Noname manuscript No.
(will be inserted by the editor)

Monoidal strengthening and unique lifting in MIQCPs

Antonia Chmiela · Gonzalo Muñoz · Felipe Serrano

Received: date / Accepted: date

Abstract Using the recently proposed maximal quadratic-free sets and the well-known monoidal strengthening procedure, we show how to improve intersection cuts for quadratically-constrained optimization problems by exploiting integrality requirements. We provide an explicit construction that allows an efficient implementation of the strengthened cuts along with computational results showing their improvements over the standard intersection cuts. We also show that, in our setting, there is *unique lifting* which implies that our strengthening procedure is generating the best possible cut coefficients for the integer variables.

Keywords MIQCP · monoidal strengthening · unique lifting.

1 Introduction

In recent years, we have seen multiple efforts in generating valid linear inequalities to quadratically constrained quadratic programs (QCQPs) which, using

A. Chmiela
Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany
E-mail: chmiela@zib.de

G. Muñoz
Universidad de O'Higgins, Rancagua, Chile
E-mail: gonzalo.munoz@uoh.cl

F. Serrano
Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany
E-mail: serrano@zib.de

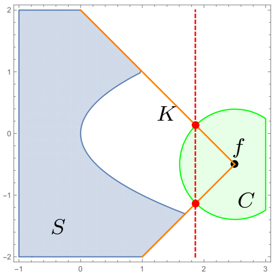


Fig. 1.1: An intersection cut (red) separating f from S (blue). The cut is computed using the intersection points of an S-free set C (green) and the rays of a simplicial cone $K \supseteq S$ (boundary in orange) with apex $f \notin S$. Figure obtained from [8].

an epigraph formulation, we can assume have the form

$$\min \quad \bar{c}^\mathsf{T} s \tag{1.1a}$$
$$\text{s.t.} \quad s \in S \subseteq \mathbb{R}^p, \tag{1.1b}$$

where

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

One of the approaches to generate such valid inequalities has been the *intersection cut* paradigm [21,1,15] which works as follows. We assume we have $f \notin S$, a basic feasible solution of a *linear programming* (LP) relaxation of (1.1). Additionally, we assume we have a simplicial conic relaxation $K \supseteq S$ with apex f, and an S-free set C—a convex set satisfying $\operatorname{int}(C) \cap S = \emptyset$—such that $f \in \operatorname{int}(C)$. Using these ingredients, we can find a cutting plane separating f from S. In Figure 1.1 we show a simple intersection cut in the case when all p rays of K intersect the boundary of the S-free set C. In such case, the intersection cut is simply defined by the hyperplane containing all such intersection points. It is well known that one can assume C to be described as $C = \{s \in \mathbb{R}^p \,:\, \phi(s - f) \leq 1\}$ where ϕ is a sublinear function. Given the constraint $f + \sum_{i=1}^p r^i s_i \in S$ with $s_i \in \mathbb{R}_+$ and $r^i \in \mathbb{R}^p$ (e.g. the extreme rays of K), the intersection cut separating f is

$$\sum_{i=1}^{p} \phi(r^i) s_i \geq 1. \tag{1.2}$$

Muñoz and Serrano [18] recently provided a method for constructing *maximal* quadratic-free sets for any arbitrary quadratic inequality, which would ensure separation of any $f \notin S$. Subsequently, Chmiela et al. [8] showed how to implement these cuts with positive results in a broad class of problems.

One of the limitations of these cutting planes is that they do not use any integrality information: if we were to add integrality requirements to (1.1)—thus obtaining an MIQCP—the intersection cuts would be completely oblivious to this.

In this work, we remedy this via the *monoidal strengthening* framework [2]; a strengthening of intersection cuts based on integrality information. Let S be a closed set and suppose we have $f \notin S$ and a maximal S-free set $C \ni f$. In order to strengthen the intersection cut that separates f, monoidal strengthening aims at finding a monoid—a set containing the origin and closed under addition—M such that C is $S+M$-free. If such a monoid is found, then we can modify the rays associated to integer non-basic variables to obtain a stronger cut coefficient: when s^i is an integer variable, the coefficient $\phi(r^i)$ in (1.2) can be improved to $\psi(r^i) := \inf_{m \in M} \phi(r^i + m)$. We provide a more in-detail overview of this procedure in Section 2.

In this paper, we show how to apply monoidal strengthening to intersection cuts obtained from maximal quadratic-free sets [18,8]. As noted in [18], using linear transformations (diagonalization and homogenization), one can shift the focus from a generic quadratic set, $S = \{s \in \mathbb{R}^p \,:\, s^\mathsf{T} Q s + b^\mathsf{T} s + c \leq 0\}$, to one of the following two sets:

$$S^h := \{(x,y) \in \mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|\}, \tag{1.3}$$

$$S^g := \{(x,y) \in \mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|,\, a^\mathsf{T} x + d^\mathsf{T} y = -1\}. \tag{1.4}$$

where $\max\{\|a\|, \|d\|\} = 1$. Whether S gets mapped to S^h or S^g depends on whether the quadratic defining S is homogeneous or not. Thus, one of the goals of this paper will be: using C as maximal S^h- and S^g-free sets of [18], to find a monoid M such that C is $S^h + M$- or $S^g + M$-free and, subsequently, strengthen the corresponding intersection cut.

Monoidal strengthening is also related to lifting [12,10,3,13]. The function ψ is such that with it we obtain *sequence independent lifting*, i.e. we can apply the strengthening to all integer variables simultaneously. However, in general, there could be other lifting functions that produce better improvements than ψ. Our final goal is to show that, choosing ϕ carefully, the ψ function we obtain yields the best possible coefficient an integer variable can achieve. This establishes that in our case there is *unique lifting* (see Section 2 for a overview).

Contributions Our main contributions are: (1) we show that the monoidal strengthening framework does not produce any strengthening when S is defined using a homogeneous quadratic; (2) in the non-homogeneous case, we show an explicit monoid construction based on a maximal S^g-free set of [18] which can be used for monoidal strengthening; (3) we show an explicit formula for how to efficiently compute $\psi(r)$ in practice; (4) we show that in our setting there is unique lifting which, in particular, implies that ψ yields the best coefficients in the strengthening of the intersection cut; and (5) we present extensive computational results that show the impact of this strengthening procedure.

2 Monoidal strengthening and unique lifting

Let S be a closed set and $f \in \mathbb{R}^n$ such that $f \notin S$. Let ϕ be a sublinear function, and let $C = \{x \in \mathbb{R}^n : \phi(x - f) \leq 1\}$ be an S-free set. Given the constraint $f + \sum_{i=1}^{p} r^i s_i \in S$ with $s_i \in \mathbb{R}_+$ and $r^i \in \mathbb{R}^n$, the intersection cut separating f is $\sum_{i=1}^{p} \phi(r^i) s_i \geq 1$.

Monoidal strengthening leverages the fact that some of the s_i are integer. The idea is to take the relation $f + \sum_{i=1}^{p} r^i s_i \in S$ and modify it in the following way. Assume that all s_i are restricted to be integer. The above relation implies that $f + \sum_{i=1}^{p} (r^i + m_i) s_i \in S + \sum_{i=1}^{p} m_i s_i$. The points $\sum_{i=1}^{p} m_i s_i$ form a monoid $M = \{m : m = \sum_{i=1}^{p} m_i s_i, s_i \in \mathbb{Z}_+\}$. Thus, we obtain the new relation: $f + \sum_{i=1}^{p} (r^i + m_i) s_i \in S + M$. If it turns out that C is still $S + M$ free, then we can use the function ϕ to generate a new cut. The above is summarized in the following result.

Theorem 2.1 ([2] Theorem 1) *Let M be a monoid such that C is $S + M$-free and let $I = \{i \in [p] : s_i \in \mathbb{Z}\}$ be the index set of the integer variables. Then,*

$$\sum_{i \notin I} \phi(r^i) s_i + \sum_{i \in I} \inf_{m \in M} \phi(r^i + m) s_i \geq 1$$

is valid and dominates the intersection cut.

The burden of the above technique is to find a monoid M such that C stays $S+M$-free. Note that, equivalently, we can find a monoid M such that $C - M$ is (possibly non-convex) S-free. Note that $C - M$ is the union of the displacements of C by the elements of $-M$. Thus, to start constructing a possible monoid to use for strengthening we can start with the simpler problem of finding a direction, r such that $C - r\mathbb{Z}_+$ is S-free (we follow this approach in Section 4).

Monoidal strengthening was introduced in [2]. In that paper, Balas and Jeroslow show how to recover GMI (Gomory Mixed Integer) cuts [16] using monoidal strengthening. For GMI cuts, the set S is $\mathbb{Z}$ and since S itself is a monoid (it is a group), then one can apply monoidal strengthening, and recover the GMI cuts, using $M = \mathbb{Z}$. However, the main application of monoidal strengthening in [2] is to disjunctive cuts. Loosely speaking, they consider the following "basic truncated" disjunctive set, $S = \{y \in \mathbb{R}^p : y_i \geq b_i, \bigvee_i y_i \geq 1\}$, with $b_i < 1$. As we will see in Theorem 3.1, monoidal strengthening is not applicable when the S-free set is maximal and S is a cone, therefore, the condition $y_i \geq b_i$ is needed for the application of monoidal strengthening. In such a setting, $C = \{y \in \mathbb{R}^p : y_i \leq 1\}$ is a maximal S-free set that yields the standard disjunctive cut and $M = \{m \in \mathbb{R}^p : m_i = (1 - b_i) z_i, \sum_i z_i \geq 0, z_i \in \mathbb{Z}\}$ is a monoid such that $C - m$ is S-free for every $m \in M$.

