

Quadratic Programs with Hollows

Boshi Yang · Kurt Anstreicher · Samuel Burer

Received: date / Accepted: date

Abstract Let \mathcal{F} be a quadratically constrained, possibly nonconvex, bounded set, and let $\mathcal{E}_1, \dots, \mathcal{E}_l$ denote ellipsoids contained in \mathcal{F} with non-intersecting interiors. We prove that minimizing an arbitrary quadratic $q(\cdot)$ over $\mathcal{G} := \mathcal{F} \setminus \bigcup_{k=1}^l \text{int}(\mathcal{E}_k)$ is no more difficult than minimizing $q(\cdot)$ over \mathcal{F} in the following sense: if a given semidefinite-programming (SDP) relaxation for $\min\{q(x) : x \in \mathcal{F}\}$ is tight, then the addition of l linear constraints derived from $\mathcal{E}_1, \dots, \mathcal{E}_l$ yields a tight SDP relaxation for $\min\{q(x) : x \in \mathcal{G}\}$. We also prove that the convex hull of $\{(x, xx^T) : x \in \mathcal{G}\}$ equals the intersection of the convex hull of $\{(x, xx^T) : x \in \mathcal{F}\}$ with the same l linear constraints. Inspired by these results, we resolve a related question in a seemingly unrelated area, mixed-integer nonconvex quadratic programming.

Keywords nonconvex quadratic programming · semidefinite programming · convex hull

Mathematics Subject Classification (2010) 90C20 · 90C22 · 90C25 · 90C26 · 90C30

1 Introduction

Let

$$\mathcal{F} := \{x \in \mathbb{R}^n : x^T A_i x + 2a_i^T x + \alpha_i \leq 0 \ (i = 1, \dots, m)\}$$

B. Yang

Department of Mathematical Sciences, Clemson University, Clemson, SC, 29634, USA.
E-mail: boshiy@clemson.edu

K. Anstreicher

Department of Management Sciences, University of Iowa, Iowa City, IA, 52242-1994, USA.
E-mail: kurt-anstreicher@uiowa.edu

S. Burer

Department of Management Sciences, University of Iowa, Iowa City, IA, 52242-1994, USA.
E-mail: samuel-burer@uiowa.edu

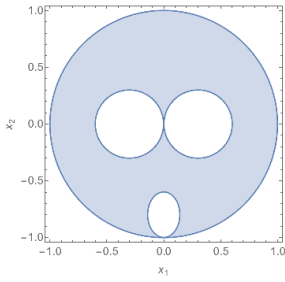


Fig. 1: The unit disk with three non-intersecting hollows (an “alien face”).

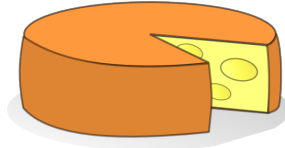


Fig. 2: An ellipsoid cut by two parallel planes with non-intersecting hollows (a “Swiss cheese wheel”). A slice is removed only to depict the hollows inside.

denote a bounded, full-dimensional, quadratically constrained set in \mathbb{R}^n , which may in general be nonconvex. Also, let $\mathcal{E}_k := \{x \in \mathbb{R}^n : x^T W_k x + 2w_k^T x + \omega_k \leq 0\}$, for $k = 1, \dots, l$, denote full-dimensional ellipsoids, each specified by a positive definite symmetric matrix $W_k \in \mathbb{R}^{n \times n}$, vector $w_k \in \mathbb{R}^n$ and scalar $\omega_k \in \mathbb{R}$. If each $\mathcal{E}_k \subseteq \mathcal{F}$ and the interiors of no two ellipsoids intersect, we say that the set

$$\mathcal{H} := \{x \in \mathbb{R}^n : x^T W_k x + 2w_k^T x + \omega_k \geq 0 \ (k = 1, \dots, l)\}$$

induces non-intersecting hollows in \mathcal{F} . Geometrically, the set $\mathcal{G} := \mathcal{F} \cap \mathcal{H}$ results by deleting l disjoint, open ellipsoids from \mathcal{F} . See Figures 1 and 2 for examples.

In this note, we study the relationship between the two optimization problems

$$\begin{aligned} v(q, \mathcal{F}) &:= \min\{q(x) : x \in \mathcal{F}\} \\ v(q, \mathcal{G}) &:= \min\{q(x) : x \in \mathcal{G}\} \end{aligned}$$

where $q(x) := x^T Q x + 2c^T x$ is a general, possibly nonconvex quadratic.

Optimizing $q(\cdot)$ over \mathcal{G} certainly cannot be easier than optimizing $q(\cdot)$ over \mathcal{F} , and at least in some cases appears to be more difficult; for example, \mathcal{G} is nonconvex even when \mathcal{F} is convex. On the other hand, there are reasons to suspect that the complexity of optimizing $q(\cdot)$ over \mathcal{G} should be closely related to that of optimizing $q(\cdot)$ over \mathcal{F} . To optimize over \mathcal{G} , one can first optimize over \mathcal{F} . If the resulting optimal x^* is in \mathcal{G} , then clearly x^* is optimal over \mathcal{G} . On the other hand, if $x^* \notin \mathcal{G}$, then because \mathcal{H} induces non-intersecting hollows in \mathcal{F} , x^* must lie in the interior of \mathcal{F} within exactly one of the deleted ellipsoids. It then follows that $q(\cdot)$ must be convex and that the global minimum over \mathcal{G} is found on the boundary of that deleted ellipsoid, in which case the global minimum can be found by solving an instance of the equality-constrained

trust-region subproblem [10]. Our note formalizes this intuition by studying semidefinite relaxations and reformulations of $v(q, \mathcal{F})$ and $v(q, \mathcal{G})$.

