{(\langle A_i,\xi\rangle)_i:\ \begin{array}{l}\xi\in\mathbb{S}^{n_1}_-\\ \langle A(f),\xi\rangle=0\end{array}\right\}+\left\{\eta\times 0:\ \begin{array}{l}\eta\in\mathbb{R}^{m_I}_-\\ \langle f,\eta\times 0\rangle=0\end{array}\right\}\\
&=\left\{(\langle A_i,\xi\rangle)_i:\ \begin{array}{l}\xi\in\mathbb{S}^{n_1}_-\\ \operatorname{range}(\xi)\subseteq\ker(A(f))\end{array}\right\}+\left\{\eta\times 0:\ \begin{array}{l}\eta\in\mathbb{R}^{m_I}_-\\ \operatorname{supp}(\eta)\cap\operatorname{supp}(f)=\varnothing\end{array}\right\}.\end{aligned}$$

Here, the last equality follows from the fact that $\Gamma^\circ\supseteq\mathbb{R}^{m_I}_-\times\{0\}$ so that $f_i\geq 0$ for all $i\in[m_I]$. ■

Corollary 2. *Suppose Assumptions 1 and 2 hold. Let $x \in \overline{\mathcal{S}}$ and $f \in \operatorname{rint}(\mathcal{F}(x))$. Then*

$$\mathcal{G}(x)^{\perp} = \left\{ \gamma : \begin{array}{l} A(\gamma) \mid_{\ker(A(f))} = 0 \\ \gamma_i = 0, \, \forall i \in [m_I] \cap \operatorname{zeros}(f) \end{array} \right\}.$$

Lemma 7. *Suppose Assumptions 1 and 2 hold. Let $x \in \overline{\mathcal{S}}$ and $f \in \operatorname{rint}(\mathcal{F}(x))$. Then*

$$\mathcal{R}(x) = \left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \left\langle A(\gamma)x^{(1)}, y^{(1)} \right\rangle + \langle b(\gamma), y \rangle = 0, \, \forall \gamma \in \mathcal{G}(x)^{\perp} \end{array} \right\}.$$

Proof. It suffices to show that $y^{(1)} \in \ker(A(f))$ if and only if $y^{(1)\top} A(\gamma) y^{(1)} = 0$ for all $\gamma \in \mathcal{G}(x)^{\perp}$.

($\Rightarrow$) By Corollary 2, we have that $A(\gamma) \mid_{\ker(A(f))} = 0$ for all $\gamma \in \mathcal{G}(x)^{\perp}$. Then clearly $y^{(1)\top} A(\gamma) y^{(1)} = 0$ for all $y^{(1)} \in \ker(A(f))$.

($\Leftarrow$) From $f \in \operatorname{rint}(\mathcal{F}(x))$ and the definition of $\mathcal{F}(x)$, we deduce that $f \in \mathcal{G}(x)^{\perp}$. Moreover, from $f \in \mathcal{F}(x)$ and $\mathcal{F}(x) \subseteq \Gamma$, we conclude $A(f) \succeq 0$. By definition of $\mathcal{R}(x)$, we have that $y^{(1)\top} A(f) y^{(1)} = 0$ holds for all $y \in \mathcal{R}(x)$. This together with $A(f) \succeq 0$ implies that $y^{(1)} \in \ker(A(f))$. ■

B.2 α^+ and α^- are well-defined

Lemma 8. *Suppose Assumption 1 holds. Let $x \in \overline{\mathcal{S}}$ and consider a bounded rounding direction $y \in \mathcal{R}(x)$. Then α^+ and α^- are well-defined.*

Proof. As $\overline{\mathcal{S}}$ is closed, the set $\left\{\alpha \in \mathbb{R} : x + \alpha y \in \overline{\mathcal{S}}\right\}$ is also closed. It suffices to show that this set is bounded: By Assumption 1, there exists $\gamma^* \in \Gamma$ such that $A(\gamma^*) \succ 0$. Then

$$\left\{\alpha \in \mathbb{R} : x + \alpha y \in \overline{\mathcal{S}}\right\} \subseteq \left\{\alpha \in \mathbb{R} : q(\gamma^*, x + \alpha y) \leq 0\right\}.$$

Noting that $A(\gamma^*) \succ 0$ and $y^{(1)} \neq 0$, and examining the expression of $q(\gamma^*, x + \alpha y)$ as a function of α, we deduce that $\alpha \mapsto q(\gamma^*, x + \alpha y)$ is a strongly convex quadratic function in α. Thus, the set on the right is bounded. ■

B.3 Finite termination

We will use the quantity $\dim(\mathcal{G}(x))$ to measure the progress of Procedure 1.