Monoidal strengthening is also related to lifting. A common strategy for constructing a cut generating function for the infinite model [11], $W(S) = \{(x, y) : f + \sum_{r\in\mathbb{R}^n} rx(r) + \sum_{r\in\mathbb{R}^n} ry(r) \in S, x(r) \geq 0, y(r) \in \mathbb{Z}_+, x, y \text{ have finite support}\}$, is to construct a cut generating function for the continuous set and then lift it. Let C be an S-free set and let ϕ be a sublinear function such that $C = \{x : \phi(x - f) \leq 1\}$. Then, (ϕ, ϕ) is a valid cut generating pair for $W(S)$, i.e., $\sum_{r\in\mathbb{R}^n} \phi(r)x(r) + \sum_{r\in\mathbb{R}^n} \phi(r)y(r) \geq 1$ is valid for $W(S)$. If M is a monoid such that C is $S + M$ free, and $\psi(r) = \inf_{m\in M} \phi(r + m)$, then (ϕ, ψ) becomes a valid pair. A nice property of ψ is that it is subadditive[1]. This implies that with ψ we obtain *sequence independent lifting*, i.e., we can lift all integer variables at the same time. This is readily seen from Theorem 2.1, where in the final cut ψ is applied to each integer variable.

In the case where $S = \mathbb{Z}^n \cap P$ where P is a polyhedron, a simple monoid M that satisfies that C is $S + M$-free is $M = \mathbb{Z}^n \cap \mathrm{lin}(\mathrm{conv}(P))$. This monoid works because $\mathbb{Z}^n \cap \mathrm{lin}(\mathrm{conv}(P)) \subseteq S$. Due to the simplicity of the monoid, the function ψ is also known as *trivial lifting* [12,10,3,13]

However, the monoidal lifting function ψ is, in general, just one possible way of lifting. We can define the best possible coefficient that a particular integer variable can achieve, with the so-called *lifting function*. The lifting function is given by

$$\pi(r) = \sup\left\{\frac{1 - \phi(s)}{\sigma} : f + s + \sigma r \in S, \sigma \in \mathbb{Z}_{\geq 1}\right\}.$$

In general, π is not subadditive so we do not have sequence independent lifting with it [5]. When it is subadditive, we say that there is *unique lifting*, because π dominates any other lifting. For the case $S = \mathbb{Z}^n \cap P$ it is well understood when we have unique lifting [3]. Roughly speaking, there is a region R_ϕ, called the *lifting region* [10], which contains the origin and has non-empty interior, such that π and all valid minimal[2] liftings coincide with the continuous cut generating function ϕ. Intuitively, this is saying that if r is short enough, then the contribution of $ry(r)$ in $W(S)$ is independent, as far as lifting is concern, of whether $y(r) \in \mathbb{Z}_+$ or $y(r) \in \mathbb{R}_+$. In [10] they show that if $R_\phi + M = \mathbb{R}^n$, where $M = \mathbb{Z}^n \cap \mathrm{lin}(\mathrm{conv}(P))$, then there is unique lifting. The idea is that one can apply monoidal strengthening to a minimal valid lifting π' to show that, just like the monoidal lifting function, $\pi'(r + m) = \pi'(r)$. Thus, if $R_\phi + M = \mathbb{R}^n$, then for $r \in \mathbb{R}^n$ there are $r_0 \in R_\phi, m \in M$ such that $r = r_0 + m$ and so $\pi'(r) = \pi'(r_0 + m) = \pi'(r_0) = \phi(r_0)$. Hence all minimal valid liftings are the same and we have unique lifting.

Unfortunately, as we will see in Section 4, the monoid that we construct does not have a "periodic" structure that allows us to use a similar reasoning as

[1] A function ψ is *subadditive* if $\psi(x + y) \leq \psi(x) + \psi(y)$

[2] A cut generating function is called minimal if it is not point-wise dominated by another cut generating function.

above. To prove unique lifting in our setting, i.e., that we have sequence independent lifting with π, we will prove that π is subadditive and we will achieve this by applying the following result.

Lemma 2.1 *Let M be a monoid such that C is $S+M$-free and*

$$\pi_1(r) = \sup\{1 - \phi(s) \,:\, f + s + r \in S\}.$$

If π_1 is subadditive, then $\pi = \pi_1$ and we have unique lifting.

Proof By definition $\pi(r) = \sup_{\sigma \in \mathbb{Z}_{\geq 1}} \frac{\pi_1(\sigma r)}{\sigma}$. Since, $\pi_1(r) \leq \sup_{\sigma \in \mathbb{Z}_{\geq 1}} \frac{\pi_1(\sigma r)}{\sigma}$, we always have that $\pi_1(r) \leq \pi(r)$.

Now, given that π_1 is subadditive, $\pi_1(\sigma r) \leq \sigma \pi_1(r)$. Therefore, $\frac{\pi_1(\sigma r)}{\sigma} \leq \pi_1(r)$ and we conclude that $\pi(r) \leq \pi_1(r)$. Thus, $\pi = \pi_1$.

It only remain to prove that π_1 is subadditive, which we do by showing that $\pi_1 = \psi$ in Section 6.

3 The homogeneous case: S^h

In this section, we analyze the case of S^h defined in (1.3) and show that the monoidal strengthening framework does not produce any improvements when the cuts are created using *maximal* S^h-free sets. The main reason behind this fact is that S^h is a cone, and consequently every maximal S^h-free set is a convex cone [7, Corollary 3][3]; we show below why this is not a good setting for monoidal strengthening. In fact, the results in this section apply to a generic *closed cone* S and are stated with respect to such set.

As mentioned before, for a given S-free set C, we are interested in finding a monoid M such that $C - M$ is S-free. With the following results, we show that in this setting M can be assumed to be a convex cone. We remind the reader that cone$(\cdot)$ is the cone generated by a set, which may not be convex.

Proposition 3.1 *Let M be a monoid. Then,* $\operatorname{cl}\operatorname{cone}(M)$ *is a convex cone.*

Proof Let $m_1, m_2 \in M$, and consider $\mu_1, \mu_2 \in \mathbb{Q}$ non-negative. We aim at showing that

$$\mu_1 m_1 + \mu_2 m_2 \in \operatorname{cone}(M).$$

[3] This citation deals with a particular set S, but the proof can be easily extended to any conic set S.

Since μ_1, μ_2 are rational, we can take $p \in \mathbb{N} \setminus \{0\}$ such that $p\mu_1, p\mu_2 \in \mathbb{N}$. Then,

$$\mu_1 m_1 + \mu_2 m_2 = \frac{1}{p}(\underbrace{p\mu_1 m_1 + p\mu_2 m_2}_{w})$$

Since M is a monoid, we have that $w \in M$. Thus, $\mu_1 m_1 + \mu_2 m_2 \in \operatorname{cone}(M)$. Using density of the rationals we conclude that $\operatorname{cl}\operatorname{cone}(M)$ is a convex cone.

Note that a convex cone is, in particular, a monoid. The next results shows that one can consider $\operatorname{cl}\operatorname{cone}(M)$ instead of M in the monoidal strengthening setting.

Lemma 3.1 *Let S be a closed cone and let C be a full-dimensional convex S-free cone. If M is a monoid such that $C - M$ is S-free, then* $\operatorname{cl}\operatorname{cone}(M)$ *is a monoid such that $C - \operatorname{cl}\operatorname{cone}(M)$ is S-free.*

Proof The fact that $\operatorname{cl}\operatorname{cone}(M)$ is a monoid follows trivially from the previous discussion. By contradiction, suppose $C - \operatorname{cl}\operatorname{cone}(M)$ is not S-free. This implies that there exists $s \in S$ such that $s \in \operatorname{int}(C - \operatorname{cl}\operatorname{cone}(M))$. Then, there exists $c \in \operatorname{int}(C)$ and $\bar{m} \in \operatorname{cl}\operatorname{cone}(M)$ such that $s = c - \bar{m}$. Furthermore, we can take $m \in \operatorname{cone}(M)$ sufficiently close to $\bar{m}$ such that

$$c' := c - \bar{m} + m \in \operatorname{int}(C).$$

Thus, we have $s = c' - m$. Note that it must hold that $m \neq 0$ as C is S-free. Since $m \in \operatorname{cone}(M)$, we must have $m = \lambda m'$ for some $\lambda > 0$ and $m' \in M$. This implies

$$\frac{s}{\lambda} = \frac{c'}{\lambda} - m'.$$

Both S and C are cones, thus $s' := s/\lambda \in S$ and $c' := c/\lambda \in \operatorname{int}(C)$. This contradict the S-freeness of $C - M$.

The fact that we can restrict to monoids that are cones has the following consequence when C is a maximal S-free set.

Theorem 3.1 *Let $S, C \subseteq \mathbb{R}^n$ where S is a closed cone and C is a convex maximal S-free set. Let $M \subseteq \mathbb{R}^n$ be monoid such that $C - M$ is S-free, then $C - M = C$. In particular, this implies that the cut obtained from monoidal strengthening would be the same as the standard intersection cut obtained through C.*

Proof Since M is a monoid, $0 \in M$ and thus $C \subseteq C - M$. By Lemma 3.1 $\operatorname{cl}\operatorname{cone}(M)$ is a monoid such that $C - \operatorname{cl}\operatorname{cone}(M)$ is S-free. Note that $C - \operatorname{cl}\operatorname{cone}(M)$ is convex, and thus the maximality of C implies that $C - \operatorname{cl}\operatorname{cone}(M) \subseteq C$. Since $C - M \subseteq C - \operatorname{cl}\operatorname{cone}(M)$, we conclude $C - M = C$.

The last theorem says that there is not much to be gained from the monoidal strengthening framework when S is a cone. This negative property, nonetheless, has a positive effect in detecting "non-maximality" of an S-free set, as stated in the next result.

Proposition 3.2 *Let S be a closed cone and let C be a full dimensional closed convex S-free cone. If there exists $r \notin -C$ such that C is $S+\operatorname{cone}(r)$-free, then C is not a maximal S-free set. Furthermore, $C+\operatorname{cone}(-r)$ is S-free and strictly contains C.*

Proof Suppose $C + \operatorname{cone}(-r)$ is not S-free and let $x \in \operatorname{int}(C + \operatorname{cone}(-r)) \cap S$. Then, $x = y - \lambda r$ for some $y \in \operatorname{int}(C)$ and $\lambda > 0$ ([19, Corollary 6.6.2]). This implies that $y = x + \lambda r \in S + \operatorname{cone}(r)$, which is a contradiction with C being $S + \operatorname{cone}(r)$-free.

Since $r \notin -C$, we conclude that $C \subsetneq C + \operatorname{cone}(-r)$.

The next example illustrates an application of the last proposition, in a connection with the work of [7].