The most basic semidefinite-programming (SDP) relaxation of the set \mathcal{F} is the *Shor relaxation*:

$$\mathcal{S}(\mathcal{F}) := \left\{ (x, X) : \begin{array}{l} A_i \bullet X + 2a_i^T x + \alpha_i \leq 0 \quad (i = 1, \dots, m) \\ Y(x, X) \succeq 0 \end{array} \right\}$$

where

$$Y(x, X) := \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

is an $(n+1) \times (n+1)$ symmetric matrix¹. Note that $\mathcal{S}(\mathcal{F})$ may be an unbounded set even when \mathcal{F} is bounded. On the other hand, the tightest convex relaxation of \mathcal{F} in the space of variables (x, X) is the convex hull

$$\mathcal{C}(\mathcal{F}) := \text{conv} \{ (x, xx^T) : x \in \mathcal{F} \}$$

which is compact because \mathcal{F} is. Clearly $\mathcal{C}(\mathcal{F}) \subseteq \mathcal{S}(\mathcal{F})$, and we call any closed, convex set $\mathcal{R}(\mathcal{F})$ a *valid SDP relaxation of \mathcal{F}* if $\mathcal{C}(\mathcal{F}) \subseteq \mathcal{R}(\mathcal{F}) \subseteq \mathcal{S}(\mathcal{F})$.² In particular, both $\mathcal{C}(\mathcal{F})$ and $\mathcal{S}(\mathcal{F})$ are valid SDP relaxations of \mathcal{F} . Furthermore, any valid SDP relaxation $\mathcal{R}(\mathcal{F})$ of \mathcal{F} gives rise to a relaxation of $v(q, \mathcal{F})$,

$$v(q, \mathcal{R}(\mathcal{F})) := \min \{ Q \bullet X + 2c^T x : (x, X) \in \mathcal{R}(\mathcal{F}) \}$$

such that $v(q, \mathcal{F}) \geq v(q, \mathcal{C}(\mathcal{F})) \geq v(q, \mathcal{R}(\mathcal{F})) \geq v(q, \mathcal{S}(\mathcal{F}))$. In fact, the first inequality is tight. (See also [23] for an alternative proof.)

Proposition 1 *The equality $v(q, \mathcal{F}) = v(q, \mathcal{C}(\mathcal{F}))$ holds for all quadratic functions $q(\cdot)$.*

Proof We have $v(q, \mathcal{C}(\mathcal{F})) \leq v(q, \mathcal{F})$ by construction. To show the reverse inequality, note that because $\mathcal{C}(\mathcal{F})$ is convex and the objective $Q \bullet X + 2c^T x$ is linear, a solution of the problem defining $v(q, \mathcal{C}(\mathcal{F}))$ must occur at an extreme point of $\mathcal{C}(\mathcal{F})$. However all extreme points of $\mathcal{C}(\mathcal{F})$ are of the form (x, xx^T) , $x \in \mathcal{F}$. It follows that $v(q, \mathcal{C}(\mathcal{F})) = Q \bullet xx^T + 2c^T x = q(x)$ for some $x \in \mathcal{F}$, and therefore $v(q, \mathcal{F}) \leq v(q, \mathcal{C}(\mathcal{F}))$.

¹ More formally, $\text{proj}_x(\mathcal{S}(\mathcal{F}))$ is a relaxation of \mathcal{F} , where $\text{proj}_x(\cdot)$ denotes projection onto the x coordinates. We ignore this distinction between $\mathcal{S}(\mathcal{F})$ and $\text{proj}_x(\mathcal{S}(\mathcal{F}))$ to reduce notation.

² To be usable in practice, a valid SDP relaxation $\mathcal{R}(\mathcal{F})$ should have a known positive semidefinite (PSD) representation [16, Section 6.4]. However, it is convenient in this note to consider $\mathcal{R}(\mathcal{F})$ to be a valid SDP relaxation regardless of whether or not an explicit PSD representation for $\mathcal{R}(\mathcal{F})$ is known. We also apply this terminology to $\mathcal{C}(\mathcal{F})$, which in fact may not have an explicit PSD representation—although the PSD constraint is always valid for $\mathcal{C}(\mathcal{F})$.

With respect to \mathcal{G} , we also define $\mathcal{S}(\mathcal{G})$, $\mathcal{C}(\mathcal{G})$, and $\mathcal{R}(\mathcal{G})$ similarly. Specifically, the Shor relaxation is

$$\mathcal{S}(\mathcal{G}) := \left\{ (x, X) : \begin{array}{l} A_i \bullet X + 2a_i^T x + \alpha_i \leq 0 \quad (i = 1, \dots, m) \\ W_k \bullet X + 2w_k^T x + \omega_k \geq 0 \quad (k = 1, \dots, l) \\ Y(x, X) \succeq 0 \end{array} \right\}$$

and we also write $\mathcal{S}(\mathcal{G}) = \mathcal{S}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$, where

$$\mathcal{L}(\mathcal{H}) := \{(x, X) : W_k \bullet X + 2w_k^T x + \omega_k \geq 0 \quad (k = 1, \dots, l)\}.$$

We prove two main results. First, we show that for a valid SDP relaxation $\mathcal{R}(\mathcal{F})$, if the SDP optimal value $v(q, \mathcal{R}(\mathcal{F}))$ equals the original optimal value $v(q, \mathcal{F})$, then defining $\mathcal{R}(\mathcal{G}) := \mathcal{R}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$, the relaxed value $v(q, \mathcal{R}(\mathcal{G}))$ equals $v(q, \mathcal{G})$; see Theorem 1. In words, if an SDP relaxation has no gap over \mathcal{F} , then the SDP relaxation obtained by simply adding the l linear constraints $W_k \bullet X + 2w_k^T x + \omega_k \geq 0$ also has no gap over \mathcal{G} . Second, we establish that the convex hulls $\mathcal{C}(\mathcal{F})$ and $\mathcal{C}(\mathcal{G})$ are related according to the equation $\mathcal{C}(\mathcal{G}) = \mathcal{C}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$; see Corollary 1. That is, the same linear constraints $W_k \bullet X + 2w_k^T x + \omega_k \geq 0$ are precisely what is required to capture $\mathcal{C}(\mathcal{G})$ from $\mathcal{C}(\mathcal{F})$.

