A geometric view of SDP exactness in QCQPs and its applications

Alex L. Wang and Fatma Kilinc-Karzan

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

November 29, 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 QCQPs and their (projected) SDP relaxations. We present sufficient (and in some cases, also necessary) conditions for objective value exactness (the condition that the objective values of the QCQP and its SDP relaxation coincide) and convex hull exactness (the condition that the convex hull of the QCQP epigraph coincides with the epigraph of its SDP relaxation). Our conditions for exactness are based on geometric properties of \( \Gamma \), the cone of convex Lagrange multipliers, and its relatives \( \Gamma^P \) and \( \Gamma^\circ \). These tools form the basis of our main message: questions of exactness can be treated systematically whenever \( \Gamma \), \( \Gamma^P \), or \( \Gamma^\circ \) is well-understood. As further evidence of this message, we apply our tools to address questions of exactness for a prototypical QCQP involving a binary on-off constraint, quadratic matrix programs, the QCQP formulation of the partition problem, and random and semi-random QCQPs.

1 Introduction

Quadratically constrained quadratic programs (QCQPs) are a fundamental class of nonconvex optimization problems of the form

\[
\text{Opt} := \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : \begin{array}{l}
q_i(x) \leq 0, \forall i \in [m] \\
q_i(x) = 0, \forall i \in [m] + 1, m \end{array} \right\},
\]

where \( q_{\text{obj}}, q_1, \ldots, q_m : \mathbb{R}^n \to \mathbb{R} \) are each (possibly nonconvex) quadratic functions. For each \( i \in [m] \), we will write \( q_i(x) = x^\top A_i x + 2b_i^\top x + c_i \) for \( A_i \in \mathbb{S}^n \), \( b_i \in \mathbb{R}^n \), and \( c_i \in \mathbb{R} \). Similarly, write \( q_{\text{obj}}(x) = x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} \).

These optimization problems arise naturally in a variety of application areas (see [4, 9, 49]). Indeed, one fundamental reason for the ubiquity of QCQPs is their expressiveness—any polynomial optimization problem or \( \{0,1\} \)-integer program may be reformulated as a QCQP.

Although QCQPs are NP-hard in general, they admit a natural tractable convex relaxation known as the standard semidefinite program (SDP) relaxation [43],

\[
\text{Opt}_{\text{SDP}} := \inf_{x \in \mathbb{R}^n} \left\{ \langle A_{\text{obj}}, X \rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \begin{array}{l}
\exists X \succeq xx^\top : \\
\langle A_i, X \rangle + 2b_i^\top x + c_i \leq 0, \forall i \in [m] \\
\langle A_i, X \rangle + 2b_i^\top x + c_i = 0, \forall i \in [m] + 1, m \end{array} \right\}.
\]

This relaxation is also referred to as the Shor SDP relaxation. In contrast to the vast literature on the approximation quality of this relaxation [9, 34, 37, 50], the question of when exactness occurs in this relaxation is much more limited and recent.

One interesting line of work has offered deterministic conditions under which the SDP relaxation of a general QCQP is exact for various definitions of exactness. In their celebrated paper, Fradkov and Yakubovich...
prove the S-lemma, which implies that the problem of minimizing an arbitrary quadratic objective function over the unit ball (or any single quadratic constraint) can be solved via SDP techniques. Specifically, the S-lemma implies that objective value exactness—the condition that the optimal value of the QCQP and the optimal value of its SDP relaxation coincide—holds for QCQPs with a single constraint; see also [48]. In contrast, Burer and Ye [16] study diagonal QCQPs—those QCQPs for which $A_{\text{obj}}, A_1, \ldots, A_m$ are diagonal matrices—with a general number of constraints and give sufficient conditions for objective value exactness. Wang and Kilinc-Karzan [47, 49] 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 projective convex Lagrange multipliers $\Gamma_P$ (see Section 2). 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 [20, 46]. (See [2, 49] and references therein for additional work in this direction.)

While the framework presented by Wang and Kilinc-Karzan [49] can at once cover and extend many existing results on objective value and convex hull exactness [14, 16, 23, 29, 32, 36, 48, 52], it is still quite limited. In particular, the assumption that the set of Lagrange multipliers $\Gamma_P$ is polyhedral is rarely satisfied outside of simultaneously diagonalizable QCQPs and precludes the results in [49] from being applicable to a wider range of interesting QCQPs.

Additional work in this direction [19] studies objective value exactness from an algebraic point of view. Specifically, Cifuentes et al. [19] consider QCQPs with fixed equality constraints and study the semialgebraic region of objective functions for which objective value exactness holds. As an example of their results, they give a formula for the degree of the algebraic boundary of this region in the setting of Euclidean distance minimization problems.

A related line of work has explored sufficient conditions for the rank-one-generated (ROG) property [2, 11–13, 27]. Recall that a conic subset of the positive semidefinite cone is said to be ROG if it is the convex hull of its rank-one elements. This property can be thought of as the SDP–QCQP analogue to the integrality property in the context of linear program relaxations of integer linear programs [2] and can be shown to imply both convex hull exactness and objective value exactness. Research in this direction has established explicit descriptions of the ROG cones related to quadratic programs over low-dimensional polytopes [13] and ellipsoids with missing caps [15]. Other work in this direction [11, 27] explores the ROG property from an algebro-geometric perspective and establishes results related to the degree and representation of such sets. More recently, Argue et al. [2] gave general sufficient conditions for this property and completely characterized the ROG cones defined by at most two linear matrix inequalities.

SDP exactness has been studied in the context of quadratic matrix programs (QMPs) as well. A QMP is an optimization problem over a matrix variable $X \in \mathbb{R}^{r \times k}$, where the objective function and constraints are each of the form

$$\text{tr}(X^T AX) + 2\text{tr}(B^T X) + c$$

for $A \in \mathbb{S}^r$, $B \in \mathbb{R}^{r \times k}$ and $c \in \mathbb{R}$, and can be thought of as a natural generalization to QCQPs. This class of problems has been used to model robust least squares problems, the orthogonal Procrustes problem [7], and sphere packing [8]. QMPs and their SDP relaxations were first studied by Beck [7], Beck et al. [8] who showed that objective value exactness holds as long as the number of constraints is small compared to $k$. Similarly, Wang and Kilinc-Karzan [49] show that both objective value exactness and convex hull exactness hold for (vectorized reformulations of) QMPs whenever the number of constraints is small enough and $\Gamma_P$ is polyhedral.

Finally, a number of exciting results have shown that various random QCQPs have exact SDP relaxations with high probability. For example, such results have been proved for signal-recovery tasks such as phase retrieval [17], sensor-network localization [42], max-likelihood angular synchronization [3], and clustering [1,
In these settings, the goal is to recover some ground-truth solution (the solution to some QCQP) via observations (constraints in a QCQP). These results then show that once an application-specific signal-to-noise ratio is large enough (for example, given enough observations/constraints), that the SDP relaxation is exact. In contrast, a second line of work [16, 33] addresses random QCQPs which do not assume the existence of a ground-truth solution. In this direction, it is shown that when the number of constraints is small enough that the SDP relaxation has a rank-one optimal solution.

1.1 Overview and outline of the paper

In this paper, we vastly generalize the framework first introduced in [47, 49] by eliminating its reliance on the polyhedrality assumption. Specifically, we give a broad set of sufficient conditions for both convex hull exactness and objective value exactness that are phrased in terms of the set of projective convex Lagrange multipliers $\Gamma_P$ (or the closely related sets $\Gamma$ and $\Gamma^\circ$; see Section 2). In particular, these sufficient conditions can be checked in a systematic manner whenever $\Gamma$, $\Gamma_P$, or $\Gamma^\circ$ is sufficiently simple. Furthermore, we show that our sufficient conditions for convex hull exactness are additionally necessary under a technical assumption (see Assumption 2). We complement our high-level theory with a number of explicit examples illustrating our tools on QCQPs from various settings, including a basic QCQP originating from modeling big-M constraints, quadratic matrix programs, the partition QCQP, and two random QCQP models.

Collectively, these results and examples offer evidence for the main message of this paper that questions of exactness can be treated systematically whenever the convex Lagrange multipliers are well-understood.

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 regarding Lagrangian aggregation and the SDP relaxation of a QCQP. We then define and examine a number of faces of the cone of convex Lagrange multipliers $\Gamma$ and its polar cone $\Gamma^\circ$ that play key roles in our analysis.

2. In Section 3, we present a sufficient condition for convex hull exactness that generalizes [49, Theorem 1]. This sufficient condition (Theorem 1) is based on an analysis of the “rounding directions” inside $S_{\text{SDP}}$ and is performed in the original space. Specifically, we show that convex hull exactness holds as long as certain systems of equations (that depend on $\Gamma$, $\Gamma_P$, or $\Gamma^\circ$) contain nontrivial solutions. In contrast to [49, Theorem 1], our sufficient condition does not make any assumptions on the geometry of $\Gamma$ and can be used to cover additional interesting QCQPs (see Section 4). One of our main technical contributions (Theorem 2) shows that our sufficient condition for convex hull exactness is in fact also necessary under the assumption that $\Gamma^\circ$ is facially exposed (see Assumption 2 and its surrounding discussion). We end Section 3 by revisiting the polyhedral setting. We derive necessary and sufficient conditions for convex hull exactness (Theorem 3) and compare it to the sufficient condition presented in [49, Theorem 1]. To the best of our knowledge, this is the first necessary and sufficient condition for convex hull exactness even in the context of diagonal QCQPs (where $\Gamma$, $\Gamma_P$, and $\Gamma^\circ$ are automatically polyhedral).

3. In Section 4, we present example applications of our general results from Section 3 to a prototypical set containing big-M constraints, quadratic matrix programs, and the partition problem. In all of these applications, the resulting $\Gamma$ sets are non-polyhedral, and thus the sufficient conditions from [49] that work under the polyhedrality assumption of $\Gamma$ fail to cover these applications.

In Section 4.1, 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 set occurs as a substructure commonly studied in sparse regression applications. The convex hull characterization of this set is well-known in the literature and is often shown as a consequence of the perspective formulation trick due to Ceria and Soares [18] (see also [21, 24, 26]).

In Section 4.2, 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 strengthens separate results first presented in [49] and [7]; see Remark 8. In contrast to the ad hoc proof given in [49], the proof we present in Section 4.2 follows the outline of our general framework.
In Section 4.3, we consider the QCQP formulation of 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 [31] 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.

4. In Section 5, we present a number of sufficient conditions for objective value exactness. In fact, our sufficient conditions further imply optimizer exactness, i.e., that the optimizers of the QCQP and its (projected) SDP relaxation coincide. Section 5.1 presents a general sufficient condition (Theorem 4) for objective value exactness based on a primal analysis. Similarly, Section 5.2 presents a general sufficient condition (Theorem 5) for objective value exactness based on a dual analysis. These results recover known sufficient conditions [16, 49] for objective value exactness and explain the roles played by polyhedrality in prior settings. We additionally specialize these abstract conditions to derive more concrete conditions (see Corollaries 5 to 8) for objective value exactness.

5. In Section 6, we present example applications of our general results from Section 5 to two models of random QCQPs. The results in this section show that ideas from Section 5 can be applied even when \( \Gamma, \Gamma_P, \) or \( \Gamma^\circ \) is only known approximately. The models in this section are inspired by recent work on objective value exactness [16, 33] where random QCQPs have been used as a testing ground for understanding the strength or explanatory power of various sufficient conditions. In Section 6.2, we consider a fully random model of QCQPs and show that objective value exactness (in fact optimizer exactness) holds with probability \( 1 - o(1) \) in the regime where \( n \) (the number of variables) diverges to \( +\infty \). In Section 6.3, we consider a semi-random model of QCQPs where, for each quadratic function, the quadratic terms are randomly generated and the linear and constant terms can be chosen adversarially. In this setting, we show that a perturbed notion of exactness holds again with probability \( 1 - o(1) \) as \( n \to +\infty \).

1.2 Notation

For \( x, y \in \mathbb{R} \), let \( [x \pm y] := [x - y, x + y] \), \( x_+ := \max(0, x) \) and \( x_+^2 := (x_+)^2 \). Let \( n \) a positive integer. Let \([n] := \{1, \ldots, n\}\) and for \( i \in [n] \), let \( e_i \) denote the \( i \)th standard basis vector in \( \mathbb{R}^n \). Let \( 0_n \) denote the zero vector in \( \mathbb{R}^n \). For \( \delta \geq 0 \) and \( x \in \mathbb{R}^n \), let \( B_n(x, \delta) := \{ y \in \mathbb{R}^n : \|x - y\| \leq \delta \} \). When \( n \) is clear from context, we will simply write \( 0 \) and \( B(x, \delta) \). Let \( \mathbb{R}^n_+ \) (resp. \( \mathbb{R}^n_{++} \)) denote the entrywise nonnegative (resp. positive) vectors in \( \mathbb{R}^n \). Similarly define \( \mathbb{R}^n_- \) and \( \mathbb{R}^n_{--} \). Let \( \mathbb{S}^{n-1} \subseteq \mathbb{R}^n \) denote the unit sphere. Let \( \mathbb{S}^n \) denote the vector space of \( n \times n \) real symmetric matrices and \( \mathbb{S}^n_+ \) the cone of positive semidefinite matrices. For \( M \in \mathbb{S}^n \), we write \( M \succeq 0 \) (resp. \( M \succeq 0 \)) to denote that \( M \) is positive semidefinite (resp. positive definite). Let \( \lambda_{\min}(M) = \lambda_1(M) \leq \cdots \leq \lambda_n(M) = \lambda_{\max}(M) \) denote the spectrum of \( M \) and let \( \ker(M) \) denote the kernel of \( M \). For \( x \in \mathbb{R}^n \), let \( \text{Diag}(x) \in \mathbb{S}^n \) denote the diagonal matrix with \( \text{Diag}(x)_{i,i} = x_i \) for all \( i \in [n] \). Let \( \mathbb{E} \) denote an arbitrary Euclidean space. Given \( \mathcal{M} \subseteq \mathbb{E} \), let \( \text{int}(\mathcal{M}), \text{bd}(\mathcal{M}), \text{conv}(\mathcal{M}), \text{clconv}(\mathcal{M}), \text{cone}(\mathcal{M}), \text{clcone}(\mathcal{M}), \text{span}(\mathcal{M}), \mathcal{M}^\perp, \text{rint}(\mathcal{M}) \) and \( \text{dim}(\mathcal{M}) \) denote the interior, boundary, convex hull, closed convex hull, conic hull, closed conic hull, span (linear hull), orthogonal complement, relative interior, and dimension of \( \mathcal{M} \) respectively. Let \( K \subseteq \mathbb{E} \) be a cone. Let \( K^\circ \) denote the polar cone of \( K \). The notation \( F \subseteq K \) denotes that \( F \) is a face of \( K \). By convention, faces of cones are always nonempty. \( C_{\infty}(\mathbb{R}^n) \) denotes the smooth functions with compact support on \( \mathbb{R}^n \). Let \( \nabla \) denote the gradient operator. Let \( N(\mu, \Sigma) \) denote the multivariate normal distribution with mean \( \mu \) and covariance \( \Sigma \).
2 Preliminaries

2.1 Setup

We will consider quadratically constrained quadratic programs (QCQPs) in $\mathbb{R}^n$ defined by $m$-many quadratic constraints
\[
\text{Opt} := \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : \begin{array}{l}
q_i(x) \leq 0, \forall i \in [m] \\
q_i(x) = 0, \forall i \in [m+1, m] \end{array} \right\}.
\] (1)

Here, $m_I$ is the number of inequality constraints and $m_E := m - m_I$ is the number of equality constraints. For each $i \in [m]$, we will write $q_i(x) = x^T A_i x + 2b_i^T x + c_i$ for some $A_i \in \mathbb{S}^n$, $b_i \in \mathbb{R}^n$, and $c_i \in \mathbb{R}$. Similarly, we will write $q_{\text{obj}}(x) = x^T A_{\text{obj}} x + 2b_{\text{obj}}^T x + c_{\text{obj}}$.

We will also consider the epigraph, $\mathcal{S}$, of this QCQP, i.e.,
\[
\mathcal{S} := \left\{ (x, t) \in \mathbb{R}^n \times \mathbb{R} : \begin{array}{l}
q_{\text{obj}}(x) \leq 2t \\
q_i(x) \leq 0, \forall i \in [m_I] \\
q_i(x) = 0, \forall i \in [m_I+1, m] \end{array} \right\}.
\]

2.2 Aggregation and the (projected) SDP relaxation

It is well known in the QCQP literature [9, 25, 49] that the SDP relaxation of a QCQP is equivalent (under a minor assumption) to the double-Lagrangian-dual. We will state this formally in Lemma 1 but will first need to introduce notation related to Lagrangian aggregation.

