Maximal quadratic-free sets

Gonzalo Muñoz\textsuperscript{1} \cdot Felipe Serrano\textsuperscript{2}

Received: 15 June 2020 / Accepted: 4 November 2021 / Published online: 30 November 2021
© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2021

Abstract
The intersection cut paradigm is a powerful framework that facilitates the generation of valid linear inequalities, or cutting planes, for a potentially complex set $S$. The key ingredients in this construction are a simplicial conic relaxation of $S$ and an $S$-free set: a convex zone whose interior does not intersect $S$. Ideally, such $S$-free set would be maximal inclusion-wise, as it would generate a deeper cutting plane. However, maximality can be a challenging goal in general. In this work, we show how to construct maximal $S$-free sets when $S$ is defined by a general quadratic inequality. Our maximal $S$-free sets are such that efficient separation of a vertex in LP-based approaches to quadratically constrained problems is guaranteed.

Mathematics Subject Classification 90C20 \cdot 90C26 \cdot 90C30

1 Introduction
Cutting planes have been at the core of the development of tractable computational techniques for integer-programming for decades. Their rich theory and remarkable empirical performance have constantly caught the attention of the optimization community, and has recently seen renewed efforts on their extensions to the nonlinear setting.

Consider a generic optimization problem, which we assume to have linear objective without loss of generality:

$$\begin{align*}
\min & \quad c^T s \\
\text{s.t.} & \quad s \in S \subseteq \mathbb{R}^p.
\end{align*}$$

\textsuperscript{1} Universidad de O’Higgins, Rancagua, Chile
\textsuperscript{2} Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany
We assume that $S$ is a nonconvex set which admits a readily available linear relaxation. A common framework for tackling this optimization problem is to first find $\bar{s}$, an extreme point optimum of an LP relaxation of (1.1), and check if $\bar{s} \in S$. If so, then (1.1) is solved. Otherwise, try to find a cutting plane: a linear inequality separating $\bar{s}$ from $S$. Such inequality is used to refine the LP relaxation of (1.1).

One way of finding such a cutting plane is through the intersection cut \cite{4,27,48} framework. We refer the reader to \cite{20} for the necessary background on this procedure. For the purposes of this article, it suffices to know that to compute an intersection cut we need $\bar{s} \notin S$ as above, a simplicial conic relaxation of $S$ with apex $\bar{s}$, and an $S$-free set $C$—a convex set satisfying $\text{int}(C) \cap S = \emptyset$—such that $\bar{s} \in \text{int}(C)$. Here $\text{int}(C)$ denotes the interior of $C$. In this work, we focus on the construction of $S$-free sets $C$ that contain a given $\bar{s} \notin S$ in their interior.

A particularly important case is obtained when (1.1) is a quadratic problem, that is,

$$S = \{ s \in \mathbb{R}^p : s^T Q_i s + b_i^T s + c_i \leq 0, i = 1, \ldots, m \}$$

for certain $p \times p$ matrices $Q_i$, not necessarily positive semi-definite. Note that if $\bar{s} \notin S$, then there exists $i \in \{1, \ldots, m\}$ such that

$$\bar{s} \notin S_i := \{ s \in \mathbb{R}^p : s^T Q_i s + b_i^T s + c_i \leq 0 \},$$

and constructing an $S_i$-free set containing $\bar{s}$ would suffice to ensure separation. Thus, slightly abusing notation, given $\bar{s}$ we focus on a systematic way of constructing $S$-free sets containing $\bar{s}$, where $S$ is defined using a single quadratic inequality:

$$S = \{ s \in \mathbb{R}^p : s^T Q s + b^T s + c \leq 0 \}.$$

As a final note, if we consider the simplest form of intersection cuts, where the cuts are computed using the intersection points of the $S$-free set and the extreme rays of the simplicial conic relaxation of $S$ (i.e., using the gauge function), then the larger the $S$-free set the better the cut. In other words, if two $S$-free sets $C_1, C_2$ are such that $C_1 \subsetneq C_2$, the intersection cut derived from $C_2$ is stronger than the one derived from $C_1$ \cite[Remark 6.6]{21}. Therefore, we aim at computing maximal $S$-free sets, that is, $S$-free sets that are not strictly contained in any other $S$-free set.

**Remark 1.1** A maximal $S_i$-free set will not, in general, be a maximal $\bigcap_{i=1}^{m} S_i$-free set. The scope of this paper is to construct maximal $S_i$-free sets. Constructing maximal $\bigcap_{i=1}^{m} S_i$-free sets is an open problem for $m \geq 2$.

---

\footnote{This citation deals with $S$ being the lattice, but the argument extends trivially to any closed $S$.}
1.1 Motivation

Our main motivation comes from Bienstock et al. [10,11]. In this work, the authors construct maximal outer-product-free sets, which are $S$-free sets with

$$S := \{X \in \mathbb{S}_+^n : \text{rank}(X) = 1\}.$$  

Here $\mathbb{S}_+^n$ is the space of symmetric $n \times n$ positive semi-definite matrices. Using these sets, the authors obtain intersection cuts for any polynomial optimization problem, relying on an extended formulation. In the case of QCQPs, this extended formulation amounts to introducing a new variable $X_{i,j}$ in place of each bilinear term $x_i x_j$ and requiring $X_{i,j} = x_i x_j$.

The derivation of one family of maximal outer-product-free sets involves the construction of maximal $S'$-free sets, where

$$S' := \{(x, y, w, z) \in \mathbb{R}^4 : xy = wz\}.$$  

This set arises from the condition $X_{i,j} X_{k,l} = X_{i,l} X_{k,j}$ that any matrix $X \in S$ must satisfy. The set $S'$ is quadratically-defined, and one family of maximal $S'$-free sets constructed by Bienstock et al. are actually maximal $S'$-free sets. This means that even though a rank-1 requirement involves multiple conditions of the form $X_{i,j} X_{k,l} = X_{i,l} X_{k,j}$, each one of these quadratic requirements, and their corresponding maximal quadratic-free sets, yield $S$-free sets.

The computational results of Bienstock et al. suggest that this path can be an important new avenue for deriving cutting planes in non-convex optimization. Moreover, from all the families of $S$-free sets proposed by Bienstock et al., the ones where maximality was guaranteed (which include the sets constructed from $S'$) exhibited the best empirical performance.

The natural question that arises is whether it is possible to generalize the construction of Bienstock et al. for $S'$ to arbitrary quadratics. In this work we present such generalization.

We also note that the well-known split cuts [21] can be viewed from a “quadratic-free set” perspective. A disjunctive set $D := \{s \in \mathbb{R}^p : a^T s \leq b \lor a^T s \geq b'\}$, with $b < b'$ (or “split”) can be alternatively described as the following quadratic set

$$D = \{s \in \mathbb{R}^p : (a^T s - b)(a^T s - b') \geq 0\}.$$  

The set $D$ is reverse-convex, thus the unique maximal $D$-free set is $\{s \in \mathbb{R}^p : a^T s \geq b \land a^T s \leq b'\}$. Our procedure below is always guaranteed to find a quadratic-free set containing $\tilde{s} \notin D$, therefore in this case we would obtain the unique maximal $D$-free set which would yield a split cut.

1.2 Contributions

The main contribution is an explicit construction of maximal $S$-free sets when $S$ is defined using a non-convex quadratic inequality (Theorems 6.1 and 6.3). We begin
by constructing maximal $S$-free sets when $S$ is defined by a homogeneous quadratic inequality. For non-homogeneous quadratics we rely on the fact that such inequality can be represented using a homogeneous quadratic inequality intersected with an affine equality. As an intermediate step, we also derive maximal $S$-free sets for sets $S$ defined as the intersection of a homogeneous quadratic inequality and a homogeneous linear inequality. This case is of independent interest as well.

In order to show our results, we state and prove a criterion for maximality of $S$-free sets which generalizes a criterion proven by Dey and Wolsey (the ‘only if’ of [24, Proposition A.4]) in the case of maximal lattice-free sets (Definition 3.2 and Theorem 3.1). We also develop a new criterion that can handle a special phenomenon that arises in our setting and also in nonlinear integer programming: the boundary of a maximal $S$-free set may not even intersect $S$. Instead, the intersection might be “at infinity”. We formalize this in Definition 3.3 and show the criterion in Theorem 3.3.

The main results of this paper were presented in [39] without proofs. Here, we present an extended version of this work: we include all proofs and intermediate steps that yield more nuanced results than in [39].

1.3 Literature review

The history of intersection cuts dates back to the 60’s. They were originally introduced in the nonlinear setting by Tuy [48] for the problem of minimizing a concave function over a polytope. Later on, they were introduced in integer programming by Balas [4] and have been largely studied since. The more modern form of these cuts deduced from an arbitrary $S$-free set is due to Glover [27], although the term $S$-free was coined by Dey and Wolsey [24].

While the origin of intersection cuts was in nonlinear optimization, most developments have been in mixed-integer linear programming. See e.g. [6,19,22] for in-depth analyses of the relation of intersection cuts using maximal $\mathbb{Z}^n$-free sets and the generation of facets of conv(S). We also refer the reader to [2,3,5,14,23,29] and references therein. For extensions to the mixed-integer conic case see [1,8,9,31,32,36,37,44,49]. Recently, Towle and Luedtke [47] proposed a method for constructing valid cutting planes with a similar approach to intersection cuts, but allowing $\bar{x}$ to be outside the $S$-free set.

Lately, there has been several developments of intersection cuts in a nonlinear setting. Fischetti et al. [25] applied intersection cuts to bilevel optimization. Bienstock et al. [10,11] studied outer-product-free sets; these can be used for generating intersection cuts for polynomial optimization when using an extended formulation. Serrano [42] showed how to construct a concave underestimator of any factorable function and from them one can build intersection cuts for factorable mixed integer nonlinear programs. Fischetti and Monaci [26] constructed bilinear-free sets through a bound disjunction and McCormick inequalities [35].

Of all these intersection cut approaches in nonlinear settings, the only one that ensures maximality of the corresponding $S$-free sets is the work of Bienstock et al. [10,11]. While their approach can also be used to generate cutting planes in our quadratic setting, the definition of $S$ differs: Bienstock et al. use a moment-based extended
formulation of polynomial optimization problems [33,34,45] and from there define $S$ as the set of matrices which are positive semi-definite and of rank 1, which the authors refer to as outer-products. While this approach is our main motivation, the relationship with our methods in terms of the resulting cutting planes is unclear at this point: in a quadratic setting, the approach of Bienstock et al. would compute a cutting plane in extended space of dimension proportional to $n^2$, whereas our approach can construct a maximal $S$-free set in the original space. The quadratic dimension increase can be a drawback in some applications, however stronger cuts can be derived from extended formulations in some cases [12]. We refer to the survey [13] for other efforts of extending cutting planes to the nonlinear setting.

Moving away from the traditional intersection cut approach, one could rely on a second-order cone instead of a simplicial cone to compute a cut, by means of the results in [16,31]. Another alternative can be conceived from a recent result showing that the convex hull of a single quadratic constraint over a polytope is second-order cone representable [41]. However, it is not clear at this point how to use this approach to efficiently generate cutting planes. These questions are subject of future work.

1.4 Notation

We mostly follow standard notation. $\| \cdot \|$ is the euclidean norm in $\mathbb{R}^n$ and given a positive definite matrix $A$, we denote $\| \cdot \|_A$ the norm defined by $\sqrt{x^TAx}$. $D_1(0)$ denotes the euclidean unit sphere, i.e., $D_1(0) = \{ y \in \mathbb{R}^n : \|y\| = 1 \}$. Given a vector $v$ and a set $C$, we denote the distance between $v$ and $C$ as $\text{dist}(v, C) = \inf_{x \in C} \|v - x\|$. We denote the set $\{ v + x : x \in C \}$ as $v + C$. For vectors $\{ v^1, \ldots, v^k \} \subseteq \mathbb{R}^n$, we denote by $\text{lin}\{ v^1, \ldots, v^k \}$ the subspace generated by them. We denote an inequality $\alpha^T x \leq \beta$ by $(\alpha; \beta)$. We denote $\text{proj}_H C$ the projection of $C \subseteq \mathbb{R}^n$ onto the subspace $H$. $(\cdot)^c$, $\text{conv}(\cdot)$, $\text{int}(\cdot)$ and $\text{rec}(\cdot)$ denote the complement, convex hull, interior and recession cone of a set, respectively.

1.5 Outline

The rest of the paper is organized as follows. In Sect. 2 we describe our high-level strategy and state our main results. In Sect. 3 we prove our maximality criteria of $S$-free sets. In Sect. 4 we show how to construct maximal $S$-free sets when $S$ is defined by a homogeneous quadratic function. Section 5 presents the construction of maximal $S$-free sets when $S$ is defined by a homogeneous quadratic function and a homogeneous linear inequality constraint. The construction of a maximal $S$-free set when $S$ is defined by a non-homogeneous quadratic function is presented in Sect. 6. In Sect. 7, we present an explicit cut computation example, discuss computational aspects of generating cuts using our maximal quadratic-free sets, and show a maximal quadratic-free set which is not generated by our approach. In Sect. 8 we conclude.
2 Preliminaries

2.1 Canonical representations of quadratic sets

To ease the presentation of our construction of maximal quadratic-free sets, it is convenient to first map a general quadratic set to what we refer to as a canonical representation. Consider a non-convex quadratic set

\[ S = \{ s \in \mathbb{R}^p : s^T Q s + b^T s + c \leq 0 \}. \]

Using, e.g., homogenization and diagonalization, any set \( S \) can be mapped via an affine one-to-one map to a set of either of the following forms

\[ \left\{ (x, y, z) \in \mathbb{R}^{n+m+l} : \|x\| \leq \|y\| \right\} \quad (2.1) \]

\[ \left\{ (x, y, z) \in \mathbb{R}^{n+m+l} : \|x\| \leq \|y\|, \ a^T x + d^T y + h^T z = -1 \right\} \quad (2.2) \]

Whether the set is mapped to (2.1) or (2.2) depends on whether the quadratic inequality defining \( S \) is homogeneous or not. Since the transformation is affine and one-to-one, constructing a maximal (2.1)-free or maximal (2.2)-free set is sufficient to generate a maximal \( S \)-free set.

**Remark 2.1** The choice of the map bringing a non-homogeneous quadratic to the form (2.2) is not unique. Different choices can produce different vectors \( a, d, h \) and, perhaps surprisingly, different maximal \( S \)-free sets. The effect of different transformations is subject of ongoing work.

Additionally, below (see Remark 3.2) we argue why the \( z \) variables can be ignored in the construction of (2.1)-free or (2.2)-free sets. Thus, throughout this work we explicitly construct maximal \( S^h \)- and \( S^g \)-free sets, where

\[ S^h = \{ (x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\| \} \quad (2.3) \]

\[ S^g = \{ (x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\|, \ a^T x + d^T y = -1 \}. \quad (2.4) \]

We refer to these as canonical representations of a quadratic set. The reader might wonder why we use norms in our canonical representations as opposed to their squares. This is a subtle point which we discuss in Appendix A.1.

As we mentioned in the introduction, we also assume we have a point \( \bar{s} \notin S \) which we would like our maximal \( S \)-free set to contain in its interior. We note that by transforming the set \( S \) into a canonical representation, \( \bar{s} \) will be mapped to a point \( (\bar{x}, \bar{y}) \) such that

\[ \|\bar{x}\| > \|\bar{y}\| \quad \text{in the case of } S^h, \quad (2.5) \]

\[ \|\bar{x}\| > \|\bar{y}\| \land a^T \bar{x} + d^T \bar{y} = -1 \quad \text{in the case of } S^g. \quad (2.6) \]

Our \( S^h \)- and \( S^g \)-free sets contain such point \( (\bar{x}, \bar{y}) \) in its interior.
Another simplification that makes the analysis cleaner is the fact that it cannot be that \(a = d = 0\). Therefore, through a further change of variables we may assume without loss of generality that \(\max\{\|a\|, \|d\|\} = 1\).

Our overall strategy begins with the construction of maximal \(S^h\)-free sets. Then, before moving to \(S^g\), we study the following set

\[
S_{\leq 0} = \{(x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\|, a^T x + d^T y \leq 0\}.
\] (2.7)

This set is interesting on its own, and provides an important intermediate step into our construction of maximal \(S^g\)-free sets. As before, we require our maximal \(S_{\leq 0}\)-free sets to contain a point \((\bar{x}, \bar{y})\) in its interior, where

\[
\|\bar{x}\| > \|\bar{y}\| \land a^T \bar{x} + d^T \bar{y} \leq 0.
\] (2.8)

### 2.2 Summary of constructed sets

Using the techniques we develop in Sect. 3, we show the following.

**Theorem 2.1** Consider \(S^h\) defined as (2.3), \(\lambda \in D_1(0)\) and

\[
C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} : \lambda^T x \geq \|y\|\}.
\]

Then, \(C_\lambda\) is a maximal \(S^h\)-free set. Furthermore, if \(\lambda = \frac{x}{\|x\|}\) with \((\bar{x}, \bar{y})\) satisfying (2.5), \(C_\lambda\) contains \((\bar{x}, \bar{y})\) in its interior.

A geometrical interpretation of \(C_\lambda\) is as follows. Note that \(C_\lambda\) has a lineality space given by \(L = \{(x, y) : \lambda^T x = 0, y = 0\}\), and projecting \(C_\lambda\) onto \(L^\perp\) yields a second-order cone.

We also remark that each \(\lambda \in D_1(0)\) in Theorem 2.1 generates a different maximal \(S^h\)-free set; we only use \(\lambda = \frac{x}{\|x\|}\) to ensure that \(C_\lambda\) contains \((\bar{x}, \bar{y})\) in its interior. This implies that, even from a fixed canonical representation, we can generate more than one \(S^h\)-free set. These are infinitely-many if \(n \geq 2\). In the latter case, we can even generate infinitely-many \(S^h\)-free sets that contain \((\bar{x}, \bar{y})\) in their interiors by considering values of \(\lambda\) in a neighborhood of \(\frac{x}{\|x\|}\).