Lemma 9. *Suppose Assumption 1 holds. Let $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$ and let $y \in \mathcal{R}(x)$ be a bounded rounding direction. Then $\dim(\mathcal{G}(x + \alpha^+ y)) < \dim(\mathcal{G}(x))$. Similarly, $\dim(\mathcal{G}(x - \alpha^- y)) < \dim(\mathcal{G}(x))$.*

Proof. First, recall that as Γ° is conic, the only face of dimension 0 is the trivial face $\{0\}$. Moreover, $x \notin \mathcal{S}$ implies that $q(x) \neq 0$, and thus $\dim(\mathcal{G}(x)) \geq 1$.

As $y \in \mathcal{R}(x)$, we have that by definition $q(x + \alpha y) \in \operatorname{span}(\mathcal{G}(x))$ for all $\alpha \in \mathbb{R}$. In particular, $x + \alpha y \in \overline{\mathcal{S}}$ if and only if $q(x + \alpha y) \in \mathcal{G}(x)$. By our choice of α^+, we have that $q(x + \alpha y) \in \operatorname{rbd}(\mathcal{G}(x))$. Finally, as $\mathcal{G}(x)$ has dimension at least one, $\operatorname{rint}(\mathcal{G}(x))$ and $\operatorname{rbd}(\mathcal{G}(x))$ are disjoint. We conclude that $\mathcal{G}(x + \alpha^+ y)$ is a proper face of $\mathcal{G}(x)$, and thus $\dim(\mathcal{G}(x + \alpha^+ y)) < \dim \mathcal{G}((x))$. The same reasoning holds for $x - \alpha^- y$ as well. ■

B.4 Completeness

Lemma 10. *Suppose Assumption 2 holds. Let $x \in \overline{\mathcal{S}}$ and suppose $[x \pm y] \subseteq \overline{\mathcal{S}}$. Then $y \in \mathcal{R}(x)$.*

Proof. Note that for any $\ell \in \mathbb{R}^m$, the map $\alpha \mapsto \langle \ell, q(x + \alpha y) \rangle$ is a univariate quadratic function in α. Consequently, $\alpha \mapsto \langle \ell, q(x + \alpha y) \rangle$ is identically zero if and only if it is zero on some open interval.

In particular, it suffices to show that $q(x + \alpha y) \in \mathcal{G}(x)$ for all $\alpha \in (-1, 1)$.

Let $f \in \operatorname{rint}(\mathcal{F}(x))$ so that by Assumption 2 and Fact 2, we have $\mathcal{G}(x) = \Gamma^\circ \cap f^\perp$. As $f \in \Gamma$, we have that $\langle f, q(x+\alpha y)\rangle$ is a convex quadratic function in α. Furthermore, as $[x \pm y] \subseteq \overline{\mathcal{S}}$, we have that $\langle f, q(x+\alpha y)\rangle \leq 0$ for all $\alpha \in (-1,1)$. Noting that $f \in \mathcal{F}(x) = \Gamma \cap q(x)^\perp$, we have that $\langle f, q(x)\rangle = 0$, i.e., the convex quadratic map $\alpha \mapsto \langle f, q(x+\alpha y)\rangle$ maps zero to zero. We deduce that $\langle f, q(x+\alpha y)\rangle = 0$ for all $\alpha \in (-1,1)$. We conclude that $q(x+\alpha y) \in \Gamma^\circ \cap f^\perp = \mathcal{G}(x)$ for all $\alpha \in (-1,1)$. ∎

Corollary 3. *Suppose Assumption 2 holds. Let $x \in \overline{\mathcal{S}}$. Then $y \in \mathcal{R}(x)$ if and only if for all $\epsilon > 0$ small enough, the interval $[x \pm \epsilon y]$ is contained in $\overline{\mathcal{S}}$.*

The following lemma states that when $x \in \operatorname{conv}(\mathcal{S}) \setminus \mathcal{S}$, there exists an interval $[x \pm y] \subseteq \operatorname{conv}(\mathcal{S})$ where $y^{(1)} \neq 0$. Along with Lemma 10, this lemma will allow us to conclude the completeness of our rounding procedure under Assumption 2.