Let $q : \mathbb{R}^n \to \mathbb{R}^{1+m}$ be indexed by \{obj\} $\cup$ [m] where $q(x)_{\text{obj}} = q_{\text{obj}}(x)$ and $q(x)_i = q_i(x)$ for $i \in [m]$. Let $e_{\text{obj}}, e_1, \ldots, e_m$ denote the corresponding unit vectors in $\mathbb{R}^{1+m}$. We will work extensively with the aggregated quadratic functions $(\langle \gamma_{\text{obj}}, \gamma \rangle, q(x))$ for $(\gamma_{\text{obj}}, \gamma) \in \mathbb{R}^{1+m}$. For notational convenience, define $A(\gamma_{\text{obj}}, \gamma) := \gamma_{\text{obj}} A_{\text{obj}} + \sum_{i \in [m]} \gamma_i A_i$. Similarly define $b(\gamma_{\text{obj}}, \gamma)$, and $c(\gamma_{\text{obj}}, \gamma)$. We will at times work in the projective version of the dual space where the distinguished variable $\gamma_{\text{obj}}$ is taken to be one. Let $A[\gamma] := A(1, \gamma)$ and similarly define $b[\gamma]$ and $c[\gamma]$. Set $[\gamma, q(x)] := (\langle \gamma, \gamma \rangle, q(x))$. Note that
\[
\langle \gamma_{\text{obj}}, \gamma \rangle, q(x) = \gamma_{\text{obj}}q_{\text{obj}}(x) + \sum_{i=1}^m \gamma_i q_i(x) = x^T A(\gamma_{\text{obj}}, \gamma) x + 2b(\gamma_{\text{obj}}, \gamma)^T x + c(\gamma_{\text{obj}}, \gamma),
\]
and
\[
[\gamma, q(x)] = q_{\text{obj}}(x) + \sum_{i=1}^m \gamma_i q_i(x) = x^T A[\gamma] x + 2b[\gamma]^T x + c[\gamma].
\]

We recall and extend the following definition from [49].

Definition 1. The cone of convex Lagrange multipliers for (1) is
\[
\Gamma := \left\{ (\gamma_{\text{obj}}, \gamma) \in \mathbb{R} \times \mathbb{R}^m : \begin{array}{l}
A(\gamma_{\text{obj}}, \gamma) \succeq 0 \\
\gamma_{\text{obj}} \geq 0, \forall i \in [m_I] \\
\gamma_i \geq 0, \forall i \in [m_I] \end{array} \right\}.
\]

The set of projective convex Lagrange multipliers for (1) is
\[
\Gamma_P := \{ \gamma \in \mathbb{R}^m : (1, \gamma) \in \Gamma \} = \left\{ \gamma \in \mathbb{R}^m : \begin{array}{l}
A[\gamma] \succeq 0 \\
\gamma_i \geq 0, \forall i \in [m_I] \end{array} \right\}.
\]

Note that given $(\gamma_{\text{obj}}, \gamma) \in \Gamma$, the quadratic function $x \mapsto (\langle \gamma_{\text{obj}}, \gamma \rangle, q(x))$ is convex. Similarly, given $\gamma \in \Gamma_P$, the quadratic function $x \mapsto (\gamma, q(x))$ is convex.

We will make the following blanket assumption for the remainder of the paper. This assumption can be interpreted as a dual strict feasibility condition and is standard in the literature [7, 10, 16, 49, 51].
**Assumption 1.** There exists \((\gamma^*_{\text{obj}}, \gamma^*) \in \Gamma\) such that \(A(\gamma^*_{\text{obj}}, \gamma^*) \succ 0\). Equivalently, there exists \(\gamma^* \in \Gamma_P\) such that \(A[\gamma^*] \succ 0\). □

**Remark 1.** Note that under Assumption 1, we have that \(\Gamma\) is the closed cone generated by its slice at \(\gamma_{\text{obj}} = 1\), i.e., \(\Gamma = \text{clcone}\{(1, \gamma) : \gamma \in \Gamma_P\}\). (See discussion following [49, Assumption 2]) □

Recall that the (projected) SDP relaxation of \(S\) is given by

\[
\mathcal{S}_{\text{SDP}} := \left\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \exists X \succeq xx^\top : \begin{array}{l}
\langle A_{\text{obj}}, X \rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} \leq 2t \\
\langle A, X \rangle + 2b_i^\top x + c_i \leq 0, \forall i \in [m] \\
\langle A_i, X \rangle + 2b_i^\top x + c_i = 0, \forall i \in [m + 1, m] \end{array} \right\},
\]

(2)

and \(\text{Opt}_{\text{SDP}} := \inf_{(x,t) \in \mathcal{S}_{\text{SDP}}} 2t\). By taking \(X = xx^\top\) in (2), we see that \(\text{Opt} \geq \text{Opt}_{\text{SDP}}\) and \(\text{conv}(S) \subseteq \mathcal{S}_{\text{SDP}}\). The following lemma states that under Assumption 1, we can rewrite \(\mathcal{S}_{\text{SDP}}\) in terms of \(\Gamma\). This lemma follows from a straightforward duality argument.

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

\[
\mathcal{S}_{\text{SDP}} = \left\{(x, t) \in \mathbb{R}^{n+1} : [\gamma, q(x)] \leq 2t, \forall \gamma \in \Gamma_P \right\} = \left\{(x, t) \in \mathbb{R}^{n+1} : \langle q_{\text{obj}}(\gamma), q(x) \rangle \leq 2\gamma_{\text{obj}} t, \forall (\gamma_{\text{obj}}, \gamma) \in \Gamma \right\} = \left\{(x, t) \in \mathbb{R}^n : q(x) - 2te_{\text{obj}} \in \Gamma^\circ \right\}.
\]

Here, \(\Gamma^\circ\) denotes the polar cone of \(\Gamma\).

**Proof.** Fix \((x, t) \in \mathbb{R}^{n+1}\). Note that

\[
\sup_{\gamma \in \Gamma_P} \gamma, q(x) = \sup_{\gamma \in \Gamma_P} \left\{\begin{array}{l}
\gamma, q(x) : A[\gamma] \succeq 0 \\
\gamma_i \geq 0, \forall i \in [m] \end{array} \right\} = \inf_{\xi \in \mathbb{R}^m} \left\{q_{\text{obj}}(x) + \langle A_{\text{obj}}, \xi \rangle : \begin{array}{l}
q_i(x) + \langle A_i, \xi \rangle \leq 0, \forall i \in [m] \\
q_i(x) + \langle A_i, \xi \rangle = 0, \forall i \in [m + 1, m] \\
\xi \succeq 0 \end{array} \right\},
\]

where the second equation follows from the strong conic duality theorem and Assumption 1. Taking \(X := xx^\top + \xi\), we deduce that the first equality in Lemma 1 holds.

Note that by Assumption 1, \(\Gamma = \text{clcone}\{(1, \gamma) : \gamma \in \Gamma_P\}\) so that \([\gamma, q(x)] \leq 2t\) for all \(\gamma \in \Gamma_P\) if and only if \(\langle (\gamma_{\text{obj}}, \gamma), q(x) \rangle \leq 2\gamma_{\text{obj}} t\) for all \((\gamma_{\text{obj}}, \gamma) \in \Gamma\); this gives the second equality. The third equality holds by definition of the polar cone. □

**Corollary 1.** Suppose Assumption 1 holds. Then

\[
\text{Opt}_{\text{SDP}} = \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma_P} [\gamma, q(x)].
\]

**Corollary 2.** Suppose Assumption 1 holds. Then, \(\mathcal{S}_{\text{SDP}}\) is closed.

**Remark 2.** In comparison with (2), the expressions for \(\mathcal{S}_{\text{SDP}}\) given in Lemma 1 make the roles played by \(\Gamma\), \(\Gamma_P\), and \(\Gamma^\circ\) explicit. In particular, these expressions for \(\mathcal{S}_{\text{SDP}}\) lend themselves to a clean analysis whenever the corresponding dual set \(\Gamma, \Gamma_P\), or \(\Gamma^\circ\) is sufficiently simple. □

**Remark 3.** Phrased differently, one may minimize \(\text{Opt}_{\text{SDP}}\) in the form (2) by minimizing over \(x \in \mathbb{R}^n\) the value of an inner minimization problem over the matrix variables \(X \succeq xx^\top \in S^n\). Writing \(X = xx^\top + \xi\) and taking the SDP dual in the \(\xi\) variable then results in the same saddle-point structure \(\text{Opt}_{\text{SDP}} = \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma_P} [\gamma, q(x)]\) observed in Corollary 1. In other words, \(\Gamma_P\) is simply the set of feasible solutions to this partial dual of (2). □
We will additionally associate faces of $\Gamma$. The next fact follows from Definition 4.

Similarly, define the conjugate face of $F$ for a face $F \subseteq \Gamma$.

Definition 3. For a face $G \subseteq \Gamma^o$, we say that $G$ is exposed if there exists $(\gamma_{\text{obj}}, \gamma) \in \Gamma$ such that $G = \Gamma^o \cap (\gamma_{\text{obj}}, \gamma)^\perp$.

We will additionally associate faces of $\Gamma$ and $\Gamma^o$ to points $(x, t) \in S_{\text{SDP}}$ as follows.

Definition 4. Given $(x, t) \in S_{\text{SDP}}$, let $G(x, t) \subseteq \Gamma^o$ denote the minimal face of $\Gamma^o$ containing $q(x) - 2te_{\text{obj}}$ and define $F(x, t) := G(x, t)^\Delta$.

The next fact follows from Definition 4.

Fact 1. Given $(x, t) \in S_{\text{SDP}}$, we have that $q(x) - 2te_{\text{obj}} \in \text{rint}(G(x, t))$ and $F(x, t) = \Gamma \cap (q(x) - 2te_{\text{obj}})^\perp$.  

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

Example 1. Consider the following QCQP epigraph,

$$S := \left\{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : \begin{array}{l} q_{\text{obj}}(x) \leq 2t \\ q_1(x) \leq 0 \\ q_2(x) \leq 0 \end{array}\right\},$$

where $q_{\text{obj}}(x) := 2x_1x_2 - x_2 - 1/4$, $q_1(x) := x_1^2 - x_2^2 - x_1 + x_2 - 1$, and $q_2(x) := x_1^2 + x_2^2 - 1$. Through a straightforward calculation, we obtain

$$\Gamma = \left\{(\gamma_{\text{obj}}, \gamma) \in \mathbb{R}^3 : \begin{array}{c} \gamma_2 \geq \sqrt{\gamma_{\text{obj}}^2 + \gamma_1^2} \\ \gamma_{\text{obj}}, \gamma_1, \gamma_2 \geq 0 \end{array}\right\},$$

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

$$S_{\text{SDP}} = \left\{(x, t) \in \mathbb{R}^2 : q_2(x) \geq \sqrt{ (q_{\text{obj}}(x) - 2t)^2 + q_1(x)^2 } \right\}. $$

See Figure 1 for the plots of the sets corresponding to $S$, $\Gamma$, $\Gamma^o$, and $S_{\text{SDP}}$. 

2.3 Faces of $\Gamma$ and $\Gamma^o$

In this section we define key faces of $\Gamma$ and $\Gamma^o$ 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 [5, 6, 38] for a more in-depth treatment of general convex cones and their faces.

Recall the following definitions.

Definition 2. Given a face $G \subseteq \Gamma^o$ and $(g_{\text{obj}}, g) \in \text{rint}(G)$, the conjugate face of $G$ is

$$G^\Delta := \Gamma \cap G^\perp = \Gamma \cap (g_{\text{obj}}, g)^\perp.$$ 

Similarly, define the conjugate face of $F$ for a face $F \subseteq \Gamma$.

Definition 3. For a face $G \subseteq \Gamma^o$, we say that $G$ is exposed if there exists $(\gamma_{\text{obj}}, \gamma) \in \Gamma$ such that $G = \Gamma^o \cap (\gamma_{\text{obj}}, \gamma)^\perp$.

We will additionally associate faces of $\Gamma$ and $\Gamma^o$ to points $(x, t) \in S_{\text{SDP}}$ as follows.

Definition 4. Given $(x, t) \in S_{\text{SDP}}$, let $G(x, t) \subseteq \Gamma^o$ denote the minimal face of $\Gamma^o$ containing $q(x) - 2te_{\text{obj}}$ and define $F(x, t) := G(x, t)^\Delta$.

The next fact follows from Definition 4.

Fact 1. Given $(x, t) \in S_{\text{SDP}}$, we have that $q(x) - 2te_{\text{obj}} \in \text{rint}(G(x, t))$ and $F(x, t) = \Gamma \cap (q(x) - 2te_{\text{obj}})^\perp$.  

Figure 1: The sets $S$, $S_{\text{SDP}}$, $\Gamma$, and $\Gamma^o$ from Example 1 are shown in blue, green, orange, and yellow respectively. By Lemma 1, $(x, t) \in S_{\text{SDP}}$ if and only if $q(x) - 2te_{\text{obj}} \in \Gamma^o$. 

7
3 Convex hull exactness

In this section, we present necessary and sufficient conditions for convex hull exactness, i.e., the property that \( \text{conv}(S) = S_{\text{SDP}} \). These results form the basis of our assertion that exactness can be treated systematically whenever \( \Gamma, \Gamma_p, \) or \( \Gamma^o \) is well-understood.

We begin by rephrasing convex hull exactness as a question regarding the existence of certain “rounding directions.” The following result follows from basic convex analysis.

**Lemma 2.** Suppose Assumption 1 holds. Then, \( \text{conv}(S) = S_{\text{SDP}} \) if and only if for every \((x,t) \in S_{\text{SDP}} \setminus S\), there exists a nonzero \((x',t') \in \mathbb{R}^{n+1} \) and \( \alpha > 0 \) such that

\[
[(x,t) \pm \alpha (x',t')] \subseteq S_{\text{SDP}}.
\]

**Proof.** Note that \( S_{\text{SDP}} \) is a closed convex set containing no lines. Also, one can easily check that \((0_n,1)\) is indeed a recessive direction of \( S_{\text{SDP}} \). Furthermore, \((0_n,1)\) is the only recessive direction of \( S_{\text{SDP}} \). To see this, let \( \gamma^* \) be such that \( A[\gamma^*] > 0 \) (which exists by Assumption 1) and consider any \((x',t')\) where \( x' \) is nonzero. Then, for any \((\tilde{x},t) \in S_{\text{SDP}} \) and all \( \alpha > 0 \) large enough, \( 2(t + \alpha t') < |\gamma^*|, q(\tilde{x} + \alpha x')| \). Therefore, we deduce by [40, Theorem 18.5], that \( S_{\text{SDP}} \) is the sum of the convex hull of its extreme points and the direction \((0_n,1)\). In particular \( \text{conv}(S) = S_{\text{SDP}} \) if and only if \((x,t)\) is not extreme for every \((x,t) \in S_{\text{SDP}} \setminus S\). By definition, \((x,t)\) is not extreme if and only if there exists \((x',t')\) and \( \alpha > 0 \) such that \([x,t] \pm \alpha (x',t') \] \subseteq S_{\text{SDP}}.

We capture the relevant set in Lemma 2 in the following definition.

**Definition 5.** The subspace of rounding directions at \((x,t) \in S_{\text{SDP}} \) is

\[
\mathcal{R}(x,t) := \{(x',t') \in \mathbb{R}^{n+1} : \exists \alpha > 0 \text{ s.t. } [(x,t) \pm \alpha (x',t')] \subseteq S_{\text{SDP}}\}.
\]

This set is nontrivial if it contains a nonzero element.

Note that \( \mathcal{R}(x,t) \) is in fact a subspace so that its name is justified. Indeed, \( \mathcal{R}(x,t) \) is a convex cone as \( S_{\text{SDP}} \) is convex. Furthermore, it holds that \(-\mathcal{R}(x,t) = \mathcal{R}(x,t)\).