We provide two proofs of Corollary 1. The first proof depends on a third result proved in this note, which provides an alternative characterization of $\mathcal{C}(\mathcal{F})$ and has, to our knowledge, not appeared in the literature. The second proof, in contrast, connects better with existing proof techniques for studying convex hulls such as $\mathcal{C}(\mathcal{F})$. Section 3 provides counterexamples showing the necessity of the assumptions that each $\mathcal{E}_k \subseteq \mathcal{F}$ and that the interiors of $\{\mathcal{E}_k\}$ are non-intersecting.

Our results are related to a number of prior works concerning the tightness of SDP relaxations. It is well known that the Shor relaxation is tight in the convex programming case, corresponding to $A_i \succeq 0$, $i = 1, \dots, m$ and $Q \succeq 0$. A classical nonconvex problem with a tight Shor relaxation is the trust-region subproblem (TRS) [10, 14, 19], whose feasible set is the unit ball with arbitrary $q(\cdot)$. The *generalized trust-region subproblem* [21] removes a concentric ball from the feasible set of TRS and yet still has a tight Shor relaxation obtained by adding a single linear constraint to the SDP relaxation of TRS [18, 24]. Other extensions to TRS are also known to have tight SDP relaxations, sometimes with additional valid inequalities added to the Shor relaxation: TRS with a single linear cut [7, 22]; TRS with multiple, non-intersecting linear cuts [9]; TRS with a homogeneous quadratic objective and an additional concentric, ellipsoidal constraint [24]; TRS with an additional ellipsoidal constraint and satisfying various conditions on the quadratic function and/or at local minimizers [2, 13]; and TRS with a general quadratic constraint in place of the unit ball constraint [15]. Many, but not all, of these results are based on characterizing the convex hull $\mathcal{C}(\mathcal{F})$ for the various feasible sets \mathcal{F} under consideration. Characterizing $\mathcal{C}(\mathcal{F})$ has also been studied for some low-dimensional polyhedral \mathcal{F} , e.g., triangles and convex quadrilaterals in \mathbb{R}^2 and tetrahedra in \mathbb{R}^3 [1,

[6]. Other authors have considered valid cuts of the form $\|x-c\|_2 \geq r$ for mixed-integer nonlinear programs [11] and valid linear cuts for the optimization of a convex quadratic over the deletion of an ellipsoid [4].

Our results may be applied uniformly to most of the above problems. For example, since TRS with non-intersecting linear cuts has a bounded feasible region and a tight SDP relaxation [9], our result implies that TRS with non-intersecting linear cuts and hollows also enjoys a tight SDP relaxation.

We note here that the *non-intersecting* assumption can be checked by solving $lm + l(l+1)/2$ trust-region subproblems. For each hollow \mathcal{E}_k , $\mathcal{E}_k \subseteq \mathcal{F}$ if and only if

$$\max \{x^T A_i x + 2a_i^T x + \alpha_i : x^T W_k x + 2w_k^T x + \omega_k \leq 0\} \leq 0$$

for all $i = 1, \dots, m$. For any two hollows \mathcal{E}_j and \mathcal{E}_k , the interiors of \mathcal{E}_j and \mathcal{E}_k do not intersect if and only if

$$\min \{x^T W_j x + 2w_j^T x + \omega_j : x^T W_k x + 2w_k^T x + \omega_k \leq 0\} \geq 0.$$

As an addendum to the results of Section 2, we consider and resolve in Section 4 an open question in the area of mixed-integer quadratic nonconvex programming, namely to characterize the closure of $\text{conv}\{(x, xx^T) : x \in \mathbb{R}^n, x_1 \in \mathbb{Z}\}$. This convex hull is closely related to optimizing a general multivariate quadratic function with a single integer variable. Although the results of Section 2 are not directly applicable, we discuss how the results inspire a conjecture for the convex hull, which we then prove directly.

2 Exact Representations with Hollows

In this section, we present the main results of the note. The first theorem proves that a tight SDP relaxation of $v(q, \mathcal{F})$ gives rise to a tight relaxation of $v(q, \mathcal{G})$.

Theorem 1 *Let $\mathcal{R}(\mathcal{F})$ be a valid SDP relaxation of \mathcal{F} , and let $q(\cdot)$ be given. If $v(q, \mathcal{R}(\mathcal{F})) = v(q, \mathcal{F})$ and \mathcal{H} induces non-intersecting hollows in \mathcal{F} , then $\mathcal{R}(\mathcal{G}) := \mathcal{R}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$ is a valid SDP relaxation of $\mathcal{G} := \mathcal{F} \cap \mathcal{H}$ and $v(q, \mathcal{R}(\mathcal{G})) = v(q, \mathcal{G})$.*

Proof By construction, it is clear that $\mathcal{R}(\mathcal{G})$ is a valid SDP relaxation of \mathcal{G} . Since $q(\cdot)$ is fixed in this proof, we write $v(\cdot) := v(q, \cdot)$ for simplicity. Clearly $v(\mathcal{G}) \geq v(\mathcal{R}(\mathcal{G}))$ since $\mathcal{R}(\mathcal{G})$ is a valid relaxation of \mathcal{G} . So it remains to prove the reverse inequality.

If $v(\mathcal{F})$ is attained at some $x^* \in \text{bd}(\mathcal{F})$, then because \mathcal{H} induces non-intersecting hollows in \mathcal{F} , we have $x^* \in \mathcal{G}$. Hence

$$v(\mathcal{G}) \leq q(x^*) = v(\mathcal{F}) = v(\mathcal{R}(\mathcal{F})) \leq v(\mathcal{R}(\mathcal{G}))$$

as desired. So assume $v(\mathcal{F})$ is attained only at some $x^* \in \text{int}(\mathcal{F})$. Then $Q \succ 0$, and x^* is the unique global minimum of $q(\cdot)$ over \mathbb{R}^n . If $x^* \in \mathcal{G}$ also, then a similar argument as above shows $v(\mathcal{G}) \leq v(\mathcal{R}(\mathcal{G}))$. On the other hand, if $x^* \notin \mathcal{G}$, then $x^* \in \text{int}(\mathcal{E}_k)$ for some k , in which case $Q \succ 0$ implies that $v(\mathcal{G})$ is attained on $\text{bd}(\mathcal{E}_k)$. Hence,