Lemma 11. *Let $x \in \operatorname{conv}(\mathcal{S}) \setminus \mathcal{S}$. Then there exists $y \in \mathbb{R}^n$ with $y^{(1)} \neq 0$ such that $[x \pm y] \subseteq \operatorname{conv}(\mathcal{S})$.*

Proof. Suppose $x \in \operatorname{conv}(\mathcal{S}) \setminus \mathcal{S}$. By Carathéodory's Theorem, there exists a convex decomposition $x = \sum_{j\in[k]} \lambda_j x_j$ where $\lambda_j \in (0,1)$ are convex combination weights and $x_j \in \mathcal{S}$ for all $j \in [k]$.

We claim there exists $j^* \in [k]$ such that $x_{j^*}^{(1)} \neq x^{(1)}$: Indeed, supposing otherwise, we have

$$q_i(x) = x^{(1)\top} A_i x^{(1)} + 2b_i^\top x + c_i = \sum_j \lambda_j \left(x_j^{(1)\top} A_i x_j^{(1)} + 2b_i^\top x_j + c_i \right) = \sum_j \lambda_j q_i(x_j),$$

for all $i \in [m]$. Here, the second equality follows from the assumption that $x_j^{(1)} = x^{(1)}$ for all $j \in [k]$. In particular, $q_i(x) \leq 0$ for all $i \in [m_I]$ and $q_i(x) = 0$ for all $i \in [m_I + 1, m]$, contradicting the assumption that $x \in \operatorname{conv}(\mathcal{S}) \setminus \mathcal{S}$.

Fix such a $j^* \in [k]$ and define $y = x_{j^*} - x$. By construction, $y^{(1)} \neq 0$. Additionally,

$$\left[x - \frac{\lambda_{j^*}}{1-\lambda_{j^*}} y, x + y \right] = \left[\frac{\sum_{j\neq j^*} \lambda_j x_j}{\sum_{j \neq j^*} \lambda_j}, x_{j^*} \right] \subseteq \operatorname{conv}(\mathcal{S}).$$ ∎

Corollary 4. *Suppose Assumption 2 holds. If $x \in \operatorname{conv}(\mathcal{S}) \setminus \mathcal{S}$, then $\mathcal{R}(x)$ contains a bounded rounding direction.*

B.5 Extension for linear optimization over quadratically constrained sets

In this section, we consider a minor modification of Procedure 1. Consider the following linear optimization problem over a quadratically constrained set and its SDP relaxation

$$\inf_{x\in\mathbb{R}^n} \left\{ w^\top x : x \in \mathcal{S} \right\} \geq \inf_{x\in\mathbb{R}^n} \left\{ w^\top x : x \in \overline{\mathcal{S}} \right\}.$$

We say that *objective value exactness* holds if equality holds in the above relation. The following procedure takes a point $x \in \overline{\mathcal{S}}$ and either outputs $x' \in \mathcal{S}$ such that $w^\top x' \leq w^\top x$ or claims to find a point $x'' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$.

Procedure 2. *Suppose Assumption 1 holds. Let $x \in \overline{\mathcal{S}}$ and $w \in \mathbb{R}^n$.*

topsep=0pt,pirsep=0pt,itemsep=0pt If $x \in \mathcal{S}$, output x.

topsep=0pt,piirsep=0pt,iitemsep=0pt Else, if $\mathcal{R}(x)$ contains a bounded rounding direction y:

- *(a) Define α^+ and α^- (see Definition 7).*
- *(b) Set $x' = x + \alpha^+ y$ or $x' = x - \alpha^- y$ so that $w^\top x' \leq w^\top x$*
- *(c) Recursively run this procedure on x' and w and return its output.*

topsep=0pt,piiirsep=0pt,iiitemsep=0pt Else, output "Failed at $x \in \overline{\mathcal{S}}$" and stop.

Note that as Procedure 1 makes two recursive calls at each step, it may need to compute $\mathcal{R}(x)$ for exponentially many (in m) choices of x. In contrast, the procedure above makes only one recursive call at each step and thus will terminate after computing $\mathcal{R}(x)$ for at most linearly many (in m) choices of x.

The following theorem follows as a corollary to Lemma 9 and Corollary 4.