**Remark 4.** One may compare our rounding directions to other similar definitions from elementary convex analysis [28, Section 5.1]. Fix a point \((x,t) \in S_{\text{SDP}} \) and \((x',t') \in \mathbb{R}^{n+1} \). Recall that \((x',t')\) is a feasible direction if there exists \( \alpha > 0 \) such that \([(x,t), (x,t) + \alpha (x',t')] \subseteq S_{\text{SDP}} \). In particular, feasible directions are a unidirectional notion, whereas rounding directions are bidirectional. Next, recall that \((x',t')\) is a tangent direction if it is a limit of feasible directions. Again, tangent directions are unidirectional.

**Remark 5.** Suppose \((x,t) \in S_{\text{SDP}} \) and \( 2t > \sup_{\gamma \in \Gamma_p} [\gamma, q(x)] \). Then, by Lemma 1 there exists \( \alpha > 0 \) such that \([(x,t) \pm \alpha (0_n,1)] \subseteq S_{\text{SDP}} \). In particular, it suffices to verify the condition of Lemma 2 for points \((x,t) \in S_{\text{SDP}} \setminus S\) for which \( 2t = \sup_{\gamma \in \Gamma_p} [\gamma, q(x)] \).

3.1 Sufficient conditions for convex hull exactness

In this section we identify a particular subset of the rounding directions at \((x,t) \in S_{\text{SDP}} \). This then leads to a sufficient condition for convex hull exactness, i.e., the condition that \( \text{conv}(S) = S_{\text{SDP}} \).

**Definition 6.** Given \((x,t) \in S_{\text{SDP}} \), define

\[
\mathcal{R}'(x,t) := \{(x',t') \in \mathbb{R}^{n+1} : q(x + \alpha x') - 2(t + \alpha t')e_{\text{obj}} \in \text{span}(G(x,t)), \forall \alpha \in \mathbb{R} \}\.
\]

**Lemma 3.** Suppose Assumption 1 holds and \((x,t) \in S_{\text{SDP}} \). Then, \( \mathcal{R}'(x,t) \subseteq \mathcal{R}(x,t) \).

**Proof.** Let \((x,t) \in S_{\text{SDP}} \) and \((x',t') \in \mathcal{R}'(x,t) \). Then, by continuity and the fact that \( q(x) - 2te_{\text{obj}} \in \text{rint}(G(x,t)) \), there exists \( \alpha > 0 \) such that

\[
q(x + \epsilon x') - 2(t + \epsilon t')e_{\text{obj}} \in G(x,t) \subseteq \Gamma^o
\]

for all \( \epsilon \in [\pm \alpha] \). By the third characterization of \( S_{\text{SDP}} \) in Lemma 1, we have that \([x,t] \pm \alpha (x',t') \] \subseteq S_{\text{SDP}}. \( \square \)
Lemmas 2 and 3 immediately imply the following sufficient condition for convex hull exactness.

**Theorem 1.** Suppose Assumption 1 holds and that for all \((x, t) \in S_{\text{SDP}} \setminus S\), the set \(R'(x, t)\) is nontrivial. Then, \(\text{conv}(S) = S_{\text{SDP}}\).

We will see a number of applications of Theorem 1 in Section 4.

In Lemma 4 below, we will record an alternate description of \(R'(x, t)\). We will require the following observation.

**Observation 1.** Suppose Assumption 1 holds. Let \((x, t) \in S_{\text{SDP}}\) where \(2t = \sup_{\gamma \in \Gamma_p}[\gamma, q(x)]\). Then, \(\text{span}(G(x, t)) \not\supseteq \mathbb{R} \times 0_m\). In particular, \(G(x, t)^\perp = \text{span} \left(G(x, t)^\perp \cap \{\gamma_{\text{obj}} = 1\}\right)\).

**Proof.** Suppose \(\text{span}(G(x, t)) \supseteq \mathbb{R} \times 0_m\) so that \((0_n, 1) \in R'(x, t)\). By Lemma 3, there exists \(\alpha > 0\) such that \((x, t - \alpha) \in S_{\text{SDP}}\). This contradicts \(2t = \sup_{\gamma \in \Gamma_p}[\gamma, q(x)]\).

We deduce that \(\text{span}(G(x, t)) \supseteq \mathbb{R} \times 0_m\). Equivalently, \(G(x, t)^\perp \not\subseteq 0 \times \mathbb{R}^m\) and there exists \((1, \bar{\gamma}) \in G(x, t)^\perp\). Then, for any \((\gamma_{\text{obj}}, \gamma) \in G(x, t)^\perp\), we can write \((\gamma_{\text{obj}}, \gamma)\) as a linear combination of
\[
(\gamma_{\text{obj}}, \gamma) + (1 - \gamma_{\text{obj}})(1, \gamma) \quad \text{and} \quad (1, \gamma).
\]

**Lemma 4.** Suppose Assumption 1 holds and let \((x, t) \in S_{\text{SDP}}\). Then,
\[
R'(x, t) = \left\{ (x', t') \in \mathbb{R}^{n+1} : (x')^\top A(\gamma_{\text{obj}}, \gamma)x' = 0, \forall (\gamma_{\text{obj}}, \gamma) \in G(x, t)^\perp \right\}
\]
If furthermore \(2t = \sup_{\gamma \in \Gamma_p}[\gamma, q(x)]\), then
\[
R'(x, t) = \left\{ (x', t') \in \mathbb{R}^{n+1} : (x')^\top A[\gamma]x' = 0, \forall (\gamma) \in G(x, t)^\perp \right\}.
\]

**Proof.** Note that \((x', t') \in R'(x, t)\) if and only if for all \((\gamma_{\text{obj}}, \gamma) \in G(x, t)^\perp\), we have that
\[
\langle (\gamma_{\text{obj}}, \gamma), q(x + \alpha x') - 2(t + \alpha t')e_{\text{obj}} \rangle = \alpha^2 (x')^\top A(\gamma_{\text{obj}}, \gamma)(x') + 2\alpha \left(A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), x' - t'\right) + \langle (\gamma_{\text{obj}}, \gamma), q(x) - 2te_{\text{obj}} \rangle
\]
is identically zero in \(\alpha\). This occurs if and only if for all \((\gamma_{\text{obj}}, \gamma) \in G(x, t)^\perp\), we have
\[
(x')^\top A(\gamma_{\text{obj}}, \gamma)x' = 0, \quad \text{and} \quad \langle A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), x' - t' \rangle = 0.
\]
This proves the first assertion. The second assertion follows from the first and Observation 1. ■

### 3.2 Necessary conditions for convex hull exactness

In Section 3.1, we gave a sufficient condition for convex hull exactness by identifying a subset of directions \(R'(x, t) \subseteq R(x, t)\) and invoking Lemma 2. In this section, we show that under a technical assumption (Assumption 2), we have \(R'(x, t) = R(x, t)\). This then leads to a necessary and sufficient condition for convex hull exactness under the technical assumption.

**Assumption 2.** Suppose \(\Gamma^\circ\) is facially exposed, i.e., every face of \(\Gamma^\circ\) 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 [39] for a longer discussion of this assumption and its connections to the **nice cones**. In general, all **nice cones** are facially exposed. Our analysis will be based on the following property of exposed faces \(G \subseteq \Gamma^\circ\) (see [6, Definition 2.A.9] and its surrounding discussion):

**Fact 2.** A face \(G \subseteq \Gamma^\circ\) is exposed if and only if \(G = (G^\wedge)^\wedge\).
We are now ready to prove a partial converse to Lemma 3.

**Lemma 5.** Suppose Assumptions 1 and 2 hold and let \((x, t) \in S_{SDP}\). Then \(R'(x, t) = R(x, t)\).

**Proof.** Fix \((x', t') \in R(x, t)\). As \(R(x, t)\) is a convex cone, we may without loss of generality assume that \([(x, t) \pm (x', t')] \subseteq S_{SDP}\). Our goal is to show that \((x', t') \in R'(x, t)\), i.e., that
\[
q(x + \alpha x') - 2(t + \alpha t')e_{obj} \in \text{span}(G(x, t)), \forall \alpha \in \mathbb{R}.
\]
As each coordinate of this vector is quadratic in \(\alpha\), it suffices to show instead that
\[
q(x + \alpha x') - 2(t + \alpha t')e_{obj} \in G(x, t), \forall \alpha \in [-1, 1].
\]
Let \((f_{obj}, f) \in \text{rint}(F(x, t))\) so that by Assumption 2 and Fact 2, we may write \(G(x, t) = \Gamma^\circ \cap (f_{obj}, f)^\perp\). As \([(x, t) \pm (x', t')] \subseteq S_{SDP}\), we immediately have that \(q(x + \alpha x') - 2(t + \alpha t')e_{obj} \in \Gamma^\circ\) for all \(\alpha \in [-1, 1]\). It remains to verify that the map
\[
\alpha \mapsto \langle (f_{obj}, f), q(x + \alpha x') - 2(t + \alpha t')e_{obj} \rangle
\]
evaluates to zero on \(\alpha \in [-1, 1]\). Again, as \([(x, t) \pm (x', t')] \subseteq S_{SDP}\), this map is nonpositive for all \(\alpha \in [-1, 1]\).

Next, note that \((f_{obj}, f) \in F(x, t) = \Gamma \cap (q(x) - 2te_{obj})^\perp\) so that this map evaluates to zero at \(\alpha = 0\). Finally, \((f_{obj}, f) \in \Gamma\) implies that this map is also convex. We conclude that this map is identically zero. 

The following necessary and sufficient condition for convex hull exactness then follows from Lemma 5.

**Theorem 2.** Suppose Assumptions 1 and 2 hold. Then, \(\text{conv}(S) = S_{SDP}\) if and only if for all \((x, t) \in S_{SDP} \setminus S\), the set \(R'(x, t)\) is nontrivial.

To close this subsection, we give a compact description of \(R'(x, t)\) under Assumption 2.

**Proposition 1.** Suppose Assumptions 1 and 2 hold. Let \((x, t) \in S_{SDP}\) where \(2t = \sup_{\gamma \in \Gamma^\circ} [\gamma, q(x)]\) and let \((1, f) \in \text{rint}(F(x, t))\). Then,
\[
R'(x, t) = \left\{ (x', t') \in \mathbb{R}^{n+1} : \langle A[\gamma], x' \rangle = 0, \forall (1, \gamma) \in G(x, t)^\perp \right\}.
\]

**Proof.** Let \((1, f) \in \text{rint}(F(x, t))\). By Lemma 4, it suffices to show that \(x' \in \ker(A[f])\) if and only if
\[
(x')^\top A[\gamma] x' = 0, \forall (1, \gamma) \in G(x, t)^\perp.
\]
The reverse direction holds immediately as \((1, f) \in F(x, t) \subseteq G(x, t)^\perp\) and \(A[f] \succeq 0\).

To see the forward direction: Let \(x' \in \ker(A[f])\) and set \(v_{obj} = (x')^\top A_{obj} x'\). Similarly, set \(v_i = (x')^\top A_i x'\). Then,
\[
\langle (v_{obj}, v), (\gamma_{obj}, \gamma) \rangle = (x')^\top A(\gamma_{obj}, \gamma) x' \geq 0, \forall (\gamma_{obj}, \gamma) \in \Gamma.
\]
Thus \((-v_{obj}, -v) \in \Gamma^\circ\). On the other hand, \(\langle (v_{obj}, v), (1, f) \rangle = (x')^\top A[f] x' = 0\). We deduce that \((-v_{obj}, -v) \in \Gamma^\circ \cap (1, f)^\perp = F(x, t)^\Delta = G(x, t)\). In particular, \((x')^\top A(\gamma_{obj}, \gamma) x' = \langle (v_{obj}, v), (\gamma_{obj}, \gamma) \rangle = 0\) for all \((\gamma_{obj}, \gamma) \in G(x, t)^\perp\).

### 3.3 Revisiting the setting of polyhedral \(\Gamma\)

Wang and Kılınç-Karzan [49] give sufficient conditions for convex hull exactness under the assumption that \(\Gamma\) is polyhedral. This assumption holds, for example, when the set of quadratic forms \(\{A_{obj}, A_1, \ldots, A_m\}\) is simultaneously diagonalizable. Specializing Theorem 2 to this setting, we prove the following necessary and sufficient counterpart to [49, Theorem 1].
Theorem 3. Suppose Assumption 1 holds and that $\Gamma$ is polyhedral. Then, $\text{conv}(\mathcal{S}) = \mathcal{S}_{\text{SDP}}$ if and only if
\[
\left\{(x', t') \in \mathbb{R}^{n+1} : x' \in \ker(A[f]) \right\}
\]
is nontrivial for every $F \subseteq \Gamma$ which is exposed by some vector $q(x) - 2te_{\text{obj}}$ for $(x, t) \in \mathcal{S}_{\text{SDP}} \setminus \mathcal{S}$. Here, $f$ is any vector such that $(1, f) \in \text{rint}(F)$.

Proof. We begin by noting that when $\Gamma$ is polyhedral, so too is $\Gamma^\circ$ so that Assumption 2 holds. Next, we claim that for every face $G \subseteq \Gamma^\circ$ we have $G^\perp = \text{span}(G^\wedge)$. By definition, $\text{span}(G^\wedge) = \text{span}(\Gamma \cap G^\perp) \subseteq G^\perp$. On the other hand, as $\Gamma$ and $\Gamma^\circ$ are polyhedral, we have that [44, Theorem 3]
\[
\dim(G) + \dim(G^\wedge) = m.
\]
Rearranging this equation, we have $\dim(G^\wedge) = m - \dim(G) = \dim(G^\perp)$. We conclude that $G^\perp = \text{span}(G^\wedge)$.

Let $(x, t) \in \mathcal{S}_{\text{SDP}}$ such that $2t = \sup_{\gamma \in \Gamma^\circ} [\gamma, q(x)]$ and let $(1, f) \in \text{rint}(F(x, t))$. Then, Observation 1 and Proposition 1 imply that
\[
\mathcal{R}'(x, t) = \left\{(x', t') \in \mathbb{R}^{n+1} : x' \in \ker(A[f]) \right\}
\]
\[
= \left\{(x', t') \in \mathbb{R}^{n+1} : x' \in \ker(A[f]) \right\}
\]
\[
= \left\{(x', t') \in \mathbb{R}^{n+1} : x' \in \ker(A[f]) \right\}
\]
Here, the second line follows because we have shown $G^\perp = \text{span}(G^\wedge)$ holds for every face $G \subseteq \Gamma^\circ$ and by definition $F(x, t) = G(x, t)^\wedge$. The third line follows from the fact that $(1, f) \in \text{rint}(F(x, t))$ implies $\ker(A[f]) \subseteq \ker(A[\gamma])$ for every $(1, \gamma) \in F(x, t)$. The result then follows from Theorem 2. \hfill \blacksquare

Remark 6. The main difference between Theorem 3 and [49, Theorem 1] is that Theorem 3 only considers certain (a fortiori semidefinite) faces of $\Gamma$ whereas [49, Theorem 1] imposes a constraint on every semidefinite face of $\Gamma$. This idea of restricting the analysis to certain faces of $\Gamma$ was previously investigated by [16, 33] who used it to provide sufficient conditions for objective value exactness. \hfill \blacksquare

4 Applications: Convex hull exactness

In this section, we apply the results of Section 3 to a number of problems. These examples provide further evidence towards the message that exactness can be treated systematically whenever $\Gamma, \Gamma_P$, or $\Gamma^\circ$ is well-understood.

4.1 Mixed binary programming

To begin, we apply our results to a well-studied prototypical set involving a convex quadratic function, a binary variable and a big-M relation. The example in this subsection highlights the systematic nature of our approach.

Consider the epigraph set
\[
\mathcal{S} = \left\{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : q_{\text{obj}}(x) := x_2^2 \leq 2t, q_1(x) := x_1(x_1 - 1) = 0, q_2(x) := \sqrt{2}x_2(x_1 - 1) = 0 \right\}.
\]
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 $t$ is the epigraph variable corresponding to $x_2^2$. The normalization of $q_2(x)$ is not important here and is made only for notational convenience in the calculations.
It is well-known that \( \text{conv}(S) \) is given by the perspective reformulation of \( S \) (see e.g., [24, 26]), i.e.,
\[
\text{conv}(S) = \{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : x_2^2 - 2tx_1 \leq 0, 0 \leq x_1 \leq 1\}.
\] (4)

We give an alternative proof of (4). We will show that \( \text{conv}(S) = S_{\text{SDP}} \), the projected SDP relaxation, using Theorem 2. Then, using an explicit description of \( \Gamma^0 \), we will give a description of \( \text{conv}(S) = S_{\text{SDP}} \) in the original space.