For the intermediate set \(S_{\leq 0}\) the situation is more complex. We show

**Theorem 2.2** Consider a non-convex set \(S_{\leq 0}\) defined as (2.7), such that \(\max\{\|a\|, \|d\|\} = 1\), and \((\bar{x}, \bar{y})\) satisfying (2.8). Define \(\lambda = \frac{x}{\|x\|}\). Then

\[
C_{\phi_{\lambda,a,d}} := \{(x, y) \in \mathbb{R}^n : \phi_{\lambda,a,d}(y) \leq \lambda^T x\}
\] (2.9)

with

\[
\phi_{\lambda,a,d}(y) = \begin{cases} 
\|y\|, & \text{if } \lambda^T a\|y\| + d^T y \leq 0 \\
\sqrt{\left(\|y\|^2 - (d^T y)^2\right)(1 - (\lambda^T a)^2)} - d^T y \lambda^T a, & \text{otherwise.}
\end{cases}
\] (2.10)
is maximal $S_{\leq 0}$-free and contains $(\bar{x}, \bar{y})$ in its interior.

**Remark 2.2** In this last theorem (and in the following theorem as well) we are only considering $\lambda = \bar{x}/\|\bar{x}\|$. This is a slight simplification of our results; later we argue that having $\lambda$ in a neighborhood of $\bar{x}/\|\bar{x}\|$ suffices to ensure maximality, and in some cases we only need $\lambda \in D_1(0)$, much like in Theorem 2.1. In particular, this implies that infinitely-many maximal $S_{\leq 0}$-free can be generated when $n \geq 2$.

At this point it may be unclear that $C_{\phi_{\lambda,a,d}}$ is maximal $S_{\leq 0}$-free and where the function $\phi_{\lambda,a,d}$ comes from. Moreover, it is not even clear why $C_{\phi_{\lambda,a,d}}$ is convex. In Sect. 5 we provide all technical details that clarify this, including a proof of convexity of $\phi_{\lambda,a,d}$. Surprisingly, the construction of the set $C_{\phi_{\lambda,a,d}}$ and its maximality proof, involve a case distinction which yields two completely different paths that reach the same expression.

The same case distinction is also needed for $S^g$, although they do not reach the same expression. These are summarized in the following Theorem.

**Theorem 2.3** Consider a non-convex set $S^g$ defined as (2.4), satisfying that $\max\{\|a\|, \|d\|\} = 1$, and $(\bar{x}, \bar{y})$ satisfying (2.6). Let $H = \{(x, y) \in \mathbb{R}^{n+m} : a^T x + d^T y = -1\}$, and $\lambda = \bar{x}/\|\bar{x}\|$. Then, either $C_{\phi_{\lambda,a,d}}$ defined in (2.9) or

$$C_{\phi_{\lambda,a,d}}^g = \begin{cases} (x, y) : \phi_{\lambda,a,d} \left( y - \frac{d}{1 - \|d\|^2} \right) \leq \lambda^T \left( x + \frac{a}{1 - \|d\|^2} \right) \text{ otherwise} \end{cases}$$

is maximal $S^g$-free with respect to $H$ and contains $(\bar{x}, \bar{y})$ in its interior.

First of all, note that in this case we use the term maximal with respect to $H$. We define this properly in Definition 3.1, and it is meant to avoid trivialities that can arise since $S^g$ is contained in the hyperplane $H$.

As in the case of Theorems 2.2, 2.3 is a slight simplification of more nuanced results. In Sect. 6 we provide all these details. In particular, we show precisely in which case $C_{\phi_{\lambda,a,d}}$ or $C_{\phi_{\lambda,a,d}}^g$ is the desired set. We also note that a similar discussion to Remark 2.2 holds for the role of $\lambda$ in $S^g$-free sets.

### 3 Techniques for proving maximality

In this section we describe some sufficient conditions to prove that a convex set $C$ is maximal $S$-free. We remark that in this section we do not assume $S$ comes from a quadratic inequality, and all results hold in general.

Since we apply our criteria to various sets, some of which are contained in a hyperplane (such as $S^g$), we need the following definition. See also [5].

**Definition 3.1** Given $S, C, H \subseteq \mathbb{R}^n$ where $S$ is closed, $C$ is closed and convex and $H$ is an affine hyperplane, we say that $C$ is $S$-free with respect to $H$ if $C \cap H$ is $S \cap H$-free w.r.t the induced topology in $H$. We say $C$ is maximal $S$-free with respect to $H$, if for any $C' \supseteq C$ that is $S$-free with respect to $H$ it holds that $C' \cap H \subseteq C \cap H$. 

Springer
The next definition is motivated by the sufficient (and necessary) condition for a full dimensional $\mathbb{Z}^n$-free set to be maximal: it must be a polyhedron containing a point of $\mathbb{Z}^n$ in the relative interior of each facet [21, Theorem 6.18]. For a general set $S$, non-polyhedral maximal $S$-free sets may exist, which may not even have facets. However, a similar property can be captured.

**Definition 3.2** Given a convex set $C \subseteq \mathbb{R}^n$ and a valid inequality $(\alpha; \beta)$, we say that a point $x_0 \in \mathbb{R}^n$ exposes $(\alpha; \beta)$ with respect to $C$ if: $\alpha^T x_0 = \beta$ and if $(\gamma; \delta)$ is any other non-trivial valid inequality for $C$ such that $\gamma^T x_0 = \delta$, then there exists a $\mu > 0$ such that $\gamma = \mu \alpha$ and $\beta = \mu \delta$. We omit saying “with respect to $C$” if it is clear from context. We also say that $(\alpha; \beta)$ is exposed by $x_0$ or that $x_0$ is an exposing point of $(\alpha; \beta)$.

To get some intuition, if $C$ is a polyhedron and $x \in C$ exposes an inequality, then that inequality is a facet and $x$ is in the relative interior of the facet. Points that expose inequalities are also called smooth points [28]. A related concept is that of blocking points [7]. Some remarks are in order.

**Remark 3.1** If there is a point exposing a valid inequality of $C$, then $C$ is full dimensional. The reader should keep this in mind throughout the paper.

In general, a point $x \notin C$ can expose a valid inequality of $C$. Consider $C = \{x \in \mathbb{R}^2 : x_1 + x_2 \geq 1\}$. Then, $(0, 0) \notin C$ exposes $x_1 + x_2 \geq 0$.

The name “exposing” comes from the concept of exposed point from convex analysis. For the interested reader, we would like to point that if $C$ is convex with 0 in its interior and the valid inequality $\alpha^T x \leq 1$ has an exposing point, then $\alpha$ is an exposed point of the polar of $C$.

**Lemma 3.1** Let $K, K' \subseteq \mathbb{R}^n$ be convex sets such that $K \subseteq K'$. If $\alpha^T x \leq \beta$ is valid for $K$, not valid for $K'$, and exposed by $x_0 \in K$ with respect to $K$, then $x_0 \in \text{int}(K')$.

**Proof** As $x_0 \in K$ exposes $\alpha^T x \leq \beta$, it holds that $\alpha^T x_0 = \beta$ and, thus, $x_0$ is in the boundary of $K$. Suppose $x_0$ is not in the interior of $K'$. Then it must be in the boundary of $K'$ and there is a valid inequality for $K'$, $\gamma^T x \leq \delta$, such that $\gamma^T x_0 = \delta$.

As $K \subseteq K'$, $\gamma^T x \leq \delta$ is also valid for $K$. Given that $(\gamma; \delta)$ is tight at $x_0$ and $x_0$ exposes $(\alpha; \beta)$, we conclude that there is a $\mu > 0$ such that $\gamma = \mu \alpha$ and $\beta = \mu \delta$. However, since $\alpha^T x \leq \beta$ is not valid for $K'$, it follows that $\gamma^T x \leq \delta$ cannot be valid for $K'$. This contradiction proves the claim.

This last lemma yields our first maximality criterion.

**Theorem 3.1** Let $S \subseteq \mathbb{R}^n$ be a closed set and $C \subseteq \mathbb{R}^n$ an $S$-free set. Assume that $C = \{x \in \mathbb{R}^n : \alpha^T x \leq \beta, \forall (\alpha, \beta) \in \Gamma\}$ for some $\Gamma \subseteq \mathbb{R}^{n+1}$ such that for every $(\alpha, \beta) \in \Gamma$ there is an $x \in S \cap C$ that exposes $(\alpha; \beta)$. Then, $C$ is maximal $S$-free.

**Proof** It suffices to show that for every $\bar{x} \notin C$, $S \cap \text{int}(\text{conv}(C \cup \{\bar{x}\}))$ is nonempty. Let $\bar{x} \notin C$ and let $(\alpha; \beta) \in \Gamma$ be a separating inequality, i.e., $\alpha^T \bar{x} > \beta$. Let $C' = \text{conv}(C \cup \{\bar{x}\})$. By hypothesis, there is an $x_0 \in S \cap C$ that exposes $(\alpha; \beta)$. Since $(\alpha; \beta)$ is valid for $C$ and not for $C'$, Lemma 3.1 implies that $x_0 \in \text{int}(C')$. \( \square \)
With minor modifications one can also get the following sufficient condition for maximality with respect to a hyperplane.

**Theorem 3.2** Let $S \subseteq \mathbb{R}^n$ be a closed set, $H$ be an affine hyperplane, and $C \subseteq \mathbb{R}^n$ be an $S$-free set. Assume that $C = \{ x \in \mathbb{R}^n : \alpha^T x \leq \beta, \forall (\alpha, \beta) \in \Gamma \}$ for some $\Gamma \subseteq \mathbb{R}^{n+1}$ such that for every $(\alpha, \beta) \in \Gamma$ there is an $x \in S \cap C \cap H$ that exposes $(\alpha; \beta)$. Then, $C$ is maximal $S$-free with respect to $H$.

There is another phenomenon that does not occur when $S = \mathbb{Z}^n$. If $S$ is a quadratic set, the inequalities of a maximal $S$-free set might not be exposed by any point of $S$. For instance, consider $S = \{(x, y) \in \mathbb{R}^2 : x^2 + 1 \leq y^2 \}$. The boundary of $S$ is a hyperbola with asymptotes $x = \pm y$. Thus, $C = \{(x, y) \in \mathbb{R}^2 : x \geq |y|\}$ is a maximal $S$-free set, because its inequalities are asymptotes of $S$, but they are not exposed by points of $S$. This phenomenon also occurs when $S = \mathbb{Z}^n \cap K$, with $K$ convex [38]. However, in that case, maximal $S$-free sets are polyhedral and their constructions rely on the concept of a facet (see e.g. [38, Theorem 3.2]). In our case, we extend the definition of what it means for an inequality to be exposed in order to handle a situation like the one above. We do this by interpreting that asymptotes are exposed “at infinity”.

**Definition 3.3** Given a convex set $C \subseteq \mathbb{R}^n$ with non-empty recession cone and a valid inequality $\alpha^T x \leq \beta$, we say that a sequence $(x_n) \subseteq \mathbb{R}^n$ exposes $(\alpha; \beta)$ at infinity with respect to $C$ if

- $\|x_n\| \to \infty$,
- $\frac{x_n}{\|x_n\|} \to d \in \text{rec}(C)$,
- $d$ exposes $\alpha^T x \leq 0$ with respect to $\text{rec}(C)$, and
- there exists $y$ such that $\alpha^T y = \beta$ and $\text{dist}(x_n, y + \text{lin}\{d\}) \to 0$.

As before, we omit saying “with respect to $C$” if it is clear from context.

Using this definition, we can prove an analogous result to Lemma 3.1.

**Lemma 3.2** Let $K, K' \subseteq \mathbb{R}^n$ be convex sets such that $K \subseteq K'$. If $\alpha^T x \leq \beta$ is valid for $K$, not valid for $K'$, and exposed at infinity by $(x_n)_n$ with respect to $K$, then there exists a $k$ such that $x_k \in \text{int}(K')$.

**Proof** Suppose that for all $k$, $x_k$ is not in the interior of $K'$. Then, for each $k$ there exists a non-trivial valid inequality for $K'$, $\gamma_k^T x \leq \delta_k$, such that $\gamma_k^T x_k \geq \delta_k$. We can assume without loss of generality that $\|(\gamma_k, \delta_k)\| = 1$. Hence, going through a subsequence if necessary, there exist $\gamma \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ such that $\gamma_k \to \gamma$ and $\delta_k \to \delta$ when $k \to \infty$ and $\|(\gamma, \delta)\| = 1$. Note that the inequality $(\gamma; \delta)$ is valid for $K'$. The idea is to show that $(\gamma; \delta)$ is the same as $(\alpha; \beta)$.

Since $(\gamma; \delta)$ is valid for $K' \supseteq K$, it follows that $\gamma^T x \leq 0$ is valid for $\text{rec}(K)$. As $d = \lim_{k \to \infty} \frac{x_k}{\|x_k\|} \in \text{rec}(K)$ (see Definition 3.3), we have that $\gamma^T d \leq 0$. On the other hand,

$$\frac{\delta_k}{\|x_k\|} \leq \frac{\gamma_k^T x_k}{\|x_k\|} \implies 0 \leq \gamma^T d.$$

@ Springer
We conclude that $\gamma^T d = 0$. As $d$ exposes $\alpha^T x \leq 0$ with respect to $\text{rec}(K)$, there exists a $\mu \geq 0$ such that $\gamma = \mu \alpha$. Note that we cannot conclude that $\mu > 0$ since, at this point, we do not know that $(\gamma; \delta)$ is a non-trivial inequality (e.g. it could be $0^T x \leq 1$).

Let $y$ be such that $\alpha^T y = \beta$ and $\text{dist}(x_k, y + \text{lin}(d)) \to 0$, which exists by Definition 3.3. Let $w_k = x_k - d^T x_k d$. We have that

$$\text{dist}(x_k, y + \text{lin}(d)) = \text{dist}(x_k - y, \text{lin}(d))$$

$$= \|x_k - y - d^T(x_k - y)d\| = \|w_k - (y - d^T yd)\|.$$ 

Thus, $w_k \to y - d^T yd$ as $k \to \infty$.

Since each $(\gamma_k; \delta_k)$ is valid for $K'$, $\gamma_k^T d \leq 0$. Additionally, for large enough $k$ it must hold that $d^T x_k > 0$. Therefore,

$$\delta_k \leq \gamma_k^T x_k = \gamma_k^T(d^T x_k d + w_k) \leq \gamma_k^T w_k.$$ 

Computing the limit when $k \to \infty$ we get,

$$\delta \leq \mu \alpha^T(y - d^T yd) = \mu \alpha^T y = \mu \beta.$$ 

If $\mu = 0$, then $\gamma = 0$ and $\delta \leq 0$. As $\|(\gamma, \delta)\| = 1$, it follows that $\delta = -1$, which cannot be since $(\gamma; \delta)$ is a valid inequality for $K'$ and $K'$ is, by hypothesis, non-empty. We conclude that $\mu > 0$ and that $\mu \alpha^T x \leq \mu \beta$ is valid for $K'$, which implies that $\alpha^T x \leq \beta$ is valid for $K'$, contradicting the hypothesis of the theorem.

With the previous results it is straightforward to prove the following generalization of Theorem 3.2.

**Theorem 3.3** Let $S \subseteq \mathbb{R}^n$ be a closed set, $H$ be an affine hyperplane, and $C \subseteq \mathbb{R}^n$ be an $S$-free set. Assume that $C = \{x \in \mathbb{R}^n : \alpha^T x \leq \beta, \forall (\alpha; \beta) \in \Gamma\}$ for some $\Gamma \subseteq \mathbb{R}^{n+1}$ such that for every $(\alpha; \beta) \in \Gamma$ there is, either, an $x \in S \cap C \cap H$ that exposes $(\alpha; \beta)$, or sequence $(x_n)_n \subseteq S \cap H$ that exposes $(\alpha; \beta)$ at infinity. Then, $C$ is maximal $S$-free with respect to $H$.

Another useful result in the study of maximal $S$-free sets is the following (see also [21, Lemma 6.17]). It states that in some cases we can project $S$ into a lower dimensional space and find maximal sets that are free for the projection.

**Theorem 3.4** Let $C$ be a full dimensional closed set with lineality space $L$. Let $S \subseteq \mathbb{R}^n$ be closed. Then, $C$ is maximal $S$-free if and only if $(C \cap L^\perp)$ is maximal $\text{cl}(\text{proj}_{L^\perp} S)$-free.

**Proof** ($\Rightarrow$) If $C \cap L^\perp$ is not maximal, let $K \subseteq L^\perp$ be a $\text{cl}(\text{proj}_{L^\perp} S)$-free set that contains it. Then, $K + L \supseteq C$. Since $C$ is maximal $S$-free, there exists an $x \in S$ such that $x \in \text{int}(K + L) = \text{int}(K) + \text{int}(L)$ ([40, Corollary 6.6.2]). That is, $x = k + \ell$ with $k \in \text{int}(K)$ and $\ell \in L$. Thus, $x - \ell \in K \subseteq L^\perp$ which implies that $x - \ell \in \text{proj}_{L^\perp} S$ and contradicts the fact that $K$ is $\text{cl}(\text{proj}_{L^\perp} S)$-free.

($\Leftarrow$) By contradiction, suppose that $C$ is not maximal $S$-free and let $K \supseteq C$ be a closed $S$-free set. Then $K \cap L^\perp \supseteq C \cap L^\perp$, which implies that $K \cap L^\perp$ is not $\text{cl}(\text{proj}_{L^\perp} S)$-free.
This implies that $\exists \tilde{s} \in \text{cl}(\text{proj}_{L^\perp} S) \cap \text{int}(K \cap L^\perp)$. Moreover, we can further assume $\tilde{s} \in \text{proj}_{L^\perp} S \cap \text{int}(K \cap L^\perp)$, as any sequence contained in $\text{proj}_{L^\perp} S$ converging to an element of $\text{cl}(\text{proj}_{L^\perp} S) \cap \text{int}(K \cap L^\perp)$ must have an element in $\text{proj}_{L^\perp} S \cap \text{int}(K \cap L^\perp)$.