Example 3.1 Consider the set $S = \{(a, b, c, d) \in \mathbb{R}^4 : ad = bc,\ a \geq 0\}$. Although this set does not fall into either the forms S^h or S^g which are our main objects of interest, it is still a closed conic set to which the results of this section apply. The set S is studied in [7], and appears when using lifted variables $X_{i,j}$ representing bilinear terms $x_i x_j$. In this case, the lifted variables must obviously satisfy

$$X_{i,j} X_{k,l} = X_{i,l} X_{k,j}.$$

Furthermore, if $i = j$, it must additionally hold that $X_{i,j} \geq 0$.

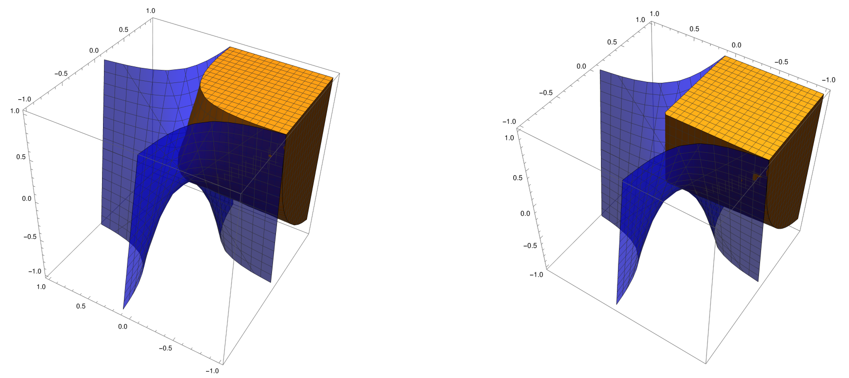
Let $C_\theta = \{(a, b, c, d) \in \mathbb{R}^4 : \cos(\theta)(a+d)+\sin(\theta)(b-c) \geq \sqrt{(a - d)^2 + (b + c)^2}\}$. In [7, Theorem 7], the authors show that C_θ is maximal S-free for specific values of θ; these values always satisfy $\cos(\theta) = 0$ or $\sin(\theta) = 0$.

Here, we prove that if θ is such that $\cos(\theta) \neq 0$ and $\sin(\theta) \neq 0$, then C_θ is not maximal S-free. More specifically, we show that $C'_\theta = C_\theta + \operatorname{cone}(e_4)$ is S-free and strictly contains C_θ, where $e_4 = (0, 0, 0, 1)$.

In virtue of Proposition 3.2, it suffices to show that $-e_4 \notin -C_\theta$ and that C_θ is $S + \operatorname{cone}(-e_4)$-free. The fact that $-e_4 \notin -C_\theta$ can be easily verified as our assumptions imply $\cos(\theta) < 1$, thus we focus on the second property.

Consider $(a, b, c, d) \in \operatorname{int}(C_\theta)$, thus $ad > bc$. If $a < 0$ then $(a, b, c, d) + \lambda e_4 \notin S$ $\forall\lambda$, since the first component does not depend on λ. If $a \geq 0$ then $(a, b, c, d) + \lambda e_4 \notin S$ $\forall\lambda \geq 0$ since

$$a(d + \lambda) \geq ad > bc.$$



(a) Slices of S (blue) and C_θ (orange)

(b) Slices of S (blue) and C'_θ (orange)

Fig. 3.1: Three-dimensional slices of S, C_θ and C'_θ in Example 3.1 given by $a = 1/10$.

Thus C_θ is $S + \text{cone}(-e_4)$-free. Since $-e_4 \notin -C_\theta$, Proposition 3.2 implies that C'_θ is S-free and strictly contains C_θ. In Figure 3.1 we show a 3-dimensional slice of the 4-dimensional sets S and C_θ, for $\theta = \pi/4$, showing how the S-free was enlarged. We remark that one can actually show that C'_θ is *maximal* S-free, but in the interest of space we leave this proof to the reader, which is based on the maximality criteria of [18]. Note that maximality is not seen in Figure 3.1, since maximality is not preserved when taking slices.

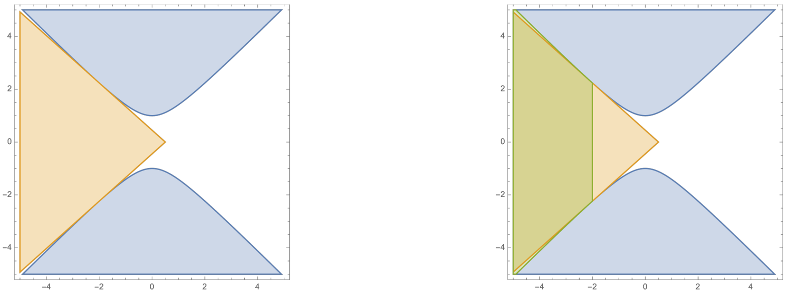
4 The non-homogeneous case: S^g

In this case, the monoidal strengthening framework does produce improvements. The intuition for our construction is as follows. Consider the maximal S-free set C represented in Figure 4.1a. The set is maximal because it has two *exposing points* [18], that is, the points of the facets of C that are tangent to S. We see that a way of translate C such that the translation is S-free is by moving the apex of C to a point not in S and to the left of the exposing points (see Figure 4.1b). This is the basic idea behind our monoid construction, and below we show how to formalize it.

4.1 A technical consideration for S^g

Before motivating the construction the monoid, we need to provide some details on the construction of maximal S^g-free presented in [18]. This construction starts from the maximal S^h-free set

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



(a) S (blue) with maximal S-free set C (orange). In this case the two inequalities of C intersect S.

(b) Set of points not in S and "to the left of the exposing points" (green). Note that the green region is not contained on the orange region: see the top left and bottom left.

Fig. 4.1: Constuction of the monoid for a maximal S-free set.

where λ is a vector in the unit sphere, and then modifies it to account for the hyperplane $H = \{(x, y) \in \mathbb{R}^{n+m} : a^\mathsf{T} x + d^\mathsf{T} y = -1\}$ (recall that $S^g = S^h \cap H$). Note that C_λ can be equivalently described as $C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} : \beta^\mathsf{T} y \leq \lambda^\mathsf{T} x \quad \forall \beta \in D_1\}$ where D_1 is the unit sphere of appropriate dimension. The proof that C_λ is maximal S^h-free boils down to noting that for each $\hat{\beta}$, the vector $(\lambda, \hat{\beta}) \in S^h \cap C_\lambda$ is tight for the inequality $\hat{\beta}^\mathsf{T} y \leq \lambda^\mathsf{T} x$ and for no other. This means that each inequality indexed by β has an exposing point in S^h.

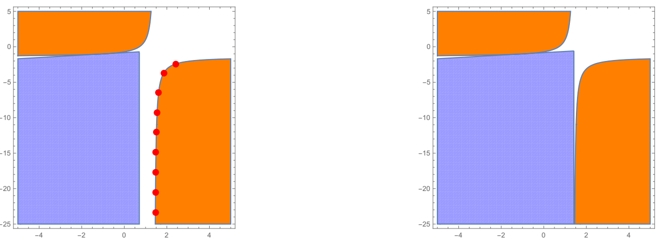
Moving to S^g, the set $C_\lambda \cap H$ is clearly S^g-free[4], but it is not necessarily maximal. The maximal S^g-free constructed in [18] first identifies the inequalities of C_λ for which an exposing point can be found in H and keeps them; these exposing points are

$$-\frac{1}{a^\mathsf{T}\lambda + d^\mathsf{T}\beta}(\lambda, \beta)$$

and they expose the inequalities given by β such that $\|\beta\| = 1$ and $a^\mathsf{T}\lambda + d^\mathsf{T}\beta < 0$. The inequalities in C_λ that correspond to β such that $a^\mathsf{T}\lambda + d^\mathsf{T}\beta \geq 0$ do not have such an exposing point and are relaxed adequately. Maximality of the resulting set is shown using the exposing points above and, for each relaxed inequality, a diverging sequence in S^g that approaches the inequality indefinitely (an exposing sequence). This is due to the fact that these relaxed inequalities may have never intersect S^g; we illustrate this procedure in Figure 4.2.

In our current monoid construction, we require that all exposing points of the maximal S^g-free set are bounded. This requirement translates to $a^\mathsf{T}\lambda + d^\mathsf{T}\beta < 0$ for all β with $\|\beta\| = 1$. This, in turn, reduces to $\|d\| < -a^\mathsf{T}\lambda$. Note that this condition implies that C_λ is maximal S^g-free with respect to H [18]. Additionally, this implies that we can assume $\|a\| = \max\{\|a\|, \|d\|\} = 1$.

[4] Note that S^g is contained on a halfspace, so S^g-freeness is with respect to the induced topology in H.



(a) S^g (orange), C_λ (blue) and an exposing sequence (red points) used to relax one of the inequalities of C_λ.

(b) Resulting maximal S^g-free set after one of the facets of C_λ is relaxed. The vertical facet does not intersect S^g.

Fig. 4.2: Plots illustrating how C_λ is modified to obtain a maximal S^g-free set. Figures obtained from [18].

4.2 Monoid construction

Using the considerations of the previous section, we can formalize the notion of “left of the exposing points”: we construct a halfspace that contains the exposing points and the directions of lineality of $C_\lambda \cap H$ as we want the apex to be on one side of the halfspace. This hyperplane is given by $\{(x, y) \in \mathbb{R}^{n+m} : (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x \geq 0\}$.

When translating C_λ by a vector m we can modify m by a vector in the lineality space of C_λ without changing the translation. Thus, we restrict to vectors m that live in a subspace that contains the exposing points and is orthogonal to the lineality space of $C_\lambda \cap H$. This subspace is given by $\langle\{\lambda, a\}\rangle \times \mathbb{R}^m$. Thus, we have the following set representing the points “left of the exposing points”:

$$L = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m \,:\, a^\mathsf{T} x + d^\mathsf{T} y = -1, \|x\| \geq \|y\|, (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x \geq 0\}. \tag{4.2}$$

To obtain the translation we find the apex of $C_\lambda \cap H$ in the space $\langle\{\lambda, a\}\rangle \times \mathbb{R}^m$. This point is given by

$$\nu = (x_0, 0) := \left(\frac{-1}{1 - (\lambda^\mathsf{T} a)^2} a + \frac{\lambda^\mathsf{T} a}{1 - (\lambda^\mathsf{T} a)^2} \lambda, 0\right). \tag{4.3}$$

Thus, $L - \nu$ is a candidate to represent the translations of C that would result in an S^g-free set. Recall that the translations we consider for C are given by “minus the monoid” and that a monoid must contain the origin, therefore our candidate for a monoid is M where

$$\begin{aligned} M = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m \ :& \, a^\mathsf{T} x + d^\mathsf{T} y = 0, \|x - x_0\| \geq \|y\|, \\ & (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x \leq -1\} \cup \{(0, 0)\}. \end{aligned} \tag{4.4}$$

We now state the necessary results for the correctness of our construction.