A simple computation shows that in this setting, we have
\[
\Gamma = \{(\gamma_{\text{obj}}, \gamma) \in \mathbb{R}^3 : \gamma_{\text{obj}} + \gamma_1 \geq \sqrt{(\gamma_{\text{obj}} - \gamma_1)^2 + (\sqrt{2}\gamma_2)^2}\}
\]
and
\[
\Gamma^0 = \{(\ell_{\text{obj}}, \ell) \in \mathbb{R}^3 : -\ell_{\text{obj}} - \ell_1 \geq \sqrt{(\ell_{\text{obj}} - \ell_1)^2 + (\sqrt{2}\ell_2)^2}\}.
\]

In words, \( \Gamma \) and \( \Gamma^0 \) are both (rotated) second-order cones and Assumptions 1 and 2 hold.

It remains to show that for all \((x, t) \in S_{\text{SDP}} \setminus S \), the set \( \mathcal{R}(x, t) \) is nontrivial. To this end, let \((x, t) \in S_{\text{SDP}} \setminus S \). Recall that \( \Gamma^0 \) has three types of 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) \( \mathcal{G}(x, t) = \{0\} \), (ii) \( \mathcal{G}(x, t) = \Gamma^0 \), and (iii) \( \mathcal{G}(x, t) \) is a one-dimensional face of \( \Gamma^0 \).

In case (i), \( q(x) - 2te_{\text{obj}} = 0 \) implying that \((x, t) \in S \), a contradiction. In case (ii), \( \text{span}(\mathcal{G}(x, t)) = \mathbb{R}^3 \) so that \( \mathcal{R}(x, t) = \mathbb{R}^3 \) and is nontrivial. In the final case, a mechanical but slightly tedious application of Proposition 1 (see Appendix A) gives
\[
\mathcal{R}'(x, t) = \left\{ \begin{pmatrix} 2t \\ -x_2 \\ 0 \end{pmatrix}, \begin{pmatrix} -x_2 \\ t \\ -x_2 \end{pmatrix}, \begin{pmatrix} x_2(x_1 - 1 + 2t) \\ -x_2^2 + x_1 - 2tx_1 - 2t \\ 2x_2 \end{pmatrix} \right\}^\perp.
\]

Finally, one may verify that \((x, t) \in \mathcal{R}'(x, t) \) is nonzero.

Remark 7. Here, the motivation for the final step of checking that \((x, t) \in \mathcal{R}'(x, t) \) is as follows: One can show that in case (iii), the first three vectors in (5) span the 2-dimensional subspace orthogonal to \((x, t) \). In particular, \( \mathcal{R}'(x, t) \) is nontrivial if and only if \((x, t) \in \mathcal{R}'(x, t) \). \( \square \)

We conclude that
\[
\text{conv}(S) = S_{\text{SDP}} = \left\{ (x, t) \in \mathbb{R}^3 : -(q_{\text{obj}}(x) - 2t) - q_1(x) \geq \sqrt{(q_{\text{obj}}(x) - 2t - q_1(x))^2 + 2q_2(x)^2} \right\}.
\]

This example highlights the systematic nature of the approach outlined in Theorem 2 for proving convex hull exactness. In contrast to ad hoc proofs of convex hull exactness which may rely on guessing and verifying a nonzero rounding direction, the system of equations defining \( \mathcal{R}'(x, t) \) gives a principled way of deducing a direction. While guessing such a rounding direction may be possible in low-dimensional settings (for example, the setting of the current subsection), this becomes more difficult in higher-dimensional settings where \( S \) and \( S_{\text{SDP}} \) are difficult to visualize. We illustrate this in the following subsection.

4.2 Quadratic matrix programs

Quadratic matrix programs (QMPs) [7, 49] are a generalization of QCQPs where the decision variable \( x \in \mathbb{R}^n \) is replaced by a decision matrix \( X \in \mathbb{R}^{r \times k} \). These problems find a variety of application and have been used to model robust least squares problems, the orthogonal Procrustes problem [7], and certain sphere packing
problems \[8\]. Formally, a QMP is an optimization problem in the variable \(X \in \mathbb{R}^{r \times k}\), where the constraints and objective function are each of the form
\[
\text{tr}(X^T A X) + 2 \text{tr}(B^T X) + c
\]
for some \(A \in \mathbb{S}^r\), \(B \in \mathbb{R}^{r \times k}\), and \(c \in \mathbb{R}\).

Alternatively, letting \(x \in \mathbb{R}^n\) (resp. \(b \in \mathbb{R}^n\)) denote the vector formed by stacking the columns of \(X\) (resp. \(B\)) on top of each other, we can rewrite the above expression as
\[
x^T (I_k \otimes A)x + 2 \langle b, x \rangle + c.
\]

We will choose to view QMPs as the special class of QCQPs where the quadratic forms \(A_{\text{obj}}, A_1, \ldots, A_m\) are each of the form \(I_k \otimes A\) for some \(A \in \mathbb{S}^r\).

The following lemma establishes that if the number of constraints is small compared to \(k\) (originally the width of the matrix variable), then convex hull exactness holds.

**Proposition 2.** Suppose Assumption 1 holds and that \(A_{\text{obj}} = I_k \otimes A_{\text{obj}}, A_1 = I_k \otimes A_1, \ldots, A_m = I_k \otimes A_m\) for some \(A_{\text{obj}}, A_1, \ldots, A_m \in \mathbb{S}^r\). Furthermore, suppose \(k \geq m\). Then, \(R(x, t)\) is nontrivial for every \((x, t) \in \mathcal{S}_{\text{SDP}} \setminus \mathcal{S}\). In particular, convex hull exactness holds, i.e., \(\text{conv}(\mathcal{S}) = \mathcal{S}_{\text{SDP}}\).

**Proof.** Fix \((x, t) \in \mathcal{S}_{\text{SDP}} \setminus \mathcal{S}\). Based on Theorem 1 and Lemma 4, our goal is to prove that
\[
R'(x, t) = \left\{ (x', t') \in \mathbb{R}^{r+1} : \begin{array}{c}
x'^T A(\gamma_{\text{obj}}, \gamma)x' = 0, \forall (\gamma_{\text{obj}}, \gamma) \in \mathcal{G}(x, t) \\
(\langle A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), x' \rangle - \gamma_{\text{obj}}t') = 0, \forall (\gamma_{\text{obj}}, \gamma) \in \mathcal{G}(x, t)
\end{array} \right\}
\]
is nontrivial. We claim that it suffices to show how to construct a nonzero \(y \in \mathbb{R}^r\) such that
\[
y^T A(\gamma_{\text{obj}}, \gamma)y = 0, \forall (\gamma_{\text{obj}}, \gamma) \in \mathcal{G}(x, t)\perp.
\]
To see that this suffices, note that for any \(w \in \mathbb{R}^k\), the vector \(x' := w \otimes y\) satisfies the first constraint in \((6)\) since for \((\gamma_{\text{obj}}, \gamma) \in \mathcal{G}(x, t)\perp\), we have
\[
\langle A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), w \otimes y \rangle - \gamma_{\text{obj}}t' = 0, \forall (\gamma_{\text{obj}}, \gamma) \in \mathcal{G}(x, t)\perp.
\]
This is a system of \(\dim(\mathcal{G}(x, t)\perp)\)-many homogeneous linear equations in the variables \((w, t) \in \mathbb{R}^{k+1}\). Note that as \(\mathcal{G}(x, t) \supseteq q(x) - 2te_{\text{obj}}\), which is nonzero by assumption, we have that \(\dim(\mathcal{G}(x, t)\perp) \leq m\). As \(k+1 > m\) by assumption, we deduce that this system has a nontrivial solution. Thus, we conclude that \((6)\) is nontrivial if there exists a nonzero \(y \in \mathbb{R}^r\) satisfying \((7)\).

It remains to construct \(y\). By definition of \(\mathcal{S}_{\text{SDP}}\), there exists \(Y \succeq 0\) such that
\[
\begin{cases}
q_{\text{obj}}(x) + \langle A_{\text{obj}}, Y \rangle \leq 2t, \\
q_i(x) + \langle A_i, Y \rangle \leq 0, \forall i \in [m_1], \text{ and} \\
q_i(x) + \langle A_i, Y \rangle = 0, \forall i \in [m_1 + 1, m].
\end{cases}
\]
Without loss of generality, \(Y = (\frac{1}{2}I_k) \otimes Y\). As \((x, t) \notin \mathcal{S}\), we have that \(Y \in \mathbb{S}^{r+1}_+ \setminus \{0\}\) and we may pick a nonzero \(y \in \mathbb{R}^r\) such that \(yy^T \preceq Y\). For notational convenience, let \(t_{\text{obj}} := y^T A_{\text{obj}}y\) and \(t_i := y^T A_i y\) for \(i \in [m]\). Note that for any \((\gamma_{\text{obj}}, \gamma) \in \Gamma\), we have \(A(\gamma_{\text{obj}}, \gamma) \succeq 0\), or equivalently \(A(\gamma_{\text{obj}}, \gamma) \succeq 0\). Thus, \(yy^T \preceq Y\) implies that \(\langle (\gamma_{\text{obj}}, \gamma), (t_{\text{obj}}, t_i) \rangle = y^T A(\gamma_{\text{obj}}, \gamma)y \leq \langle A(\gamma_{\text{obj}}, \gamma), Y \rangle\). Also, from \(Y = (\frac{1}{2}I_k) \otimes Y\) and the relation between the matrices \(A_{\text{obj}}, A_i\), and \(A_{\text{obj}}, A_i\), we have \(\langle A(\gamma_{\text{obj}}, \gamma), Y \rangle = \langle A(\gamma_{\text{obj}}, \gamma), Y \rangle\). We deduce that for \((\gamma_{\text{obj}}, \gamma) \in \Gamma\),
\[
\left\langle \left( \begin{array}{c} \gamma_{\text{obj}} \\ \gamma \end{array} \right), q(x) - 2te_{\text{obj}} + \left( t_{\text{obj}} \right) \right\rangle \preceq \left\langle \left( \begin{array}{c} \gamma_{\text{obj}} \\ \gamma \end{array} \right), q(x) - 2te_{\text{obj}} + \left( \langle A_{\text{obj}}, Y \rangle \right) \right\rangle \leq 0,
\]
where the last inequality follows from (8) and \((\gamma_{\text{obj}}, \gamma) \in \Gamma\). This then shows that \(q_{\text{obj}}(x) - 2t e_{\text{obj}} + (\ell_{\text{obj}}, \ell) \in \Gamma^o\). Moreover, because \(A(\gamma_{\text{obj}}, \gamma) \geq 0\), we have

\[
0 \geq -y^\top A(\gamma_{\text{obj}}, \gamma)y = \left(\begin{array}{c} \gamma_{\text{obj}} \\ \gamma \end{array}\right) ^\top \left(\begin{array}{c} -\ell_{\text{obj}} \\ -\ell \end{array}\right),
\]

which implies \(- (\ell_{\text{obj}}, \ell) \in \Gamma^o\). We have shown that \(q_{\text{obj}}(x) - 2t e_{\text{obj}} + (\ell_{\text{obj}}, \ell) \) and \(- (\ell_{\text{obj}}, \ell) \) both lie in \(\Gamma^o\). Then, as \(q_{\text{obj}}(x) - 2t e_{\text{obj}} \in \text{rint}(G(x,t))\), we deduce that \((\ell_{\text{obj}}, \ell) \in \text{span}(G(x,t))\). In particular, \(y^\top A(\gamma_{\text{obj}}, \gamma)y = ((\gamma_{\text{obj}}, \gamma), (\ell_{\text{obj}}, \ell)) = 0\) for all \((\gamma_{\text{obj}}, \gamma) \in \Gamma^0\).

\[\square\]

**Remark 8.** SDP exactness in the context of QMPs was previously studied by Beck [7], Beck et al. [8], Wang and Kılınç-Karzan [49]. Specifically, Beck [7] shows that objective value exactness holds whenever \(k \geq m + 2\). Proposition 2 strengthens both of these results by showing that convex hull exactness holds whenever \(k \geq m\). \[\square\]

### 4.3 The partition problem

We next consider the partition QCQP and its SDP relaxation. Recall the partition QCQP: Given \(a \in \mathbb{R}^n\), we want to minimize

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

Note that \(\text{Opt} = 0\) if and only if the vector \(a\) can be partitioned into two sets of equal weight. Thus, deciding whether \(\text{Opt} = 0\) is NP-hard [30]. In this section, we will first give an explicit description of \(\mathcal{S}_{\text{SDP}}\) under a minor assumption. This explicit \(\mathcal{S}_{\text{SDP}}\) description will then let us conclude that \(\text{conv}(\mathcal{S}) \neq \mathcal{S}_{\text{SDP}}\) under the same minor assumption.

**Assumption 3.** \(a \in \mathbb{R}_{++}^n\) and \(n \geq 2\). \[\square\]

**Remark 9.** Assumption 3 is essentially without loss of generality: It is straightforward to derive a closed form description of \(\mathcal{S}_{\text{SDP}}\) when \(n = 1\). Similarly, one can relate \(\mathcal{S}_{\text{SDP}}\) corresponding to an arbitrary \(a \in \mathbb{R}^n\) with the set \(\mathcal{S}_{\text{SDP}}\) corresponding to some \(a' \in \mathbb{R}_{++}^{n'}\) for \(n' := |\{i \in [n] : a_i \neq 0\}|\). \[\square\]

**Proposition 3.** Suppose Assumption 3 holds. Then,

\[
\Gamma = \left\{ (\gamma_{\text{obj}}, \gamma) \in \mathbb{R} \times \mathbb{R}^n : \gamma_{\text{obj}} a a^\top + \text{Diag}(\gamma) \succeq 0 \right\}, \text{ and}
\]

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

See Appendix A for a proof of this statement.

Recall from [31] 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 Proposition 3. (See Appendix A.)

**Corollary 3.** Suppose Assumption 3 holds. Then, \(\text{Opt}_{\text{SDP}} = 0\) if and only if \(a\) is balanced.

As a consequence of Corollary 3 (and the NP-hardness of deciding whether \(\text{Opt} = 0\) 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 [31].

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

**Corollary 4.** Suppose Assumption 3 holds. Then, \(\text{conv}(\mathcal{S}) \neq \mathcal{S}_{\text{SDP}}\).

The proof of Corollary 4 follows from the observation that \(\text{conv}(\mathcal{S})\) is polyhedral and that \(\mathcal{S}_{\text{SDP}}\) is not polyhedral. See Appendix A for details.
5 Objective value exactness

In this section, we present sufficient conditions for objective value exactness, i.e., the property that $\text{Opt} = \text{Opt}_{\text{SDP}}$. In fact, all of our sufficient conditions imply the stronger condition, which we refer to as optimizer exactness, that the optimizers of the QCQP and its SDP relaxation coincide, i.e.,

$$\min_{(x,t)\in\mathcal{S}_{\text{SDP}}} 2t = \min_{(x,t)\in\mathcal{S}} 2t.$$ 

We begin by presenting sufficient conditions stemming from a primal analysis. These sufficient conditions generalize [49, Theorem 3]. Our second set of sufficient conditions are based on a dual analysis and require the additional assumption that the dual optimum is achieved. These conditions imply further that the optimizers are unique.

5.1 Sufficient conditions based on a primal analysis

We begin by presenting a very general sufficient condition for optimizer exactness.

**Theorem 4.** Suppose Assumption 1 holds. Furthermore, suppose that for all $(x,t) \in \mathcal{S}_{\text{SDP}} \setminus \mathcal{S}$, there exists closed cones $K_1, K_2 \subseteq \mathbb{R}^{1+m}$ and $(x',t') \in \mathbb{R}^{n+1}$ satisfying

$$\begin{cases}
K_1 \subseteq (q(x) - 2te_{\text{obj}})^+ \\
-K_2 \cap (q(x) - 2te_{\text{obj}})^\circ = \{0\} \\
K_1 + K_2 \supseteq \Gamma \\
(x')^T A(\gamma_{\text{obj}}, \gamma)x' = 0, \forall (\gamma_{\text{obj}}, \gamma) \in K_1 \\
\langle A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), x' \rangle - \gamma_{\text{obj}}t' \leq 0, \forall (\gamma_{\text{obj}}, \gamma) \in K_1 \\
t' < 0
\end{cases}.$$  

(9)

Then, optimizer exactness holds, i.e., $\min_{(x,t)\in\mathcal{S}_{\text{SDP}}} 2t = \min_{(x,t)\in\mathcal{S}} 2t$.