$$\begin{aligned} v(\mathcal{G}) &= \min\{q(x) : x \in \text{bd}(\mathcal{E}_k)\} \\ &= \min\{q(x) : x^T W_k x + 2w_k^T x + \omega_k = 0\} \\ &= \min\{q(x) : x^T W_k x + 2w_k^T x + \omega_k \geq 0\} \\ &= \min\{Q \bullet X + 2c^T x : W_k \bullet X + 2w_k^T x + \omega_k \geq 0, Y(x, X) \succeq 0\} \\ &\leq v(\mathcal{R}(\mathcal{G})) \end{aligned}$$

where the third equality comes from $Q \succ 0$, the fourth equality comes from the fact that the Shor relaxation with one linear constraint is exact (when it is feasible and its optimal value is attained, which occurs in this case because the dual SDP is interior feasible since $Q \succ 0$) [17], and the inequality comes from the fact that $\mathcal{R}(\mathcal{G})$ is a tightening of the preceding feasible set.

Our next theorem establishes a relationship between tight SDP relaxations and the convex hull $\mathcal{C}(\mathcal{F})$. It requires a classical separation result for nonempty closed convex sets.

Lemma 1 (cf. [12]) *Let $K \subseteq \mathbb{R}^p$ be a nonempty, closed, and convex set, and suppose $z \notin K$. Then there exists $s \in \mathbb{R}^p$ such that $s^T z > \sup\{s^T y : y \in K\}$.*

Lemma 1 can also be stated in minimization form as follows: there exists $t \in \mathbb{R}^p$ such that $t^T z < \inf\{t^T y : y \in K\}$ [12, Section 4.1]. More generalized results can also be found in [20, Section 11].

Theorem 2 *Let $\mathcal{R}(\mathcal{F})$ be a valid SDP relaxation of \mathcal{F} . The equality $v(q, \mathcal{F}) = v(q, \mathcal{R}(\mathcal{F}))$ holds for all quadratic functions $q(\cdot)$ if and only if $\mathcal{R}(\mathcal{F}) = \mathcal{C}(\mathcal{F})$.*

Proof The *if* direction follows by Proposition 1. To prove the contrapositive of the *only if* direction, first recall that $\mathcal{R}(\mathcal{F}) \supseteq \mathcal{C}(\mathcal{F})$. If there exists $(\bar{x}, \bar{X}) \in \mathcal{R}(\mathcal{F}) \setminus \mathcal{C}(\mathcal{F})$, then the minimization form of Lemma 1 implies the existence of (\bar{Q}, \bar{c}) and corresponding $\bar{q}(x) = x^T \bar{Q} x + 2\bar{c}^T x$ such that

$$v(\bar{q}, \mathcal{R}(\mathcal{F})) \leq \bar{Q} \bullet \bar{X} + 2\bar{c}^T \bar{x} < v(\bar{q}, \mathcal{C}(\mathcal{F})) = v(\bar{q}, \mathcal{F}).$$

An interesting application of Theorem 2 occurs in [9], where it is shown that $\mathcal{R}(\mathcal{F}) = \mathcal{C}(\mathcal{F})$ for a certain SDP relaxation $\mathcal{R}(\mathcal{F})$ when \mathcal{F} corresponds to the TRS with additional nonintersecting linear constraints. This result is partially extended to the case where linear constraints are allowed to intersect on the boundary of the unit ball defining TRS in [9, Section 5], where it is argued that $v(q, \mathcal{F}) = v(q, \mathcal{R}(\mathcal{F}))$ continues to hold for any $q(\cdot)$. Applying Theorem 2, it follows that in fact $\mathcal{R}(\mathcal{F}) = \mathcal{C}(\mathcal{F})$ must also hold when the linear constraints are permitted to intersect on the boundary of the unit ball.

As a corollary of Theorems 1 and 2, we now state our second main result of the note, which gives a description of the convex hull $\mathcal{C}(\mathcal{G})$ in terms of $\mathcal{C}(\mathcal{F})$ and $\mathcal{L}(\mathcal{H})$.

Corollary 1 *If \mathcal{H} induces non-intersecting hollows in \mathcal{F} , then $\mathcal{C}(\mathcal{G}) = \mathcal{C}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$.*

Proof Applying Theorem 1 with $\mathcal{R}(\mathcal{F}) = \mathcal{C}(\mathcal{F})$ and $\mathcal{R}(\mathcal{G}) = \mathcal{C}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$, we see that $v(q, \mathcal{R}(\mathcal{G})) = v(q, \mathcal{G})$ for any $q(\cdot)$. Then Theorem 2 implies $\mathcal{R}(\mathcal{G}) = \mathcal{C}(\mathcal{G})$.

We finally provide an alternative proof of Corollary 1, which connects better with existing proof techniques involving sets such as $\mathcal{C}(\mathcal{F})$ and $\mathcal{C}(\mathcal{G})$.

Proof We first prove the corollary for $l = 1$. The containment $\mathcal{C}(\mathcal{G}) \subseteq \mathcal{C}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$ is easy because $\mathcal{G} \subseteq \mathcal{F}$ and $\mathcal{C}(\mathcal{G}) \subseteq \mathcal{L}(\mathcal{H})$. For the reverse containment, let (x, X) be an extreme point of $\mathcal{C}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$. If $W_1 \bullet X + 2w_1^T x + \omega_1 > 0$, then (x, X) is in fact an extreme point of $\mathcal{C}(\mathcal{F})$, and so $X = xx^T$. It follows that $(x, X) \in \mathcal{C}(\mathcal{G})$. So assume $W_1 \bullet X + 2w_1^T x + \omega_1 = 0$, and consider the following lemma [22]:

Let V be a symmetric matrix, and suppose $Y \succeq 0$ with $V \bullet Y = 0$ and $\text{rank}(Y) = s$. Then there exists a rank-1 decomposition $Y = \sum_{p=1}^s y^p (y^p)^T$ such that, for all p , it holds that $y^p \neq 0$ and $(y^p)^T V y^p = 0$.