By the definition of orthogonal projection, there must exist $\tilde{s} \in \text{proj}_{L^\perp} S \cap \text{int}(K \cap L^\perp)$, as any sequence contained in $\text{proj}_{L^\perp} S$ converging to an element of $\text{cl}(\text{proj}_{L^\perp} S) \cap \text{int}(K \cap L^\perp)$ must have an element in $\text{proj}_{L^\perp} S \cap \text{int}(K \cap L^\perp)$.

By the definition of orthogonal projection, there must exist $s \in S$ and $\ell \in L$ such that $\tilde{s} = s - \ell$. Thus, we obtain $s - \ell \in \text{int}(K \cap L^\perp)$, i.e.

$$s \in \text{int}(K \cap L^\perp) + L.$$ 

Since the lineality space of $K$ must contain $L$, we conclude $s \in \text{int}(K)$; a contradiction with $K$ being $S$-free.

\[\Box\]

**Remark 3.2** As we anticipated, Theorem 3.4 allows us to ignore the $z$ variables in (2.1) and (2.2). In the former, we can clearly see that $C \times \mathbb{R}^l$ is maximal (2.1)-free, where $C$ is any maximal $S^h$-free. The story for (2.2) is similar. If $h \neq 0$, then $C \times \mathbb{R}^l$ is maximal $S$-free (with respect to the corresponding hyperplane), where $C$ is any maximal $S^h$-free.

### 4 Maximal quadratic-free sets for homogeneous quadratics

We first construct maximal $S^h$-free sets and prove Theorem 2.1. Recall that

$$S^h = \{(x, y) \in \mathbb{R}^{n+m} : \|x\| - \|y\| \leq 0\}.$$ 

A simple way of obtaining an $S^h$-free set is via a concave underestimator of the function $f(x, y) = \|x\| - \|y\|$ (see [42]). It is not hard to see that $\tilde{x}^T x \|\tilde{x}\| - \|y\|$ is a concave underestimator of $f(x, y)$, which is tight at $\tilde{x} \neq 0$. Thus,

$$\{(x, y) \in \mathbb{R}^{n+m} : \frac{\tilde{x}^T x}{\|\tilde{x}\|} - \|y\| \leq 0\} \supseteq S^h.$$ 

The set on the left is reverse-convex, which yields the $S^h$-free set

$$C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} : \lambda^T x \geq \|y\|\}, \quad (4.1)$$

where $\lambda = \frac{\tilde{x}}{\|\tilde{x}\|}$. $C_\lambda$ turns out to be maximal, even if we consider any other $\lambda \in D_1(0)$. We note that in [10], the authors use a similar technique and reformulate a 4-variable homogeneous quadratic condition of outer-product-free sets in the form $\|x\| \leq \|y\|$ (see the discussion in Sect. 1.1). This allows them to construct maximal outer-product-free sets that are of the form $C_\lambda$.

We now prove that $C_\lambda$ is maximal $S^h$-free. The idea is to exploit that every inequality describing $C_\lambda$ is exposed by a point in $S^h \cap C_\lambda$ and use Theorem 3.1. We begin with a lemma whose technical proof is in the appendix. We recall that a function is sublinear if and only if it is convex and positively homogeneous.
Lemma 4.1 Let \( \phi : \mathbb{R}^n \to \mathbb{R} \) be a sublinear function, \( \lambda \in D_1(0) \), and let

\[
C = \{(x, y) : \phi(y) \leq \lambda^T x\}.
\]

Let \((\bar{x}, \bar{y}) \in C\) be such that \(\phi\) is differentiable at \(\bar{y}\) and \(\phi(\bar{y}) = \lambda^T \bar{x}\). Then \((\bar{x}, \bar{y})\) exposes the valid inequality \(-\lambda^T x + \nabla \phi(\bar{y})^T y \leq 0\).

We now prove Theorem 2.1, which we restate here to aid the reader.

Theorem 2.1 Consider \( S^h \) defined as (2.3), \( \lambda \in D_1(0) \) and \( C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} : \lambda^T x \geq \|y\|\} \). Then, \( C_\lambda \) is a maximal \( S^h \)-free set. Furthermore, if \( \lambda = \frac{\bar{x}}{\|\bar{x}\|} \) with \((\bar{x}, \bar{y})\) satisfying (2.5), \( C_\lambda \) contains \((\bar{x}, \bar{y})\) in its interior.

Proof \( S^h \)-freeness follows by construction. To show that \( C_\lambda \) is maximal, we first notice that

\[
C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} : -\lambda^T x + \beta^T y \leq 0, \ \forall \beta \in D_1(0)\}.
\]

Since the norm function \( \|\cdot\| \) is sublinear, differentiable everywhere but in the origin, and \( \|\beta\| = 1 = \lambda^T \lambda \), Lemma 4.1 shows that \((\lambda, \beta) \in S^h \cap C_\lambda\) exposes \((-\lambda, \beta); 0\). From Theorem 3.1 we conclude that \( C_\lambda \) is maximal \( S^h \)-free. The fact that \((\bar{x}, \bar{y}) \in \text{int}(C_\lambda)\) when \( \lambda = \frac{\bar{x}}{\|\bar{x}\|} \), can be verified directly. \[ \square \]

Example 4.1 When \( n = m = 2 \), Theorem 2.1 states that

\[
\{(x, y) \in \mathbb{R}^4 : \|y_1, y_2\| \leq \lambda_1 x_1 + \lambda_2 x_2\}
\]

is always maximal \( S^h \)-free, when \( \|(\lambda_1, \lambda_2)\| = 1 \). Up to a change of variables, this is one of families of sets presented in [10], although maximality is proved with respect to a different feasible set. In Examples 5.1 and 5.2 we show more examples of sets of the form \( C_\lambda \).

5 Homogeneous quadratics with a single homogeneous linear constraint

We now construct maximal \( S_{\leq 0} \)-free sets with \((\bar{x}, \bar{y})\) in its interior, where

\[
S_{\leq 0} = \{(x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\|, a^T x + d^T y \leq 0\},
\]

\[
\max\{\|a\|, \|d\|\} = 1, \text{ and } (\bar{x}, \bar{y}) \text{ satisfies (2.8)}. \text{ The maximal } S_{\leq 0} \text{-free sets we construct, as described in Theorem 2.2, are of the form } C_{\phi_{\lambda, a, d}} := \{(x, y) \in \mathbb{R}^n : \phi_{\lambda, a, d}(y) \leq \}
\]

\[ \square \] Springer
\[ \lambda^T x \}, \text{ with} \]

\[ \phi_{\lambda, a, d}(y) = \begin{cases} \|y\|, & \text{if } \lambda^T a \|y\| + d^T y \leq 0 \\ \sqrt{(\|y\|^2 - (d^T y)^2)(1 - (\lambda^T a)^2)} - d^T y \lambda^T a, & \text{otherwise.} \end{cases} \]

The derivation of this closed-form expression follows two different paths depending on whether \( \|a\| = 1 \) or \( \|a\| = 1 \). These require significantly different analyses, with \( \phi_{\lambda, a, d} \) arising from two seemingly unrelated expressions. Therefore, the proof of Theorem 2.2 is provided in separate cases.

In the previous section, the strategy for proving maximal \( S^h \)-freeness of \( C_\lambda \) was to use the outer-description

\[ C_\lambda = \{ (x, y) \in \mathbb{R}^{n+m} : -\lambda^T x + \beta^T y \leq 0, \ \forall \beta \in D_1(0) \}, \]

and to find an exposing point in \( S^h \cap C_\lambda \) for each of the inequalities defining \( C_\lambda \). As \( S_{\leq 0} \subseteq S^h \), \( C_\lambda \) is clearly \( S_{\leq 0} \)-free. However, if we try to prove it is maximal following the same technique, we find that it is not clear that some inequalities have exposing points in \( S_{\leq 0} \cap C_\lambda \). The exposing point of inequality \((-\lambda, \beta); 0\), which is \((\lambda, \beta)\), is in \( S_{\leq 0} \) if and only if \( a^T \lambda + d^T \beta \leq 0 \). Let

\[ G(\lambda) = \{ \beta : \|\beta\| = 1, a^T \lambda + d^T \beta \leq 0 \}. \]

It is natural to ask, then, if

\[ C_1 = \{ (x, y) \in \mathbb{R}^{n+m} : -\lambda^T x + \beta^T y \leq 0, \ \forall \beta \in G(\lambda) \} \]

is maximal \( S_{\leq 0} \)-free. Intuitively, \( C_1 \) is obtained from \( C_\lambda \) by removing from its description all inequalities that do not have an exposing point in \( a^T \lambda + d^T \beta \leq 0 \). It is reasonable to expect maximality, as, by construction, every inequality has a point exposing it. Indeed,

**Proposition 5.1** If \( C \) is any \( S_{\leq 0} \)-free set such that \( C_\lambda \subseteq C \), then \( C \subseteq C_1 \).

**Proof** Suppose, by contradiction, that \( C \not\subseteq C_1 \). This implies that there must exist \( \beta_0 \in G(\lambda) \) such that \(-\lambda^T x + \beta_0^T y \leq 0 \) is not valid for \( C \). As \( C_\lambda \subseteq C_1 \), \(-\lambda^T x + \beta_0^T y \leq 0 \) is valid for \( C_\lambda \).

As we saw in Theorem 2.1, \( (\lambda, \beta_0) \in C_\lambda \) exposes \(-\lambda^T x + \beta_0^T y \leq 0 \), and since \( C_\lambda \subseteq C \), Lemma 3.1 implies that \( (\lambda, \beta_0) \in \text{int}(C) \). However, since \( \beta_0 \in G(\lambda) \), we have \( (\lambda, \beta_0) \in S_{\leq 0} \). This contradicts the \( S_{\leq 0} \)-freeness of \( C \). \( \Box \)

In light of Proposition 5.1, we just need for \( C_1 \) to be \( S_{\leq 0} \)-free for it to be maximal. \( C_1 \) is clearly convex. Additionally, note that

\[ C_1 = \{ (x, y) \in \mathbb{R}^{n+m} : \max_{\beta \in G(\lambda)} y^T \beta \leq \lambda^T x \}, \quad (5.1) \]

and so to prove \( S_{\leq 0} \)-freeness, it is enough to show that for every \( (x, y) \in S_{\leq 0} \), \( \max_{\beta \in G(\lambda)} y^T \beta \geq \lambda^T x \). The case distinction discussed above arises when trying to
prove this inequality. As we show below, $C_1$ is only $S_{\leq 0}$-free when $\|a\| \leq \|d\| \land m > 1$, and the remaining case needs a significantly different approach. The following remark dismisses a simple case.  

**Remark 5.1** If $m = 1$ and $\|a\| < \|d\|$ then $S_{\leq 0}$ is convex. To see this, assume that $d > 0$ and let $(x, y) \in S_{\leq 0}$ with $y \neq 0$. Then, $\frac{dy}{dx} \leq \frac{a^T x}{\|a\|} \leq \|a\| \|x\| \leq \|a\| |y| < \|d\| |y|$. This can only happen if $y < 0$. Therefore, $S_{\leq 0}$ is the second order cone $\{(x, y) : \|x\| \leq -y\}$. The case $d < 0$ is analogous. We remark that the assumption $\|a\| < |d|$ is fundamental for the argument. As we show in Example 5.3, $S_{\leq 0}$ is not necessarily convex if $\|a\| = |d|$.

We divide the remaining cases as follows:

**Case 1:** $\|a\| \leq \|d\| = 1 \land m > 1$.

**Case 2:** $\|d\| \leq \|a\| = 1$.

### 5.1 Case 1: $\|a\| \leq \|d\| = 1 \land m > 1$

**Proposition 5.2** Let $(\bar{x}, \bar{y})$ be as in (2.8), i.e. $\|\bar{x}\| > \|\bar{y}\| \land a^T \bar{x} + d^T \bar{y} \leq 0$, and $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. If $\|a\| \leq \|d\| = 1$ and $m > 1$, then $C_1$ is maximal $S_{\leq 0}$-free and contains $(\bar{x}, \bar{y})$ in its interior. Moreover, $C_1 = C_{\phi, a, d}$.

**Proof** As discussed above, in order to establish that $C_1$ is maximal $S_{\leq 0}$-free, it is enough to show that

$$\max_{\beta \in G(\lambda)} y^T \beta \geq \lambda^T x$$

for every $(x, y) \in S_{\leq 0}$. (5.2)

As the objective function of $\max_{\beta \in G(\lambda)} y^T \beta$ is linear and $m > 1$, we can replace the $\|\beta\| = 1$ constraint with an inequality and obtain

$$\max_{\beta \in G(\lambda)} y^T \beta = \max\{y^T \beta : \|\beta\| \leq 1, a^T \lambda + d^T \beta \leq 0\}.$$ (5.3)

As $G(\lambda)$ is constructed from an infeasible point $(\bar{x}, \bar{y}) \notin S_{\leq 0}$ such that $a^T \bar{x} + d^T \bar{y} \leq 0$, i.e., $\|\bar{y}\| < \|\bar{x}\|$, we have $\|\bar{y}\|/\|\bar{x}\| < 1$. Moreover, perturbing the latter we can argue that the rightmost optimization problem in (5.3) has a strictly feasible point. Thus, Slater’s condition holds and so does strong duality.

The conic dual of this problem is given by $\inf \{\mu - \lambda^T a \theta : \theta \geq 0, \mu \geq \|y - \theta d\|\}$, and the inequality involving $\mu$ must be tight in an optimal solution, thus

$$\max\{y^T \beta : \|\beta\| \leq 1, a^T \lambda + d^T \beta \leq 0\} = \inf_{\theta \geq 0} \|y - d\theta\| - \lambda^T a \theta.$$ (5.4)

Using (5.4), (5.2) is equivalent to

$$\inf_{\theta \geq 0} \|y - d\theta\| - \lambda^T a \theta \geq \lambda^T x$$

for every $(x, y) \in S_{\leq 0}$.
We now prove that if \((x, y) \in S_{\leq 0}\), then \(\lambda^T(x + a\theta) \leq \|y - d\theta\|\), which implies the result. By Cauchy-Schwarz and \(\|\lambda\| = 1\), we have that \(\lambda^T(x + a\theta) \leq \|x + a\theta\|\). Furthermore, \(\|x + a\theta\|^2 = \|x\|^2 + 2\theta a^T x + \|a\|^2\). Since \(\theta \geq 0\), \(\theta a^T x \leq -\theta d^T y\). Together with \(\|x\|^2 \leq \|y\|^2\) they imply
\[
\|x + a\theta\|^2 \leq \|y\|^2 - 2\theta d^T y + \|a\|^2 \theta^2 = \|y - d\theta\|^2 + (\|a\|^2 - \|d\|^2) \theta^2 \leq \|y - d\theta\|^2,
\]
where the last inequality follows since \(\|d\| \geq \|a\|\).

We have shown that \(\|x + a\theta\| \leq \|y - d\theta\|\). Hence, \(\lambda^T(x + a\theta) \leq \|y - d\theta\|\) as we wanted to show, which implies that \(C_1\) is \(S_{\leq 0}\)-free. Proposition 5.1 implies the maximality of \(C_1\), and \((\tilde{x}, \tilde{y}) \in \text{int}(C_1)\) since \(C_\lambda \subseteq C_1\).

To obtain the closed-form expression given by \(C_{\phi_{\lambda, a, d}}\), we use Proposition A1 (after relabeling variables) to show that
\[
\max_{\beta} \{y^T \beta : \|\beta\| \leq 1, a^T \lambda + d^T \beta \leq 0\} = \phi_{\lambda, a, d}(y) \tag{5.6}
\]

\[\square\]

**Remark 5.2** Although it is not needed in this case, convexity of \(\phi_{\lambda, a, d}\) in the case \(\|a\| \leq \|d\| = 1\) is implied by (5.6), since it is maximum of linear functions. This is in contrast with the case \(\|d\| \leq \|a\| = 1\) below, where proving convexity of \(\phi_{\lambda, a, d}\) is crucial, and harder to obtain.

**Remark 5.3** Setting \(\lambda\) to \(\tilde{x}/\|\tilde{x}\|\) is important to obtain a strictly feasible point for (5.3) and to ensure that \(C_1\) contains \((\tilde{x}, \tilde{y})\) in its interior. However, we can also consider \(\lambda\) in a neighborhood of \(\tilde{x}/\|\tilde{x}\|\) such that \(\|\lambda\| = 1\) and obtain the same result. This implies that, when \(n \geq 2\), we can obtain infinitely-many maximal \(S_{\leq 0}\)-free sets, as mentioned in Sect. 2.

The proof of Proposition 5.2 heavily relies on our assumptions \(\|a\| \leq \|d\|\) (to show (5.5)) and \(m > 1\) (to show (5.3)), so the natural question is whether these conditions are actually necessary for our statement to be true. Before moving on to the next case, we argue through illustrative examples why they indeed are.

**Example 5.1** Consider the following set of the type \(S_{\leq 0}\), which we denote \(S^1_{\leq 0}\):
\[
S^1_{\leq 0} = \{(x, y_1, y_2) \in \mathbb{R}^3 : |x| \leq \|y\|, \ ax + d^T y \leq 0\}
\]
with \(a = \frac{1}{\sqrt{2}}\) and \(d = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T\). Let us consider the point \((\tilde{x}, \tilde{y}) = (-1, 0, 0)^T\), clearly satisfying the linear inequality, but not in \(S^1_{\leq 0}\). In Fig. 1 we show \(S^1_{\leq 0}\), the \(S^1_{\leq 0}\)-free set given by \(C_\lambda\), and the set \(C_1\) for \(\lambda = \frac{\tilde{x}}{\|\tilde{x}\|}\). Since in this case \(|a| = \frac{1}{\sqrt{2}} \leq 1 = \|d\|\) and \(m > 1\), we know \(C_1\) is maximal \(S^1_{\leq 0}\)-free.
Example 5.2  Consider the set \( S_{\leq 0}^2 \), defined as

\[
S_{\leq 0}^2 = \{ (x_1, x_2, y) \in \mathbb{R}^3 : \|x\| \leq |y|, \quad a^T x + dy \leq 0 \}
\]

with \( a = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T \) and \( d = \frac{1}{\sqrt{2}} \), and \((\bar{x}, \bar{y}) = (-1, -1, 0)^T\). This point satisfies the linear inequality in \( S_{\leq 0}^2 \), but it is not in \( S^2_{\leq 0} \). Let \( \lambda = \frac{\bar{x}}{\|\bar{x}\|} \).