**Proof.** Let $(x,t) \in \mathcal{S}_{\text{SDP}} \setminus \mathcal{S}$. It suffices to show that $(x,t) \notin \min_{(x,t)\in\mathcal{S}_{\text{SDP}}} 2t$. Let $K_1, K_2, x', t'$ denote the quantities furnished by the assumption.

We claim that for all $\alpha > 0$ small enough, $(x + \alpha x', t + \alpha t') \in \mathcal{S}_{\text{SDP}}$. Indeed, for all $\alpha > 0$ small enough and $(\gamma_{\text{obj}}, \gamma) \in K_1$, we have

$$\langle \gamma_{\text{obj}}, q(x + \alpha x') - 2(t + \alpha t')e_{\text{obj}} \rangle = \alpha^2 (x')^T A(\gamma_{\text{obj}}, \gamma)x' + 2\alpha \langle A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), x' \rangle - \gamma_{\text{obj}}t' \leq 0.$$

Next, set $\mathcal{B} := K_2 \cap \mathcal{S}^{(1+m)-1}$ so that cone($\mathcal{B}$) = $K_2$. By definition of $K_2$ and $\mathcal{B}$, we have $-\mathcal{B} \cap (q(x) - 2te_{\text{obj}})^\circ = \emptyset$ so that the map

$$\alpha \mapsto \max_{(\gamma_{\text{obj}}, \gamma) \in \mathcal{B}} \langle \gamma_{\text{obj}}, q(x + \alpha x') - 2(t + \alpha t')e_{\text{obj}} \rangle$$

is negative at $\alpha = 0$. Note also that this map is a continuous function of $\alpha$. Then, by continuity, this map is negative for some $\alpha > 0$.

Finally, by linearity and the fact that $K_1 + K_2 \supseteq \Gamma$, we deduce that $(x + \alpha x', t + \alpha t') \in \mathcal{S}_{\text{SDP}}$ for some $\alpha > 0$. This shows $(x,t) \notin \min_{(x,t)\in\mathcal{S}_{\text{SDP}}} 2t$. \[\blacksquare\]
We next recover more concrete sufficient conditions by picking $K_1$ and $K_2$ appropriately. The following corollary recovers the sufficient condition for objective value exactness (in the setting of polyhedral $\Gamma$) presented in [49, Theorem 3].

**Corollary 5.** Suppose Assumption 1 holds and that $\Gamma$ is polyhedral. Furthermore, suppose that for all $(x,t) \in \mathcal{S}_{SDP} \setminus \mathcal{S}$, there exists $(x',t') \in \mathbb{R}^{n+1}$ satisfying

\[
\begin{aligned}
(x')^T A_{\text{obj}} x' &= 0, \quad \forall (\gamma_{\text{obj}}, \gamma) \in \mathcal{F}(x,t) \\
\langle b(\gamma_{\text{obj}}, \gamma), x' \rangle - \gamma_{\text{obj}} t' &\leq 0, \quad \forall (\gamma_{\text{obj}}, \gamma) \in \mathcal{F}(x,t) \\
t' &< 0
\end{aligned}
\]

Then, $\arg\min_{(x,t) \in \mathcal{S}} 2t = \arg\min_{(x,t) \in \mathcal{S}_{SDP}} 2t$.

**Proof.** Let $(x,t) \in \mathcal{S}_{SDP} \setminus \mathcal{S}$. Since $\Gamma$ is polyhedral, we can write $\Gamma := \text{cone}\left\{ (\gamma_{\text{obj}}^{(i)}, \gamma^{(i)}) \right\}_{i \in [T]}$ for a finite set of generators. Take,

\[
K_1 = \text{cone}\left\{ (\gamma_{\text{obj}}^{(i)}, \gamma^{(i)}) : \left\langle (\gamma_{\text{obj}}^{(i)}, \gamma^{(i)}), q(x) - 2te_{\text{obj}} \right\rangle = 0 \right\} = \mathcal{F}(x,t)
\]

and

\[
K_2 = \text{cone}\left\{ (\gamma_{\text{obj}}^{(i)}, \gamma^{(i)}) : \left\langle (\gamma_{\text{obj}}^{(i)}, \gamma^{(i)}), q(x) - 2te_{\text{obj}} \right\rangle < 0 \right\}.
\]

Note that $K_1$ and $K_2$ are polyhedral and thus closed. Moreover, the first three requirements of (9) are satisfied for this choice of $K_1$ and $K_2$. Moreover, note that for every $(\gamma_{\text{obj}}, \gamma) \in \mathcal{F}(x,t) \subseteq \Gamma$ we have $A(\gamma_{\text{obj}}, \gamma) \succeq 0$ and for any $A \succeq 0, x^T Ax = 0$ implies $Ax = 0$. Thus, from $K_1 = \mathcal{F}(x,t)$, we deduce $\langle A(\gamma_{\text{obj}}, \gamma)x, x' \rangle = 0$ for every $(\gamma_{\text{obj}}, \gamma) \in K_1$ so that the last three requirements of (9) coincide with (10).

The following corollary derives a sufficient condition for objective value exactness without the assumption that $\Gamma$ is polyhedral. In words, this assumption supposes that for any $(x,t) \in \mathcal{S}_{SDP} \setminus \mathcal{S}$, there exists a direction $(x',t') \in \mathbb{R}^{n+1}$ such that $q(x + \alpha x') - 2(t + \alpha t')e_{\text{obj}}$ varies only along the line containing $q(x) - 2te_{\text{obj}}$. In particular, by picking $\alpha$ appropriately, we can achieve $q(x + \alpha x') - 2(t + \alpha t')e_{\text{obj}} = 0$.

**Corollary 6.** Suppose Assumption 1 holds. Furthermore, suppose that for all $(x,t) \in \mathcal{S}_{SDP} \setminus \mathcal{S}$, there exists $(x',t') \in \mathbb{R}^{n+1}$ satisfying

\[
\begin{aligned}
(x')^T A_{\text{obj}} x' &= 0, \quad \forall (\gamma_{\text{obj}}, \gamma) \in (q(x) - 2te_{\text{obj}})^\perp \\
\langle A(\gamma_{\text{obj}}, \gamma)x + b(\gamma_{\text{obj}}, \gamma), x' \rangle - \gamma_{\text{obj}} t' &= 0, \quad \forall (\gamma_{\text{obj}}, \gamma) \in (q(x) - 2te_{\text{obj}})^\perp \\
t' &< 0
\end{aligned}
\]

Then, $\arg\min_{(x,t) \in \mathcal{S}_{SDP}} 2t = \arg\min_{(x,t) \in \mathcal{S}} 2t$.

**Proof.** Take $K_1 = (q(x) - 2te_{\text{obj}})^\perp$ and $K_2 = -\text{cone}(q(x) - 2te_{\text{obj}})$. Then, $K_1$ and $K_2$ are both closed convex cones and we can easily observe that the first three requirements in (9) are automatically satisfied for this choice of $K_1$ and $K_2$. The last three requirements in (9) coincide with (11).

### 5.2 Sufficient conditions based on a dual analysis

Next, we give a strengthened sufficient condition for objective value exactness depending on a dual analysis. To this end, we define the concave extended-real valued function $d : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ by

\[
d(\gamma) := \inf_{x \in \mathbb{R}^n} [\gamma, q(x)].
\]
Remark 10. Recall here that by Corollary 1, we can write $\text{Opt}_{\text{SDP}}$ in the saddle-point form $\text{Opt}_{\text{SDP}} = \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma_p} [\gamma, q(x)]$ given in (3). Whence, by coercivity [22, Proposition VI.2.3] we can also write $\text{Opt}_{\text{SDP}} = \sup_{\gamma \in \Gamma_p} d(\gamma)$. □

The following theorem states that if $d(\gamma)$ is maximized at a point $\gamma$ where $A[\gamma] > 0$ (e.g., on the interior of $\Gamma_p$), then optimizer exactness holds. This theorem can be interpreted as the observation that if the dual to (2) in $\mathbb{S}^{n+1}$ has a rank-$n$ optimizer, then (2) has a unique rank-1 solution. This is well-known and has been vastly explored in the literature. We state it as a theorem not because it is new or difficult to prove but because of its importance in deriving additional sufficient conditions (see Corollaries 7 and 8).

**Theorem 5.** Suppose Assumption 1 holds and that $\sup_{\gamma \in \Gamma_p} d(\gamma)$ is achieved at some $\gamma^*$ for which $A[\gamma^*] > 0$ (e.g., $\gamma^* \in \text{int}(\Gamma)$). Then, $\arg \min_{(x,t) \in \mathcal{S}} 2t = \arg \min_{(x,t) \in \text{SDP}} 2t$. Furthermore, the minimizers of these two optimization problems are unique.

**Proof.** It suffices to show that $\min_{(x,t) \in \text{SDP}} 2t$ has a unique solution $(x^*, t^*)$ and that $(x^*, t^*) \in \mathcal{S}$. Let $(x^*, t^*) \in \arg \min_{(x,t) \in \text{SDP}} 2t$ so that $x^* \in \arg \min_x \sup_{\gamma \in \Gamma_p} [\gamma, q(x)]$ and $2t^* = \sup_{\gamma \in \Gamma_p} [\gamma, q(x^*)]$. By the Saddle Point Theorem applied to (3), we have

$$0_n = \nabla_x [\gamma^*, q(x^*)] = 2 \left( A[\gamma^*]x^* + b[\gamma^*] \right).$$

Because $A[\gamma^*] > 0$, which implies that $x^* = -A[\gamma^*]^{-1}b[\gamma^*]$. This proves uniqueness of $(x^*, t^*)$.

Note that there exists $\alpha > 0$ such that $[\gamma^*, \gamma^* + \alpha e_i] \subseteq \Gamma_p$ for all $i \in [m_I]$ and $[\gamma^* - \alpha e_i] \subseteq \Gamma_p$ for all $i \in [m_I + 1, m_I]$. Then, by the Saddle Point Theorem we have

$$q_i(x^*) = \nabla_{\gamma_i} [\gamma^*, q(x^*)] \leq 0, \forall i \in [m_I],$$

and

$$q_i(x^*) = \nabla_{\gamma_i} [\gamma^*, q(x^*)] = 0, \forall i \in [m_I + 1, m_I].$$

We deduce that $q_{\text{obj}}(x^*) \leq \sup_{\gamma \in \Gamma_p} [\gamma, q(x^*)] = 2t^*$. Hence, we conclude that $(x^*, t^*) \in \mathcal{S}$. ■

**Remark 11.** Note that for any $\gamma$ for which $A[\gamma] > 0$, the dual function $d(\gamma)$ is the sum of a linear function $c[\gamma]$ and a concave function $-b[\gamma]^{\top}A[\gamma]^{-1}b[\gamma]$, i.e.,

$$d(\gamma) = -b[\gamma]^{\top}A[\gamma]^{-1}b[\gamma] + c[\gamma].$$

We will use this structure to derive more concrete sufficient conditions ensuring that $d(\gamma)$ is maximized at some point $\gamma^*$ for which $A[\gamma^*] > 0$. □

The following sufficient condition can be interpreted as requiring $-b[\gamma]^{\top}A[\gamma]^{-1}b[\gamma]$ (and hence also $d(\gamma)$) to diverge to $-\infty$ as $\gamma$ approaches a point $\gamma \in \Gamma_p$ for which $A[\gamma] \not\succ 0$.

**Corollary 7.** Suppose Assumption 1 holds and that $\sup_{\gamma \in \Gamma_p} d(\gamma)$ is achieved. Furthermore, suppose that for all $\gamma \in \Gamma_p$, we have

$$A[\gamma] \not\succ 0 \implies \exists v \in \ker(A[\gamma]) \text{ s.t. } \langle v, b[\gamma] \rangle \neq 0.$$

Then, $\arg \min_{(x,t) \in \mathcal{S}} 2t = \arg \min_{(x,t) \in \text{SDP}} 2t$. Furthermore, the minimizers of these two optimization problems are unique.

**Proof.** Let $x^* \in \arg \max_{\gamma \in \Gamma_p} d(\gamma)$. By Theorem 5, it suffices to show that $A[\gamma^*] > 0$. Suppose otherwise, so that $A[\gamma^*] \not\succ 0$. Then, the assumptions of the corollary furnish a $v \in \ker(A[\gamma^*])$ such that $\langle v, b[\gamma^*] \rangle \neq 0$. Let $(x^*, t^*) \in \arg \min_{(x,t) \in \text{SDP}} 2t$. By the Saddle Point Theorem applied to (3), we deduce

$$0 = \langle v, 0_n \rangle = \langle v, \nabla_x [\gamma^*, q(x^*)] \rangle = 2 \langle v, A[\gamma^*]x^* + b[\gamma^*] \rangle \neq 0,$$

a contradiction. ■
Remark 12. Burer and Ye [16] study diagonal QCQPs and show [16, Theorem 1] that objective value exactness holds whenever certain systems of equations are infeasible. Specifically, their sufficient condition for diagonal QCQPs can be rewritten as the condition that for any $i \in [n]$, the system $\{ \gamma \in \Gamma_p, e_i \in \ker(A[\gamma]), b[\gamma]_i = 0 \}$ is infeasible. Corollary 7 generalizes [16, Theorem 1] by considering general matrices $A_i$ as opposed to diagonal matrices considered in [16]. \hfill \Box

Alternatively, one may impose the slightly weaker condition that $-b[\gamma]^T A[\gamma]^{-1} b[\gamma]$ gets “sufficiently steep near points at which $A[\gamma] \neq 0$” compared to $\|(c_1, \ldots, c_m)\|_2$.

Corollary 8. Suppose Assumption 1 holds and that $\sup_{\gamma \in \Gamma_p} d(\gamma)$ is achieved. Furthermore, suppose that for all $\gamma \in \Gamma_p$ such that $A[\gamma] \neq 0$, there exists $\delta \in \mathbb{R}^m$ such that $\gamma_\delta := \gamma + \delta \in \text{int}(\Gamma_p)$ and $\gamma_\delta$ for which $d(\gamma) \leq d(\gamma_\delta)$. As $d(\gamma)$ is concave and $\gamma_\delta, \gamma_\delta^*$, $\gamma_2b$ all lie on a line, it suffices in turn to show that $d(\gamma_\delta) \leq d(\gamma_2b)$. Finally, as $\gamma_\delta$ and $\gamma_2b$ both lie in $\text{int}(\Gamma_p)$, we may expand

$$
(d(\gamma_\delta)) - (d(\gamma_2b)) = (-b[\gamma_\delta]^T A[\gamma_\delta]^{-1} b[\gamma_\delta] + c[\gamma_\delta]) - (-b[\gamma_2b]^T A[\gamma_2b]^{-1} b[\gamma_2b] + c[\gamma_2b])
$$

$$
\leq - \|\delta\|_2 \sqrt{\sum_{i=1}^{m} c_i^2} - \sum_{i=1}^{m} d_i c_i \leq 0.
$$

Applying Theorem 5 concludes the proof. \hfill \Box

6 Applications: Objective value exactness

In this section, we apply the results of Section 5 to random and semi-random QCQPs. Again, these examples offer further evidence that questions of exactness can be treated systematically whenever $\Gamma$, $\Gamma_p$, or $\Gamma^c$ is well-understood. In fact, the results in this section show that the ideas of Section 5 can be applied (at times with additive errors) even when the dual set $\Gamma$, $\Gamma_p$, or $\Gamma^c$ is not known exactly, but only approximately. The random and semi-random QCQPs considered in this section are motivated by recent work [16, 33], which has treated random QCQPs as a testing ground for understanding the strength or explanatory power of various sufficient conditions for objective value exactness.

We will fix $m$, the number of quadratic constraints, and take $n$, the number of variables, to $+\infty$ independently. We will abbreviate “with probability $1 - o(1)$ as $n \to +\infty$” as “asymptotically almost surely” (a.a.s.).

The random and semi-random QCQPs we will consider in this section will involve data generated according to the normalized Gaussian Orthogonal Ensemble (NGOE). We collect some basic facts on the NGOE in the following section.