Theorem 4.1 *Let M be defined as in* (4.4) *with* $\|d\| < -\lambda^\mathsf{T} a$ *and* $\|a\| = \|\lambda\| = 1$. *The set M is a monoid.*

Theorem 4.2 *Let S^g and C_λ be defined as in* (1.4) *and* (4.1) *respectively, and* $H = \{(x, y) \in \mathbb{R}^{n+m} : a^\mathsf{T} x + d^\mathsf{T} y = -1\}$. *Let M be defined as in* (4.4) *with* $\|d\| < -\lambda^\mathsf{T} a$ *and* $\|a\| = \|\lambda\| = 1$. *The set* $C_\lambda \cap H - M$ *is S^g-free.*

The proofs of these results are highly technical, thus, for the sake of readability we have relegated them to Appendix A.

For the rest of the paper, we assume $\|d\| < -\lambda^\mathsf{T} a < 1$ and $\|a\| = \|\lambda\| = 1$, i.e., we are in the case where monoidal strengthening is actually possible. Furthermore, we denote by S^g and C_λ the sets defined by (1.4) and (4.1), respectively, and let $H = \{(x, y) \in \mathbb{R}^{n+m} : a^\mathsf{T} x + d^\mathsf{T} y = -1\}$.

5 Solving the monoidal strengthening problem

In order to strengthen the cut using Theorem 2.1 and the monoid constructed in the previous section, we need to solve $\psi(r) = \inf_{m \in M} \phi(r + m)$, where ϕ is a sublinear function such that $C_\lambda \cap H = \{z : \phi(z - f) \leq 1\}$. From now on, $\lambda = \frac{f_x}{\|f_x\|}$, where f is the point we want to separate, i.e., $f \notin S^g$. Furthermore, we restrict $C_\lambda \cap H$ to $\langle \{\lambda, a\} \rangle \times \mathbb{R}^m$ because any representation of $C_\lambda \cap H$ is invariant in the directions of the lineality space of $C_\lambda \cap H$, namely, $\langle \{\lambda, a\} \rangle^\perp \times \{0\}$. Thus, we define $C = C_\lambda \cap H \cap \langle \{\lambda, a\} \rangle \times \mathbb{R}^m$. Likewise, we restrict all rays to $\langle \{\lambda, a\} \rangle \times \mathbb{R}^m$.

We work with the so-called minimal representation of $C - f$.

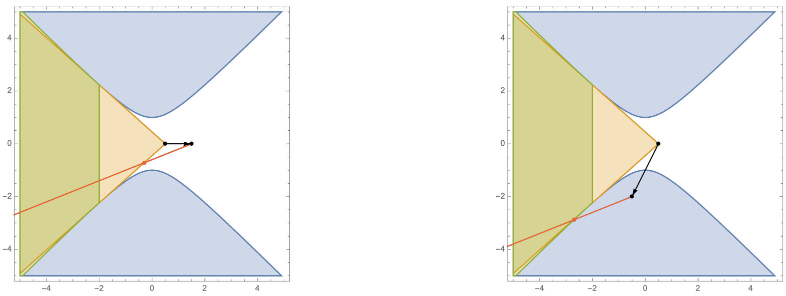
Proposition 5.1 *The minimal representation of $C - f$ is*

$$\phi(x, y) = \begin{cases} \sup_{\|\beta\|=1} \frac{\beta^\mathsf{T} y - \lambda^\mathsf{T} x}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} & \text{if } a^\mathsf{T} x + d^\mathsf{T} y = 0 \text{ and } x \in \langle \{\lambda, a\} \rangle \\ +\infty & \text{otherwise.} \end{cases} \tag{5.1}$$

Proof See Appendix A.7.

The monoidal problem is equivalent to $\psi(r) = \inf\{\tau : \phi(r + m) \leq \tau, m \in M\}$. In order to understand this problem better, we need to understand the set $\{z : \phi(z) \leq \tau\}$.

Proposition 5.2 *Let ϕ be the minimal representation of $C - f$ given in* (5.1). *Then* $\{z : \phi(z) \leq \tau\} = C - \nu - \tau(f - \nu)$, *where ν is defined in* (4.3) *(the apex of $C_\lambda \cap H$ in the space $\langle \{\lambda, a\} \rangle \times \mathbb{R}^m$).*



(a) Line $\{l(\tau) : \tau > 0\}$ (red) intersects $L \cup C$ (green and orange) at $l(\tau_2)$ (red point)

(b) Line $\{l(\tau) : \tau > 0\}$ (red) intersects $L \cup C$ (green and orange) at $l(\tau_1)$ (red point)

Fig. 5.1: Solving the monoidal strengthening problem.

Proof See Appendix A.3.

Using the above proposition, the monoidal problem is equivalent to $\psi(r) = \inf\{\tau : r + m \in C - \nu - \tau(f - \nu), m \in M\}$. The constraints $r + m \in C - \nu - \tau(f - \nu)$ and $m \in M$ are equivalent to $r + \nu + \tau(f - \nu) \in C - M$, thus we have that

$$\psi(r) = \inf\{\tau : r + \nu + \tau(f - \nu) \in C - M\}. \tag{5.2}$$

In other words, solving the monoidal strengthening problem reformulates to finding the first intersection point between the line $l(\tau) = r + \nu + \tau(f - \nu)$, and the set $C - M$.

As shown in Proposition A.1, we have $C - M = L \cup C$. Thus,

$$\psi(r) = \inf\{\tau : l(\tau) \in L \cup C\} = \min\{\tau_1, \tau_2\}$$

where $\tau_1 = \inf\{\tau : l(\tau) \in L\}$ and $\tau_2 = \inf\{\tau : l(\tau) \in C\}$. Note that τ_2 corresponds to the normal intersection cut coefficient $\phi(r)$. The following proposition shows how to evaluate $\psi(r)$.

Proposition 5.3 *Let $\bar{\tau}$ be the largest root of the univariate quadratic equation $\|l_x(\tau)\|^2 = \|l_y(\tau)\|^2$. If the root exists and $l(\bar{\tau}) \in L$, then $\psi(r) = \bar{\tau}$. Otherwise, $\psi(r) = \phi(r)$.*

Proof See Appendix A.5

Figure 5.1a is an example where $\psi(r) = \phi(r)$. In this particular example, this is due to the fact that $\|l_x(\tau)\|^2 = \|l_y(\tau)\|^2$ does not have roots. Figure 5.1b is an example where $\psi(r)$ is given by the largest root of $\|l_x(\tau)\|^2 = \|l_y(\tau)\|^2$.

In Section 7, we show how to explicitly compute $\psi(r)$ when dealing with a general quadratic constraint.

Remark 5.1 If we choose ϕ as the gauge centered at f, i.e., $\phi(r) = \inf\{\tau : f + \frac{r}{\tau} \in C, \tau > 0\}$, instead of the minimal representation, it can be shown that $\psi(r)$ is equivalent to (5.2) but with the additional constraint $\tau > 0$. This results in a possibly weaker intersection cut and weaker strengthening.

6 Unique lifiting

In this section, we show that we have unique lifting. To show that we have unique lifting using Lemma 2.1, we have to show that π_1 is subadditive. As mentioned at the end of Section 2, we do so by proving that π_1 is equal to the monoidal strengthening problem, i.e., $\pi_1 = \psi$.

Recall that, for a ray r, the lifting function is given by

$$\pi(r) = \sup\left\{\frac{1-\phi(s)}{\sigma} \,:\, f + s + \sigma r \in S^g, \sigma \in \mathbb{Z}_{\geq 1}\right\},$$

where ϕ is the minimal representation of $C - f$, see Proposition 5.1. Evaluating $\pi(r)$ is equivalent to solving

$$\pi(r) = \sup\left\{\frac{\tau}{\sigma} \,:\, f + s + \sigma r \in S^g, \phi(s) \leq 1 - \tau, \sigma \in \mathbb{Z}_{\geq 1}\right\} \tag{6.1}$$

By Proposition 5.2, $\{z : \phi(z) \leq 1 - \tau\} = C - f + \tau(f - \nu)$. Using this equivalence, the constraints $f + s + \sigma r \in S^g$ and $\phi(s) \leq 1 - \tau$ reformulate to $\sigma r + \nu + \tau(f - \nu) \in S^g - (C - \nu)$. Since $C - \nu = \mathrm{rec}(C)$, (6.1) can be simplified to

$$\pi(r) = \sup\left\{\frac{\tau}{\sigma} \,:\, \sigma r + \nu + \tau(f - \nu) \in S^g - \mathrm{rec}(C), \sigma \in \mathbb{Z}_{\geq 1}\right\}.$$

Then, looking at the problem with $\sigma = 1$ fixed, yields

$$\pi_1(r) = \sup\left\{\tau \,:\, r + \nu + \tau(f - \nu) \in S^g - \mathrm{rec}(C)\right\}.$$

Evaluating π_1 reduces to finding the largest intersection point between the line $r + \nu + \tau(f - \nu)$ and the set $S^g - \mathrm{rec}(C)$.

After all this manipulations, we can see that the lifting problem $\pi_1(r)$ and the monoidal problem $\psi(r)$ have a similar structure. Indeed, we can show that they are the same problem.