In this case \( a^T \lambda = 0 \), and as a consequence \( G(\lambda) = \{-1\} \). This implies that \( C_1 = \{ (x_1, x_2, y) \in \mathbb{R}^3 : \frac{1}{\sqrt{2}} (x_1 + x_2) - y \leq 0 \} \). In contrast, a simple calculation using (2.10) yields

\[
\phi_{\lambda, a, d}(y) = \begin{cases} 
-y, & \text{if } y \leq 0 \\
\frac{y}{\sqrt{2}}, & \text{if } y > 0,
\end{cases}
\]

which implies

\[
C_{\phi_{\lambda, a, d}} = \{ (x, y) : \frac{1}{\sqrt{2}} (x_1 + x_2) - y \leq 0, \quad \frac{1}{\sqrt{2}} (x_1 + x_2) + \frac{1}{\sqrt{2}} y \leq 0 \}.
\]

Clearly, \( C_1 \neq C_{\phi_{\lambda, a, d}} \). This does not contradict Proposition 5.2 since \( \|a\| = 1 > \frac{1}{\sqrt{2}} = |d| \). Furthermore, we have no guarantee on the \( S^2_{\leq 0}\)-freeness of \( C_1 \) and, indeed, it is not. Figure 2 shows \( S^2_{\leq 0} \), the \( S^2_{\leq 0}\)-free set given by \( C_\lambda \), and the set \( C_1 \), which is clearly not \( S^2_{\leq 0}\)-free. The set \( C_{\phi_{\lambda, a, d}} \) is depicted in Fig. 3, see Example 5.4.

Example 5.3  Let \( n = 2, m = 1, a = (-\frac{3}{5}, \frac{4}{5})^T, d = 1, (\bar{x}, \bar{y}) = -(4, 3, 1), \) and \( \lambda = \frac{\bar{x}}{\|\bar{x}\|} \). Clearly, \( \|d\| = \|a\| = 1 \), and all hypotheses of Proposition 5.2 are satisfied, but \( m > 1 \). Note that \( (x, y) = (3, -4, 5) \in S_{\leq 0} \), and \( \lambda^T x + y = 0 + 5 > 0 \). Since \( G(\lambda) = \{-1\} \), the inequality implies that \( (x, y) \notin \text{int}(C_1) \). Thus, \( C_1 \) is not \( S_{\leq 0}\)-free.
5.2 Case 2: \( \|d\| \leq \|a\| = 1 \)

As we have seen in Example 5.2, when \( \|a\| \leq \|d\| \) does not hold, \( C_1 \) is not necessarily \( S_{\leq 0} \)-free. As we anticipated, we show below that a maximal \( S_{\leq 0} \)-free set is given by \( C_{\varphi \lambda, a, d} \), which in this case does not correspond to \( C_1 \).

5.2.1 Projecting-out the lineality space, and back

As before, we look for a maximal \( S_{\leq 0} \)-free set \( C \) containing \( C_\lambda \). The lineality space of \( C_\lambda \) is \( L = \{(x, y) : \lambda^T x = 0, y = 0\} \) and it must be contained in the lineality space of \( C \). By Theorem 3.4, \( \text{proj}_{L^\bot} C \) is maximal \( \text{proj}_{L^\bot} S_{\leq 0} \)-free, thus, we look for \( C \) by studying maximal \( \text{proj}_{L^\bot} S_{\leq 0} \)-free sets. \(^2\) Note that

\[
L^\bot = \text{lin}\{\lambda\} \times \mathbb{R}^m \quad \text{and} \quad \text{proj}_{L^\bot} S_{\leq 0} = \{(\lambda^T x, y) : \|x\| \leq \|y\|, a^T x + d^T y \leq 0\}. 
\]

Since we are interested in an \( S_{\leq 0} \)-free set containing \((\bar{x}, \bar{y})\), the following result guides us to where in \((\text{proj}_{L^\bot} S_{\leq 0})^c\) the projection of \((\bar{x}, \bar{y})\) is, thus motivating our construction of \( S_{\leq 0} \)-free set.

**Lemma 5.1** Consider \((\bar{x}, \bar{y})\) such that \( \|\bar{x}\| > \|\bar{y}\| \) and \( a^T \bar{x} + d^T \bar{y} \leq 0 \) and \( \lambda = \frac{\bar{x}}{\|\bar{x}\|} \). The projection of \((\bar{x}, \bar{y})\) onto \( L^\bot \) is given by \((\lambda^T \bar{x}, \bar{y})\) and satisfies

\[
\lambda^T \bar{x} > \max_x \{\lambda^T x : (x, \bar{y}) \in S_{\leq 0}\}. 
\]

**Proof** The fact that the projection of \((\bar{x}, \bar{y})\) onto \( L^\bot \) is given by \((\lambda^T \bar{x}, \bar{y})\) can be easily verified. Additionally,

\[
\max_x \{\lambda^T x : \|x\| \leq \|\bar{y}\|, a^T x + d^T \bar{y} \leq 0\} \leq \|\bar{y}\| < \|\bar{x}\| = \lambda^T \bar{x}. 
\]

\(^2\) The original motivation to look into the projection was to plot high-dimensional examples. These plots suggested that the complement of \( \text{proj}_{L^\bot} S_{\leq 0} \) is formed by two disjoint convex sets. This motivated the definition and study of (5.7) below.

@Springer
This lemma shows that the epigraph of the function
\[ y \mapsto \max_x \{ \lambda^T x : \|x\| \leq \|y\|, a^T x + d^T y \leq 0 \}, \quad (5.7) \]
contains the corresponding projection of \((\bar{x}, \bar{y})\) in its interior. In addition, such epigraph clearly does not contain an element of \(\text{proj}_{S \perp} S_{\leq 0}\) in its interior. This gives us a candidate for a maximal \(S_{\leq 0}\)-free set containing \((\bar{x}, \bar{y})\). Going back to the original space, the candidate becomes
\[ C_2 = \{(x, y) \in \mathbb{R}^{n+m} : (5.7) \leq \lambda^T x\}. \quad (5.8) \]
At this point, this is only a motivation. We can ensure \(\text{int}(C_2) \cap S_{\leq 0} = \emptyset\) directly, but we do not yet know if \(C_2\) is convex. In fact it is not, in general.

### 5.2.2 Convexity of \(C_2\)

We begin by showing that the function in (5.7) is defined over all \(\mathbb{R}^m\).

**Proposition 5.3** If \(\|d\| \leq \|a\|\), then for every \(y\) the set \(\{(x, y) : \|x\| \leq \|y\|, a^T x \leq -d^T y\}\) is not empty.

**Proof** It can be directly verified that \(x = -d^T y \frac{a}{\|a\|^2}\) belongs to the set. \(\square\)

We now provide a full characterization of the function defined in (5.7), which implies the convexity of \(C_2\).

**Proposition 5.4** Let \(\lambda \in \mathcal{D}_1(0), a \in \mathbb{R}^n\) and \(d \in \mathbb{R}^m\) such that \(\|d\| \leq \|a\| = 1\) (i.e. \(a \in \mathcal{D}_1(0)\)). Then, (5.7) = \(\phi_{\lambda, a, d}\). Furthermore, \(\phi_{\lambda, a, d}\) is sublinear and

- if \(\|d\| = 1 \wedge m > 1\), then \(\phi_{\lambda, a, d}\) is differentiable \(\mathbb{R}^m \setminus d\mathbb{R}_+\),
- otherwise \(\phi_{\lambda, a, d}\) is differentiable in \(\mathbb{R}^m \setminus \{0\}\).

**Remark 5.4** We recall the definition of \(\phi_{\lambda, a, d}\) to aid the reader.

\[ \phi_{\lambda, a, d}(y) = \begin{cases} \|y\|, & \text{if } \lambda^T a \|y\| + d^T y \leq 0 \\ \sqrt{\left(\|y\|^2 - (d^T y)^2\right)(1 - (\lambda^T a)^2)} - d^T y \lambda^T a, & \text{otherwise.} \end{cases} \]

The discussion of the convexity of \(\phi_{\lambda, a, d}\) in Remark 5.2, does not help in this case. Remark 5.2 shows convexity of \(\phi_{\lambda, a, d}\) when \(\|a\| \leq \|d\| = 1\), in which case \(\phi_{\lambda, a, d}\) is the maximum of linear functions in (5.6). In our current case, \(\phi_{\lambda, a, d}\) is no longer equal to (5.6) (see Example 5.2).

**Proof** We leave the proof that (5.7) = \(\phi_{\lambda, a, d}\) when \(\|d\| \leq \|a\| = 1\) in the appendix (Proposition A1). The fact that \(\phi_{\lambda, a, d}\) is positively homogeneous can be verified directly. Thus convexity and differentiability remains.

First, note that if \(\lambda = a\), then \(\phi_{\lambda, a, d}(y) = -d^T y\). This function is clearly sublinear and differentiable everywhere. On the other hand, if \(\lambda = -a\), then \(\phi_{\lambda, a, d}(y) = \|y\|\).
This function is clearly sublinear and differentiable everywhere but the origin. We now consider \( \lambda \neq \pm a \). Let

\[
A_1 := \{ y : \lambda^T a \|y\| + a^T y \leq 0 \}, \quad A_2 := \{ y : \lambda^T a \|y\| + a^T y \geq 0 \},
\]

and let \( \phi_{\lambda,a,d}^1 \) and \( \phi_{\lambda,a,d}^2 \) be the restriction of \( \phi_{\lambda,a,d} \) to \( A_1 \) and \( A_2 \), respectively.

To show that \( \phi_{\lambda,a,d} \) is convex we are going to use \([46, \text{Theorem } 3]\). In our particular case, since \( \phi_{\lambda,a,d} \) is positively homogeneous, this theorem implies that we just need to check that \( \phi_{\lambda,a,d} \) is convex on each convex subset of \( A_1 \) and \( A_2 \), \( \phi_{\lambda,a,d}^1 \) on \( A_1 \cap A_2 \), and that

\[
\phi_{\lambda}^1(y; \rho) + \phi_{\lambda}^1(y; -\rho) \geq 0, \text{ for all } \rho \in \mathbb{R}^m \setminus \{0\}, y \in A_1 \cap A_2.
\]

(5.10)

Here, \( \phi_{\lambda,a,d}^1(y; \rho) \) is the directional derivative of \( \phi_{\lambda,a,d} \) at \( y \) in the direction \( \rho \).

Clearly, \( \phi_{\lambda,a,d} \) is convex in each convex subset of \( A_1 \). The function \( \phi_{\lambda,a,d}^2 \) is of the form \( c_1 \|y\|_W - c_2 d^T y \), where \( W = I - d d^T \geq 0 \) and \( c_1, c_2 \in \mathbb{R}_+ \). Thus, \( \phi_{\lambda,a,d} \) is convex on each convex subset of \( A_2 \).

It is not hard to see that \( \phi_{\lambda,a,d}^1(y) = \phi_{\lambda,a,d}^2(y) \) for \( y \in A_1 \cap A_2 \).

Let us verify (5.10) for \( y \neq 0 \). For this, first notice that \( \phi_{\lambda,a,d}^1(y) \) is differentiable whenever \( y \neq 0 \). Likewise, \( \phi_{\lambda,a,d}^2(y) \) is differentiable whenever \( y \neq 0 \) if \( \|d\| < 1 \) or whenever \( y \notin d \mathbb{R}_+ \) if \( \|d\| = 1 \). However, if \( y \in A_1 \cap A_2 \setminus \{0\} \) and \( \|d\| = 1 \), then \( y \notin d \mathbb{R}_+ \), thus \( \phi_{\lambda,a,d}^2 \) is differentiable in a neighborhood of \( y \). Furthermore,

\[
\nabla \phi_{\lambda,a,d}^2(y) = \frac{(1 - (\lambda^T a)^2)(I - dd^T)y}{\sqrt{\|y\|^2 - (d^T y)^2}(1 - (\lambda^T a)^2)} - \lambda^T ad = \frac{y}{\|y\|} - \nabla \phi_{\lambda,a,d}^1(y).
\]

Therefore, \( \phi_{\lambda,a,d} \) is differentiable for \( y \neq 0 \) if \( \|d\| < 1 \) or for \( y \notin d \mathbb{R}_+ \) if \( \|d\| = 1 \). Thus, (5.10) holds with equality for \( y \in A_1 \cap A_2 \setminus \{0\} \).

It remains to verify (5.10) for \( y = 0 \). Let \( \rho \) be such that \( \rho \in A_1 \) and \( -\rho \in A_2 \). As \( \phi_{\lambda,a,d} \) is positively homogeneous, \( \phi_{\lambda,a,d}^1(0; \cdot) = \phi_{\lambda,a,d}^2(\cdot) \). Hence,

\[
\phi_{\lambda,a,d}^1(0; \rho) = \|\rho\| \text{ and } \phi_{\lambda,a,d}^1(0; -\rho) = \sqrt{1 - (\lambda^T a)^2} \sqrt{\|\rho\|^2 - (d^T \rho)^2} + d^T \rho \lambda^T a.
\]

We need to prove that

\[
\sqrt{1 - (\lambda^T a)^2} \sqrt{\|\rho\|^2 - (d^T \rho)^2} + d^T \rho \lambda^T a + \|\rho\| \geq 0.
\]

By Cauchy-Schwarz, \(|d^T \rho \lambda^T a| \leq \|d\| \|\rho\| < \|\rho\| \). Thus, \( d^T \rho \lambda^T a + \|\rho\| > 0 \). Since \( \sqrt{1 - (\lambda^T a)^2} \sqrt{\|\rho\|^2 - (d^T \rho)^2} \geq 0 \), the inequality follows.

We have proved that \( \phi_{\lambda,a,d} \) is convex and differentiable in \( \mathbb{R}^m \setminus \{0\} \) if \( \|d\| < 1 \) and in \( \mathbb{R}^m \setminus \mathbb{R}_+ \) if \( \|d\| = 1 \). It remains to show that if \( m = 1 \) and \( \|d\| = 1 \), then
Let us recall the set $\phi_{\lambda}$, Fig. 3 (orange) and the maximal $S_{\leq 0}$-free $C_2$ (blue) (colour figure online) in $D$.

In the cases where $\phi_{\lambda}$, figure online) just showed is $C_\lambda$ is $S_{\leq 0}$-free but not maximal, and $C_1$ is not $S_{\leq 0}$-free. Figure 3 depicts $C_2$, which we just showed is $S_{\leq 0}$-free and, due to Proposition 5.4, $C_2 = \phi_{\lambda,a,d}$. Thus, an explicit description of $C_2$ is in Example 5.2. Maximality of $C_2$ remains to be proved, and it is the subject of the next section.

5.2.3 Maximality proof

**Proposition 5.5** Let $\lambda \in D(1(0))$, $a \in \mathbb{R}^n$ and $d \in \mathbb{R}^m$ such that $\|d\| \leq \|a\| = 1$ (i.e. $a \in D(1(0))$). Then $C_2 = \phi_{\lambda,a,d}$ is maximal $S_{\leq 0}$-free. Additionally, if $(\bar{x}, \bar{y})$ satisfies (2.8), i.e. $\|\bar{x}\| > \|\bar{y}\| \land a^T \bar{x} + d^T \bar{y} \leq 0$, then $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ ensures $(\bar{x}, \bar{y}) \in \text{int}(C_2)$.

**Proof** We already know $C_2$ is convex. Moreover, $C_2$ is $S_{\leq 0}$-free since every $(x, y) \in S_{\leq 0}$ is feasible for the optimization problem (5.7). We now focus on proving maximality. In the cases where $\phi_{\lambda,a,d}$ is differentiable in $\mathbb{R}^m \setminus \{0\}$, we have

$$C_2 = \{(x, y) \in \mathbb{R}^{n+m} : \nabla \phi_{\lambda,a,d}(\beta)^T y \leq \lambda^T x, \forall \beta \in D(1(0))\}.$$ 

Let $\beta \in D(1(0))$ and let $x_\beta$ be the optimal solution of the problem (5.7). That is, $\lambda^T x_\beta = \phi_{\lambda,a,d}(\beta)$. By Lemma 4.1, the inequality $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq 0$ is exposed by $(x_\beta, \beta)$.

The only remaining case is $\|d\| = 1 \land m > 1$, where $\phi_{\lambda,a,d}$ is only differentiable in $D(1(0)) \setminus \{d\}$. Since in this case $m > 1$ we can safely remove a single inequality from the outer-description of $C_2$ without affecting it, i.e.,

$$C_2 = \{(x, y) \in \mathbb{R}^{n+m} : \nabla \phi_{\lambda,a,d}(\beta)^T y \leq \lambda^T x, \forall \beta \in D(1(0)) \setminus \{d\}\}.$$ 

Using the same argument as above we can find an exposing point of each inequality $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq 0$ for $\beta \in D(1(0)) \setminus \{d\}$. The fact that $(\bar{x}, \bar{y}) \in \text{int}(C_2)$ when $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ follows directly since $C_\lambda \subseteq C_2$. 

$\square$
6 Non-homogeneous quadratics

We now consider the set \( S^g \) defined in (2.4), that is,
\[
S^g = \{ (x, y) \in \mathbb{R}^{n+m} : \| x \| \leq \| y \|, a^T x + d^T y = -1 \},
\]
and consider \((\bar{x}, \bar{y}) \notin S^g\) satisfying (2.6) (i.e. only violating the nonlinear constraint in \( S^g \)). As in Sect. 5, we begin by dismissing a simple case.