6.1 Preliminaries on the (normalized) Gaussian Orthogonal Ensemble

Here, we recall the normalized GOE and a few of its basic properties.
Definition 7. Let $A \in \mathbb{S}^n$ be a random matrix where: each diagonal entry $A_{i,i}$ is i.i.d. $N(0, 1/2n)$; each superdiagonal entry $A_{i,j}$ is i.i.d. $N(0, 1/4n)$; and each subdiagonal entry $A_{i,j}$ is defined by symmetry. We will refer to this distribution as the normalized Gaussian Orthogonal Ensemble (NGOE). We will write $A \sim \text{NGOE}(n)$ to denote the fact that $A$ is drawn according to this distribution. □

Remark 13. A different procedure for generating the same distribution is: sample $M \in \mathbb{R}^{n \times n}$ with every entry i.i.d. $N(0, 1/2n)$ and return $A = (M + M^\top)/2$. □

The NGOE is a very well-understood distribution [45]. We will only need a few basic facts. The first two facts state that the NGOE is invariant under various notions of rotation.

Fact 3. Fix $U \in \mathbb{R}^{n \times n}$ orthogonal and let $A \sim \text{NGOE}(n)$. Then, $U^\top A U \sim \text{NGOE}(n)$.

Fact 4. Fix $U \in \mathbb{R}^{k \times k}$ orthogonal and let $A_1, \ldots, A_k \overset{i.i.d.}{\sim} \text{NGOE}(n)$. Define $\tilde{A}_i := \sum_{j=1}^k U_{i,j} A_j$. Then, $\tilde{A}_1, \ldots, \tilde{A}_k \overset{i.i.d.}{\sim} \text{NGOE}(n)$.

Define also the normalized semicircular measure

$$\mu_{\text{usc}} := \frac{2}{\pi} \sqrt{(1 - x^2)_+}.$$ 

The next fact states that the NGOE obeys the semicircle law.

Fact 5. For any $\psi \in C_c^\infty(\mathbb{R})$ and $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P} \left( \left| \int \psi \, d\mu_n - \int \psi \, d\mu_{\text{usc}} \right| > \epsilon \right) = 0.$$ 

Here, $\mu_n$ is the random measure constructed by sampling $A \sim \text{NGOE}(n)$ and setting $\mu_n := \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j(A)}$, where $\delta_{\lambda_j(A)}$ is the Dirac measure at $\lambda_j(A)$.

Finally, we recall that the operator norm of $A \sim \text{NGOE}(n)$ is $\approx 1$ asymptotically almost surely.

Fact 6. Fix $\epsilon > 0$ and let $A \sim \text{NGOE}(n)$. Then, $-\lambda_{\min}(A), \lambda_{\max}(A) \in [1 \pm \epsilon]$ a.a.s.

6.2 Exactness in the fully Gaussian setting

This subsection considers random Euclidean distance minimization problems of the form

$$\inf_{x \in \mathbb{R}^n} \left\{ \|x\|^2 : q_i(x) = 0, \forall i \in [m] \right\}. \quad (12)$$

In words, we are looking for minimum norm solutions to random quadratic systems.

We will sample each quadratic constraint $q_i(x) = x^\top A_i x + 2b_i^\top x + c_i$ independently where $A_i \sim \text{NGOE}(n)$, $b_i \sim N(0, I_n/n)$, and $c_i \sim N(0, 1)$. Here, the normalization on the $A_i$s and $b_i$s are chosen so that $\|A_i\|_2 \approx 1$ and $\|b_i\|_2 \approx 1$.

Below, we will show that for any fixed $m$ and $n \to \infty$, (12) has an exact SDP relaxation a.a.s.. Specifically, we will apply ideas from Corollary 7 to prove:

Proposition 4. Let $A_1, \ldots, A_m \overset{i.i.d.}{\sim} \text{NGOE}(n)$, $b_1, \ldots, b_m \overset{i.i.d.}{\sim} N(0, I_n/n)$ and $c_1, \ldots, c_m \overset{i.i.d.}{\sim} N(0, 1)$ be independent. Then, a.a.s., optimizer exactness holds in (12), i.e., $\arg \min_{(x,t) \in S} 2t = \arg \min_{(x,t) \in S_{\text{SDP}}} 2t$.

We will highlight the very simple geometric ideas underlying the proof of this result and defer proofs of the more technical lemmas to Appendix B.

We will Proposition 4 using Theorem 5; specifically, we will show that $d(\gamma)$ is maximized on the interior of $\Gamma_p$. As a first step, we observe that $\Gamma_p$ contains the unit ball (shrunk by $\epsilon$) a.a.s.. The following lemma follows from an $\epsilon$-net argument, concavity of $\lambda_{\min}(A[\gamma])$ as a function of $\gamma$, and Facts 5 and 6.
Lemma 6. Fix $r \geq 0$ and $\epsilon > 0$. Let $A_1, \ldots, A_m \overset{i.i.d.}{\sim} \text{NGOE}(n)$. Then, a.a.s.,

$$\lambda_{\min}(A[\gamma]) \in [1 - r \pm \epsilon], \ \forall \gamma \in rS^{m-1}. $$

In particular, $\text{int}(\Gamma_P) \supseteq B(0, 1 - \epsilon)$ a.a.s.

Recall Remark 11 that for $\gamma \in \text{int}(\Gamma_P)$, we can write

$$d(\gamma) = -b[\gamma]^T A[\gamma]^{-1} b[\gamma] + c[\gamma].$$

The next lemma notes that the first term in $d(\gamma)$, i.e., $-b[\gamma]^T A[\gamma]^{-1} b[\gamma]$, concentrates to a sphere cap and follows from Fact 5.

Lemma 7. Fix $r \in (0, 1)$ and $\epsilon > 0$. Let $A_1, \ldots, A_m \overset{i.i.d.}{\sim} \text{NGOE}(n)$. Then, a.a.s.,

$$-b[\gamma]^T A[\gamma]^{-1} b[\gamma] \in [\phi(r) \pm \epsilon], \ \forall \gamma \in rS^{m-1},$$

where $\phi(r) := 2(\sqrt{1 - r^2} - 1)$.

We are now ready to prove Proposition 4. The proof will observe that the gradient of $-b[\gamma]^T A[\gamma]^{-1} b[\gamma]$ gets “arbitrarily steep at the boundary of $\Gamma_P$” so that any maximizer of $d(\gamma)$ must lie in $\text{int}(\Gamma_P)$. One may compare the proof of Proposition 4 to Corollary 8.

Proof of Proposition 4. For convenience, let $c \in \mathbb{R}^m$ denote the vector with $i$th coordinate $c_i$. Fix $\delta > 0$ and let $M > 0$ such that $\Pr_{c}[\|c\|_2 \leq M]\geq 1 - \delta/2$. Let $0 < r_1 < r_2 < 1$ and $\epsilon \in (0, 1 - r_2)$ such that

$$\frac{\phi(r_1) - \phi(r_2) - 2\epsilon}{r_2 - r_1} \geq M.$$

In the remainder of the proof, we will condition on the events that $\|c\|_2 \leq M$,

$$\lambda_{\min}(A[\gamma]) \geq 1 - r_2 - \epsilon, \ \forall \gamma \in r_2S^{m-1},$$

$$-b[\gamma]^T A[\gamma]^{-1} b[\gamma] \geq \phi(r_1) - \epsilon, \ \forall \gamma \in r_1S^{m-1},$$

$$-b[\gamma]^T A[\gamma]^{-1} b[\gamma] \leq \phi(r_2) + \epsilon, \ \forall \gamma \in r_2S^{m-1}.$$  

By Lemmas 6 and 7, this holds with probability $1 - \delta$ for all $n$ large enough.

Let $\gamma \in \Gamma_P \setminus B(0, r_2)$ and let $\gamma^{(1)}, \gamma^{(2)}$ denote the projections of $\gamma$ onto $B(0, r_1)$ and $B(0, r_2)$ respectively. We claim that $d(\gamma^{(2)}) \geq d(\gamma)$. By concavity of $d(\gamma)$, it suffices to show that $d(\gamma^{(1)}) \geq d(\gamma^{(2)})$. We compute,

$$d(\gamma^{(1)}) = -b \left[ \gamma^{(1)} \right]^T A \left[ \gamma^{(1)} \right]^{-1} b \left[ \gamma^{(1)} \right] + \langle c, \gamma^{(1)} \rangle$$

$$\geq d(\gamma^{(2)}) + (\phi(r_1) - \epsilon) - (\phi(r_2) + \epsilon) + \langle c, \gamma^{(1)} - \gamma^{(2)} \rangle$$

$$\geq d(\gamma^{(2)}) + (\phi(r_1) - \phi(r_2) - 2\epsilon) - M(r^{(2)} - r^{(1)})$$

We conclude that $d(\gamma)$ is maximized on the interior of $\Gamma_P$.  

6.3 Almost exactness in a semi-random setting

This section considers semi-random QCQPs of the form

$$\inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : \begin{array}{l} q_i(x) = 0, \forall i \in [m] \\ \|x\|_2^2 \leq 1 \end{array} \right\},$$

(13)

20
For notational convenience, define \( q_{m+1}(x) := \|x\|^2 - 1 \).

We will consider the following semi-random model: First, \( A_{\text{obj}}, A_1, \ldots, A_m \) are independently sampled from \( \text{NGOE}(n) \). Then, \( b_{\text{obj}}, b_1, \ldots, b_m \) and \( c_{\text{obj}}, c_1, \ldots, c_m \) are chosen arbitrarily (possibly adversarially depending on the \( A_i \)).

Below, we will show that for any fixed \( m \), (13) has an “almost” exact SDP relaxation a.a.s. Specifically, we will apply ideas from Corollary 6 to prove:

**Proposition 5.** Fix \( \epsilon > 0 \) and let \( A_{\text{obj}}, \ldots, A_m \) i.i.d. \( \sim \) \( \text{NGOE}(n) \). Then, a.a.s., for all \( b_{\text{obj}}, \ldots, b_m \in \mathbb{R}^n \) and \( c_{\text{obj}}, \ldots, c_m \), we have

\[
\text{Opt} \geq \text{Opt}_{\text{SDP}} \geq \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) - \epsilon : q_i(x) \in [\pm \epsilon], \forall i \in [m] \right\}.
\]

In a slight departure from previous notation, we will write our dual vector as \( (\gamma_{\text{obj}}, \gamma, \gamma_{m+1}) \in \mathbb{R}^{1+m+1} \) where \( \gamma_{m+1} \in \mathbb{R} \) corresponds to the constraint \( \|x\|^2 \leq 1 \). As in Section 6.2, we will emphasize the main ideas in the proof of Proposition 5 and leave the proofs of more technical lemmas to Appendix B.

The following lemma says that in this random model, \( \Gamma \) will again converge to the second-order cone. This lemma follows from Lemma 6.

**Lemma 8.** Fix \( r \geq 0 \) and \( \epsilon > 0 \). Let \( A_{\text{obj}}, A_1, \ldots, A_m \) i.i.d. \( \sim \) \( \text{NGOE}(n) \). Then, a.a.s.,

\[
\lambda_{\min}(A(\gamma_{\text{obj}}, \gamma, 1)) \in [1 - r \pm \epsilon], \forall (\gamma_{\text{obj}}, \gamma) \in r\mathbb{S}^m.
\]

In particular, a.a.s.,

\[
\left\{ (\gamma_{\text{obj}}, \gamma, \gamma_{m+1}) : \|A(\gamma_{\text{obj}}, \gamma)\|_2 \leq (1 - \epsilon)\gamma_{m+1} \right\} \subseteq \Gamma \\
\subseteq \left\{ (\gamma_{\text{obj}}, \gamma, \gamma_{m+1}) : \|A(\gamma_{\text{obj}}, \gamma)\|_2 \leq (1 + \epsilon)\gamma_{m+1} \right\}.
\]

The following lemma says that a version of Corollary 6 with errors holds in this setting. This lemma follows from an \( \epsilon \)-net argument along with Fact 5.

**Lemma 9.** Fix \( \epsilon > 0 \) and \( N \in \mathbb{N} \). Then, a.a.s., for every \( (\gamma_{\text{obj}}, \gamma) \in \mathbb{S}^m \), there exists an \( N \)-dimensional vector space \( W \subseteq \mathbb{R}^n \) such that

\[
w^\top A(\gamma_{\text{obj}}, \gamma, 1)w \in [\pm \epsilon] \|w\|^2_2, \forall w \in W.
\]

With Lemmas 8 and 9, we may now prove Proposition 5.

**Proof of Proposition 5.** Without loss of generality, we assume \( \epsilon \in (0, 1/2) \) and \( b_{\text{obj}}, b_1, \ldots, b_m, c_{\text{obj}}, c_1, \ldots, c_m \) are picked so that the SDP relaxation is feasible, i.e.,

\[
\infty > \inf_{x \in \mathbb{R}^n} \sup_{(\gamma, \gamma_{m+1}) \in \Gamma^\circ} \left[ (\gamma, \gamma_{m+1}), q(x) \right].
\]

Let \( x^* \) denote an optimizer of (14) with value \( 2t^* \). Consider the vector \( q(x^*) - 2t^* e_{\text{obj}} \in \mathbb{R}^{1+m+1} \). Without loss of generality, we may assume that \( q(x^*) - 2t^* e_{\text{obj}} \) is both nonzero and on the boundary of \( \Gamma^\circ \). By Lemma 8 and the assumption that \( q(x^*) - 2t^* e_{\text{obj}} \in \text{bd}(\Gamma^\circ) \), we have

\[
\tau := \sqrt{(q_{\text{obj}}(x^*) - 2t^*)^2 + \sum_{i=1}^{m} q_i(x^*)^2} \in [1 \pm \epsilon]q_{m+1}(x^*).
\]

Next, as \( q(x^*) - 2t^* e_{\text{obj}} \) is nonzero, we have that \( 0 < q_{m+1}(x^*) = 1 - \|x^*\|^2 \), i.e., \( \|x^*\|^2 < 1 \). Hence, by definition of \( \tau \), we have \( |\tau| \leq 1 + \epsilon \).
Set \((f_{\text{obj}}, f, f_{m+1}) := \left( \frac{q_{\text{obj}}(x^*) - 2tf^*}{\tau}, \frac{q_{1}(x^*)}{\tau}, \ldots, \frac{q_{m}(x^*)}{\tau}, 1 \right)\) so that \(\|(f_{\text{obj}}, f)\|_2 = 1\).

Note that by Lemma 9, there exists a subspace \(W\) of dimension \(m + 3\) such that
\[
\begin{align*}
 w^\top A(f_{\text{obj}}, f, f_{m+1})w & \in [\pm \epsilon] \|w\|_2^2, \quad \forall w \in W.
\end{align*}
\]

By a dimension counting argument, there exists a unit \(w \in W\) satisfying
\[
\langle A(\gamma_{\text{obj}}, \gamma, \gamma_{m+1})x^* + b(\gamma_{\text{obj}}, \gamma, \gamma_{m+1}), w \rangle = 0, \quad \forall(\gamma_{\text{obj}}, \gamma, \gamma_{m+1}) \in \mathbb{R}^{1+m+1}.
\]

Then, for this vector \(w\) we have
\[
\begin{align*}
 w^\top A(f_{\text{obj}}, f, 1)w & \in [\pm \epsilon], \\
 w^\top A(0, 0, 1)w &= 1, \quad \text{and} \\
 w^\top A(\gamma_{\text{obj}}, \gamma, 1)w & \geq 0, \quad \forall(\gamma_{\text{obj}}, \gamma) \in (1 - \epsilon)S^m.
\end{align*}
\]

Here, the first two relations follow from \(\|w\|_2^2 = 1\). The third relation follows from Lemma 8, which implies that \(A(\gamma_{\text{obj}}, \gamma, 1) \geq 0\) for all \((\gamma_{\text{obj}}, \gamma) \in (1 - \epsilon)S^m\).