Theorem 5. *Suppose Assumption 1 holds. Then given $x \in \overline{\mathcal{S}}$ and $w \in \mathbb{R}^n$, Procedure 2 will terminate in a finite number of steps and either output $x' \in \overline{\mathcal{S}}$ with $w^\top x' \leq w^\top x$ or fail at some (possibly different) point $x'' \in \overline{\mathcal{S}}$. If Procedure 2 fails and Assumption 2 holds, then $x'' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$.*

C Supplementary material for Section 4

C.1 Proof of Lemma 3

Lemma 3. *Suppose Assumption 1 holds. Let $A_i = I_k \otimes \mathbb{A}_i$ for each $i \in [m]$ and assume $k \geq m+1$. Then, for all $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$, the set $\mathcal{R}(x)$ contains a bounded rounding direction. In particular, $\operatorname{conv}(\mathcal{S}) = \overline{\mathcal{S}}$.*

Proof. For notational convenience, let $\mathbb{A}(\gamma) \coloneqq \sum_{i=1}^m \gamma_i \mathbb{A}_i$.

Fix $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$. We will construct the desired bounded rounding direction $y \in \mathcal{R}(x)$. Note that

$$\begin{aligned} 0 = \sup_{\gamma \in \Gamma} \langle \gamma, q(x) \rangle &= \sup_{\gamma \in \mathbb{R}^m} \left\{ \langle \gamma, q(x) \rangle : \begin{array}{l} A(\gamma) \succeq 0 \\ \gamma_i \geq 0,\ \forall i \in [m_I] \end{array} \right\} \\ &= \inf_{\xi \in \mathbb{S}^{n_1/k}} \left\{ 0 : \begin{array}{l} q_i(x) + \langle \mathbb{A}_i, \xi \rangle \leq 0,\ \forall i \in [m_I] \\ q_i(x) + \langle \mathbb{A}_i, \xi \rangle = 0,\ \forall i \in [m_{I+1}, m] \\ \xi \succeq 0 \end{array} \right\}. \end{aligned}$$

Then, by Assumption 1 and strong conic duality theorem, there exists $\xi \in \mathbb{S}^{n_1/k}$ feasible in the final program. As $x \notin \mathcal{S}$, we have $\xi \neq 0$. Let $z \in \operatorname{range}(\xi)$ be a nonzero vector such that $\xi \succeq zz^\top$. Consider $\ell \in \mathbb{R}^m$, where $\ell_i \coloneqq \left\langle \mathbb{A}_i, zz^\top \right\rangle$.

Then for any $\gamma \in \Gamma$, since $A(\gamma) \succeq 0$ we have

$$\langle \gamma, q(x) + \ell \rangle \leq \sum_{i=1}^m \gamma_i \left(q_i(x) + \langle \mathbb{A}_i, \xi \rangle \right) b \leq 0.$$

Similarly, for any $\gamma \in \Gamma$,

$$-\langle \gamma, \ell \rangle = -\left\langle \mathbb{A}(\gamma), zz^\top \right\rangle \leq 0.$$

We deduce that both $q(x) + \ell, -\ell \in \Gamma^\circ$, whence using Fact 1 we conclude that both are also in $\mathcal{G}(x)$. Then $\operatorname{span}\{q(x), \ell\} \subseteq \operatorname{span}(\mathcal{G}(x))$.

Finally, set $y = (y^{(1)}, y^{(2)}) \coloneqq (w \otimes z, 0)$ for some $w \in \mathbf{S}^{k-1}$ to be chosen later. Note that, for all $i \in [m]$,

$$\begin{aligned} q_i(x + \alpha y) &= \alpha^2 y^{(1)\top} A_i y^{(1)} + 2\alpha \left\langle A_i x^{(1)} + b_i^{(1)}, y^{(1)} \right\rangle + q_i(x) \\ &= \left(q_i(x) + \alpha^2 \ell_i \right) + 2\alpha \left\langle A_i x^{(1)} + b_i^{(1)}, w \otimes z \right\rangle. \end{aligned}$$

We now fix w to be a solution of the following system (which is feasible as $k \geq m+1$):

$$\begin{cases} \left\langle A_i x^{(1)} + b_i^{(1)}, w \otimes z \right\rangle = 0,\ \forall i \in [m] \\ w \in \mathbf{S}^{k-1} \end{cases}.$$

Consequently, we have constructed y such that $q(x + \alpha y) = q(x) + \alpha^2 \ell \in \operatorname{span}(\mathcal{G}(x))$ for all $\alpha \in \mathbb{R}$. Furthermore, $y^{(1)} = w \otimes z \neq 0$ as desired. ■