Remark 6.1 When \( \| a \| \leq \| d \| \land m = 1, S^g \) is convex. Indeed, as \( d \neq 0 \) (if not, then \( a = 0 \) and \( S^g = \emptyset \)) we can write \( y = \frac{1}{d}(-1 - a^T x) \) and consequently
\[
S^g = \{ (x, y) \in \mathbb{R}^{n+1} : \| x \|^2 \leq \frac{1}{d^2}(1 + 2a^T x + (a^T x)^2), a^T x + d^T y = -1 \}
\]
\[
= \{ (x, y) : x^T \left( I - \frac{1}{d^2}aa^T \right) x - \frac{1}{d^2}(1 + 2a^T x) \leq 0, a^T x + d^T y = -1 \}.
\]
Since \( I - \frac{1}{d^2}aa^T \succeq 0 \) whenever \( |d| \geq \| a \| \), the set \( S^g \) is convex.

Similarly to Sect. 5, we distinguish the following cases:

Case 1: \( \| a \| \leq \| d \| = m > 1 \). \hspace{1cm} Case 2: \( \| d \| < \| a \| = 1 \).

Since \( S^g \subsetneq S_{\leq 0} \), then \( C_1 (C_2) \) is \( S^g \)-free in Case 1 (Case 2) as per Sect. 5. It is natural to wonder whether these sets are already maximal for \( S^g \).

6.1 Case 1: \( \| a \| \leq \| d \| = m > 1 \)

The intuition for this case is as follows. We showed maximality of \( C_1 \) with respect to \( S_{\leq 0} \) by exploiting that \( C_1 \) is defined by the inequalities of \( C_\lambda \) exposed by elements in \( S_{\leq 0} \). Similarly, we study which inequalities of \( C_1 \) are exposed by a point of \( S^g \). Recall that
\[
C_1 = \{ (x, y) \in \mathbb{R}^{n+m} : -\lambda^T x + \beta^T y \leq 0, \forall \beta \in G(\lambda) \},
\]
where \( G(\lambda) = \{ \beta \in \mathbb{R}^m : \| \beta \| = 1, a^T \lambda + d^T \beta \leq 0 \} \).

Consider an inequality in the definition of \( C_1 \) given by \((-\lambda, \beta); 0\) such that \( a^T \lambda + d^T \beta < 0 \). Then, the point \((\lambda, \beta) \in S_{\leq 0}\) can be scaled by \( \mu = \frac{-1}{a^T \lambda + d^T \beta} \) to the exposing point \( \mu(\lambda, \beta) \in S^g \). Thus, almost every inequality describing \( C_1 \) is exposed by points of \( S^g \). As for the inequalities \((-\lambda, \beta); 0\) with \( a^T \lambda + d^T \beta = 0 \), we will show that we can drop them from the description of \( C_1 \), without changing it. To do so, we use the following intermediate lemma.

Lemma 6.1 Consider \( \alpha \in \mathbb{R}^m \), with \( m > 1 \) and \( \alpha_0 \in \mathbb{R} \) such that
\[
\{ \beta \in \mathbb{R}^m : \alpha^T \beta < \alpha_0, \| \beta \| \leq 1 \} \neq \emptyset.
\]
Additionally, consider \( \beta_0 \) such that \( \| \beta_0 \| = 1 \land \alpha^T \beta_0 = \alpha_0 \). Then, there exists a sequence \((\beta_i)_{i \in \mathbb{N}}\) such that \( \| \beta_i \| = 1, \alpha^T \beta_i < \alpha_0 \) and \( \lim_{i \to \infty} \beta_i = \beta_0 \).

**Proof** From assumption (6.1), for \( \epsilon > 0 \) small enough, the following optimization problem is feasible

\[
\nu(\epsilon) := \max\{\beta^T \beta : \alpha^T \beta \leq \alpha_0 - \epsilon, \| \beta \| \leq 1\}.
\]

Let \( \beta_\epsilon \) be an optimal solution of (6.2). Since \( m > 1 \), we may assume \( \| \beta_\epsilon \| = 1 \) and feasibility implies \( \alpha^T \beta_\epsilon < \alpha_0 \). Since (6.2) is a convex problem, \( \nu(\epsilon) \) is continuous (see e.g. [15, Section 5.6.1]) and, additionally, \( \nu(0) = \| \beta_0 \| = 1 \), therefore \( \beta_0^T \beta_\epsilon \to 1 \) when \( \epsilon \to 0 \). As \( \beta_0^T \beta_\epsilon = 1 - \frac{1}{2} \| \beta_0 - \beta_\epsilon \|^2 \), we conclude that \( \beta_\epsilon \to \beta_0 \) when \( \epsilon \to 0 \). The definition of the sequence using \( \beta_\epsilon \) follows directly. \(\Box\)

**Theorem 6.1** Consider a non-convex set \( S^g \) defined as (2.4), with \( \| a \| \leq \| d \| = 1 \land m > 1 \), and \((\tilde{x}, \tilde{y})\) satisfying (2.6), i.e. \( \| \tilde{x} \| > \| \tilde{y} \| \land a^T \tilde{x} + d^T \tilde{y} = -1 \). Let \( \lambda = \frac{\tilde{x}}{\| \tilde{x} \|}, \) and \( H = \{ (x, y) \in \mathbb{R}^{n+m} : a^T x + d^T y = -1 \} \). Then, \( C_1 \) is maximal \( S_{\leq 0}\)-free with respect to \( H \) and contains \((\tilde{x}, \tilde{y})\) in its interior.

**Proof** By Proposition 5.2, we know that \( C_1 \) is maximal \( S_{\leq 0}\)-free. Thus, \( C_1 \) is \( S_{\leq 0}\)-free with respect to \( H \). From the proof of Proposition 5.2 we know that \{\( \beta \in \mathbb{R}^m : \| \beta \| \leq 1, a^T \lambda + d^T \beta < 0 \)\} \( \neq \emptyset \). Since \( m > 1 \), Lemma 6.1 implies that every \( \beta \) such that \( a^T \lambda + d^T \beta = 0 \) can be obtained as a limit, thus

\[
C_1 = \{ (x, y) \in \mathbb{R}^{n+m} : -\lambda^T x + \beta^T y \leq 0, \forall \beta \in \hat{G}(\lambda) \},
\]

where \( \hat{G}(\lambda) = \{ \beta \in \mathbb{R}^m : \| \beta \| = 1, a^T \lambda + d^T \beta < 0 \} \). Consider \( \beta_0 \in \hat{G}(\lambda) \). As we saw in Proposition 5.1, \((\lambda, \beta_0) \in C_1 \cap S_{\leq 0}\) exposes the inequality \((\lambda, \beta_0)\). As \( C_1 \cap S_{\leq 0}\) is a (non-convex) cone, we have that for any \( \mu > 0 \), \( \mu (\lambda, \beta_0) \in C_1 \cap S_{\leq 0}\) exposes the inequality \((\lambda, \beta_0)\). Since \( a^T \lambda + d^T \beta_0 < 0 \), \( \mu = -\frac{1}{a^T \lambda + d^T \beta_0} > 0 \) and so

\[
-\frac{(\lambda, \beta_0)}{a^T \lambda + d^T \beta_0} \in S_{\leq 0} \cap H \cap C_1,
\]

exposes the inequality \((\lambda, \beta_0)\). The claim now follows from Theorem 3.2. \(\Box\)

The above theorem states that obtaining a maximal \( S^g \)-free set in this case amounts to simply using the maximal \( S_{\leq 0}\)-free set \( C_1 \), and then intersecting with \( H \) (recall that \( S^g = S_{\leq 0} \cap H \)). Note that in this case \( \lambda \) can also be fixed to other values and produce different maximal \( S^g \)-free set, as per our discussion in Remark 5.3.

The next case is considerably different.

**6.2 Case 2:** \( \| d \| < \| a \| = 1 \)

Unfortunately, in this case the maximal \( S_{\leq 0}\)-free set \( C_2 \) is not maximal \( S^g \)-free, as the following example shows.
Example 6.1 We continue with $S^2_{\leq 0}$ defined in Example 5.2. In Fig. 3 we showed how $C_2$ gives us a maximal $S^2_{\leq 0}$-free set. If we now consider $H = \{(x, y) \in \mathbb{R}^{n+m} : a^T x + d^T y = -1\}$ with $a = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$ and $d = \frac{1}{\sqrt{2}}$, we do not obtain that $C_2 \cap H$ is maximal $S^2_{\leq 0} \cap H$-free, as shown by Fig. 4.

Figure 4 displays an interesting feature though: the inequalities defining $C_2$ seem to have the correct “slope” and just need to be translated. In what follows, we show that this idea yields maximal $S^g$-free sets.

6.2.1 Set-up

Recall that $C_2 = C_{\phi_{\lambda},a,d}$ and note that, since $\|d\| < \|a\| = 1$, $\phi_{\lambda,a,d}$ is differentiable in $D_1(0)$ (see Proposition 5.4). Thus we can write

\[
C_2 = \{(x, y) : \phi_{\lambda,a,d}(y) \leq \lambda^T x\} = \{(x, y) : -\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq 0, \forall \beta \in D_1(0)\}.
\]

We denote by $r(\beta)$ the maximum we can relax each inequality of $C_2$ such that

\[
C^g_{\phi_{\lambda,a,d}} = \{(x, y) : -\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq r(\beta), \forall \beta \in D_1(0)\}, \tag{6.3}
\]

is $S^g$-free. Note that when $\beta$ satisfies $\lambda^T a + d^T \beta < 0$, the inequalities of $C_2$ are the same as the ones of $C_1$ and, just like in Sect. 6.1, they have exposing points in $S^g$. An inequality of this type can be seen in Fig. 4b: it is the inequality of $C_2$ tangent to $S$ at one of its exposing points. Thus, we expect that $r(\beta) = 0$ when $\lambda^T a + d^T \beta < 0$. In the following we find $r(\beta)$ when $\lambda^T a + d^T \beta \geq 0$ and show maximality of the resulting set.

As in Sect. 5.2, not all statement in this section require $\lambda = \frac{\hat{x}}{\|\hat{x}\|}$. However, we assume $\lambda \neq \pm a$. This assumption is not restrictive when constructing maximal $S^g$-free sets, as the following remark shows.
Remark 6.2 If $\lambda = -a$, then for every $\beta \in D_1(0)$ it holds that $\lambda^T a + d^T \beta < 0$. In this case $r(\beta)$ will be simply defined as $0$ everywhere and $C^g_{\phi_{\lambda,a,d}} = C_2$. This means all inequalities defining $C^g_{\phi_{\lambda,a,d}}$ have an exposing point in $S^g$ and maximality follows directly.

On the other hand, if we take $\lambda = \bar{x} \parallel \bar{x} \parallel$ with $(\bar{x}, \bar{y}) \in H$ and $\parallel \bar{x} \parallel > \parallel \bar{y} \parallel$, we have that if additionally $\lambda = a$

$$a^T \bar{x} + d^T \bar{y} = -1 \iff \parallel \bar{x} \parallel + d^T \bar{y} = -1 \Rightarrow \parallel \bar{y} \parallel + d^T \bar{y} < -1.$$ 

The latter cannot be, as $\parallel d \parallel < 1$.

Remark 6.3 The assumption $\lambda \neq \pm a$ has an unexpected consequence: as $\lambda \neq \pm a$ and $\parallel a \parallel = \parallel \lambda \parallel = 1$, it must hold that $n \geq 2$. This implicit assumption, however, does not present an issue: whenever $n = 1$ either $\lambda = a$ or $\lambda = -a$. By Remark 6.2, if we use $\lambda = \bar{x} \parallel \bar{x} \parallel$, then $\lambda = -a$. Thus, $C^g_{\phi_{\lambda,a,d}} = C_2$ and maximality holds.

6.2.2 Construction of $r(\beta)$

Let $\beta \in D_1(0)$ be such that $\lambda^T a + d^T \beta \geq 0$. Then, the face of $C_2$ defined by the valid inequality $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq 0$ does not intersect $S^g$. See Lemma A1 for a proof of this statement. In particular, the inequality is not exposed by any point in $S^g \cap C_2$. However, it is exposed by $(x_\beta, \beta) \in S_{\leq 0}$, where $x_\beta$ is the optimal solution of (5.7) (see the proof of Proposition 5.5). Note that $(x_\beta, \beta) \in H_0 = \{ (x, y) : a^T x + d^T y = 0 \}$, as otherwise we can scale it so that it belongs to $S^g$.

The quantity $r(\beta)$ is the amount we need to relax the inequality in order to be an “asymptote”, and we compute it as follows. We first find a sequence of points, $(x_n, y_n)_{n \in \mathbb{N}}$, in $S_{\leq 0}$ that converge to $(x_\beta, \beta)$, enforcing that no element of the sequence belongs to $H_0$. To do so, we take $y_n = \beta$ and $x_n$ such that $\parallel x_n \parallel = 1, a^T x_n + d^T \beta < 0$, and $x_n \to x_\beta$. Such a sequence always exists: since we are assuming $\parallel a \parallel = 1$ and $\parallel d \parallel < 1$ it follows that $-a \in \{ x \in \mathbb{R}^n : a^T x + d^T \beta < 0, \parallel x \parallel \leq 1 \}$.

The sequence $x_n$ is then obtained in virtue of Lemma 6.1 (using $n$ in place of $m$, which satisfies $n > 1$ as mentioned in Remark 6.3). Then, we scale the sequence:

$$z_n = -\frac{(x_n, \beta)}{a^T x_n + d^T \beta} \in S^g.$$ (6.4)

This last scaled sequence diverges, as the denominator goes to 0. The idea is that the violation $(-\lambda, \nabla \phi_{\lambda,a,d}(\beta))^T z_n$ given by this sequence will give us, in the limit, the maximum relaxation that will ensure $S^g$-freeness. We illustrate such a sequence with our running example.
Fig. 5 Projection onto \((x_1, x_2)\) of \(S_{\leq 0}^2 \cap H\) (orange) and \(C_2\) (blue), along with the first two coordinates of the sequence \((z_n)_{n \in \mathbb{N}}\) defined in Example 6.2 for several values of \(n\) (red). The sequence is diverging “downwards” (colour figure online).

**Example 6.2** We continue with Example 6.1. From Examples 5.2 and 5.4,

\[ C_2 = \left\{ (x, y) : \frac{1}{\sqrt{2}}(x_1 + x_2) - y \leq 0, \frac{1}{\sqrt{2}}(x_1 + x_2) + \frac{1}{\sqrt{2}}y \leq 0 \right\}. \quad (6.5) \]

It is not hard to check that \(-\left(1, 1, \sqrt{2}\right) \in S_{\leq 0}^2 \cap H \cap C_2\) exposes the leftmost inequality in (6.5). This is the tangent point in Fig. 4b we discussed above. On the other hand, the rightmost inequality in (6.5), which is obtained from \(\beta = 1\), does not have an exposing point in \(S_{\leq 0}^2 \cap H \cap C_2\), and corresponds to an inequality we should relax as per our discussion. This inequality, however, is exposed by \((x_\beta, \beta) = (0, -1, 1) \in S_{\leq 0}^2 \cap C_2\).

Consider the following sequence

\[ (x_n, \beta) = \left( \frac{1}{\sqrt{n^2 + 1}}, -\frac{n}{\sqrt{n^2 + 1}}, 1 \right) \in S_{\leq 0}^2. \]

Clearly \(x_n \to x_\beta = (0, -1)\) and

\[ a^T x_n + d^T \beta = \frac{1}{\sqrt{2}} \left( -\frac{1}{\sqrt{n^2 + 1}} - \frac{n}{\sqrt{n^2 + 1}} + 1 \right) < 0. \]

Now we scale the sequence to obtain \(z_n \in S_{\leq 0}^2 \cap H\) as in (6.4). Continuing with Fig. 4, in Fig. 5 we plot the first two components of the sequence \((z_n)_{n \in \mathbb{N}}\) along with \(S_{\leq 0}^2 \cap H \cap C_2\). From this figure we can anticipate where our argument is going: the sequence \((z_n)_{n \in \mathbb{N}}\) moves along the boundary of \(S_{\leq 0}^2 \cap H\) towards an “asymptote” from where we can deduce \(r(\beta)\).

Motivated by the discussion above, we define

\[ r(\beta) = \lim_{n \to \infty} (-\lambda, \nabla \phi_{\lambda, a, d}(\beta))^T z_n = -\lim_{n \to \infty} \frac{-\lambda^T x_n + \nabla \phi_{\lambda, a, d}(\beta)^T \beta}{a^T x_n + d^T \beta}. \]

In what follows, we show that this limit is well-defined, and that such definition of \(r(\beta)\) yields the desired maximal \(S^8\)-free set.
6.2.3 Computing $r(\beta)$

Since $x_\beta \in H_0$ and it is the optimal solution of (5.7)

$$\nabla \phi_{\lambda,a,d}(\beta)^T \beta = \phi_{\lambda,a,d}(\beta) = \lambda^T x_\beta,$$

and $d^T \beta = -a^T x_\beta$.