We apply this lemma with

$$V := \begin{pmatrix} \omega_1 & w_1^T \\ w_1 & W_1 \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}$$

and $Y := Y(x, X)$, in which case

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = Y = \sum_{p=1}^s (y^p)(y^p)^T = \sum_{p=1}^s \begin{pmatrix} x_0^p \\ x^p \end{pmatrix} \begin{pmatrix} x_0^p \\ x^p \end{pmatrix}^T$$

with each $y^p \neq 0$, $(y^p)^T V y^p = 0$, $x_0^p \in \mathbb{R}$ and $x^p \in \mathbb{R}^n$. Suppose some $x_0^p = 0$. Then $(x^p)^T W_1 x^p = 0$, which would imply $x^p = 0$ because $W_1 \succ 0$, a contradiction. Hence, in fact each $x_0^p \neq 0$. Then defining $\bar{x}^p := x^p / x_0^p$, we have the convex combination

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \sum_{p=1}^s (x_0^p)^2 \begin{pmatrix} 1 \\ \bar{x}^p \end{pmatrix} \begin{pmatrix} 1 \\ \bar{x}^p \end{pmatrix}^T$$

where each $\bar{x}^p \in \text{bd}(\mathcal{E}_1) = \text{bd}(\mathcal{H}) \subseteq \mathcal{G}$. It follows that $(x, X) \in \mathcal{C}(\mathcal{G})$.

The result for general $l > 1$ can be obtained by induction on l . Consider

$$\mathcal{F}_l := \mathcal{F} \cap \{x \in \mathbb{R}^n : x^T W_k x + 2w_k^T x + \omega_k \geq 0 \ (k = 1, \dots, l-1)\},$$

and

$$\mathcal{H}_l := \{x \in \mathbb{R}^n : x^T W_l x + 2w_l^T x + \omega_l \geq 0\}.$$

With the non-intersecting assumption, \mathcal{H}_l induces a non-intersecting hollow in \mathcal{F}_l , which completes the proof.

We close this section with an observation concerning conditions that guarantee that our various SDP relaxations satisfy *strong duality*, not just enjoy a *zero duality gap*. By assumption, \mathcal{F} is full-dimensional, and except for pathological cases (such as when \mathcal{F} and \mathcal{H} are given by $x^T x \leq 1$ and $x^T x \geq 1$, respectively) $\mathcal{G} = \mathcal{F} \cap \mathcal{H}$ will be full-dimensional as well. In this case, one can prove that both $\mathcal{C}(\mathcal{F})$ and $\mathcal{C}(\mathcal{G})$ are full-dimensional in the space (x, X) . Hence, strong duality holds for any SDP relaxation based on $\mathcal{R}(\mathcal{F})$ or $\mathcal{R}(\mathcal{G})$ in the sense that the primal and dual SDP values equal one another and the dual value is attained. Another variation of strong duality occurs when the dual feasible set contains an interior point, in which case the SDP values are equal and the primal attains its optimal value. Because we have assumed that \mathcal{F} is bounded, this can happen, for example, when a redundant constraint $x^T x \leq \mu$ is added to \mathcal{F} , which in turn translates into a primal constraint $\text{trace}(X) \leq \mu$, which in turn guarantees that the dual SDP has an interior feasible solution.

3 Counterexamples

Theorem 1 assumes that \mathcal{H} induces non-intersecting hollows in \mathcal{F} , i.e., that each $\mathcal{E}_k \subseteq \mathcal{F}$ and all $\mathcal{E}_1, \dots, \mathcal{E}_l$ have disjoint interiors. We now provide two examples showing that both conditions are necessary for Theorem 1, and hence also for Corollary 1.

Counterexample 1 Let $\mathcal{F} := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$ be the unit ball, and define $q(x) := 4x_1^2 - (x_2 + 0.5)^2$. Because of the simplicity of the quadratics involved, it is straightforward to compute $v(q, \mathcal{F}) = -2.25$ with optimal solution $x = (0, 1)$. Moreover, the SDP relaxation over $\mathcal{S}(\mathcal{F})$ is tight with $v(q, \mathcal{S}(\mathcal{F})) = -2.25$.

Now let $\mathcal{H} := \{x \in \mathbb{R}^2 : 2x_1^2 + (x_2 - 0.4)^2 \geq 0.9\}$, and define $\mathcal{G} := \mathcal{F} \cap \mathcal{H}$, which is depicted in Figure 3. Note that the corresponding ellipsoid defined by $2x_1^2 + (x_2 - 0.4)^2 \leq 0.9$ crosses the boundary of \mathcal{F} . It is not difficult to check that $v(q, \mathcal{G}) = -0.25$ with optimal solution $(0, -1)$. However, the SDP relaxation over $\mathcal{S}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$ has optimal value -1.575 , which shows that Theorem 1 does not hold.

It is worthwhile to note that counterexamples similar to Counterexample 1 can also be constructed for the case when an excluded ellipsoid crosses the linear portion of the boundary of a feasible set of problem TRS with an added linear inequality constraint, as discussed at the end of Section 1, e.g., for the set $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, x_2 \leq 0.5\}$.

Counterexample 2 Let $\mathcal{F} := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 4\}$, and define $q(x) := 2x_1^2 + (x_2 - 0.1)^2$, which is strictly convex. One can verify that $v(q, \mathcal{F}) = 0$ with optimal solution $x = (0, 0.1)$. Moreover, the Shor relaxation $\mathcal{S}(\mathcal{F})$ is tight with $v(q, \mathcal{S}(\mathcal{F})) = 0$.