C.2 Proof of Lemma 4

Lemma 4. *Suppose Assumption 3 holds. Then,*

$$\Gamma = \left\{(\mu,\tau)\in\mathbb{R}\times\mathbb{R}^n:\ \begin{array}{l}\mu aa^\top + \operatorname{diag}(\tau)\succeq 0\\ \mu\geq 0\end{array}\right\},\ \textit{and}$$

$$\overline{\mathcal{S}} = \left\{(x,t)\in[-1,1]^n\times\mathbb{R}:\ (a^\top x)^2 + \max_{i\in[n]}\left(a_i\sqrt{1-x_i^2} - \sum_{j\neq i} a_j\sqrt{1-x_j^2}\right)_+^2 \leq t\right\}.$$

The first identity follows immediately from the definition of Γ. The remainder of this section will be devoted to proving the second more difficult identity.

Lemma 12. *Let $\mathcal{T} := \left\{\tau\in\mathbb{R}^n:\ aa^\top + \operatorname{Diag}(\tau)\succeq 0\right\}$. Then, $\Gamma = \operatorname{clcone}(1\times\mathcal{T})$. In particular, $(x,t)\in\overline{\mathcal{S}}$ if and only if*

$$\sup_{\tau\in\mathcal{T}}\left\langle\begin{pmatrix}1\\ \tau\end{pmatrix}, q(x,t)\right\rangle \leq 0.$$

Proof. The forward direction is immediate. In the other direction, let $\gamma\in\Gamma$. Either $\gamma_1>0$ or $\gamma_1=0$. In the first case, we have that $(\gamma_2,\dots,\gamma_{n+1})/\gamma_1\in\mathcal{T}$, whence $\gamma\in\operatorname{cone}(\mathcal{T})$. In the second case, we have that $(\gamma_2,\dots,\gamma_{n+1})\geq 0$. Note that $\tau = (\gamma_2,\dots,\gamma_{n+1})/\epsilon\in\mathcal{T}$ for all $\epsilon>0$. Then, $(\epsilon,\gamma_2,\dots,\gamma_{n+1})\in\operatorname{cone}(\mathcal{T})$ and $\gamma = (0,\gamma_2,\dots,\gamma_{n+1})\in\operatorname{clcone}(\mathcal{T})$. ■

The following lemma allows us to decompose $\mathcal{T}$ as the union of n-many sets $\mathcal{N}_i := \mathcal{T}\cap\{\tau\in\mathbb{R}^n:\ \tau_i<0\}$ for $i\in[n]$ and $\mathcal{T}\cap\mathbb{R}^n_+ = \mathbb{R}^n_+$.

Lemma 13. *Suppose Assumption 3 holds. Then, for $i\in[n]$*

$$\mathcal{N}_i = \left\{\tau\in\mathbb{R}^n:\ \begin{array}{l}\tau_j>0,\ \forall j\neq i\\ 0<-\tau_i\leq\frac{a_i^2}{1+\sum_{j\neq i}a_j^2/\tau_j}\end{array}\right\}.$$

Proof. We begin by observing that

$$\left\{\tau\in\mathbb{R}^n:\ \begin{array}{l}\tau_i<0\\ aa^\top+\operatorname{Diag}(\tau)\succeq 0\end{array}\right\} = \left\{\tau\in\mathbb{R}^n:\ \begin{array}{l}\tau_i<0\\ \tau_j>0,\ \forall j\neq i\\ aa^\top+\operatorname{Diag}(\tau)\succeq 0\end{array}\right\}.$$

One containment is immediate. In the other direction, given τ on the left and $j\neq i$, we can consider the 2×2 principal submatrix of $aa^\top+\operatorname{Diag}(\tau)$ corresponding to the (i,j) indices. We have that

$$\begin{pmatrix}a_i^2+\tau_i & a_ia_j\\ a_ia_j & a_j^2+\tau_j\end{pmatrix}\succeq 0.$$

Then as $\tau_i<0$, we have that $\tau_j>0$ (here, we have used the assumption that $a\in\mathbb{R}^n_{++}$).