Thus,

$$r(\beta) = -\lim_{n \to \infty} \frac{-\lambda^T x_n + \lambda^T x_\beta}{a^T x_n - a^T x_\beta} = \lim_{n \to \infty} \frac{\lambda^T (x_n - x_\beta)}{a^T (x_n - x_\beta)}.$$

From the explicit expression of $x_\beta$ (see (A.3)), we can see that $x_\beta \in \Lambda := \text{lin}\{\lambda, a\}$. Note that $\Lambda$ is 2-dimensional, since $\lambda \neq \pm a$ (see Remark 6.2). Furthermore, we can assume that $x_n$ also belongs to $\Lambda$ as any other component of $x_n$ is irrelevant for the value of the limit. Indeed, as $\mathbb{R}^n = \Lambda \oplus \Lambda^\perp$, then $x_n = x_n^\parallel + x_n^\perp$, where $x_n^\parallel \in \Lambda$ and $x_n^\perp \in \Lambda^\perp$, and

$$\frac{\lambda^T (x_n - x_\beta)}{a^T (x_n - x_\beta)} = \frac{\lambda^T (x_n^\parallel - x_\beta)}{a^T (x_n^\parallel - x_\beta)}.$$

To compute the limit observe that

$$\frac{\lambda^T (x_n - x_\beta)}{a^T (x_n - x_\beta)} = \frac{\lambda^T \frac{x_n-x_\beta}{\|x_n-x_\beta\|}}{a^T \frac{x_n-x_\beta}{\|x_n-x_\beta\|}}.$$

Notice that $\frac{x_n-x_\beta}{\|x_n-x_\beta\|}$ converges, as $x_n \in \Lambda$, $\|x_n\| = 1$, and $x_n \to x_\beta$. Let $\hat{x}$ be the limit of the ratio $\frac{x_n-x_\beta}{\|x_n-x_\beta\|}$ and note that $\hat{x}$ is orthogonal to $x_\beta$. Indeed,

$$x_\beta^T \hat{x} = \lim_{n \to \infty} x_\beta^T \frac{x_n-x_\beta}{\|x_n-x_\beta\|} = \lim_{n \to \infty} \frac{x_\beta^T x_n - 1}{\|x_n-x_\beta\|} = \lim_{n \to \infty} -\frac{\|x_n-x_\beta\|^2}{2\|x_n-x_\beta\|} = 0.$$

Hence,

$$r(\beta) = \lim_{n \to \infty} \frac{\lambda^T (x_n - x_\beta)}{a^T (x_n - x_\beta)} = \frac{\lambda^T \hat{x}}{a^T \hat{x}}.$$

Since we are interested in the quotient of $\lambda^T \hat{x}$ and $a^T \hat{x}$, any multiple of $\hat{x}$ can be used, that is, any vector orthogonal to $x_\beta$ in $\Lambda$. Using $\lambda$ and $a$ as basis for $\Lambda$, we have that for $x \in \Lambda$ with coordinates $x_\lambda$ and $x_a$, the vector $y$ with coordinates $y_\lambda = -(x_a + x_\lambda \lambda^T a)$ and $y_a = x_\lambda + x_a \lambda^T a$ is orthogonal to $x$. Indeed,

$$x^T y = (x_\lambda \lambda + x_a a)^T (y_\lambda \lambda + y_a a) = x_\lambda y_\lambda + x_a y_a + (x_\lambda y_a + x_a y_\lambda) \lambda^T a = (x_\lambda + x_a \lambda^T a) y_\lambda + (x_a + x_\lambda \lambda^T a) y_a = 0.$$
Thus, let \( \tilde{x} = -(x_{\beta a} + x_{\beta \lambda} \lambda^T a) \lambda + (x_{\beta \lambda} + x_{\beta a} \lambda^T a) a. \) Given that \( \lambda^T a + d^T \beta \geq 0, \) from (A.3) we have

\[
x_{\beta} = \sqrt{\frac{1 - (d^T \beta)^2}{1 - (\lambda^T a)^2}} \lambda - \left( d^T \beta + \lambda^T a \sqrt{\frac{1 - (d^T \beta)^2}{1 - (\lambda^T a)^2}} \right) a.
\]

(6.6)

Note that while this last explicit formula for \( x_{\beta} \) is the one stated for the case \( \lambda^T a + d^T \beta > 0, \) it also holds when \( \lambda^T a + d^T \beta = 0. \) Therefore,

\[
\tilde{x} = (d^T \beta) \lambda + \left( \sqrt{\frac{1 - (d^T \beta)^2}{1 - (\lambda^T a)^2}} \right) - \left( d^T \beta + \lambda^T a \sqrt{\frac{1 - (d^T \beta)^2}{1 - (\lambda^T a)^2}} \right) \lambda^T a \)
\]

\[
= (d^T \beta) \lambda + \phi_{\lambda, a, d}(\beta) a.
\]

Altogether, we obtain

\[
r(\beta) = \frac{\lambda^T \tilde{x}}{\tilde{x}} = \frac{d^T \beta + \lambda^T a \phi_{\lambda, a, d}(\beta)}{\phi_{\lambda, a, d}(\beta) + d^T \beta \lambda^T a}.
\]

Note that if \( \lambda^T a + d^T \beta = 0, \) then \( r(\beta) = 0. \) We summarize the above discussion in the following result.

Lemma 6.2 Let \( \alpha, \lambda, \beta \in D_1(0), d \) such that \( \|d\| < \|a\| = 1, \) and \( \lambda \neq \pm a \) be such that \( \|d\| < \|a\| \) and \( \lambda^T a + d^T \beta \geq 0. \) Then, every sequence \((x_n)_{n \in \mathbb{N}} \subseteq \text{lin} \{\lambda, a\}\) converging to \( x_{\beta} \) such that \( \|x_n\| = 1 \) and \( a^T x_n + d^T \beta < 0, \) satisfies

\[
r(\beta) = \lim_{n \to \infty} \frac{\lambda^T (x_n - x_{\beta})}{a^T (x_n - x_{\beta})} = \frac{d^T \beta + \lambda^T a \phi_{\lambda, a, d}(\beta)}{\phi_{\lambda, a, d}(\beta) + d^T \beta \lambda^T a}.
\]

Such sequences are always guaranteed to exist.

Therefore, for \( \beta \in D_1(0), \) we define

\[
r(\beta) = \begin{cases} 
0, & \text{if } \lambda^T a + d^T \beta \leq 0 \\
\frac{d^T \beta + \lambda^T a \phi_{\lambda, a, d}(\beta)}{\phi_{\lambda, a, d}(\beta) + d^T \beta \lambda^T a}, & \text{otherwise}.
\end{cases}
\]

(6.7)

We extend \( r \) to \( y \in \mathbb{R}^m \setminus \{0\} \) by \( r(y) = r \left( \frac{y}{\|y\|} \right) \) and leave it undefined at 0.

Example 6.3 We continue with our running example in Example 6.2. In this case \( r(-1) = 0, \) and since \( \phi_{\lambda, a, d}(\beta) = 1/\sqrt{2}, \lambda^T a = 0 \) and \( d = 1/\sqrt{2} \) it can be checked that \( r(1) = 1. \) Now, let

\[
C^g_{\phi_{\lambda, a, d}} = \left\{ (x, y) : -\lambda^T x + \nabla \phi_{\lambda, a, d}(\beta)^T y \leq r(\beta), \text{ for all } \beta \in D_1(0) \right\}
\]

Springer
(a) $S^2_{\leq 0}$ (orange), $H$ (green) and $C_{\phi_{\lambda,a,d}}^{g}$ (blue). $C_{\phi_{\lambda,a,d}}^{g}$ is no longer $S^2_{\leq 0}$-free.

(b) Projection onto $(x_1, x_2)$ of $S^2_{\leq 0} \cap H$ (orange) and $C_{\phi_{\lambda,a,d}}^{g} \cap H$ (blue).

**Fig. 6** Plots of $S^2_{\leq 0}$, $H$ and $C_{\phi_{\lambda,a,d}}^{g}$ as defined in Example 6.3 showing that $C_{\phi_{\lambda,a,d}}^{g}$ is maximal $S^2_{\leq 0}$-free with respect to $H$ (colour figure online).

\[
\begin{aligned}
\{ (x, y) : & \frac{1}{\sqrt{2}}(x_1 + x_2) - y \leq 0, \frac{1}{\sqrt{2}}(x_1 + x_2) + \frac{1}{\sqrt{2}}y \leq 1 \}.
\end{aligned}
\]

Figure 6 shows the same plots as Fig. 4 with $C_{\phi_{\lambda,a,d}}^{g}$ instead of $C_{2}$.

The characterization of $r$ as a limit is going to be useful to prove maximality of $C_{\phi_{\lambda,a,d}}^{g}$. However, to show that $C_{\phi_{\lambda,a,d}}^{g}$ is $S^g$-free, we need a different interpretation of $r$, which connects it with the dual of (5.7). Under the assumptions of Lemma 6.2, we show in Proposition A1 that the lagrangian dual of (5.7) is given by

\[
\inf_{\theta} \{ \|\lambda - \theta a\|y\| - \theta d^T y : \theta \geq 0 \}.
\]

Moreover, the optimal dual solution is $\theta : \mathbb{R}^m \to \mathbb{R}_+ \cup \{+\infty\}$,

\[
\theta(y) = \begin{cases} 
0, & \text{if } \lambda^T a \|y\| + d^T y \leq 0 \\
\lambda^T a + d^T y \frac{1- (\lambda^T a)^2}{\sqrt{\|y\|^2 - (d^T y)^2}}, & \text{otherwise.}
\end{cases}
\]

The following unexpected lemma, whose proof we leave in the appendix, links $\theta(\beta)$ and $r(\beta)$.

**Lemma 6.3** For every $\beta \in D_1(0)$ and every $\mu \in \mathbb{R}$, $\mu > 0$, it holds that $r(\beta) = \theta(\mu \beta)$, where $\theta(\mu \beta)$ is the optimal dual solution of the optimization problem (5.7).

### 6.2.4 $S^g$-freeness and maximality proofs

**Theorem 6.2** Let $\lambda \in D_1(0)$ such that $\lambda \neq \pm a$, $C_{\phi_{\lambda,a,d}}^{g}$ defined in (6.3), that is,

\[
C_{\phi_{\lambda,a,d}}^{g} = \{ (x, y) : -\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq r(\beta), \text{ for all } \beta, \|\beta\| = 1 \}.
\]
with \( r(\beta) \) as in (6.7), and \( S^g \) defined as in (2.4), with \( \|d\| < \|a\| = 1 \). Then, \( C^g_{\phi_{\lambda,a,d}} \) is \( S^g \)-free.

**Proof** \( C^g_{\phi_{\lambda,a,d}} \) is clearly convex. Let \((x_0, y_0) \in S^g\) and let \( \beta_0 = \frac{m}{\|y_0\|} \). The claim will follow if we are able to show that \(-\lambda^T x_0 + \nabla \phi_{\lambda,a,d}(\beta_0)^T y_0 \geq r(\beta_0)\). Since \( x_0 \) satisfies \( \|x_0\| \leq \|y_0\| \) and \( a^T x_0 + d^T y_0 = -1 \), it follows that

\[
\lambda^T x_0 \leq \max_x \{ \lambda^T x : \|x\| \leq \|y_0\|, a^T x + d^T y_0 \leq -1 \}.
\]

By weak duality we have

\[
\max_x \{ \lambda^T x : \|x\| \leq \|y_0\|, a^T x + d^T y_0 \leq -1 \} \leq \inf_{\theta \geq 0} \|y_0\| \|\lambda - a\theta\| - (d^T y_0 + 1)\theta.
\]

Recall that \( \theta(y_0) \) is the optimal dual solution to the optimization problem (5.7), which corresponds to \( \phi_{\lambda,a,d}(y_0) \). Thus, it holds that \( \theta(y_0) \in \mathbb{R}_+ \) and \( \theta(y_0) < +\infty \) because \( \|d\| < 1 \). Consequently,

\[
\inf_{\theta \geq 0} \|y_0\| \|\lambda - a\theta\| - (d^T y_0 + 1)\theta \leq \|y_0\| \|\lambda - a\theta(y_0)\| - (d^T y_0 + 1)\theta(y_0)
\]

\[
= \phi_{\lambda,a,d}(y_0) - \theta(y_0),
\]

where the last equality follows from the strong duality. See Proposition A1. All the inequalities together show that

\[
\lambda^T x_0 \leq \phi_{\lambda,a,d}(y_0) - \theta(y_0).
\]

From Lemma 6.3 it follows \( \theta(y_0) = \theta(\beta_0) = r(\beta_0) \). Thus,

\[
-\lambda^T x_0 + \phi_{\lambda,a,d}(y_0) \geq r(\beta_0),
\]

as we wanted to establish. \( \square \)

**Theorem 6.3** Let \( \lambda, a, d, S^g \) and \( C^g_{\phi_{\lambda,a,d}} \) be as in Theorem 6.2, and \( H = \{(x, y) \in \mathbb{R}^{n+m} : a^T x + d^T y = -1\} \). Then, \( C^g_{\phi_{\lambda,a,d}} \) is maximal \( S^g \)-free with respect to \( H \). Additionally, if \( \lambda = \frac{\bar{x}}{\|\bar{x}\|} \) with \((\bar{x}, \bar{y}) \in H \) and \( \|\bar{x}\| > \|\bar{y}\| \), then \((\bar{x}, \bar{y}) \in \text{int}(C^g_{\phi_{\lambda,a,d}}) \).

**Proof** Theorem 6.2 established \( S^g \)-freeness. To show maximality we will use Theorem 3.3, that is, we will show that every inequality of \( C^g_{\phi_{\lambda,a,d}} \) is either exposed by a point in \( S^g \cap C^g_{\phi_{\lambda,a,d}} \) or exposed at infinity by a sequence in \( S^g \).

Let \( \beta_0 \in D_1(0) \) and consider the valid inequality \(-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta_0)^T y \leq r(\beta_0) \). Assume, first, that \( a^T \lambda + d^T \beta_0 < 0 \) as \( a^T \lambda + d^T \beta_0 < 0 \), we have that \( r(\beta_0) = 0 \), \( \phi_{\lambda,a,d}(\beta_0) = \|\beta_0\| = 1 \), and \( \nabla \phi_{\lambda,a,d}(\beta_0) = \beta_0 \). Hence, the inequality is \(-\lambda^T x + \beta_0^T y \leq 0 \). It is exposed by

\[
-\frac{1}{a^T \lambda + d^T \beta_0}(\lambda, \beta_0) \in S^g \cap C_2 \subseteq S^g \cap C^g_{\phi_{\lambda,a,d}}.
\]
Now, let us assume that $a^T \lambda + d^T \beta_0 \geq 0$. We will show that $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta_0)^T y \leq r(\beta_0)$ is exposed at infinity. Let $(x_n) \subseteq \text{lin}\{\lambda, a\}$ be a sequence converging to $x_{\beta_0}$ as in Lemma 6.2, that is, $\|x_n\| = 1$, $a^T x_n + d^T \beta_0 < 0$, and $r(\beta_0) = \lim_{n \to \infty} \frac{\lambda^T (x_n - x_{\beta_0})}{a^T (x_n - x_{\beta_0})}$. Consider the sequence conformed by

$$z_n = -\frac{(x_n, \beta_0)}{a^T x_n + d^T \beta_0} = \frac{(x_n, \beta_0)}{a^T (x_n - x_{\beta_0})} \in S^g,$$

where the equality above follows from $a^T x_{\beta_0} + d^T \beta_0 = 0$. We proceed to verify that $z_n$ exposes $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta_0)^T y \leq r(\beta_0)$ at infinity.

As $x_n \to x_{\beta_0}$, we have that $\|z_n\| \to \infty$. Also, $\frac{z_n}{\|z_n\|} = \frac{1}{\sqrt{2}} (x_n, \beta_0)$ converges to $v = \frac{1}{\sqrt{2}} (x_{\beta_0}, \beta_0) \in C_2 = \text{rec}(C^g_{\phi_{\lambda,a,d}})$ and exposes $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta_0)^T y \leq 0$.

Finally, we have to show that there exists a $w$ such that $(-\lambda, \nabla \phi_{\lambda,a,d}(\beta_0))^T w = r(\beta_0)$ and $\text{dist}(z_n, w + \text{lin}\{v\}) \to 0$. Let $\hat{x} = \lim_{n \to \infty} \frac{x_n - x_{\beta_0}}{\|x_n - x_{\beta_0}\|}$ and let $w = (-\hat{x}, 0)$. We have that $(-\lambda, \nabla \phi_{\lambda,a,d}(\beta_0))^T w = r(\beta_0)$. Also,

$$z_n - \frac{\sqrt{2}}{a^T (x_{\beta_0} - x_n)} v = \frac{1}{a^T (x_{\beta_0} - x_n)} (x_n - x_{\beta_0}, 0) \to -\left(\frac{\hat{x}}{a^T \hat{x}}, 0\right) = w.$$

Thus, $\text{dist}(z_n, w + \text{lin}\{v\}) \to 0$. □

### 6.2.5 A closed-form formula for $C^g_{\phi_{\lambda,a,d}}$

Since the construction of $C^g_{\phi_{\lambda,a,d}}$ involves translating some of the inequalities of $C_2$ of its outer-description, it is natural to ask if this translation yields a translation of the whole function $\phi_{\lambda,a,d}$. This would yield a closed-form formula for $C^g_{\phi_{\lambda,a,d}}$ which is much more appealing from a computational standpoint. In what follows, we ask whether there exists an $(x_0, y_0)$ such that the following equation holds

$$\{(x, y) : -\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq r(\beta), \text{ for all } \beta, \lambda^T a + d^T \beta \geq 0\}$$

$$= \{(x, y) : -\lambda^T (x - x_0) + \nabla \phi_{\lambda,a,d}(\beta)^T (y - y_0) \leq 0, \text{ for all } \beta, \lambda^T a + d^T \beta \geq 0\}.$$

In order to reach this equality it would suffice to satisfy

$$\lambda^T x_0 - \nabla \phi_{\lambda,a,d}(\beta)^T y_0 = -r(\beta). \quad (6.8)$$

Note that since $\lambda^T a + d^T \beta \geq 0$

$$\nabla \phi_{\lambda,a,d}(\beta) = \sqrt{1 - (\lambda^T a)^2 \frac{W^\beta}{\|\beta\| W}} - \lambda^T a \frac{W^\beta}{\|\beta\| W}$$

$$r(\beta) = \lambda^T a + d^T \beta \sqrt{1 - (\lambda^T a)^2 \frac{W^\beta}{\|\beta\| W}}. \quad (6.9)$$
where \( W = I - dd^T \). Thus (6.8) becomes

\[
\lambda^T (x_0 + ad^Ty_0) - \sqrt{1 - (\lambda^T a)^2} \frac{\beta^T Wy_0}{\|\beta\|_W} = -\lambda^T a - d^T \beta \frac{\sqrt{1 - (\lambda^T a)^2}}{\|\beta\|_W}.
\]

From the last expression, we see that if we are able to find \((x_0, y_0)\) such that

\[
\begin{align*}
  x_0 + ad^Ty_0 &= -a \\  d^T \beta &= \beta^T Wy_0
\end{align*}
\]

then (6.8) would hold. Note that \( d \) is an eigenvector of \( W = I - dd^T \) with eigenvalue \( 1 - \|d\|^2 \). Thus, with \( y_0 = \frac{d}{1-\|d\|^2} \), we can easily check that (6.10b) holds. With \( y_0 \) defined, in order to satisfy (6.10a) it suffices to set

\[
x_0 = -a(d^Ty_0 + 1) = -\frac{a}{1-\|d\|^2}.
\]

Finally, noting that when \( \lambda^T a \|y\| + d^Ty \leq 0 \) we have \( \phi_{\lambda,a,d}(y) = \|y\| \), we arrive to the following expression for \( C_{\phi_{\lambda,a,d}}^g \),

\[
C_{\phi_{\lambda,a,d}}^g = \left\{ (x, y) : \begin{array}{ll}
\|y\| \leq \lambda^T x & \text{if } \lambda^T a \|y\| + d^Ty \leq 0 \\
\phi_{\lambda,a,d} \left( y - \frac{d}{1-\|d\|^2} \right) \leq \lambda^T \left( x + \frac{a}{1-\|d\|^2} \right) & \text{otherwise}
\end{array} \right\}.
\]

(6.11)

This concludes the construction and maximality proofs of all our quadratic-free sets.