Theorem 6.1 *The functions ψ and π_1 are equal.*

Proof We will need that

$$H \setminus (S^g - \mathrm{rec}(C)) = \mathrm{ri}(C - M) \tag{6.2}$$

which we proceed to show by contradiction. Assume that the intersection of $S^g - \text{rec}(C)$ and $\text{ri}(C - M)$ is non-empty. Then, there exists $s \in S^g$ and $r \in \text{rec}(C)$ such that $s - r \in \text{ri}(C - M)$. This implies that

$$s \in \text{ri}(C - M) + r = \text{ri}(C + r - M) = \text{ri}(C - M),$$

which contradicts the S^g-freeness of $C - M$ established in Theorem 4.2. This proves the relation (6.2).

Let $r \in \mathbb{R}^{n+m}$. Proposition 5.3 shows that $\psi(r)$ is finite, so let $\tau^* = \psi(r) = \inf\{\tau : l(\tau) \in C - M\}$, which implies that $l(\tau^*) \notin \text{ri}(C - M)$. To show that $\tau^* = \sup\{\tau : l(\tau) \in S^g - \text{rec}(C)\}$, we need to prove that $l(\tau^*) \in S^g - \text{rec}(C)$ and $l(\tau^* + \epsilon) \notin S^g - \text{rec}(C)$ for all $\epsilon > 0$. By (6.2), we have $l(\tau^*) \in S^g - \text{rec}(C)$.

It remains to prove that $l(\tau^* + \epsilon) \notin S^g - \text{rec}(C)$ which is equivalent to $l(\tau^* + \epsilon) \in \text{ri}(C - M)$. Since $l(\tau^* + \epsilon) - l(\tau^*) = \epsilon(f - \nu)$, $f - \nu \in \text{ri}(\text{rec}(C))$, and $\epsilon > 0$, we have that $l(\tau^* + \epsilon) \in l(\tau^*) + \text{ri}(\text{rec}(C))$. Therefore, there exists $\bar{\epsilon} > 0$ such that $B_{\bar{\epsilon}}(0) + l(\tau^* + \epsilon) \subseteq l(\tau^*) + \text{ri}(\text{rec}(C))$. Furthermore, $l(\tau^*) \in C - M$, thus $B_{\bar{\epsilon}}(0) + l(\tau^* + \epsilon) \subseteq C - M + \text{ri}(\text{rec}(C)) = C - M$. This shows that $l(\tau^* + \epsilon) \in \text{ri}(C - M)$.

Finally, we conclude that we have unique lifting by Lemma 2.1.

7 Monoidal strengthening for a general quadratic constraint

In this section, we present how to explicitly apply monoidal strengthening for enforcing an arbitrary quadratic constraint. From now on, instead of looking at S to be of the form S^g, we consider S to be defined by a general quadratic constraint, i.e., $S = \{s \in \mathbb{R}^p : s^\mathsf{T} Q s + b^\mathsf{T} s + c \leq 0\}$ with $Q \in \mathbb{R}^{p \times p}$, $b \in \mathbb{R}^p$ and $c \in \mathbb{R}$. In [8] the authors show that S can be rewritten as

$$S = \{s \in \mathbb{R}^p : \|x(s)\|^2 - \|y(s)\|^2 + \bar{b}^\mathsf{T} z(s) + \kappa \leq 0\}$$

and for the non-homogeneous case, i.e., $\bar{b} \neq 0$ or $\kappa \neq 0$, they showed how to write it in the form

$$S = \{s \in \mathbb{R}^p : \|\hat{x}(s)\|^2 - \|\hat{y}(s)\|^2 \leq 0, a^\mathsf{T} \hat{x}(s) + d^\mathsf{T} \hat{y}(s) = -1\}.$$

The values of a and d depend on $\bar{b}$ and κ. Since we need $\|d\| < -\lambda^\mathsf{T} a < 1$ to apply monoidal strengthening, the transformations from [8] show that this can only be achieved if $\bar{b} = 0$[5] and $\kappa > 0$.

[5] This means that there are no purely linear variables which discards the application of the monoidal strengthening developed in this paper for quadratic constraints coming from the epigraph reformulation of the objective function.

More specifically, in our implementation, we transform S using the eigenvalue decomposition $Q = V\Theta V^\mathsf{T}$. Let θ_i, $i \in [p]$, be the eigenvalues of Q, and let $I_+ = \{i : \theta_i > 0\}$, $I_- = \{i : \theta_i < 0\}$ and $I_0 = \{i : \theta_i = 0\}$ be the index sets of the positive, negative and zero eigenvalues, respectively. We assume that $I_+ = \{1, \dots, |I_+|\}$, $I_- = \{|I_+|+1, \dots, |I_+|+|I_-|\}$, and $I_0 = \{|I_+|+|I_-|, \dots, p\}$. Furthermore, denote by v_i the i-th eigenvector of Q, that is, the i-th column of V. Then, S can be rewritten as

$$S = \{s \in \mathbb{R}^p \,:\, \|(\hat{x}(s), \zeta)\|^2 - \|\hat{y}(s)\|^2 \leq 0, a^\mathsf{T}(x(s), \zeta) + d^\mathsf{T} y(s) = -1\},$$

where $a = -e_{-1}$, $d = 0$, and

$$\begin{aligned}
\hat{x}_i(s) &= \sqrt{\frac{\theta_i}{\kappa}} v_i^\mathsf{T}(s + \frac{b}{2\theta_i}), && \forall i \in I_+,\\
\hat{y}_i(s) &= \sqrt{\frac{-\theta_i}{\kappa}} v_i^\mathsf{T}(s + \frac{b}{2\theta_i}), && \forall i \in I_-,\\
\hat{z}_i(s) &= \frac{1}{\sqrt{\kappa}} v_i^\mathsf{T} s, && \forall i \in I_0,\\
\kappa &= c - \frac{1}{4} \sum_{i \in I_+ \cup I_-} \frac{(v_i^\mathsf{T} b)^2}{\theta_i}.
\end{aligned}$$

Note that z_i does not appear in the definition of S since $\bar{b} = 0$. Nonetheless, we state it here so that the mapping $T \colon s \mapsto (\hat{x}(s), \hat{y}(s), \hat{z}(s))$ is invertible.

Furthermore, the hyperplane $a^\mathsf{T}(x(s), \zeta) + d^\mathsf{T} y(s) = -1$ is equivalent to $\zeta = 1$. The maximal quadratic-free set is

$$C = \{s \in \mathbb{R}^p \,:\, \|\hat{y}(s)\| \leq \lambda^\mathsf{T}(\hat{x}(s), 1)\}.$$

Finally, denote by $\bar{s}$ the point we want to separate, i.e., $\bar{s} \notin S$. Then, $\lambda = \frac{(\hat{x}(\bar{s}), \zeta)}{\|(\hat{x}(\bar{s}), 1)\|} = \frac{(\hat{x}(\bar{s}), 1)}{\|(\hat{x}(\bar{s}), 1)\|}$ and the necessary condition $\lambda^\mathsf{T} a < 1$ translates into $\|\hat{x}(\bar{s})\| > 0$.

To solve the monoidal strengthening problem for a ray r, we need to find the roots of the quadratic function $\|(\hat{x}(s), 1)\|^2 - \|\hat{y}(s)\|^2$ along the line $r + \nu + \tau(\bar{s} - \nu)$ where ν denotes the apex of C. For this, we first need to find ν:

Proposition 7.1 *The apex ν of C is*

$$\nu = -\frac{\kappa}{\sum_{j \in I_+} \theta_j (v_j^\mathsf{T}(\bar{s} + \frac{b}{2\theta_j}))^2} \sum_{j \in I_+} v_{ij}(v_j^\mathsf{T}(\bar{s} + \frac{b}{2\theta_j})) - \sum_{j \in I_+ \cup I_-} v_{ij} \frac{v_j^\mathsf{T} b}{2\theta_j}.$$

Therefore, to perform monoidal strengthening, we need to find the largest root of the following quadratic function:

$$
\begin{aligned}
&\|(\hat{x}(r+\nu+\tau(\bar{s}-\nu)),1)\|^2 - \|\hat{y}(r+\nu+\tau(\bar{s}-\nu))\|^2 \\
&= \frac{1}{\kappa}\sum_{i\in I_+}\theta_i\left(v_i^{\mathsf{T}}(r+\nu+\tau(\bar{s}-\nu)+\frac{b}{2\theta_i})\right)^2 + \frac{1}{\kappa}\sum_{i\in I_-}\theta_i\left(v_i^{\mathsf{T}}(r+\nu+\tau(\bar{s}-\nu)+\frac{b}{2\theta_i})\right)^2 + 1 \\
&= A\tau^2 + B\tau + D
\end{aligned}
$$

with coefficients

$$
\begin{aligned}
A &= \sum_{i\in I_+\cup I_-}\theta_i\left(v_i^{\mathsf{T}}(\bar{s}-\nu)\right)^2, \\
B &= 2\sum_{i\in I_+\cup I_-}\theta_i\left(v_i^{\mathsf{T}}(\bar{s}-\nu)\right)\left(v_i^{\mathsf{T}}(r+\nu+\frac{b}{2\theta_i})\right), \\
D &= \sum_{i\in I_+\cup I_-}\theta_i\left(v_i^{\mathsf{T}}(r+\nu+\frac{b}{2\theta_i})\right)^2 + \kappa.
\end{aligned}
$$

We can use Proposition 5.3 to find the cut coefficient $\psi(r)$ by solving $A\tau^2 + B\tau + D = 0$ and computing the standard intersection cut coefficient $\phi(r)$. The latter can also be done efficiently as shown in [8].

8 Computational Results

In this section, we show results of computational experiments testing the efficacy of the monoidal strengthening procedure we propose. We embedded the computation of the monoidal strengthening cut coefficients in SCIP 8.0 [6] as a subroutine of the already implemented intersection cut generator. As the underlying LP solver, we used CPLEX 12.10.0.0. For testing, we used a Linux cluster of Intel Xeon CPU E5-2680 0 2.70GHz with 20MB cache and 64GB main memory. The time limit in all experiments was set to two hours. The test set we consider consists of the publicly available instances of the MINLPLib [17] and QPLib [14]. We selected all non-convex instances with (mixed)-integer constriants and at least one quadratic constraint of the correct case, leaving us with 95 instances. Furthermore, we filtered out all instances that are either infeasible, where no dual bound was found or where monoidal strengthening could not been applied. This leaves us with 63 instances. All experiments are run with three different permutations for each instance. We treat every instance-permutation pair as an individual instance, since permuting the constraints and variables of a problem formulation may considerably change the solving process.