Now let $\mathcal{H} := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \geq 1, x_1^2 + (x_2 - 1)^2 \geq 0.5\}$, and define $\mathcal{G} := \mathcal{F} \cap \mathcal{H}$, which is depicted in Figure 4. Clearly the two ellipsoids defining

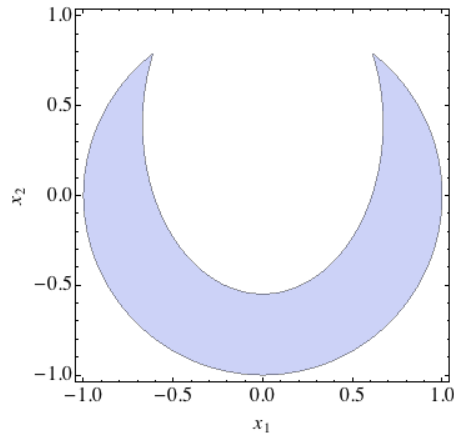


Fig. 3: The feasible region \mathcal{G} of Counterexample 1

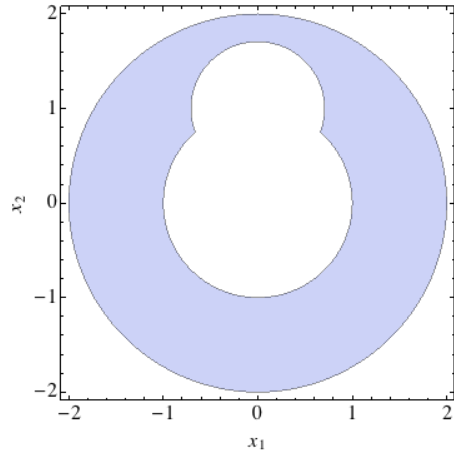


Fig. 4: The feasible region \mathcal{G} of Counterexample 2

\mathcal{H} have a nontrivial intersection. The quadratic optimal value is $v(q, \mathcal{G}) = 1.21$ with solution $(0, -1)$, but the SDP relaxation over $\mathcal{S}(\mathcal{F}) \cap \mathcal{L}(\mathcal{H})$ has optimal value 0.86, which shows that Theorem 1 does not hold.

We remark that Counterexample 2 is written to show that Theorem 1 may fail when two excluded ellipsoids have a nontrivial intersection, but it can also be interpreted as an example where a single excluded ellipsoid intersects the boundary of a nonconvex set \mathcal{F} for which $\mathcal{S}(\mathcal{F}) = \mathcal{C}(\mathcal{F})$. For this second interpretation we can start with $\mathcal{F} := \{x \in \mathbb{R}^2 : 1 \leq x_1^2 + x_2^2 \leq 4\}$, corresponding to a generalized TRS for which the Shor relaxation $\mathcal{S}(\mathcal{F})$ remains tight, and let $\mathcal{H} := \{x \in \mathbb{R}^2 : x_1^2 + (x_2 - 1)^2 \geq 0.5\}$.

4 Resolving an Open Question in Mixed-Integer Nonconvex Quadratic Programming

Burer and Letchford [8] studied optimization problems of the form

$$\begin{aligned} \inf\{q(x) : Ax = b, x \in \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}\}, \\ \inf\{q(x) : Ax = b, x \in \mathbb{Z}_+^{n_1} \times \mathbb{R}_+^{n_2}\}, \end{aligned}$$

where the overall dimension of x is $n := n_1 + n_2$. They showed that solving such problems is closely related to characterizing the closed convex hulls

$$\begin{aligned} \text{MIQ}_{n_1, n_2} &:= \overline{\text{conv}}\{(x, xx^T) : x \in \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}\}, \\ \text{MIQ}_{n_1, n_2}^+ &:= \overline{\text{conv}}\{(x, xx^T) : x \in \mathbb{Z}_+^{n_1} \times \mathbb{R}_+^{n_2}\}, \end{aligned}$$

and the authors introduced several classes of valid inequalities. Note that any inequality valid for the unconstrained case MIQ_{n_1, n_2} is automatically valid for the nonnegative case MIQ_{n_1, n_2}^+ .

For the case when there are no integer variables,

$$\begin{aligned} \text{MIQ}_{0, n} &= \{(x, X) : Y(x, X) \succeq 0\}, \\ \text{MIQ}_{0, n}^+ &= \{(x, X) : Y(x, X) \text{ is completely positive}\}, \end{aligned}$$

where *completely positive* means that $Y(x, X)$ has a Gram factorization in which the factor is component-wise nonnegative; see [3]. Moreover, for the case of a single variable, which is integer, Burer and Letchford showed that

$$\begin{aligned} \text{MIQ}_{1, 0} &= \{(x_1, X_{11}) : X_{11} - (2j-1)x_1 + j(j-1) \geq 0 \forall j \in \mathbb{Z}\}, \\ \text{MIQ}_{1, 0}^+ &= \{(x_1, X_{11}) : X_{11} - (2j-1)x_1 + j(j-1) \geq 0 \forall j \in \mathbb{Z}, x_1 \geq 0\}. \end{aligned}$$

The validity of the constraints $X_{11} - (2j-1)x_1 + j(j-1) \geq 0$ can be seen by linearizing the following single-variable splits, which are valid for $x_1 \in \mathbb{Z}$:

$$x_1 \leq j-1 \vee x_1 \geq j \iff (x_1 - (j-1))(x_1 - j) \geq 0.$$

Finally, the authors proved that two-variable splits capture $\text{MIQ}_{2, 0}$, and later Buchheim and Traversi [5] gave a simplified proof. However, the same two-variable splits are not enough to capture $\text{MIQ}_{2, 0}^+$. The mixed two-variable cases $\text{MIQ}_{1, 1}$ and $\text{MIQ}_{1, 1}^+$ were left open in [8].

For the open case $\text{MIQ}_{1, 1}$ of one integer variable and one continuous variable, it is possible to approximate $\text{MIQ}_{1, 1}$ using the results of Section 2. Consider Figure 5, which depicts one portion of the plane \mathbb{R}^2 with two sets of non-intersecting hollows—the smaller, blue discs and the larger, elongated, orange ellipsoids. Note that the both sets of hollows remove points in \mathbb{R}^2 for which x_1 is fractional, and the second set removes a larger portion than the first. By elongating the hollows even further, one can imagine approximating $\mathbb{Z} \times \mathbb{R}$ better and better. Moreover, in the limit, the non-intersecting ellipsoids will precisely enforce the one-variable splits $x_1 \leq j-1 \vee x_1 \geq j$ for all $j \in \mathbb{Z}$.