7 Remarks on computational aspects and other considerations

The focus of this work has been the construction of maximal quadratic-free sets, which is the main ingredient in the computation of intersection cuts —our main motivation. The explicit computation of such cutting planes and their computational evaluation is the subject of current work; the interested reader can find early promising results in [17,18]. Nonetheless, we would like to point out important observations regarding the numerical aspects of the intersection cut computations using our maximal quadratic-free sets, and to illustrate the resulting cutting plane our procedure would generate in an example. We also show an example of a maximal quadratic-free set that is not generated by our current framework and thus motivates future work.
7.1 Example of a cutting plane computation

Let us consider a set \( \hat{Q} \) defined as the intersection of quadratic and linear inequalities. For the intersection cut computation, we assume we have \( \vec{s} \notin \hat{Q} \) and a simplicial cone \( K \supseteq \hat{Q} \) with apex \( \vec{s} \) —the vector \( \vec{s} \) is typically a vertex of a polyhedral relaxation of \( \hat{Q} \) and \( K \) is obtained from a basis defining \( \vec{s} \). As mentioned in the introduction, since \( \vec{s} \notin \hat{Q} \), we may assume \( \vec{s} \notin Q \), where \( Q \) is defined using a single quadratic inequality of \( \hat{Q} \). Our procedure would generate a maximal \( Q \)-free set which can be used to compute a cutting plane. In what follows, we show this computation explicitly in a continuation of Example 6.3.

Remark 7.1 The computations in this example involve irrational numbers, which we handle exactly using Mathematica 12.3 [30]. In Sect. 7.2 we discuss how the presence of irrational numbers can be handled, which is part of the implementation in the follow-up work [17,18].

Let us consider \( \hat{Q} = Q \cap K \), with

\[
Q = \{ s \in \mathbb{R}^2 : s_1 s_2 + \sqrt{2}s_1 - \sqrt{2}s_2 \leq 1 \},
K = \{ s \in \mathbb{R}^2 : s_1 \geq -3, s_2 \geq -3 \}.
\]

Also consider \( \vec{s} = (-3, 3) \notin Q \), and note that \( K \) is a simplicial cone with apex \( \vec{s} \). The reader can find plots of \( Q, K \) and \( \vec{s} \) in Fig. 7a.

The first step in the cut computation is to construct a canonical representation of the set \( Q \). It can be easily verified that the following is a valid canonical description

\[
S = \{ (x, y) \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq y^2, a^T x + d^T y = -1 \},
\]

with \( a = (-1/\sqrt{2}, 1/\sqrt{2})^T \) and \( d = 1/\sqrt{2} \). Indeed, \( Q \) is the projection of \( S \) onto the \( x \) variables: to see this, it suffices to notice that \( y = -\sqrt{2} + x_1 - x_2 \), and rename the \( x \) variables to \( s \). In the \((x, y)\) variables, the point \( \vec{s} \) to separate is represented as \((\vec{x}, \vec{y}) = (-3, -3, -\sqrt{2})\).

In the second step, we construct the maximal \( S \)-free set with respect to the corresponding hyperplane \( H = \{ (x, y) \in \mathbb{R}^3 : a^T x + d^T y = -1 \} \). Note that \( \lambda = \vec{x}/\|\vec{x}\| = (1/\sqrt{2}, 1/\sqrt{2}) \), therefore the parameters \( \lambda, a, d \) are the same as discussed in Example 6.3. Following this example, we construct the following \( S \)-free set

\[
C^g_{\phi, \lambda, a, d} = \left\{ (x, y) : \frac{1}{\sqrt{2}}(x_1 + x_2) - y \leq 0, \frac{1}{\sqrt{2}}(x_1 + x_2) + \frac{1}{\sqrt{2}}y \leq 1 \right\}.
\]

Returning to the \( s \) variables, this corresponds to the following maximal \( Q \)-free set

\[
C = \left\{ s \in \mathbb{R}^2 : s_1 \leq \sqrt{2}, \left( \frac{1}{\sqrt{2}} - 1 \right)s_1 + \left( \frac{1}{\sqrt{2}} + 1 \right)s_2 + \sqrt{2} \leq 0 \right\}.
\]
In Fig. 7a we display $C$. The maximality of $C$ may not be appreciated in this plot, however, the reader can refer to Fig. 6, which shows $C$ over a larger domain.

Now that we have the $Q$-free set, the third step involves, for each extreme ray $r_j$ of $K$, the calculation of the step-lengths $\alpha^*_j \in (0, \infty]$ defined as $\alpha^*_j = \sup \{ \alpha : \bar{s} + \alpha r_j \in C \}$. The intersection cut is the hyperplane that contains all intersection points $\bar{s} + \alpha^*_j r_j$ when $\alpha^*_j < \infty$, and is parallel to all rays $r_j$ such that $\alpha^*_j = \infty$ (see [20]).

In this case, the extreme rays of $K$ are simply $r^1 = (1, 0)$ and $r^2 = (0, 1)$, and the corresponding intersection points are $s^1 = (\sqrt{2}, -3)$ and $s^2 = (-3, -3\sqrt{2} - 1)\sqrt{2+1}$. In Fig. 7b we show both intersection points and the resulting cutting plane, which is given by

$$4s_1 + \left(5 + 4\sqrt{2}\right)s_2 + 8\sqrt{2} + 15 \geq 0.$$

### 7.2 Numerical considerations

Since the expressions of our maximal quadratic-free sets are not simple in some cases, the reader might wonder how challenging the step-length calculations can be, and if there is a numerically stable way of dealing with the potential presence of irrational intersection points as in the previous example.

Firstly, note that in order to transform $S$ to either $S^h$ or $S^g$ it suffices to use a matrix factorization and homogenization. These can be computed efficiently in practice. As for the step-length calculations using our maximal $S^h$ or $S^g$-free sets, in [17,18] the authors show how to reduce them to computing the roots of single-variable quadratic equations. Additionally, the exact step-length is not needed to obtain a valid cutting plane: a step-back (i.e. a smaller step-length) can be used to obtain a valid cutting plane with a step-length that is computationally more amenable (for example, a rational number). The latter computations can be performed quickly and safely using binary search.

For details such as the explicit equations involved, and a numerically stable path for the calculation of the coefficients, see [17,18].
7.3 A different type of maximal quadratic-free set

While our approach can generate potentially infinitely-many quadratic-free sets that contain any given point in their interior, and can prove their maximality, there are maximal quadratic-free sets that our approach misses.

Consider $H = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 : -x_1 - x_2 = -1\}$, $S^g = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 : x_1^2 + x_2^2 \leq y_1^2 + y_2^2\} \cap H$, and $C = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 : |y_i| \leq x_i, i = 1, 2\}$. The set $C$ can be proven to be maximal $S^g$-free (with respect to $H$), but with the techniques developed in this paper, unfortunately, we cannot prove its maximality. The set $C$ does not have exposing points in $S^g \cap C$ (every element of $S^g \cap C$ belongs to at least two facets of $C \cap H$) and it does not have exposing sequences ($C \cap H$ is a polytope). Future work involves refining our maximality criteria so that it can handle this example and direct the construction of new families of maximal quadratic-free sets.

8 Summary and future work

In this work we have shown how to construct maximal quadratic-free sets, i.e., convex sets whose interiors do not intersect the sublevel set of a quadratic function. We strongly believe that, by carefully laying a theoretical framework for quadratic-free sets, this work provides an important contribution to the understanding and future computational development for solving non-convex quadratically constrained optimization problems.

The maximal quadratic-free sets we construct in this work allow for an efficient computation of the corresponding intersection cuts. Computing such cutting planes amount to solving one-dimensional convex optimization problems using the quadratic-free sets we show here. The closed-form expressions we provide for the sets $C_\lambda, C_{\phi_{\lambda,a,d}},$ and $C^g_{\phi_{\lambda,a,d}}$ ensure this efficient separation in LP-based methods for quadratically constrained optimization problems.

The empirical performance of the resulting intersection cuts is subject of ongoing work. Early results on these lines can be found in [17,18], where a promising computational performance of these cuts is observed in the root node experiments. This computational work shows that a careful implementation can ensure these calculations to be numerically stable as well.

Future work includes a careful comparison of the resulting intersection cuts and the related approaches mentioned in the literature review, in particular the work of [10] and [16]. Other important open questions involve a better understanding of the role different canonical representations of a quadratic set have, since we have evidence that different transformations from $S$ to a set of the type $S^g$ can result in different $S$-free sets. We would also like to perform a theoretical and empirical comparison with the method proposed by Bienstock et al. [10,11]. Finally, motivated by the example in Sect. 7.3, we would like to devise new methods for producing other families of quadratic-free sets and new criteria that can assert their maximality.

Acknowledgements We are indebted to Franziska Schlösser for several inspiring conversations. We would like to thank Stefan Vigerske, Antonia Chmiela, Ksenia Bestuzheva, Nils-Christian Kempke and Joseph Paat for helpful discussions. We would also like to thank the two anonymous reviewers for their valuable
feedback. Lastly, we would like to acknowledge the support of the IVADO Institute for Data Valorization for their support through the IVADO Post-Doctoral Fellowship program and to the IVADO-ZIB academic partnership. The described research activities are funded by the German Federal Ministry for Economic Affairs and Energy within the project EnBA-M (ID: 03ET1549D). The work for this article has been (partly) conducted within the Research Campus MODAL funded by the German Federal Ministry of Education and Research (BMBF Grant Numbers 05M14ZAM, 05M20ZBM). Financial support was also provided by the Government of Chile through the FONDECYT Grant Number 11190515.

A Appendix

A.1 Removing strict convexity matters

In Sect. 4 we construct a maximal $S^h$-free set for $S^h = \{(x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\|\}$ using a concave underestimator of the function $\|x\| - \|y\|$. It is worthwhile to show what happens if we use the same approach with an equivalent description of $S^h$, namely

$$S^h = \{(x, y) \in \mathbb{R}^{n+m} : \|x\|^2 \leq \|y\|^2\}.$$

Recall we assume we are given $(\bar{x}, \bar{y})$ such that $\|\bar{x}\|^2 > \|\bar{y}\|^2$. A concave underestimator of $f(x, y) = \|x\|^2 - \|y\|^2$, tight at $(\bar{x}, \bar{y})$, is given by $\|\bar{x}\|^2 + 2\|\bar{x}\|(x - \bar{x}) - \|y\|^2$. This concave underestimator yields the $S^h$-free set $\{ (x, y) \in \mathbb{R}^{n+m} : \|\bar{x}\|^2 + 2\|\bar{x}\|(x - \bar{x}) - \|y\|^2 \geq 0\}$. A simple example shows that such an $S^h$-free set is not maximal.

Example A.1 The case $n = m = 1$ with $\bar{x} = 3$ yields the $S^h$-free set

$$C = \{(x, y) \in \mathbb{R}^2 : -9 + 6x - y^2 \geq 0\}$$

In Fig. 8 we can see that the set is not maximal $S^h$-free.

The problem seems to be that $\|x\|^2$ is a strictly convex function. Indeed, suppose $S = \{x \in \mathbb{R}^n : f(x) \leq 0\}$ where $f$ is strictly convex. The $S$-free set obtained via a concave underestimator at $\bar{x}$ is $C = \{x \in \mathbb{R}^n : f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) \geq 0\}$. It is not hard to see that the strict convexity of $f$ implies that $C$ is not maximal $S$-free. The reason is that the linearization of $f$ at $\bar{x} \notin S$ will not support the region $S$. On the other hand, if $f$ is instead sublinear, then any linearization of $f$ will support $S$. See [43] for an extended discussion on this phenomenon.

Fig. 8 $S^h$ in Example A.1 (blue) and the $S^h$-free set constructed using a concave underestimator of $\|x\|^2 - \|y\|^2$ (orange) (colour figure online)
A.2 Technical results

Lemma 4.1 Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a sublinear function, $\lambda \in D_1(0)$, and let

$$C = \{(x, y) : \phi(y) \leq \lambda^T x\}.$$

Let $(\bar{x}, \bar{y}) \in C$ be such that $\phi$ is differentiable at $\bar{y}$ and $\phi(\bar{y}) = \lambda^T \bar{x}$. Then $(\bar{x}, \bar{y})$ exposes the valid inequality $-\lambda^T x + \nabla \phi(\bar{y})^T y \leq 0$.

Proof We need to verify both conditions of Definition 3.2. $\alpha$ and $\theta$ are such that $\theta \neq 0$ as otherwise $\alpha = \gamma = 0$ and the inequality would be trivial.

Given that any $(x, 0)$ such that $\lambda^T x = 0$ belongs to $C$, it follows that $\alpha$ is parallel to $\lambda$, i.e., there exists $v \in \mathbb{R}$ such that $\alpha = v \lambda$. Furthermore, $(\mu \lambda, 0) \in C$ for every $\mu \geq 0$, implies that $0 > \alpha^T \lambda = v$. Therefore, $\gamma = -v \nabla \phi(\bar{y})$ and the inequality reads $v \lambda^T (x-\bar{x}) - v \nabla \phi(\bar{y})^T (y-\bar{y}) \leq 0$. Dividing by $|v|$ and using that $-\lambda^T x + \nabla \phi(\bar{y})^T y \leq 0$ is tight at $(\bar{x}, \bar{y})$, we conclude that the inequality can be written as

$$-\lambda^T x + \nabla \phi(\bar{y})^T y \leq 0.$$ 

Proposition A1 Let $a, \lambda \in D_1(0), \lambda \neq \pm a$ and let $d \in \mathbb{R}^m$ be such that $\|d\| \leq 1$. The (Lagrangian) dual problem of

$$\max_{x} \{\lambda^T x : \|x\| \leq \|y\|, a^T x + d^T y \leq 0\} \quad (A.1)$$

is

$$\inf_{\theta} \{\|\lambda - \theta a\| \|y\| - \theta d^T y : \theta \geq 0\}. \quad (A.2)$$
The optimal solution to (A.1) is $x : \mathbb{R}^m \to \mathbb{R}^n$,

$$x(y) = \begin{cases} 
\lambda \|y\|, & \text{if } \lambda^Ta \|y\| + d^Ty \leq 0 \\
\sqrt{\|y\|^2 - (d^Ty)^2} \lambda - (d^Ty + \lambda^T a \sqrt{\|y\|^2 - (d^Ty)^2}) a, & \text{otherwise.}
\end{cases}$$

(A.3)

The optimal dual solution is $\theta : \mathbb{R}^m \to \mathbb{R}_+ \cup \{+\infty\}$,

$$\theta(y) = \begin{cases} 
0, & \text{if } \lambda^Ta \|y\| + d^Ty \leq 0 \\
\lambda^T a + d^Ty \sqrt{\frac{1-(\lambda^T a)^2}{\|y\|^2 - (d^Ty)^2}}, & \text{otherwise.}
\end{cases}$$

(A.4)

Here, $\frac{1}{0} = +\infty$ and $r + (+\infty) = +\infty$ for every $r \in \mathbb{R}$. Moreover, strong duality holds, that is, (A.1) = (A.2), and

$$\min_{x} \{ \|y\|, \sqrt{(\|y\|^2 - (d^Ty)^2)(1 - (\lambda^T a)^2)} - d^Ty \lambda^T a \} \quad (A.5)$$

Finally, (A.5) holds even if $\lambda = \pm a$.

Proof First, note that since $\lambda \neq \pm a$ and $\|d\| \leq 1$, $x(y)$ and $\theta(y)$ are defined for every $y \in \mathbb{R}^m$. Second, to make some of the calculations that follow more amenable, let $S(y) = \sqrt{\frac{\|y\|^2 - (d^Ty)^2}{1 - (\lambda^T a)^2}}$. The Lagrangian of (A.1) is $L : \mathbb{R}^n \times \mathbb{R}^2_+ \to \mathbb{R}$,

$$L(x, \mu, \theta) = \lambda^T x - \mu(\|x\| - \|y\|) - \theta(a^T x + d^Ty).$$

Thus, the dual function is

$$d(\mu, \theta) = \max_x L(x, \mu, \theta).$$

We have that $d(\mu, \theta)$ is infinity whenever $\mu < \|\lambda - a\theta\|$, and $\mu \|y\| - \theta d^Ty$ otherwise. Hence, the dual problem, $\min_{\theta, \mu \geq 0} d(\mu, \theta)$, is $\min_{\mu \|y\| - \theta d^Ty} : \theta \geq 0, \mu \geq \|\lambda - a\theta\|$ which is (A.2).