We consider two different settings that are both based on SCIP's default settings: ICUTS additionally generates the original intersection cuts, whereas

Table 8.1: Summary of results for branch-and-bound experiments. Rows labeled $[t, 7200]$ consider instances where one of the settings took at least t seconds. Columns labeled *relative* show the relative improvement of MONOIDAL compared to ICUTS.

		ICUTS			MONOIDAL			relative	
subset	instances	solved	time	nodes	solved	time	nodes	time	nodes
all	189	113	221.87	5282	115	214.63	5321	0.97	0.97
$[0, 7200]$	115	113	22.81	936	115	21.56	883	0.95	0.94
$[1, 7200]$	83	81	67.62	2377	83	62.40	2184	0.92	0.92
$[10, 7200]$	81	79	72.54	2574	81	66.56	2341	0.92	0.91
$[100, 7200]$	23	21	724.66	186545	23	565.24	144747	0.78	0.78
$[1000, 7200]$	10	8	2475.04	631764	10	1252.96	307639	0.51	0.49

MONOIDAL uses the strengthened cutting planes if possible. Furthermore, we restrict ICUTS and MONOIDAL to add at most 20 intersection cuts per quadratic constraint. We found this to be the best performing setting compared to default SCIP.

Summarized results can be found in Table 8.1. MONOIDAL consistently outperforms ICUTS with respect to solving time as well as number of nodes needed. On the whole test set, the strengthened intersection cuts reduce both metrics by around 3% while solving two more instances. This improvement increases when looking at harder instances: On the hardest test set $[1000, 7200]$ containing only instances for which at least one setting needs 1000 seconds or more, this MONOIDAL 49% less time and 51% less nodes.

A Missing proofs

A.1 Proof of Theorem 4.1: the set M is a monoid

As the title suggests, the main goal of this section is to prove that M, with M defined in (4.4) is indeed a monoid.

Proof (Theorem 4.1) Due to the way this proof was originally developed, it will be slightly more convenient to show that $-M$ is a monoid, which is an equivalent statement.

To show that $-M$ is a monoid, we take two vectors $(x_i, y_i) \in -M$ $i = 1, 2$ and show that their sum is in $-M$ (if one of them is the origin, the result follows trivially). Let us recall the definition of $-M$ without the origin:

$$-M_{\neq 0} = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m \ : \ a^\mathsf{T} x + d^\mathsf{T} y = 0, \|x + x_0\| \geq \|y\|, (a - \lambda^\mathsf{T} a \lambda)^\mathsf{T} x \geq 1\}. \quad \text{(A.1)}$$

where $\|d\| < -\lambda^\mathsf{T} a$ and $\|\lambda\| = \|a\| = 1$. The linear constraints in (A.1) are satisfied trivially by the sum, hence in the following, we will focus on showing that $\|x_1 + x_2 + x_0\| \geq \|y_1 + y_2\|$. We prove this by showing that the optimization problem

$$\begin{aligned} \min_{x_i, y_i} \quad & \|x_1 + x_2 + x_0\|^2 - \|y_1 + y_2\|^2 \\ \text{s.t.} \quad & (x_i, y_i) \in -M_{\neq 0} \qquad i = 1, 2 \end{aligned}$$

is non-negative. First, we expand the objective function:

$$\begin{aligned}
&\|x_1 + x_2 + x_0\|^2 - \|y_1 + y_2\|^2 \\
=&\|x_1 + x_0 + x_2 + x_0 - x_0\|^2 - \|y_1 + y_2\|^2 \\
=&\|x_1 + x_0\|^2 + \|x_2 + x_0\|^2 + \|x_0\|^2 + 2(x_1 + x_0)^\mathsf{T}(x_2 + x_0) \\
&- 2x_0^\mathsf{T}(x_1 + x_0) - 2x_0^\mathsf{T}(x_2 + x_0) - \|y_1\|^2 - \|y_2\|^2 - 2y_1^\mathsf{T} y_2 \\
\geq& 2x_1^\mathsf{T} x_2 - 2y_1^\mathsf{T} y_2 - \|x_0\|^2
\end{aligned}$$

where the last inequality follows from the constraints $\|x_i + x_0\| \geq \|y_i\|$ and $(a - \lambda^\mathsf{T} a \lambda)^\mathsf{T} x_i \geq 1 \Leftrightarrow -x_0^\mathsf{T} x_i \geq \|x_0\|^2$. Hence, showing that the above optimization problem is non-negative is equivalent to proving that

$$\begin{aligned}
\min_{x_i, y_i} \quad & x_1^\mathsf{T} x_2 - y_1^\mathsf{T} y_2 \\
\text{s.t.} \quad & (x_i, y_i) \in -M_{\neq 0} \qquad i = 1, 2
\end{aligned} \tag{P}$$

is bounded by $\frac{1}{2}\|x_0\|^2$.

We can always decompose $y_i = \omega_i d + \rho_i$ where ρ_i is orthogonal to d. Furthermore, since $x_i \in \langle\{a, \lambda\}\rangle$, x_i can be represented in terms of a and λ, i.e., $x_i = \theta_i a + \eta_i \lambda$. In the following, we will use the expansions of x_i and y_i to reformulate (P). Using that $\|a\| = 1$, the hyperplane in $-M_{\neq 0}$ becomes

$$0 = a^\mathsf{T} x_i + d^\mathsf{T} y_i = \theta_i + \eta_i \lambda^\mathsf{T} a + \omega_i \|d\|^2.$$

Furthermore, noting that $\lambda^\mathsf{T} x_0 = 0$ and $a^\mathsf{T} x_0 = -1$, we get

$$-x_0^\mathsf{T} x_i \geq \|x_0\|^2 \Leftrightarrow \theta_i \geq \|x_0\|^2$$

and the nonlinear constraint in (A.1) expands to

$$\begin{aligned}
0 \geq \|x_i + x_0\|^2 - \|y_i\|^2 &= \|x_i\|^2 + 2x_0^\mathsf{T} x_i + \|x_0\|^2 - \omega_i^2 \|d\|^2 - \|\rho_i\|^2 \\
&= \theta_i^2 + \eta_i^2 + 2\theta_i \eta_i \lambda^\mathsf{T} a - 2\theta_i + \|x_0\|^2 - \omega_i^2 \|d\|^2 - \|\rho_i\|^2.
\end{aligned}$$

Finally, replacing x_i and y_i in the objective of (P) yields

$$x_1^\mathsf{T} x_2 - y_1^\mathsf{T} y_2 = \theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \omega_1\omega_2\|d\|^2 - \rho_1^\mathsf{T}\rho_2.$$

To summarize, problem (P) can be reformulated as

$$\begin{aligned}
\min_{\theta_i, \eta_i, \omega_i, \rho_i} \quad & \theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \omega_1\omega_2\|d\|^2 - \rho_1^\mathsf{T}\rho_2 \\
\text{s.t.} \quad & 0 \leq \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - 2\theta_i + \|x_0\|^2 - \omega_i^2\|d\|^2 - \|\rho_i\|^2 \\
& \|x_0\|^2 \leq \theta_i \\
& \|d\|^2\omega_i = -\theta_i - \eta_i\lambda^\mathsf{T} a \\
& \rho_i^\mathsf{T} d = 0
\end{aligned}$$

In what follows, we consider a relaxed version of this problem by removing constraints $\rho_i^\mathsf{T} d = 0$, which leaves us with the problem:

$$\begin{aligned}
\min_{\theta_i, \eta_i, \omega_i, \rho_i} \quad & \theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \omega_1\omega_2\|d\|^2 - \rho_1^\mathsf{T}\rho_2 \\
\text{s.t.} \quad & 0 \leq \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - 2\theta_i + \|x_0\|^2 - \omega_i^2\|d\|^2 - \|\rho_i\|^2 \\
& \|x_0\|^2 \leq \theta_i \\
& \|d\|^2\omega_i = -\theta_i - \eta_i\lambda^\mathsf{T} a
\end{aligned} \tag{P_{exp}}$$

and thus it suffices to show that the value of this optimization problem is $\geq \frac{1}{2}\|x_0\|^2$ to prove that the given set is a monoid. We leave this fact as a claim that we show in Corollary A.2.

Claim Problem (P_{exp}) is lower bounded by $\frac{1}{2}\|x_0\|^2$.

In the subsequent sections, our main goal is to prove this last claim.

A.1.1 The case $\|d\| > 0$

We begin by showing that Problem (P_{exp}) is lower bounded by $\frac{1}{2}\|x_0\|^2$ in the case $\|d\| > 0$: we assume this throughout this section.

Note that we have $-2\theta_i + \|x_0\|^2 \le -\|x_0\|^2 \le 0$ since $\|x_0\|^2 \le \theta_i$ in (P_{exp}). Hence, by dropping the term $-2\theta_i + \|x_0\|^2$ in each of the nonlinear constraints, we relax the latter and thus obtain a relaxation of (P_{exp}), namely problem (P_{rel}):

$$\begin{aligned} \min_{\theta_i,\eta_i,\omega_i,\rho_i} \quad & \theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \omega_1\omega_2\|d\|^2 - \rho_1^\mathsf{T}\rho_2 \\ \text{s.t.} \quad & 0 \le \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - \omega_i^2\|d\|^2 - \|\rho_i\|^2 \\ & \|x_0\|^2 \le \theta_i \\ & \|d\|^2\omega_i = -\theta_i - \eta_i\lambda^\mathsf{T} a. \end{aligned} \tag{P_{rel}}$$

Before showing the desired lower bound, we begin by showing that (P_{exp}) is simply bounded. To do so, we will show that (P_{rel}) is bounded, and for the latter, we first show the following structural result.