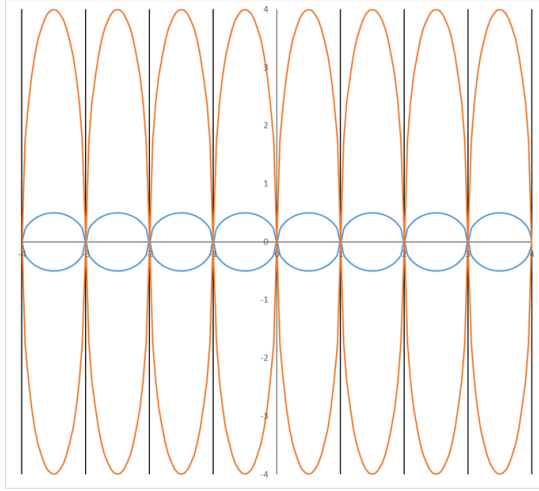


Fig. 5: Hollows to approximate the mixed-integer convex hull $\text{MIQ}_{1,1}$

Using this logic associated with Figure 5, we conjecture that $\text{MIQ}_{1,1}$ is completely characterized by $Y(x, X) \succeq 0$ and the one-variable splits. However, the results of Section 2 are not directly applicable primarily due to the unboundedness of $\mathbb{Z} \times \mathbb{R}$. While we believe a proof based on Section 2 should be possible, we have instead opted to prove the generalization of this conjecture to arbitrary n_2 using a direct approach that closely mimics the alternate proof of Corollary 1 presented in Section 2.

Lemma 2 *Let $x_1 \in \mathbb{R}$ and define $\mu := \inf_{j \in \mathbb{Z}} (x_1 - j)(x_1 - (j - 1))$. Then μ is attained, and if μ is attained at two distinct values of j , then $x_1 \in \mathbb{Z}$.*

Proof The optimal value μ is attained because $(x_1 - j)(x_1 - (j - 1))$ is a convex quadratic in j . Next, completing the square implies that the unconstrained minimum of

$$\begin{aligned} (x_1 - j)(x_1 - (j - 1)) &= x_1^2 - (2j - 1)x_1 + j(j - 1) \\ &= (j - (x_1 + \frac{1}{2}))^2 + [x_1^2 - (x_1 + \frac{1}{2})^2] \end{aligned} \quad (1)$$

occurs at $x_1 + \frac{1}{2}$. Hence, if μ is attained at two distinct values of j , say, $j_1 \neq j_2$, then geometrically, this ensures

$$\frac{1}{2}(j_1 + j_2) = x_1 + \frac{1}{2} \quad \Rightarrow \quad 2x_1 \in \mathbb{Z} \quad \Rightarrow \quad x_1 \in \mathbb{Z} \quad \text{or} \quad x_1 + \frac{1}{2} \in \mathbb{Z}.$$

However, if $x_1 + \frac{1}{2} \in \mathbb{Z}$, then (1) implies that the unconstrained minimum occurs at an integer, which geometrically implies a single unique minimum over $j \in \mathbb{Z}$, a contradiction. Hence $x_1 \in \mathbb{Z}$.

Theorem 3

$$\text{MIQ}_{1,n_2} = \left\{ (x, X) : \begin{array}{l} Y(x, X) \succeq 0 \\ X_{11} - (2j-1)x_1 + j(j-1) \geq 0 \quad \forall j \in \mathbb{Z} \end{array} \right\}.$$

Proof The containment \subseteq is clear. For the reverse containment, first let (\bar{x}, \bar{X}) be an extreme ray of the right-hand side, which implies $\bar{x} = 0$ and $\bar{X} \succeq 0$. Since $(0, \bar{X})$ is extreme, \bar{X} must be rank-1. Hence, by Theorem 1 of [8], (\bar{x}, \bar{X}) is also an extreme ray of MIQ_{1,n_2} . Next, let (\bar{x}, \bar{X}) be an extreme point of the right-hand side, and note that each linearized split can be expressed as

$$(X_{11} - x_1^2) + (x_1 - j)(x_1 - (j-1)) \geq 0.$$

Suppose that all splits are inactive at (\bar{x}, \bar{X}) . Then, by Lemma 2, there exists $\epsilon > 0$ such that $(\bar{X}_{11} - \bar{x}_1^2) + (\bar{x}_1 - j)(\bar{x}_1 - (j-1)) \geq \epsilon$ for all j . Hence, (\bar{x}, \bar{X}) is extreme for the set $\{(x, X) : Y(x, X) \succeq 0\}$, which implies $\bar{X} = \bar{x}\bar{x}^T$. In particular, $\bar{X}_{11} = \bar{x}_1^2$, and hence $(\bar{x}_1 - j)(\bar{x}_1 - (j-1)) \geq \epsilon$ for all j , which is however impossible. Thus, at least one split is active at (\bar{x}, \bar{X}) .

Suppose exactly one split is active. Using the results of [17] on the rank of extreme points for semidefinite systems, (\bar{x}, \bar{X}) satisfies $\bar{X} = \bar{x}\bar{x}^T$, and so $\bar{X}_{11} = \bar{x}_1^2$. All splits then evaluate to $(\bar{x}_1 - j)(\bar{x}_1 - (j-1)) \geq 0$, which implies $\bar{x}_1 \in \mathbb{Z}$ with two active splits, a contradiction.

Finally, by Lemma 2, if two or more splits are active, then $x_1 \in \mathbb{Z}$, which implies $(\bar{x}_1 - j)(\bar{x}_1 - (j-1)) \geq 0$ for all j . Hence either of the active splits implies $\bar{X}_{11} = \bar{x}_1^2$ since $X_{11} - x_1^2 \geq 0$ by semidefiniteness of $Y(x, X) \succeq 0$. It is now clear that $\bar{X} = \bar{x}\bar{x}^T$ because (\bar{x}, \bar{X}) is extreme and because the original variables x_2, \dots, x_n are free. Then (\bar{x}, \bar{X}) is in the left-hand side convex hull MIQ_{1,n_2} (before taking the closure).

Acknowledgements The authors would like to thank two anonymous referees and the associate editor for helpful suggestions and insights.