Let us assume that $\lambda^T a \|y\| + d^Ty \leq 0$. Clearly, $x(y) = \lambda \|y\|$ is feasible for (A.1). Its objective value is $\|y\|$. On the other hand, $\theta(y) = 0$ is always feasible for (A.1). Its objective value is also $\|y\|$, therefore, $x(y)$ is the primal optimal solution and $\theta(y)$ the dual optimal solution.

Now let us consider the case $\lambda^T a \|y\| + d^Ty > 0$. Let us check that $\theta(y)$ is dual feasible, that is, $\theta(y) \geq 0$. Note that, due to the positive homogeneity of $\theta(y)$ and the condition $\lambda^T a \|y\| + d^Ty > 0$ with respect to $y$, we can assume without loss of generality that $\|y\| = 1$. 

\[ Springer \]
Let $\alpha = \lambda^T a$ and $\beta = d^T y$. Since $\theta(d) = +\infty \geq 0$ when $\|d\| = 1$, we can assume that $y \neq d$ when $\|d\| = 1$. Note that the same does not occur when $y = -d$ since we are assuming $\lambda^T a \|y\| + d^T y > 0$. Thus, $\alpha, \beta \in (-1, 1)$.

We will prove that $\theta(y)\sqrt{1 - \beta^2} = \alpha\sqrt{1 - \beta^2} + \beta\sqrt{1 - \alpha^2} \geq 0$, which implies that $\theta(y) \geq 0$. If $\alpha, \beta \geq 0$, then we are done. As $\alpha + \beta > 0$, at least one of them must be positive. Let us assume $\alpha > 0$ and $\beta < 0$, the other case is analogous. Then, $\alpha > -\beta \geq 0$. This implies that $\alpha^2 > \beta^2$. Subtracting $\alpha^2 \beta^2$, factorizing and taking square roots we obtain the desired inequality.

Let us compute the value of the dual solution $\theta(y)$. First, $y = d$ and $\|d\| = 1$, $\theta(y) = +\infty$, which means that the optimal value is

$$\lim_{\theta \to +\infty} \|\lambda - \theta a\| - \theta = -\lambda^T a.$$ 

One way of computing this limit is to multiply and divide the expression by $\|\lambda - \theta a\| + \theta$, expand, and simplify the numerator and denominator until one obtains something simple enough.

Now assume $y \neq d$ if $\|d\| = 1$. Observe that $\|\lambda - \theta(y) a\| \|y\| - \theta(y)d^T y = \sqrt{\|\lambda - \theta(y) a\|^2 \|y\| - \theta(y)d^T y}$. We have that

$$\|\lambda - \theta(y) a\|^2 = 1 + \theta(y)(\theta(y) - 2\lambda^T a)$$

$$= 1 + (\theta(y) - \lambda^T a + \lambda^T a)(\theta(y) - \lambda^T a - \lambda^T a)$$

$$= 1 + (\theta(y) - \lambda^T a)^2 - (\lambda^T a)^2.$$ 

Replacing $\theta(y)$, we obtain

$$\|\lambda - \theta(y) a\|^2 = 1 + \frac{(d^T y)^2}{S(y)} - (\lambda^T a)^2$$

$$= \frac{1}{S(y)}(S^2(y)(1 - (\lambda^T a)^2) + (d^T y)^2) = \frac{\|y\|^2}{S^2(y)}.$$ 

Therefore,

$$\|\lambda - \theta(y) a\| \|y\| - \theta(y)d^T y = \frac{\|y\|^2}{S(y)} - d^T y \lambda^T a - \frac{(d^T y)^2}{S(y)}$$

$$= \frac{\|y\|^2 - (d^T y)^2}{S(y)} - d^T y \lambda^T a$$

$$= \sqrt{\|y\|^2 - (d^T y)^2(1 - (\lambda^T a)^2) - d^T y \lambda^T a}.$$ 

Let us now check the feasibility of $x(y)$. Let us first check that $\|x(y)\|^2 \leq \|y\|^2$. We have $\|x(y)\|^2 = S^2(y) - 2S(y)(d^T y + \lambda^T a) + (d^T y + \lambda^T a S(y))^2$. Expanding and removing common terms yields $\|x(y)\|^2 = S^2(y)(1 - (\lambda^T a)^2) + (d^T y)^2 = \|y\|^2$. Thus, the first constraint is satisfied.

To check the second constraint just notice that, as $\|a\| = 1$, $a^T x(y) = -d^T y$. 

Springer
The primal value of $x(y)$ is

$$\lambda^T x(y) = S(y)(1 - (\lambda^T a)^2) - d^T y \lambda^T a = \sqrt{(\|y\|^2 - (d^T y)^2)(1 - (\lambda^T a)^2) - d^T y \lambda^T a}.$$ 

As it coincides with the value of the dual solution, even when $y = d$ and $\|d\| = 1$, we conclude that both are optimal.

It only remains to check (A.5) for $\lambda = \pm a$. If $\lambda = -a$, then the linear constraint becomes $\lambda^T x \geq d^T y$ and the optimal solution is $x = \lambda \|y\|$. If $\lambda = a$, then the linear constraint becomes $\lambda^T x \leq -d^T y$ and $x = -d^T y \lambda$ is then optimal. In both cases (A.5) holds.

**Lemma A1** Consider the set

$$S^g = \{(x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\|, a^T x + d^T y = -1\}$$

with $a$, $d$ such that $\|a\| > \|d\|$. Let $\lambda, \beta \in D_1(0)$ be two vectors satisfying $\lambda^T a + d^T \beta \geq 0$ and consider $C_2$ defined in (5.8).

Then, the face of $C_2$ defined by the valid inequality $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\beta)^T y \leq 0$ does not intersect $S^g$.

**Proof** Recall that in this case $C_2 = C_{\phi_{\lambda,a,d}}$. By contradiction, suppose that $(\bar{x}, \bar{y}) \in C_2$ is such that

$$(\bar{x}, \bar{y}) \in S^g \land -\lambda^T \bar{x} + \nabla \phi_{\lambda,a,d}(\beta)^T \bar{y} = 0.$$ 

The latter equality and the fact that $\phi_{\lambda,a,d}$ is sublinear implies $\phi_{\lambda,a,d}(\bar{y}) = \lambda^T \bar{x}$. Moreover, $\bar{x}$ is a feasible solution of the optimization problem $\phi_{\lambda,a,d}(\bar{y})$, which implies it is an optimal solution.

By Lemma 4.1 we know $(\bar{x}, \bar{y})$ exposes the valid inequality of $C_2$ given by $-\lambda^T x + \nabla \phi_{\lambda,a,d}(\bar{y})^T y \leq 0$. By definition of exposing point this means

$$\nabla \phi_{\lambda,a,d}(\bar{y}) = \nabla \phi_{\lambda,a,d}(\beta).$$

From (6.9), since $W$ is invertible, we can see that this implies $\beta = \frac{\bar{y}}{\|\bar{y}\|}$. However, as $\lambda^T a + d^T \beta \geq 0$, the optimal solution in the definition of $\phi_{\lambda,a,d}(\bar{y}), x_0$, must satisfy $a^T x_0 + d^T \bar{y} = 0$. This contradicts $\phi_{\lambda,a,d}(\bar{y}) = \lambda^T \bar{x}$, since $\bar{x}$ is an optimal solution but $a^T \bar{x} + d^T \bar{y} = -1$.

**Lemma 6.3** For every $\beta \in D_1(0)$ and every $\mu \in \mathbb{R}, \mu > 0$, it holds that $r(\beta) = \theta(\mu \beta)$, where $\theta(\mu \beta)$ is the optimal dual solution of the optimization problem (5.7).

**Proof** From (A.4), we see that it suffices to show the result for $\mu = 1$. If $\lambda^T a + d^T \beta \leq 0$, $r(\beta) = 0 = \theta(\beta)$. Let $\beta \in D_1(0)$ be such that $\lambda^T a + d^T \beta > 0$. Then,

$$r(\beta) = \frac{d^T \beta + \lambda^T a \phi_{\lambda,a,d}(\beta)}{\phi_{\lambda,a,d}(\beta) + d^T \beta \lambda^T a} = \frac{d^T \beta + \lambda^T a \sqrt{1 - (\lambda^T a)^2} \sqrt{1 - (d^T \beta)^2} - d^T \beta (\lambda^T a)^2}{\sqrt{1 - (\lambda^T a)^2} \sqrt{1 - (d^T \beta)^2}}.$$
\[
\begin{align*}
\frac{d^T \beta \sqrt{1 - (\lambda^T a)^2}}{\sqrt{1 - (d^T \beta)^2}} + \lambda^T a &= \theta(\beta).
\end{align*}
\]

**References**

1. Andersen, K., Jensen, A.N.: Intersection cuts for mixed integer conic quadratic sets. In: Goemans, M., Correa, J. (eds.) Integer Programming And Combinatorial Optimization, pp. 37–48. Springer (2013)
2. Andersen, K., Louveaux, Q., Weismantel, R.: An analysis of mixed integer linear sets based on lattice point free convex sets. Math. Oper. Res. 35(1), 233–256 (2010)
3. Andersen, K., Louveaux, Q., Weismantel, R., Wolsey, L.A.: Inequalities from two rows of a simplex tableau. In: Integer Programming and Combinatorial Optimization, pp. 1–15. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-72792-7_1
4. Balas, E.: Intersection cuts—a new type of cutting planes for integer programming. Oper. Res. 19(1), 19–39 (1971). https://doi.org/10.1287/opre.19.1.19
5. Basu, A., Conforti, M., Cornuèjols, G., Zambelli, G.: Maximal lattice-free convex sets in linear subspaces. Math. Oper. Res. 35(3), 704–720 (2010). https://doi.org/10.1287/moor.1100.0461
6. Basu, A., Conforti, M., Cornuèjols, G., Zambelli, G.: Minimal inequalities for an infinite relaxation of integer programs. SIAM J. Discrete Math. 24(1), 158–168 (2010)
7. Basu, A., Dey, S.S., Paat, J.: Nonunique lifting of integer variables in minimal inequalities. SIAM J. Discrete Math. 33(2), 755–783 (2019). https://doi.org/10.1137/17m1117070
8. Belotti, P., Góez, J.C., Pólik, I., Ralphs, T.K., Terlaky, T.: On families of quadratic surfaces having fixed intersections with two hyperplanes. Discrete Appl. Math. 161(16–17), 2778–2793 (2013)
9. Belotti, P., Góez, J.C., Pólik, I., Ralphs, T.K., Terlaky, T.: A conic representation of the convex hull of disjunctive sets and conic cuts for integer second order cone optimization. In: Numerical Analysis and Optimization, pp. 1–35. Springer (2015)
10. Bienstock, D., Chen, C., Muñoz, G.: Outer-product-free sets for polynomial optimization and oracle-based cuts. Math. Program 1–44 (2020)
11. Bienstock, D., Chen, C., Muñoz, G.: Intersection cuts for polynomial optimization. In: Integer Programming and Combinatorial Optimization, pp. 72–87. Springer (2019). https://doi.org/10.1007/978-3-030-17953-3_6
12. Bodur, M., Dash, S., Günlük, O.: Cutting planes from extended lp formulations. Math. Program. 161(1–2), 159–192 (2017)
13. Bonami, P., Linderoth, J., Lodi, A.: Disjunctive cuts for mixed integer nonlinear programming problems. Progress in Combinatorial Optimization, pp. 521–541 (2011)
14. Borozan, V., Cornuèjols, G.: Minimal valid inequalities for integer constraints. Math. Oper. Res. 34(3), 538–546 (2009). https://doi.org/10.1287/moor.1080.0370
15. Boyd, S., Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)
16. Burer, S., Kilimç-Karzan, F.: How to convexify the intersection of a second order cone and a nonconvex quadratic. Math. Program. 162(1–2), 393–429 (2017)
17. Chmiela, A., Muñoz, G., Serrano, F.: On the implementation and strengthening of intersection cuts for qcqs. Tech. Rep. 20-29, ZIB, Takustr. 7, 14195 Berlin (2020)
18. Chmiela, A., Muñoz, G., Serrano, F.: On the implementation and strengthening of intersection cuts for qcqs. In: IPCO, pp. 134–147 (2021)
19. Conforti, M., Cornuèjols, G., Danilidis, A., Lemaréchal, C., Malick, J.: Cut-generating functions and S-free sets. Math. Oper. Res. 40(2), 276–391 (2015). https://doi.org/10.1287/moor.2014.0670
20. Conforti, M., Cornuèjols, G., Zambelli, G.: Corner polyhedron and intersection cuts. Sur. Oper. Res. Manag. Sci. 16(2), 105–120 (2011). https://doi.org/10.1016/j.sorms.2011.03.001
21. Conforti, M., Cornuèjols, G., Zambelli, G.: Integer Programming. Springer (2014)
22. Cornuèjols, G., Wolsey, L., Yıldız, S.: Sufficiency of cut-generating functions. Math. Program. 152(1–2), 643–651 (2015)
23. Dey, S.S., Wolsey, L.A.: Lifting integer variables in minimal inequalities corresponding to lattice-free triangles. In: Lodi, A., Panconesi, A., Rinaldi, G. (eds.) Integer Programming and Combinatorial Optimization, pp. 463–475. Springer (2008)
24. Dey, S.S., Wolsey, L.A.: Constrained infinite group relaxations of MIPs. SIAM J. Optim. 20(6), 2890–2912 (2010). https://doi.org/10.1137/090754388
25. Fischetti, M., Ljubić, I., Monaci, M., Sinnl, M.: Intersection cuts for bilevel optimization. In: Integer Programming and Combinatorial Optimization, pp. 77–88. Springer (2016). https://doi.org/10.1007/978-3-319-33461-5_7
26. Fischetti, M., Monaci, M.: A branch-and-cut algorithm for mixed-integer bilinear programming. Eur. J. Oper. Res. (2019). https://doi.org/10.1016/j.ejor.2019.09.043
27. Glover, F.: Convexity cuts and cut search. Oper. Res. 21(1), 123–134 (1973). https://doi.org/10.1287/opre.21.1.123
28. Goberna, M., González, E., Martínez-Legaz, J., Todorov, M.: Motzkin decomposition of closed convex sets. J. Math. Anal. Appl. 364(1), 209–221 (2010). https://doi.org/10.1016/j.jmaa.2009.10.015
29. Gomory, R.E., Johnson, E.L.: Some continuous functions related to corner polyhedra. Math. Program. 3–3(1), 23–85 (1972). https://doi.org/10.1007/bf01584976
30. Inc., W.R.: Mathematica, Version 12.3.1. https://www.wolfram.com/mathematica. Champaign, IL (2021)
31. Kilınç-Karzan, F.: On minimal valid inequalities for mixed integer conic programs. Math. Oper. Res. 41(2), 477–510 (2015)
32. Kilınç-Karzan, F., Yıldız, S.: Two-term disjunctions on the second-order cone. Math. Program. 154(1–2), 463–491 (2015)
33. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. 11(3), 796–817 (2001)
34. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Emerging Applications of Algebraic Geometry, pp. 157–270. Springer (2009)
35. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: part i—convex underestimating problems. Math. Program. 10(1), 147–175 (1976). https://doi.org/10.1007/bf01580665
36. Modaresi, S., Kilınç, M.R., Vielma, J.P.: Split cuts and extended formulations for mixed integer conic quadratic programming. Oper. Res. Lett. 43(1), 10–15 (2015)
37. Modaresi, S., Kilınç, M.R., Vielma, J.P.: Intersection cuts for nonlinear integer programming: Convexification techniques for structured sets. Math. Program. 155(1–2), 575–611 (2016)
38. Morán, D., Dey, S.S.: On maximal s-free convex sets. SIAM J. Discrete Math. 25(1), 379–393 (2011). https://doi.org/10.1137/100796947
39. Muñoz, G., Serrano, F.: Maximal quadratic-free sets. In: International Conference on Integer Programming and Combinatorial Optimization, pp. 307–321. Springer, Cham (2020)
40. Rockafellar, R.T.: Convex Analysis. Princeton University Press (1970)
41. Santana, A., Dey, S.S.: The convex hull of a quadratic constraint over a polytope. Preprint arXiv:1812.10160 (2018)
42. Serrano, F.: Intersection cuts for factorable MINLP. In: Integer Programming and Combinatorial Optimization, pp. 385–398. Springer (2019). https://doi.org/10.1007/978-3-030-17953-3_29
43. Serrano, F., Schwarz, R., Gleixner, A.: On the relation between the extended supporting hyperplane algorithm and kelley’s cutting plane algorithm. Preprint arXiv:1905.08157 (2019)
44. Shahabsafa, M., Góez, J.C., Terlaky, T.: On pathological disjunctions and redundant disjunctive conic cuts. Oper. Res. Lett. 46(5), 500–504 (2018)
45. Shor, N.Z.: Quadratic optimization problems. Soviet J. Comput. Syst. Sci. 25, 1–11 (1987)
46. Solovev, V.N.: On a criterion for convexity of a positive-homogeneous function. Math. USSR-Sbornik 46(2), 285–290 (1983). https://doi.org/10.1070/sm1983v046n02abeh002787
47. Towle, E., Luedtke, J.: Intersection disjunctions for reverse convex sets. Preprint arXiv:1901.02112 (2019)
48. Tuy, H.: Concave programming with linear constraints. In: Doklady Akademii Nauk, vol. 159, pp. 32–35. Russian Academy of Sciences (1964)
49. Yıldız, S., Kilınç-Karzan, F.: Low-complexity relaxations and convex hulls of disjunctions on the positive semidefinite cone and general regular cones. Optimization Online (2016)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.