Now fix τ with $\tau_i<0$ and $\tau_j>0$ for all $j\neq i$. By the Schur complement Lemma,

$$\begin{aligned}
aa^\top+\operatorname{Diag}(\tau)\succeq 0 &\iff aa^\top+\operatorname{Diag}(\tau-(\tau_i-1)e_i)-(1-\tau_i)e_ie_i^\top\succeq 0\\
&\iff \begin{pmatrix}\frac{1}{1-\tau_i} & e_i^\top\\ e_i & aa^\top+\operatorname{Diag}(\tau-(\tau_i-1)e_i)\end{pmatrix}\succeq 0\\
&\iff \frac{1}{1-\tau_i} - e_i^\top\left[aa^\top+\operatorname{Diag}(\tau-(\tau_i-1)e_i)\right]^{-1}e_i\geq 0
\end{aligned}$$

Finally, by the Sherman-Morrison-Woodbury formula, the expression on the left of the final line is equal to

$$\frac{1}{1-\tau_i} - \left(1 - \frac{a_i^2}{1 + a_i^2 + \sum_{j\neq i} a_j^2/\tau_j}\right) = \frac{1}{1-\tau_i} - \frac{1 + \sum_{j\neq i} a_j^2/\tau_j}{1 + \sum_{j\neq i} a_j^2/\tau_j + a_i^2}.$$

We deduce that $\tau \in \mathcal{N}_i$ if and only if the right hand side term is nonnegative. Equivalently, if and only if

$$-\tau_i \le \frac{1 + \sum_{j\neq i} a_j^2/\tau_j + a_i^2}{1 + \sum_{j\neq i} a_j^2/\tau_j} - 1 = \frac{a_i^2}{1 + \sum_{j\neq i} a_j^2/\tau_j}. \qquad \blacksquare$$

Define the following convex functions (taking values in $\mathbb{R} \cup \{+\infty\}$):

$$\begin{aligned} f(x) &\coloneqq \inf\{t : (x,t) \in \mathrm{conv}(\mathcal{S})\}, \\ \bar{f}_0(x) &\coloneqq \sup_{\tau \in \mathbb{R}^n_+} x^\top(aa^\top + \mathrm{Diag}(\tau))x - 1^\top \tau, \\ \bar{f}_i(x) &\coloneqq \sup_{\tau \in \mathcal{N}_i} x^\top(aa^\top + \mathrm{Diag}(\tau))x - 1^\top \tau, \ \forall i \in [n], \\ \bar{f}(x) &\coloneqq \sup_{\tau \in \mathcal{T}} x^\top(aa^\top + \mathrm{Diag}(\tau))x - 1^\top \tau. \end{aligned}$$

Lemma 14. *We have* $\mathrm{conv}(\mathcal{S}) = \mathrm{epi}(f)$*, and* $\overline{\mathcal{S}} = \mathrm{epi}\left(\bar{f}\right) = \bigcap_{i\in[0,n]} \mathrm{epi}\left(\bar{f}_i\right)$*. In particular,* $\mathrm{conv}(\mathcal{S}) = \overline{\mathcal{S}}$ *if and only if* $f = \bar{f}$*.*

Proof. This follows from Lemma 12, the decomposition $\mathcal{T} = \mathbb{R}^n_+ \cup \bigcup_{i\in[n]} \mathcal{N}_i$ and the observation

$$\left\langle \begin{pmatrix} 1 \\ \tau \end{pmatrix}, q(x,t) \right\rangle = x^\top(aa^\top + \mathrm{Diag}(\tau))x - 1^\top\tau - t. \qquad \blacksquare$$

Lemma 15. $\mathrm{epi}(\bar{f}_0) = \left\{ (x,t) : \begin{array}{l} t \ge (a^\top x)^2 \\ x \in [-1,1]^n \end{array} \right\}$.

Proof. Let $x \in \mathbb{R}^n$. Then

$$\begin{aligned} \sup_{\tau\in\mathbb{R}^n_+} x^\top(aa^\top + \mathrm{Diag}(\tau))x - 1^\top\tau &= x^\top aa^\top x + \sup_{\tau\in\mathbb{R}^n_+} \left(\sum_{i=1}^n \tau_i(x_i^2 - 1) \right) \\ &= \begin{cases} (a^\top x)^2 & \text{if } x \in [-1,1] \\ +\infty & \text{else.} \end{cases} \end{aligned} \qquad \blacksquare$$

Lemma 16. *Suppose Assumption 3 holds. Let* $x \in [-1,1]^n$ *and* $i \in [n]$*. Then*

$$\bar{f}_i(x) = (a^\top x)^2 + \left(a_i \sqrt{1 - x_i^2} - \sum_{j\neq i} a_j \sqrt{1 - x_j^2} \right)_+^2.$$