References

1. Anstreicher, K.M., Burer, S.: Computable representations for convex hulls of low-dimensional quadratic forms. *Mathematical Programming* **124**(1), 33–43 (2010). DOI [10.1007/s10107-010-0355-9](https://doi.org/10.1007/s10107-010-0355-9). URL <http://dx.doi.org/10.1007/s10107-010-0355-9>
2. Beck, A., Eldar, Y.C.: Strong duality in nonconvex quadratic optimization with two quadratic constraints. *SIAM Journal on Optimization* **17**(3), 844–860 (2006). DOI [10.1137/050644471](https://doi.org/10.1137/050644471). URL <http://dx.doi.org/10.1137/050644471>
3. Berman, A., Shaked-Monderer, N.: *Completely Positive Matrices*. World Scientific (2003)
4. Bienstock, D., Michalka, A.: Polynomial solvability of variants of the trust-region subproblem. In: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 380–390. DOI [10.1137/1.9781611973402.28](https://doi.org/10.1137/1.9781611973402.28). URL <http://epubs.siam.org/doi/abs/10.1137/1.9781611973402.28>
5. Buchheim, C., Traversi, E.: On the separation of split inequalities for non-convex quadratic integer programming. *Discrete Optim.* **15**, 1–14 (2015). DOI [10.1016/j.disopt.2014.08.002](https://doi.org/10.1016/j.disopt.2014.08.002). URL <http://dx.doi.org/10.1016/j.disopt.2014.08.002>

6. Burer, S.: A gentle, geometric introduction to copositive optimization. *Mathematical Programming* **151**(1), 89–116 (2015). DOI 10.1007/s10107-015-0888-z. URL <http://dx.doi.org/10.1007/s10107-015-0888-z>
7. Burer, S., Anstreicher, K.M.: Second-order-cone constraints for extended trust-region subproblems. *SIAM Journal on Optimization* **23**(1), 432–451 (2013). DOI 10.1137/110826862. URL <http://dx.doi.org/10.1137/110826862>
8. Burer, S., Letchford, A.N.: Unbounded convex sets for non-convex mixed-integer quadratic programming. *Math. Program.* **143**(1-2, Ser. A), 231–256 (2014). DOI 10.1007/s10107-012-0609-9. URL <http://dx.doi.org/10.1007/s10107-012-0609-9>
9. Burer, S., Yang, B.: The trust region subproblem with non-intersecting linear constraints. *Mathematical Programming* **149**(1-2), 253–264 (2015). DOI 10.1007/s10107-014-0749-1. URL <http://dx.doi.org/10.1007/s10107-014-0749-1>
10. Conn, A., Gould, N., Toint, P.: *Trust Region Methods*. Society for Industrial and Applied Mathematics (2000). DOI 10.1137/1.9780898719857. URL <http://epubs.siam.org/doi/abs/10.1137/1.9780898719857>
11. D’Ambrosio, C., Frangioni, A., Liberti, L., Lodi, A.: On interval-subgradient and no-good cuts. *Operations Research Letters* **38**(5), 341 – 345 (2010). DOI <http://dx.doi.org/10.1016/j.orl.2010.05.010>. URL <http://www.sciencedirect.com/science/article/pii/S0167637710000738>
12. Hiriart-Urruty, J.B., Lemaréchal, C.: *Fundamentals of Convex Analysis*, 1 edn. Springer Berlin Heidelberg (2001). DOI 10.1007/978-3-642-56468-0
13. Jeyakumar, V., Li, G.: Trust-region problems with linear inequality constraints: exact SDP relaxation, global optimality and robust optimization. *Mathematical Programming* **147**(1-2), 171–206 (2014). DOI 10.1007/s10107-013-0716-2. URL <http://dx.doi.org/10.1007/s10107-013-0716-2>
14. Martínez, J.M.: Local minimizers of quadratic functions on euclidean balls and spheres. *SIAM Journal on Optimization* **4**(1), 159–176 (1994). DOI 10.1137/0804009. URL <http://dx.doi.org/10.1137/0804009>
15. Moré, J.J.: Generalizations of the trust region problem. *Optimization Methods and Software* **2**, 189–209 (1993)
16. Nesterov, Y.E., Nemirovskii, A.S.: *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, Philadelphia (1994)
17. Pataki, G.: On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of Operations Research* **23**(2), pp. 339–358 (1998). URL <http://www.jstor.org/stable/3690515>
18. Pong, T., Wolkowicz, H.: The generalized trust region subproblem. *Computational Optimization and Applications* **58**(2), 273–322 (2014). DOI 10.1007/s10589-013-9635-7. URL <http://dx.doi.org/10.1007/s10589-013-9635-7>
19. Rendl, F., Wolkowicz, H.: A semidefinite framework for trust region subproblems with applications to large scale minimization. *Mathematical Programming* **77**(1), 273–299 (1997). DOI 10.1007/BF02614438. URL <http://dx.doi.org/10.1007/BF02614438>
20. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press (1970)
21. Stern, R.J., Wolkowicz, H.: Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM Journal on Optimization* **5**(2), 286–313 (1995). DOI 10.1137/0805016. URL <http://dx.doi.org/10.1137/0805016>
22. Sturm, J.F., Zhang, S.: On cones of nonnegative quadratic functions. *Mathematics of Operations Research* **28**(2), 246–267 (2003). DOI 10.1287/moor.28.2.246.14485. URL <http://dx.doi.org/10.1287/moor.28.2.246.14485>
23. Tunçel, L.: On the slater condition for the sdp relaxations of nonconvex sets. *Operations Research Letters* **29**(4), 181–186 (2001). DOI 10.1016/S0167-6377(01)00093-1. URL [http://dx.doi.org/10.1016/S0167-6377\(01\)00093-1](http://dx.doi.org/10.1016/S0167-6377(01)00093-1)
24. Ye, Y., Zhang, S.: New results on quadratic minimization. *SIAM Journal on Optimization* **14**(1), 245–267 (2003). DOI 10.1137/S105262340139001X. URL <http://dx.doi.org/10.1137/S105262340139001X>