Set \(v_{\text{obj}} := w^\top A_{\text{obj}}w\) and \(v \in \mathbb{R}^m\) where \(v_i := w^\top A_iw\) for \(i \in [m]\). Note that by (15), we have
\[
q(x^* + \alpha w) - 2t^* e_{\text{obj}} = (q(x^*) - 2t^* e_{\text{obj}}) + \alpha^2 (v_{\text{obj}}, v, 1).
\]

Then, by the first two lines of (16),
\[
\langle (v_{\text{obj}}, v), (f_{\text{obj}}, f) \rangle = f_{\text{obj}} w^\top A_{\text{obj}}w + \sum_{i=1}^{m} f_i w^\top A_iw
\]
\[
= w^\top A(f_{\text{obj}}, f, 1)w - w^\top w \in [-1 \pm \epsilon].
\]

Next, by the third line of (16), we have \(\|(v_{\text{obj}}, v)\|_2 \leq 1/(1 - \epsilon)\). Set \(\delta_{\text{obj}}, \delta) := (v_{\text{obj}}, v) + (f_{\text{obj}}, f)\). We will argue that \((\delta_{\text{obj}}, \delta)\) is small by bounding its components along \((f_{\text{obj}}, f)\) and orthogonal to \((f_{\text{obj}}, f)\),
\[
\|(\delta_{\text{obj}}, \delta)\|_2^2 \leq \epsilon^2 + \left(\frac{1}{(1 - \epsilon)^2} - (1 - \epsilon)^2\right) = O(\epsilon).
\]

Finally, set \(\bar{x} = x^* + \alpha w\) where \(\alpha = \sqrt{1 - \|x^*\|^2}\) and note that
\[
q(\bar{x}) - 2t^* e_{\text{obj}} = q(x^*) - 2t^* e_{\text{obj}} + (1 - \|x^*\|^2_2) (v_{\text{obj}}, v, 1)
\]
\[
= q(x^*) - 2t^* e_{\text{obj}} + (1 - \|x^*\|^2_2) e_{m+1} + \tau(v_{\text{obj}}, v, 0)
\]
\[
= q(x^*) - 2t^* e_{\text{obj}} - (\tau f_{\text{obj}}, \tau f, \|x^*\|^2_2 - 1)
\]
\[
+ \tau(\delta_{\text{obj}}, \delta, 0) + (1 - \|x^*\|^2_2 - \tau)(v_{\text{obj}}, v, 0)
\]
\[
= \tau(\delta_{\text{obj}}, \delta, 0) + (1 - \|x^*\|^2_2 - \tau)(v_{\text{obj}}, v, 0).
\]

The conclusion then follows from the bounds \(|\tau| \leq (1 + \epsilon), \|(\delta_{\text{obj}}, \delta)\|_2 = O(\sqrt{\epsilon}), \left|1 - \|x^*\|^2_2 - \tau\right| \leq \epsilon\) and \(\|(v_{\text{obj}}, v)\|_2 \leq 1/(1 - \epsilon)\).

**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. *arXiv preprint*, (arXiv:2007.07433), 2020.

[3] A. S. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Math. Program.*, 163:145–167, 2017.

[4] X. Bao, N. V. Sahinidis, and M. Tawarmalani. Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons. *Math. Program.*, 129:129, 2011.

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

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

[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 A. Nemirovski. Lectures on Modern Convex Optimization, volume 2 of *MPS-SIAM Ser. Optim.* SIAM, 2001.

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

[11] G. Blekherman, R. Sinn, and M. Velasco. Do sums of squares dream of free resolutions? *SIAM J. Appl. Algebra Geom.*, 1:175–199, 2017.

[12] S. Burer. A gentle, geometric introduction to copositive optimization. *Math. Program.*, 151:89–116, 2015.

[13] S. Burer and K. M. Anstreicher. Second-order-cone constraints for Extended Trust-Region Subproblems. *SIAM J. Optim.*, 23(1):432–451, 2013.

[14] 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.

[15] S. Burer and B. Yang. The trust region subproblem with non-intersecting linear constraints. *Math. Program.*, 149:253–264, 2014.

[16] 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.

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

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

[19] D. Cifuentes, C. Harris, and B. Sturmfels. The geometry of SDP-exactness in quadratic optimization. *Math. Program.*, 182:399–428, 2020.

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

[21] 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.
[22] I. Ekeland and R. Temam. *Convex analysis and variational problems*, volume 28 of *Classics Appl. Math.* SIAM, 1999.

[23] 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.

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

[25] T. Fujie and M. Kojima. Semidefinite programming relaxation for nonconvex quadratic programs. *J. Global Optim.*, 10(4):367–380, 1997.

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

[27] R. Hildebrand. Spectrahedral cones generated by rank 1 matrices. *J. Global Optim.*, 64:349–397, 2016.

[28] J. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*, volume 1 of *Grundlehren Text Editions*. Springer, 2004.

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

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

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

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

[33] M. Locatelli. KKT-based primal-dual exactness conditions for the Shor relaxation. *arXiv preprint*, 2011.05033, 2020.

[34] A. Megretski. Relaxations of quadratic programs in operator theory and system analysis. In *Systems, Approximation, Singular Integral Operators, and Related Topics*, pages 365–392, 2001.

[35] D. G. Mixon, S. Villar, and R. Ward. Clustering subgaussian mixtures by semidefinite programming. *arXiv preprint*, arXiv:1602.06612, 2016.

[36] 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.

[37] Y. Nesterov. Quality of semidefinite relaxation for nonconvex quadratic optimization. Technical Report 1997019, Université Catholique de Louvain, Center for Operations Research and Econometrics(CORE), 1997.

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

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

[40] R. T. Rockafellar. *Convex Analysis*. Number 28 in Princeton Mathematical Series. Princeton University Press, 1970.

[41] 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.

[42] D. Shamsi, N. Taheri, Z. Zhu, and Y. Ye. Conditions for correct sensor network localization using sdp relaxation. In *Discrete geometry and optimization*, pages 279–301. Springer, 2013.
A Deferred proofs from Section 4

A.1 Deferred proofs from Section 4.1

We compute

\[ \Gamma = \left\{ (\gamma_{\text{obj}}, \gamma_1, \gamma_2) \in \mathbb{R}_+ \times \mathbb{R}^2 : \begin{pmatrix} \gamma_1 \\ \gamma_2/\sqrt{2} \\ \gamma_{\text{obj}} \end{pmatrix} \succeq 0 \right\} \]

\[ = \left\{ (\gamma_{\text{obj}}, \gamma_1, \gamma_2) \in \mathbb{R}^3 : \frac{\gamma_{\text{obj}} + \gamma_1}{2} \geq 0, \gamma_2 \geq \gamma_1 \right\} \]

\[ = \left\{ (\gamma_{\text{obj}}, \gamma_1, \gamma_2) \in \mathbb{R}^3 : \gamma_{\text{obj}} + \gamma_1 \geq \sqrt{(\gamma_{\text{obj}} - \gamma_1)^2 + (\sqrt{2}\gamma_2)^2} \right\}. \]

The expression for \( \Gamma^o \) follows from \( \Gamma \).

Proof of (5). Let \((x, t) \in \mathcal{S}_{\text{SDP}} \setminus \mathcal{S}\) such that \(\mathcal{G}(x, t)\) is a one-dimensional face of \(\Gamma^o\). For notational convenience, let \(\ell_{\text{obj}} = q_{\text{obj}}(x) - 2t\), \(\ell_1 = q_1(x)\) and \(\ell_2 = q_2(x)\). Note that \(\mathcal{G}(x, t) = \mathbb{R}_+ (\ell_{\text{obj}}, \ell_1, \ell_2)\) so that \(\mathcal{F}(x, t) = \mathbb{R}_+ (-\ell_1, -\ell_{\text{obj}}, \ell_2)\). Furthermore, by the assumption that \((\ell_{\text{obj}}, \ell_1, \ell_2)\) is nonzero and on the boundary of \(\Gamma^o\), we have

\[ \mathcal{G}(x, t)^\perp = \text{span} \left\{ \begin{pmatrix} -\ell_1 \\ -\ell_{\text{obj}} \\ \ell_2 \end{pmatrix}, \begin{pmatrix} \ell_2 \\ -\ell_1 \\ -\ell_{\text{obj}} \end{pmatrix} \right\}. \]
We will prove Proposition 3 in the following series of lemmas. Note that the first identity of Proposition 3

We deduce that

\[
\begin{pmatrix}
-\ell_{\text{obj}}(2x_1 - 1) + \ell_2(\sqrt{2}x_2) \\
-\ell_1(2x_2) + \sqrt{2}\ell_2(x_1 - 1)
\end{pmatrix}^\top
\left[
\begin{pmatrix}
-\ell_{\text{obj}}(2x_1 - 1) + \ell_2(\sqrt{2}x_2) \\
-\ell_1(2x_2) + \sqrt{2}\ell_2(x_1 - 1)
\end{pmatrix}
\right]^\top
\]

Expanding, we deduce that

Here, the first two vectors span \(\text{span}(A(f_{\text{obj}}, f))\). The second two vectors correspond to the constraints

\[\langle A(\gamma_{\text{obj}}, \gamma)x, x'\rangle - \gamma_{\text{obj}}t = 0 \text{ for } (\gamma_{\text{obj}}, \gamma) \in \mathcal{G}(x, t)^\perp.\]

Below, we will simplify this expression. By the assumption that \((\ell_{\text{obj}}, \ell_1, \ell_2)\) is nonzero and on the boundary of \(\Gamma^o\), we have

\[-\ell_{\text{obj}} - \ell_1 = \sqrt{(\ell_{\text{obj}} - \ell_1)^2 + (\sqrt{2}\ell_2)^2}\]

where the term within the radical is nonzero. Expanding, we deduce that

\[
\begin{cases}
\ell_{\text{obj}} + \ell_1 < 0 \\
\ell_2^2 - 2\ell_{\text{obj}}\ell_1 = 0
\end{cases}
= \begin{cases}
x_2^2 - 2t + x_1(x_1 - 1) < 0 \\
(x_2^2 - 2tx_1)(x_1 - 1) = 0
\end{cases}.
\]

Note that \((0, 1, 0) \in \Gamma\) so that \(x_1 \in [0, 1]\). If \(x_1 = 1\), then \(\ell_{\text{obj}} < 0, \ell_1 = 0\) and \(\ell_2 = 0\) so that \((x_1, x_2, t) \in \mathcal{S}\), a contradiction. We deduce \(1 - x_1 > 0\) and \(x_2^2 - 2tx_1 = 0\) and that \((\ell_{\text{obj}}, \ell_1, \ell_2) = (x_1 - 1)(2t, x_1, \sqrt{2}x_2)\). Plugging this into (17) gives

\[
\mathcal{R}'(x, t) = \begin{cases}
-2t \\
x_2 \\
-x_2 \\
-t \\
x_1 \\
x_1 \\
-t \\
2tx_1 - 2t
\end{cases},
\]

A.2 Defered proofs from Section 4.3

We will prove Proposition 3 in the following series of lemmas. Note that the first identity of Proposition 3 follows from definition. To prove the second identity of Proposition 3, we will partition \(\Gamma^p\) into \(n+1\) pieces depending on the sign pattern of \(\gamma \in \Gamma^p\).

Note that \(\gamma \in \Gamma^p\) if and only if \(aa^\top + \text{Diag}(\gamma) \succeq 0\). In particular, \(\gamma \in \Gamma^p\) if \(\gamma\) is nonnegative. On the other hand, by the Eigenvalue Interlacing Theorem, \(\gamma \notin \Gamma^p\) if it has at least two negative coordinates. It remains to understand \(\mathcal{N}_i := \Gamma^p \cap \{\gamma \in \mathbb{R}^n : \gamma_i < 0, \gamma_j \geq 0, \forall j \neq i\}\). The next lemma follows from a straightforward application of the Schur Complement Lemma and the Sherman–Morrison Formula.

Lemma 10. Suppose Assumption 3 holds. Then, for any \(i \in [n]\),

\[
\mathcal{N}_i = \left\{ \gamma \in \mathbb{R}^n : \begin{array}{c}
0 > \gamma_i \geq \\
1 + \sum_{j \neq i} \frac{a_i^2}{\gamma_j}
\end{array} \right\}.
\]

Proof. Without loss of generality we assume \(i = n\). For convenience, let \(\bar{\gamma}\) and \(\bar{a}\) denote the first \(n-1\) entries of \(\gamma\) and \(a\) respectively. By Assumption 3, we have that \(\gamma_j > 0\) for all \(j < n\) (as otherwise the \(2 \times 2\) minor of \(aa^\top + \text{Diag}(\gamma)\) corresponding to \((j, n)\) is not positive semidefinite). The Schur Complement Lemma and the Sherman–Morrison Formula then imply that \(\gamma \in \mathcal{N}_n\) if and only if \(\gamma_n < 0, \bar{\gamma} > 0\) and

\[
\gamma_n + a_n^2 \geq a_n^2 \bar{a}^\top (\bar{a} \bar{a}^\top + \text{Diag}(\bar{\gamma}))^{-1} \bar{a}
\]

Rearranging terms completes the proof.
Then decomposing $\Gamma_P = \mathbb{R}_+^n \cup \bigcup_{i \in [n]} \mathcal{N}_i$, we get

$$
\mathcal{S}_{\text{SDP}} = \left\{ (x,t) \in \mathbb{R}^n : \quad 2t \geq \max_{i \in [n]} \sup_{\gamma \in \mathcal{N}_i} \left[ \gamma, q(x) \right] \right\}
$$

It remains to prove the following lemma.

**Lemma 11.** Suppose Assumption 3 holds and let $l \in [n]$. For any $x \in [\pm 1]^n$, we have

$$
\sup_{\gamma \in \mathcal{N}_l} \left[ \gamma, q(x) \right] = (a^T x)^2 + \left( a_1 \sqrt{1 - x_1^2} - \sum_{j \neq 1} a_j \sqrt{1 - x_j^2} \right)^2.
$$

We will need the following two useful facts.

**Lemma 12.** Let $\xi \in \mathbb{R}^k_+$ and $\alpha > 0$, then

$$
\sup_{\zeta \in \mathbb{R}_{++}^k} \left\{ \sum_{i=1}^k \frac{\xi_i}{\zeta_i^2} : \sum_{i=1}^k \zeta_i^2 = \alpha \right\} = -\frac{1}{\alpha} \left( \sum_{i=1}^k \sqrt{-\xi_i} \right)^2.
$$

Proof. Without loss of generality, we may assume $\xi \in \mathbb{R}^k_+$. Then by Cauchy-Schwarz, we have $-\sum_{i=1}^k \xi_i/\zeta_i^2 = -\frac{1}{\alpha} \left( \sum_{i=1}^k \xi_i/\zeta_i^2 \right) \left( \sum_{i=1}^k \zeta_i^2 \right) \geq \frac{1}{\alpha} \left( \sum_{i=1}^k \sqrt{-\xi_i} \right)^2$. Furthermore, equality holds when $\zeta_i^2 \propto \sqrt{\xi_i}$.

**Lemma 13.** Let $\alpha, \beta \geq 0$, then

$$
\sup_{x > 0} \left( \frac{\alpha}{1 + x} - \frac{\beta}{x} \right) = \left( \sqrt{\alpha} - \sqrt{\beta} \right)^2.
$$

Proof. Let $f(x) := \alpha(1 + x)^{-1} - \beta x^{-1}$. Note that $\frac{d}{dx} f(x) = -\alpha (1 + x)^{-2} + \beta x^{-2}$. There are three cases to consider. If $\beta = 0$, then $f(x) = \alpha (1 + x)^{-1}$ and $\sup_{x > 0} \alpha (1 + x)^{-1} = \alpha$. Next, suppose $0 \leq \alpha \leq \beta$, then $\frac{d}{dx} f(x) = -\alpha (1 + x)^{-2} + \beta x^{-2} \geq \beta (x^{-2} - (1 + x)^{-2}) \geq 0$ so that $\sup_{x > 0} f(x) = \lim_{x \to \infty} f(x) = 0$. Finally, suppose $0 < \beta < \alpha$. Note that $f'(x) > 0$ for all $x$ small enough. Similarly, $f'(x) < 0$ for all $x$ large enough. We deduce that $\sup_{x > 0} f(x)$ is achieved. Computing the first-order-necessary conditions, we see that $f(x)$ is maximized at $\frac{\sqrt{\beta}}{\sqrt{\alpha} - \sqrt{\beta}}$ with value $(\sqrt{\alpha} - \sqrt{\beta})^2$.