Proof. Let $\mathcal{J} := \{j \in [n] \setminus \{i\} : x_j^2 < 1\}$. By Lemma 13,

$$\begin{aligned}
\bar{f}_i(x) - (a^\top x)^2 &= \sup_\tau \left\{ x^\top \operatorname{Diag}(\tau) x - 1^\top \tau : \begin{array}{l} \tau_j > 0, \forall j \neq i \\ 0 < -\tau_i \leq \frac{a_i^2}{1+\sum_{j\neq i} a_j^2/\tau_j} \end{array} \right\} \\
&= \sup_\tau \left\{ -\tau_i(1-x_i^2) - \sum_{j\neq i} \tau_j(1-x_j^2) : \begin{array}{l} \tau_j > 0, \forall j \neq i \\ 0 < -\tau_i \leq \frac{a_i^2}{1+\sum_{j\neq i} a_j^2/\tau_j} \end{array} \right\} \\
&= \sup_\tau \left\{ \frac{a_i^2(1-x_i^2)}{1+\sum_{j\neq i} a_j^2/\tau_j} - \sum_{j\neq i} \tau_j(1-x_j^2) : \tau_j > 0, \forall j \neq i \right\} \\
&= \sup_\tau \left\{ \frac{a_i^2(1-x_i^2)}{1+\sum_{j\in\mathcal{J}} a_j^2/\tau_j} - \sum_{j\in\mathcal{J}} \tau_j(1-x_j^2) : \tau_j > 0, \forall j \in \mathcal{J} \right\}
\end{aligned}$$

Note that if $x_i^2 = 1$, then $\bar{f}_i(x) - (a^\top x)^2 = 0$. Similarly, if $\mathcal{J} = \varnothing$, then $\bar{f}_i(x) - (a^\top x)^2 = a_i^2(1-x_i^2)$. In the remainder, we assume that $x_i^2 < 1$ and $\mathcal{J} \neq \varnothing$. By Cauchy-Schwarz, we have

$$\sum_{j\in\mathcal{J}} \frac{a_j^2}{\tau_j} \sum_{j\in\mathcal{J}} \tau_j(1-x_j^2) \geq \left(\sum_{j\in\mathcal{J}} a_j \sqrt{1-x_j^2} \right)^2,$$

where equality holds for $\tau_j \propto \frac{a_j}{\sqrt{(1-x_j^2)}}$. We deduce that

$$\begin{aligned}
\bar{f}_i(x) - (a^\top x)^2 &= \sup_\tau \left\{ \frac{a_i^2(1-x_i^2)}{1+\frac{\left(\sum_{j\in\mathcal{J}} a_j\sqrt{1-x_j^2}\right)^2}{\sum_{j\in\mathcal{J}} \tau_j(1-x_j^2)}} - \sum_{j\in\mathcal{J}} \tau_j(1-x_j^2) : \tau_j > 0, \ \forall j \in \mathcal{J} \right\} \\
&= \sup_\alpha \left\{ \frac{a_i^2(1-x_i^2)}{1+\frac{\left(\sum_{j\in\mathcal{J}} a_j\sqrt{1-x_j^2}\right)^2}{\alpha}} - \alpha : \alpha > 0 \right\}.
\end{aligned}$$

Differentiating this expression with respect to α and setting equal to zero, we see that

$$\bar{f}_i(x) - (a^\top x)^2 = \begin{cases} \left(a_i\sqrt{1-x_i^2} - \sum_{j\in\mathcal{J}} a_j \sqrt{1-x_j^2} \right)^2 & \text{if } a_i\sqrt{1-x_i^2} \geq \sum_{j\in\mathcal{J}} a_j\sqrt{1-x_j^2}, \\ 0 & \text{else.} \end{cases}$$ ∎

Lemma 4 then follows from Lemmas 15 and 16.

C.3 Proof of Corollary 1

Corollary 1. *Suppose Assumption 3 holds. Then,* $\mathrm{Opt}_{\mathrm{SDP}} = 0$ *if and only if* a *is balanced.*

Proof. The value of $\mathrm{Opt}_{\mathrm{SDP}}$ is the minimum of $\bar{f}(x)$. By homogeneity, $\bar{f}(x)$ is minimized at $x = 0$. By Lemma 4, we have

$$\mathrm{Opt}_{\mathrm{SDP}} = \bar{f}(0) = \min_{i\in[n]} \left(a_i - \sum_{j\neq i} a_j \right)^2.$$

This expression is zero if and only if a is balanced. ∎

C.4 Proof of Theorem 3

Theorem 3. *Suppose Assumption 3 holds. Then,* $\mathrm{conv}(\mathcal{S}) \neq \overline{\mathcal{S}}$.