Lemma A.1 *The feasible region of (P_{rel}) is convex.*

Proof We begin by using the linear equality constraint involving ω_i and replace $\theta_i = -\eta_i\lambda^\mathsf{T} a - \omega_i\|d\|^2$. This transforms the linear inequality constraints to $\|x_0\|^2 \le -\eta_i\lambda^\mathsf{T} a - \omega_i\|d\|^2$ and the nonlinear constraints to

$$\begin{aligned} 0 &\le (-\eta_i\lambda^\mathsf{T} a - \omega_i\|d\|^2)^2 + \eta_i^2 + 2(-\eta_i\lambda^\mathsf{T} a - \omega_i\|d\|^2)\eta_i\lambda^\mathsf{T} a - \omega_i^2\|d\|^2 - \|\rho_i\|^2 \\ &= \eta_i^2(1 - \lambda^\mathsf{T} a^2) - \omega_i^2\|d\|^2(1 - \|d\|^2) - \|\rho_i\|^2 \end{aligned}$$

Convexity follows if $\eta_i \ge 0$ for all $(\theta_i, \eta_i, \omega_i, \rho_i) \in \mathcal{F}_i$ where

$$\begin{aligned} \mathcal{F}_i = \{(\theta_i, \eta_i, \omega_i, \rho_i) \, : \, & 0 \le \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - \omega_i^2\|d\|^2 - \|\rho_i\|^2, \\ & \|x_0\|^2 \le \theta_i, \|d\|^2\omega_i = -\theta_i - \eta_i\lambda^\mathsf{T} a\}. \end{aligned}$$

To show this, we prove that $\mathcal{F}_i \cap \{(\theta_i, \eta_i, \omega_i, \rho_i) \, : \, \eta_i < 0\} = \emptyset$.

Let $(\theta_i, \eta_i, \omega_i, \rho_i) \in \mathcal{F}_i$ with $\eta_i < 0$. Then, we can write the constraint $0 \le \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - \omega_i^2\|d\|^2 - \|\rho_i\|^2$ as

$$\sqrt{\omega_i^2\|d\|^2(1 - \|d\|^2) + \|\rho_i\|^2} \le -\eta_i\sqrt{(1 - \lambda^\mathsf{T} a^2)}. \tag{A.2}$$

Using the linear constraint, we also have that $0 \le -\eta_i\lambda^\mathsf{T} a - \omega_i\|d\|^2$. Multiplying this inequality by $\frac{\sqrt{1-\lambda^\mathsf{T} a^2}}{-\lambda^\mathsf{T} a} \ge 0$ and adding it to (A.2) yields

$$\sqrt{\omega_i^2\|d\|^2(1 - \|d\|^2) + \|\rho_i\|^2} \le \frac{\|d\|^2\sqrt{1 - \lambda^\mathsf{T} a^2}}{\lambda^\mathsf{T} a}\omega_i.$$

Squaring the above yields

$$\omega_i^2\|d\|^2(1-\|d\|^2)+\|\rho_i\|^2 \le \frac{\|d\|^4(1-\lambda^\mathsf{T} a^2)}{\lambda^\mathsf{T} a^2}\omega_i^2,$$

which is equivalent to

$$\omega_i^2\left(\|d\|^2(1-\|d\|^2)-\frac{\|d\|^4(1-\lambda^\mathsf{T} a^2)}{\lambda^\mathsf{T} a^2}\right)+\|\rho_i\|^2 \le 0. \tag{A.3}$$

Note that the coefficient of ω_i^2 satisfies

$$\begin{aligned}\|d\|^2(1-\|d\|^2)-\frac{\|d\|^4(1-\lambda^\mathsf{T} a^2)}{\lambda^\mathsf{T} a^2} &= \|d\|^2(1-\|d\|^2-\frac{\|d\|^2}{\lambda^\mathsf{T} a^2}+\|d\|^2)\\ &= \|d\|^2(1-\frac{\|d\|^2}{\lambda^\mathsf{T} a^2})\\ &= \frac{\|d\|^2}{\lambda^\mathsf{T} a^2}(\lambda^\mathsf{T} a^2-\|d\|^2)>0\end{aligned}$$

This and (A.3) implies that $\omega_i = 0$ and $\rho_i = 0$. But, then, the constraint $0 \le -\eta_i\lambda^\mathsf{T} a - \omega_i\|d\|^2$ becomes $0 \le -\eta_i\lambda^\mathsf{T} a$, which implies that $\eta_i \ge 0$ which contradicts the hypothesis that $\eta_i < 0$. Thus, it follows that $\eta_i \ge 0$ for all elements in $\mathcal{F}_i$, showing that (P_{rel}) is convex.

Lemma A.2 *The optimization problem (P_{rel}) is bounded.*

Proof Since Lemma A.1 shows that the feasible region of (P_{rel}) is convex, we have that $\eta_i \ge 0$. Thus, the nonlinear constraint $0 \le \theta_i^2+\eta_i^2+2\theta_i\eta_i\lambda^\mathsf{T} a-\omega_i^2\|d\|^2-\|\rho_i\|^2$ can be reformulated to

$$\sqrt{\omega_i^2\|d\|^2(1-\|d\|^2)+\|\rho_i\|^2} \le \eta_i\sqrt{(1-\lambda^\mathsf{T} a^2)}.$$

Now, replace $\theta_i = -\eta_i\lambda^\mathsf{T} a-\omega_i\|d\|^2$ in the objective function. This gives

$$\begin{aligned}&\theta_1\theta_2+\eta_1\eta_2+\theta_1\eta_2\lambda^\mathsf{T} a+\theta_2\eta_1\lambda^\mathsf{T} a-\omega_1\omega_2\|d\|^2-\rho_1^\mathsf{T}\rho_2\\ &=\eta_1\eta_2(\lambda^\mathsf{T} a^2+1-2\lambda^\mathsf{T} a^2)+\omega_1\omega_2(\|d\|^4-\|d\|^2)-\rho_1^\mathsf{T}\rho_2\\ &=\eta_1\eta_2(1-\lambda^\mathsf{T} a^2)-\|d\|^2\omega_1\omega_2(1-\|d\|^2)-\rho_1^\mathsf{T}\rho_2.\end{aligned}$$

Therefore, problem (P_{rel}) can be equivalently stated as

$$\begin{aligned}\min_{\eta_i,\omega_i,\rho_i}\quad & \eta_1\eta_2(1-\lambda^\mathsf{T} a^2)-\|d\|^2\omega_1\omega_2(1-\|d\|^2)-\rho_1^\mathsf{T}\rho_2\\ \text{s.t.}\quad & \sqrt{\omega_i^2\|d\|^2(1-\|d\|^2)+\|\rho_i\|^2}\le\eta_i\sqrt{(1-\lambda^\mathsf{T} a^2)}\\ & \|x_0\|^2 \le -\eta_i\lambda^\mathsf{T} a-\omega_i\|d\|^2.\end{aligned}$$

To prove that the above problem is bounded, we define the vectors $\nu_i = (\rho_i, \|d\|\sqrt{1-\|d\|^2}\omega_i)$. Note that $\nu_1^\mathsf{T}\nu_2 = \|d\|^2\omega_1\omega_2(1-\|d\|^2)+\rho_1^\mathsf{T}\rho_2$. By Cauchy-Schwarz, we have

$$\nu_1^\mathsf{T}\nu_2 \le \|\nu_1\|\|\nu_2\| = \sqrt{\omega_1^2\|d\|^2(1-\|d\|^2)+\|\rho_1\|^2}\sqrt{\omega_2^2\|d\|^2(1-\|d\|^2)}$$

Using the constraint $\sqrt{\omega_i^2\|d\|^2(1-\|d\|^2)+\|\rho_i\|^2}\le\eta_i\sqrt{(1-\lambda^\mathsf{T} a^2)}$, we have that

$$\nu_1^\mathsf{T}\nu_2 \le \eta_1\eta_2(1-\lambda^\mathsf{T} a^2)$$

and therefore,

$$\|d\|^2\omega_1\omega_2(1-\|d\|^2)+\rho_1\rho_2 \le \eta_1\eta_2(1-\lambda^\mathsf{T} a^2),$$

which shows that the problem is bounded by 0.

Note that the above result shows a lower bound of 0 to the value of (P_{exp}), but we require a stronger lower bound on this optimization problem. However, now that we know (P_{exp}) is bounded, we can obtain more properties for it.

Lemma A.3 *The following set has at least one extreme point*

$$F = \{(\theta, \eta, \omega, \rho) \in \mathbb{R}^3 \times \mathbb{R}^m \mid 0 \leq \theta^2 + \eta^2 + 2\theta\eta\lambda^\intercal a - 2\theta + \|x_0\|^2 - \omega^2\|d\|^2 - \|\rho\|^2, \\ \|x_0\|^2 \leq \theta,\ \|d\|^2\omega = -\theta - \eta\lambda^\intercal a\}.$$

Moreover, all extreme points satisfy the constraint $\|x_0\|^2 \leq \theta$ *with equality.*

Proof We first argue that there is at least one extreme point. Indeed, using $\|x_0\|^2 \leq \theta$ we have that $F \subseteq F'$, with

$$F' = \{(\theta, \eta, \omega, \rho) \in \mathbb{R}^3 \times \mathbb{R}^m \mid 0 \leq \theta^2 + \eta^2 + 2\theta\eta\lambda^\intercal a - \omega^2\|d\|^2 - \|\rho\|^2, \\ \|x_0\|^2 \leq \theta,\ \|d\|^2\omega = -\theta - \eta\lambda^\intercal a\}$$

In the proof of Lemma A.1 we showed that F' is convex, moreover, we showed that $\eta \geq 0$ which implies

$$F' \subseteq \{(\theta, \eta, \omega, \rho) \in \mathbb{R}^3 \times \mathbb{R}^m \mid \sqrt{\omega^2\|d\|^2(1 - \|d\|^2) + \|\rho\|^2} \leq \eta\sqrt{(1 - \lambda^\intercal a^2)}, \\ \|x_0\|^2 \leq \theta,\ \|d\|^2\omega = -\theta - \eta\lambda^\intercal a\}$$

Since $\|d\| < 1$ and we are assuming $\|d\| \neq 0$, the nonlinear constraint defines a pointed cone, which implies that F has an extreme point.