Proof of Lemma 11. Without loss of generality, $i = n$. Let $b \in \mathbb{R}_+$ where $b_j = x_j^2 - 1$. Let $\hat{\gamma}$ denote the first $n - 1$ entries of $\gamma$. Then,

$$
\sup_{\gamma \in \mathcal{N}_n} \left[ \gamma, q(x) \right] - q_{\text{obj}}(x) = \sup_{\gamma \in \mathcal{N}_n} \left( \gamma, b \right) = \sup_{\hat{\gamma} \in \mathbb{R}_{++}^{n-1}} \sum_{i=1}^{n-1} \gamma_i b_i - \frac{a_n^2 b_n}{1 + \sum_{i=1}^{n-1} a_i^2 / \gamma_i}
$$

$$
= \sup_{\alpha > 0} \left( -\frac{a_n^2 b_n}{1 + \alpha} + \sup_{\hat{\gamma} \in \mathbb{R}_{++}^{n-1}} \left\{ \sum_{i=1}^{n-1} \frac{a_i^2 b_i}{\zeta_i} : \sum_{i=1}^{n-1} \zeta_i = \alpha \right\} \right)
$$

$$
= \sup_{\alpha > 0} \left( \frac{(a_n \sqrt{-b_n})^2}{1 + \alpha} - \frac{1}{\alpha} \left( \sum_{i=1}^{n-1} a_i \sqrt{-b_i} \right)^2 \right)
$$

$$
= \left( a_n \sqrt{-b_n} - \sum_{i=1}^{n-1} a_i \sqrt{-b_i} \right)^2.
$$

Here, the second line follows from a change of variables of $\zeta_i := a_i^2 / \gamma_i$ and $\alpha := \sum_{i=1}^{n-1} \zeta_i$. The third line follows from Lemma 12 and the fourth line follows from Lemma 13.
Proof of Corollary 3. Let \( \gamma \in \Gamma_p \) and \( x \in \mathbb{R}^n \). By convexity of \([\gamma, q(x)]\) in \( x \) and the fact that \( q(x) = q(-x) \), we deduce that \([\gamma, q(0)] \subseteq [\gamma, q(x)]\). We deduce that \( \text{Opt}_{\text{SDP}} = \inf_{x} \sup_{\gamma \in \Gamma_p} |\gamma, q(x)| = \sup_{\gamma \in \Gamma_p} |\gamma, q(0)| \). By Proposition 3, we conclude that

\[
\text{Opt}_{\text{SDP}} = \max_{i \in [n]} \left( a_i - \sum_{j \neq i} a_j \right)^2.
\]

Proof of Corollary 4. Pick an open set \( U \subseteq [\pm 1]^n \) such that

\[
a_1(1 - x_1^2) > \sum_{j > 1} a_j(1 - x_j^2), \quad \forall x \in U.
\]

Then by Proposition 3, for any \( x \in U \), we have \((x, t) \in S_{\text{SDP}}\) if and only if

\[
2t \geq f(x) := (a^\top x)^2 + \left( a_1 \sqrt{1 - x_1^2} - \sum_{j > 1} a_j \sqrt{1 - x_j^2} \right)^2.
\]

Note that \( f(x) \) is smooth on \( U \) and nonlinear (for example note \( \frac{\partial^2 f(x)}{\partial x^2} \neq 0 \)). We conclude that \( S_{\text{SDP}} \neq \text{conv}(S) \) as \( \text{conv}(S) \) is polyhedral.

B Deferred proofs from Section 6

B.1 Useful lemmas

We first recall that under some minor conditions, pointwise convergence implies uniform convergence for convex functions. We extend this statement to show that pointwise a.a.s. convergence implies a.a.s. uniform convergence.

Lemma 14. Let \( \Omega \subseteq \mathbb{R}^n \) be an open set and let \( f : \Omega \to \mathbb{R} \) be a convex function. Suppose \( g_1, g_2, \cdots : \Omega \to \mathbb{R} \) is a sequence of random convex functions such that for all \( x \in \Omega \) and \( \epsilon > 0 \), we have that a.a.s.,

\[
|g_i(x) - f(x)| \leq \epsilon.
\]

Then, for any compact \( C \subseteq \Omega \) and \( \epsilon > 0 \), we have that a.a.s.,

\[
|g_i(x) - f(x)| \leq \epsilon, \quad \forall x \in C.
\]

Proof. Fix \( C \subseteq \Omega \) compact. Without loss of generality, we will assume that \( \epsilon > 0 \) satisfies \( C + B(0, 3\epsilon) \subseteq \Omega \) and that \( f \) is 1-Lipschitz on \( C + B(0, 3\epsilon) \).

Fix a finite net \( \mathcal{N} \subseteq C + B(0, 3\epsilon) \) such that for all \( x \in C + B(0, 2\epsilon) \), we have \( x \in \text{conv} (\mathcal{N} \cap B(x, \epsilon)) \). By our assumption and the fact that \( \mathcal{N} \) is finite, we have that a.a.s., \( |f(x) - g_i(x)| \leq \epsilon \) for all \( x \in \mathcal{N} \). We condition on this event in the remainder of the proof.

For any \( x \in C \), let \( x = \sum_j \lambda_j x_j \) denote the convex decomposition guaranteed by \( x \in \text{conv} (\mathcal{N} \cap B(x, \epsilon)) \). Then,

\[
g_i(x) \leq \sum_j \lambda_j g_i(x_j) \leq \sum_j \lambda_j (f(x_j) + \epsilon) \leq f(x) + 2\epsilon.
\]

Here, the last inequality follows from \( f(x_j) \leq f(x) + \|x - x_j\|_2 \leq f(x) + \epsilon \).

Let \( x \in C \) and \( x' = \mathcal{N} \cap B(x, \epsilon) \). Note that \( y := x' + (x' - x) \in C + B(0, 2\epsilon) \). By construction, there exists \( y' \in \mathcal{N} \cap B(y, \epsilon) \) such that \( g_i(y') \geq g_i(y) \). Finally,

\[
f(x) + 4\epsilon \geq f(y') + \epsilon \geq g_i(y') \geq g_i(y) \geq 2g_i(x') - g_i(x) \geq 2f(x') - \epsilon - g_i(x) \geq 2f(x) - g_i(x) - 4\epsilon.
\]

Therefore, by rearranging and combining, we deduce that a.a.s., \( |g_i(x) - f(x)| \leq 8\epsilon, \forall x \in C \).
Lemma 15. Let \( r \in [-1, 1] \), then
\[
- \int_{\sigma = -1}^{1} \frac{r^2}{1 + r \sigma} \, d\mu_{nsc}(\sigma) = 2(\sqrt{1 - r^2} - 1) = \phi(r).
\]
Proof. We begin by expanding the definition of \( \mu_{nsc} \) and substituting \( \sigma = -\cos \theta \):
\[
- \int_{\sigma = -1}^{1} \frac{r^2}{1 + r \sigma} \, d\mu_{nsc}(\sigma) = -\frac{2}{\pi} \int_{\sigma = -1}^{1} \frac{r^2(1 - \sigma^2)}{1 + r \sigma} \, d\sigma
= -\frac{2}{\pi} \int_{\theta = 0}^{\pi} \frac{r^2 \sin^2 \theta}{1 - r \cos \theta} \, d\theta
= -\frac{2}{\pi} \int_{\theta = 0}^{\pi} \frac{r^2 - r^2 \cos^2 \theta}{1 - r \cos \theta} \, d\theta
= -\frac{2}{\pi} \int_{\theta = 0}^{\pi} \left(1 + \frac{1}{1 - r \cos \theta} \right) \, d\theta - 2
= \frac{2(1 - r^2)}{\pi} \left(\int_{\theta = 0}^{\pi} \frac{1}{1 - r \cos \theta} \, d\theta\right) - 2. \tag{21}
\]
We now focus on the bracketed integral. Perform the change of variables \( \theta = 2\eta \) to get
\[
\int_{\theta = 0}^{\pi} \frac{1}{1 - r \cos \theta} \, d\theta = 2 \int_{\eta = 0}^{\pi/2} \frac{1}{1 - r \cos(2\eta)} \, d\eta. \tag{22}
\]
Recalling the identities \( \cos(2\eta) = 2 \cos^2(\eta) - 1 \) and \( \cos^{-1} \eta = \sec^2 \eta = \tan^2 \eta + 1 = \frac{d}{d\eta} \tan(\eta) \), we then have
\[
\frac{1}{1 - r \cos(2\eta)} = \frac{1}{1 + r - 2r \cos^2 \eta} = \frac{\frac{d}{d\eta} \tan \eta}{(1 + r) \tan^2 \eta + (1 - r)}.
\]
Performing one last change of variables \( t = \tan \eta \) gives
\[
2 \int_{\eta = 0}^{\pi/2} \frac{1}{1 - r \cos(2\eta)} \, d\eta = 2 \int_{\eta = 0}^{\pi/2} \frac{\frac{d}{d\eta} \tan \eta}{(1 + r) \tan^2 \eta + (1 - r)} \, d\eta
= 2 \int_{t = 0}^{\infty} \frac{1}{(1 + r)t^2 + (1 - r)} \, dt
= 2 \frac{\arctan \left( t \sqrt{\frac{1 + r}{1 - r}} \right)}{\sqrt{1 - r^2}} \bigg|_{t = 0}^{\infty}
= \frac{\pi}{\sqrt{1 - r^2}}. \tag{23}
\]
Combining (21), (22), and (23) gives the desired identity.

\[\blacksquare\]

B.2 Deferred proofs from Section 6.2

Proof of Lemma 6. Let \( \Omega = \mathbb{R}^m \) and set \( f(\gamma) := 1 - \|\gamma\|_2 \). Note that \( f \) and \( \lambda_{\min}(A[\gamma]) \) are both concave functions on \( \Omega \). We have \( \lambda_{\min}(A[0]) = 1 = f(0) \). Furthermore, for any nonzero \( \gamma \in \mathbb{R}^m \) and \( \epsilon > 0 \),
\[
\lambda_{\min}(A[\gamma]) = 1 + \|\gamma\|_2 \lambda_{\min} \left( \sum_{i=1}^{m} \frac{\gamma_i}{\|\gamma\|_2} A_i \right) \in 1 + \|\gamma\|_2 [-1 + \epsilon] = [f(\gamma) \pm \|\gamma\|_2 \epsilon], \text{ a.a.s.}
\]
Here, the inclusion holds by Facts 4 and 6. Taking \( C = rS^{m-1} \) and applying Lemma 14.

\[\blacksquare\]
Proof of Lemma 7. Fix \( r \in (0,1) \). Without loss of generality, \( r + 2\epsilon < 1 \). Set \( \Omega := \{ \gamma \in \mathbb{R}^m : \| \gamma \|_2 < r + 2\epsilon \} \).

Let \( \gamma \in \Omega \). Note that we may generate \( A[\gamma] \) and \( b[\gamma] \) via the following equivalent process: Sample \( A \sim \text{NGOE}(n) \) and \( b \sim N(0, I_n/n) \) independently and set \( A[\gamma] := I + rA \) and \( b[\gamma] := rb \). With this notation,

\[-b[\gamma]A[\gamma]^{-1}b[\gamma] = -r^2 \gamma^\top (I + rA)^{-1}b \]

Let \( \hat{A} = \sum_{i=1}^n \sigma_i v_i v_i^\top \) be the eigenvalue decomposition of \( \hat{A} \) and let \( \mu_{\hat{A}} \) denote its Empirical Spectral Distribution. By Lemma 15, we have

\[
\frac{1}{r^2} \left| -b[\gamma]^\top A[\gamma]^{-1} b[\gamma] - \phi(r) \right|
\]

where the last inequality follows from the identity \((I + r\hat{A})^{-1} = \sum_{i=1}^n \frac{1}{1 + r\sigma_i} v_i v_i^\top \) and Cauchy-Schwartz inequality. Note that by Fact 6, for all \( c \), the first term can be bounded by \( r \). We will compute the mean and variance of the first term conditioned on this event. By independence of \( b \) and \( A \),

\[
\mathbb{E}_b \left[ \sum_{i=1}^n \frac{(v_i^\top b)^2 - 1/n}{1 + r\sigma_i} \right] 1 + r\sigma_i \geq c, \forall i \right] = \sum_{i=1}^n \left[ \frac{1}{1 + r\sigma_i} \right] \mathbb{E}_b \left[ (v_i^\top b)^2 - \frac{1}{n} \right] 1 + r\sigma_i \geq c, \forall i \right] = 0, \quad \text{and}
\]

\[
\mathbb{E}_b \left[ \left( \sum_{i=1}^n \frac{(v_i^\top b)^2 - 1/n}{1 + r\sigma_i} \right)^2 \right] 1 + r\sigma_i \geq c, \forall i \right] \leq \left( \frac{1}{c} \right) \mathbb{E}_b \left[ \left( \sum_{i=1}^n (v_i^\top b)^2 - 1/n \right)^2 \right] 1 + r\sigma_i \geq c, \forall i \right] = \frac{2}{en^2}.
\]

In particular, the first term can be bounded by \( \epsilon/(2r^2) \) a.a.s.

For the second term, define the \( C_c^\infty \) function

\[
\psi(x) := \begin{cases} \frac{1}{1 + \delta}, & \text{if } |x| \leq 1 + \delta \\ 0, & \text{if } |x| \geq 1 + 2\delta \\ C_c^\infty, & \text{else.} \end{cases}
\]

By Fact 6, we have that a.a.s. \( \int \frac{1}{1 + \delta} d\mu_{\hat{A}}(\sigma) = \int \psi(\sigma) d\mu_{\hat{A}}(\sigma) \). Applying Fact 5, we conclude that the second term can be bounded by \( \epsilon/(2r^2) \) a.a.s.

Combining the two bounds shows that for any \( \gamma \in \Omega \) and \( \epsilon > 0 \), \( \| b[\gamma]^\top A[\gamma]^{-1} b[\gamma] - \phi(\gamma) \| \leq \epsilon \) a.a.s. Applying Lemma 14 with \( C = r \mathbb{S}^{m-1} \) concludes the proof.

\[\square\]

B.3 Deferred proofs from Section 6.3

Lemma 16. Fix \( \epsilon > 0 \) and \( N \in \mathbb{N} \). Let \( A \sim \text{NGOE}(n) \). Then, a.a.s. there exists a \( N \)-dimensional vector space \( W \subseteq \mathbb{R}^n \) such that

\[ w^\top Aw \in [1 \pm \epsilon] \|w\|^2, \forall w \in W. \]

Proof. Let \( \psi \) denote a \( C_c^\infty \) function from \( \mathbb{R} \) to \( [0,1] \) that takes the value one on \( [1 \pm \epsilon/2] \) and the value zero outside of \([1 \pm \epsilon]\). Note that \( \theta := \int \psi d\mu_{\text{usc}} \) is some positive constant independent of \( n \). Let \( W \) denote the vector space corresponding to the eigenvalues of \( A \) in the range \([1 \pm \epsilon]\). Clearly \( w^\top Aw \in [1 \pm \epsilon] \|w\|^2 \) for all \( w \in W \). It remains to note that by Fact 5, we have a.a.s.

\[ \frac{\dim(W)}{n} = \left\lfloor \{ i \in [n] : \lambda_i(A) \in [1 \pm \epsilon] \} \right\rfloor \geq \int \psi d\mu_{\hat{A}} \geq \int \psi d\mu_{\text{usc}} - \theta/2 = \theta/2 \]

so that \( \dim(W) \geq N \) a.a.s. \[\square\]
Proof of Lemma 9. Let $\mathcal{N}$ denote a finite $\epsilon$-net on $S^m \subseteq \mathbb{R}^{1+m}$. By Lemma 16, a.a.s., for every $(\gamma_{\text{obj}}, \gamma) \in \mathcal{N}$, there exists an $N$ dimensional subspace $W$ such that

$$w^T A(\gamma_{\text{obj}}, \gamma, 1) w \in [\pm \epsilon] \|w\|^2, \forall w \in W.$$ 

Furthermore, by Lemma 6, we have that a.a.s. $\|A(\gamma_{\text{obj}}, \gamma, 0)\|_2 \leq \|A(\gamma_{\text{obj}}, \gamma)\|_2 [1 \pm \epsilon]$ for all $(\gamma_{\text{obj}}, \gamma) \in \mathbb{R}^m$. We condition on these two events.

Now, let $(\gamma_{\text{obj}}, \gamma) \in S^m$ and let $(\gamma'_{\text{obj}}, \gamma') \in \mathcal{N} \cap B((\gamma_{\text{obj}}, \gamma), \epsilon)$. Let $W$ denote the $N$-dimensional subspace guaranteed for $(\gamma'_{\text{obj}}, \gamma')$. Then for all $w \in W$,

$$w^T A(\gamma_{\text{obj}}, \gamma, 1) w = w^T A(\gamma'_{\text{obj}}, \gamma', 1) w + w^T A(\gamma_{\text{obj}} -\gamma'_{\text{obj}}, \gamma - \gamma') w \in [\pm 3\epsilon] \|w\|^2.$$ 

$\blacksquare$