Proof. Pick an open set $U \subseteq [-1,1]^n$ such that

$$a_1(1-x_1^2) > \sum_{j=2}^{n} a_j(1-x_j^2),\ \forall x \in U.$$

Thus by Lemma 4, we have

$$\bar{f}(x) = (a^\top x)^2 + \left(a_1\sqrt{1-x_1^2} - \sum_{j=2}^{n} a_j\sqrt{1-x_j^2}\right)^2,\ \forall x \in U.$$

Finally, a straightforward calculation gives

$$\frac{\partial^2}{\partial x_1^2}\bar{f}(x) = \frac{2a_1\sum_{j=2}^{n} a_j\sqrt{1-x_j^2}}{(1-x_1^2)^{3/2}}.$$

In particular, $\frac{\partial^2 \bar{f}(x)}{\partial x_1^2} > 0$ for all $x \in U$. Finally, noting that f is piecewise linear, we conclude $\bar{f} \neq f$ and $\overline{\mathcal{S}} \neq \mathrm{conv}(\mathcal{S})$. ■

C.5 Proof of Theorem 4

Theorem 4. *Suppose Assumption 1 holds and that* Γ *is polyhedral. Then* $\mathrm{conv}(\mathcal{S}) = \overline{\mathcal{S}}$ *if and only if*

$$\left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \langle b(\gamma), y\rangle = 0,\ \forall \gamma \in \mathcal{F} \end{array} \right\}$$

contains a bounded rounding direction for every $\mathcal{F} \trianglelefteq \Gamma$ *which is exposed by some vector* $q(x)$ *for* $x \in \overline{\mathcal{S}} \setminus \mathcal{S}$. *Here,* f *is any vector in* $\mathrm{rint}(\mathcal{F})$.

Proof. We begin by noting that when Γ is polyhedral, so too is Γ° so that Assumption 2 holds.

Next, we claim that for every nonempty face $\mathcal{G} \trianglelefteq \Gamma^\circ$ we have $\mathcal{G}^\perp = \mathrm{span}(\mathcal{G}^\triangle)$. By definition,

$$\mathrm{span}(\mathcal{G}^\triangle) = \mathrm{span}(\Gamma \cap \mathcal{G}^\perp) \subseteq \mathcal{G}^\perp.$$

On the other hand, as Γ and Γ° are polyhedral, we have that

$$\dim(\mathcal{G}) + \dim(\mathcal{G}^\triangle) = m.$$

Rearranging, we have $\dim(\mathcal{G}^\triangle) = m - \dim(\mathcal{G}) = \dim(\mathcal{G}^\perp)$. We conclude that $\mathcal{G}^\perp = \mathrm{span}(\mathcal{G}^\triangle)$.

Let $x \in \overline{\mathcal{S}}$ and $f \in \mathrm{rint}(\mathcal{F}(x))$. Then by Lemma 2,

$$\begin{aligned}
\mathcal{R}(x) &= \left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \left\langle A(\gamma)x^{(1)}, y^{(1)}\right\rangle + \langle b(\gamma), y\rangle = 0,\ \forall \gamma \in \mathcal{G}(x)^\perp \end{array} \right\} \\
&= \left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \left\langle A(\gamma)x^{(1)}, y^{(1)}\right\rangle + \langle b(\gamma), y\rangle = 0,\ \forall \gamma \in \mathcal{F}(x) \end{array} \right\} \\
&= \left\{ y \in \mathbb{R}^n : \begin{array}{l} y^{(1)} \in \ker(A(f)) \\ \langle b(\gamma), y\rangle = 0,\ \forall \gamma \in \mathcal{F}(x) \end{array} \right\}.
\end{aligned}$$

Here, the second line follows as $\mathcal{F}(x) = \mathcal{G}(x)^\triangle$ and $\mathcal{G}(x)^\perp = \mathrm{span}(\mathcal{G}(x)^\triangle)$. The third line follows from the fact that $\ker(A(f)) \subseteq \ker(A(\gamma))$ for every $\gamma \in \mathcal{F}(x)$ as $f \in \mathrm{rint}(\mathcal{F}(x))$. ■