Now, by contradiction, let us assume that there is an extreme point $(\theta^*, \eta^*, \omega^*, \rho^*)$ satisfying $\|x_0\|^2 < \theta^*$. We begin by reformulating the set F. Replacing $\theta = -\eta\lambda^\intercal a - \omega\|d\|^2$ in the quadratic constraint yields

$$\begin{aligned}
0 \leq\ & (\eta\lambda^\intercal a + \omega\|d\|^2)^2 + \eta^2 - 2\eta\lambda^\intercal a(\eta\lambda^\intercal a + \omega\|d\|^2) - \omega^2\|d\|^2 - \|\rho\|^2 \\
& + 2(\eta\lambda^\intercal a + \omega\|d\|^2) + \|x_0\|^2 \\
=\ & \eta^2(\lambda^\intercal a^2 + 1 - 2\lambda^\intercal a^2) + \omega^2(\|d\|^4 - \|d\|^2) - \|\rho\|^2 + 2\eta\lambda^\intercal a + 2\omega\|d\|^2 + \|x_0\|^2 \\
=\ & \eta^2(1 - \lambda^\intercal a^2) - \omega^2\|d\|^2(1 - \|d\|^2) - \|\rho\|^2 + 2\eta\lambda^\intercal a + 2\omega\|d\|^2 + \|x_0\|^2 \\
=\ & (1 - \lambda^\intercal a^2)(\eta^2 + 2\eta\frac{\lambda^\intercal a}{1 - \lambda^\intercal a^2}) - \|d\|^2(1 - \|d\|^2)(\omega^2 + 2\omega\frac{1}{1 - \|d\|^2}) - \|\rho\|^2 + \|x_0\|^2
\end{aligned}$$

Completing the squares yields

$$\begin{aligned}
0 \leq\ & (1 - \lambda^\intercal a^2)(\eta + \frac{\lambda^\intercal a}{1 - \lambda^\intercal a^2})^2 - \|d\|^2(1 - \|d\|^2)(\omega + \frac{1}{1 - \|d\|^2})^2 - \|\rho\|^2 \\
& + \|x_0\|^2 - \frac{\lambda^\intercal a^2}{1 - \lambda^\intercal a^2} + \frac{\|d\|^2}{1 - \|d\|^2} \\
=\ & (1 - \lambda^\intercal a^2)(\eta + \frac{\lambda^\intercal a}{1 - \lambda^\intercal a^2})^2 - \|d\|^2(1 - \|d\|^2)(\omega + \frac{1}{1 - \|d\|^2})^2 - \|\rho\|^2 + 1 + \frac{\|d\|^2}{1 - \|d\|^2}
\end{aligned}$$

The latter equation follows since $\|x_0\|^2 = \frac{1}{1 - \lambda^\intercal a^2}$. Using the change of variables $\tilde{\eta} := \sqrt{1 - \lambda^\intercal a^2}(\eta + \frac{\lambda^\intercal a}{1 - \lambda^\intercal a^2})$ and $\tilde{\omega} := \sqrt{\|d\|^2(1 - \|d\|^2)}(\omega + \frac{1}{1 - \|d\|^2})$, as well as defining the constant $c := 1 + \frac{\|d\|^2}{1 - \|d\|^2}$, the constraint writes as

$$0 \leq \tilde{\eta}^2 - \tilde{\omega}^2 - \|\rho\|^2 + c,$$

giving us the equivalent set

$$\tilde{F} := \{(\theta, \tilde{\eta}, \tilde{\omega}, \rho) \in \mathbb{R}^{3+m} \mid 0 \leq \tilde{\eta}^2 - \tilde{\omega}^2 - \|\rho\|^2 + c, \|x_0\|^2 \leq \theta,$$
$$\frac{\tilde{\omega}\|d\|^2}{\sqrt{\|d\|^2(1-\|d\|^2)}} - \frac{\|d\|^2}{1-\|d\|^2} = -\theta - \frac{\lambda^\mathsf{T} a\tilde{\eta}}{\sqrt{1-\lambda^\mathsf{T} a^2}} + \frac{\lambda^\mathsf{T} a^2}{1-\lambda^\mathsf{T} a^2}\}$$

Let us call $(\theta^*, \tilde{\eta}^*, \tilde{\omega}^*, \rho^*)$ the transformed extreme point. Note that we are assuming $\|x_0\|^2 < \theta^*$, thus such extreme point must satisfy the quadratic constraint with equality.

Since $c > 0$, after replacing the value of θ given by the equality constraint onto the inequality, we can apply of [20, Lemma 3.5][6] to show that $(\theta^*, \tilde{\eta}^*, \tilde{\omega}^*, \rho^*)$ satisfying $\|x_0\|^2 < \theta^*$ cannot be an extreme point. Hence, all extreme points must satisfy $\|x_0\|^2 = \theta$, proving the statement.

The previous result provides the structure of extreme points of a set of type F, however, in Problem (P_{exp}) we have two sets of that type. In order to handle this, we note that Problem (P_{exp}) has the following form:

$$\begin{aligned} \min_{\theta_i,\eta_i,\omega_i,\rho_i} \quad & (\theta_1, \eta_1, \omega_1, \rho_1) Q (\theta_2, \eta_2, \omega_2, \rho_2)^\mathsf{T} \\ \text{s.t} \quad & (\theta_i, \eta_i, \omega_i, \rho_i) \in F_i \qquad\qquad i = 1, 2 \end{aligned}$$

where Q is some matrix, and show the following result.

Lemma A.4 *Consider a bounded, bilinear program of the form*

$$\min\{x^\mathsf{T} Ay : x \in P, y \in Q\} \tag{A.4}$$

with $x \in \mathbb{R}^n, y \in \mathbb{R}^m, A \in \mathbb{R}^{n\times m}, P \subseteq \mathbb{R}^n$ *and* $Q \subseteq \mathbb{R}^m$. *Furthermore, assume that both* $\operatorname{conv} P$ *and* $\operatorname{conv} Q$ *have at least one extreme point. Then* (A.4) *has an optimal solution* $(x^*, y^*) \in \operatorname{ext}\operatorname{conv} P \times \operatorname{ext}\operatorname{conv} Q$.

Proof Let $(\bar{x}, \bar{y})$ be an optimal solution. We fix $x = \bar{x}$ and optimize the resulting problem

$$\min\{\bar{x}^\mathsf{T} Ay : y \in Q\} = \min\{\bar{x}^\mathsf{T} Ay : y \in \operatorname{conv} Q\}.$$

The optimal solution of the above problem can be taken to be an extreme point y^* of $\operatorname{conv} Q$ and its optimal value has to be equal to the value achieved by $(\bar{x}, \bar{y})$. Finally, we repeat the argument taking as solution $(\bar{x}, y^*)$ and fixing y instead of x. Thus, there exists an optimal solution (x^*, y^*) to A.4 with $(x^*, y^*) \in \operatorname{ext}\operatorname{conv} P \times \operatorname{ext}\operatorname{conv} Q$.

Corollary A.1 *When* $\|d\| > 0$, *problem* (P_{exp}) *is lower bounded by* $\frac{1}{2}\|x_0\|^2$.

Proof Lemma A.2 shows that (P_{rel}) is bounded. Thus, it follows that (P_{exp}) is also bounded, i.e. there exists an optimal solution. By Lemmas A.4 and A.3, we know that we can assume an optimal solution satisfies $\|x_0\|^2 \leq \theta_i$ with equality. Thus, in the following we will assume that $\|x_0\|^2 = \theta_i$.

[6] In [20], the authors assume throughout that the linear constraints define a polytope, which is not our setting. Nonetheless, this particular result does not need this assumption.

Penalizing the nonlinear constraints with multipliers $-\frac{1}{2}$ yields the objective

$$\begin{aligned}
&\theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \omega_1\omega_2\|d\|^2 - \rho_1\rho_2 \\
&- \frac{1}{2}(\theta_1^2 + \eta_1^2 + 2\theta_1\eta_1\lambda^\mathsf{T} a - 2\theta_1 + \|x_0\|^2 - \omega_1^2\|d\|^2 - \|\rho_1\|^2) \\
&- \frac{1}{2}(\theta_2^2 + \eta_2^2 + 2\theta_2\eta_2\lambda^\mathsf{T} a - 2\theta_2 + \|x_0\|^2 - \omega_2^2\|d\|^2 - \|\rho_2\|^2) \\
&= -\frac{1}{2}(\theta_1 - \theta_2)^2 - \lambda^\mathsf{T} a(\theta_1 - \theta_2)(\eta_1 - \eta_2) - \frac{1}{2}(\eta_1 - \eta_2)^2 + \frac{\|d\|^2}{2}(\omega_1 - \omega_2)^2 \\
&+ \|\rho_1 - \rho_2\|^2 + \theta_1 + \theta_2 - \|x_0\|^2.
\end{aligned}$$

Since a penalization produces a relaxation, it is enough to show that the penalized problem is lower bounded by $\frac{1}{2}\|x_0\|^2$. This problem reads

$$\begin{aligned}
\min_{\theta_i,\eta_i,\omega_i,\rho_i} \quad & -\frac{1}{2}(\theta_1 - \theta_2)^2 - \lambda^\mathsf{T} a(\theta_1 - \theta_2)(\eta_1 - \eta_2) - \frac{1}{2}(\eta_1 - \eta_2)^2 \\
& + \frac{\|d\|^2}{2}(\omega_1 - \omega_2)^2 + \|\rho_1 - \rho_2\|^2 + \theta_1 + \theta_2 - \|x_0\|^2 \\
\text{s.t.} \quad & \|x_0\|^2 = \theta_i \\
& \|d\|^2\omega_i = -\theta_i - \eta_i\lambda^\mathsf{T} a
\end{aligned}$$

Replacing θ_i and ω_i gives

$$\min_{\eta_i,\rho_i} \quad -\frac{1}{2}(\eta_1 - \eta_2)^2 + \frac{\lambda^\mathsf{T} a^2}{2\|d\|^2}(\eta_1 - \eta_2)^2 + \|\rho_1 - \rho_2\|^2 + \|x_0\|^2.$$

The coefficient of $(\frac{\lambda^\mathsf{T} a^2}{2\|d\|^2} - \frac{1}{2})(\eta_1 - \eta_2)^2$ is positive since $0 < \|d\| < -\lambda^\mathsf{T} a$, that is, $\lambda^\mathsf{T} a^2 - \|d\|^2 > 0$. Therefore, the minimum of the above problem is $\|x_0\|^2 \geq \frac{1}{2}\|x_0\|^2$.

A.1.2 The case $\|d\| = 0$

In the previous section we showed that Problem (P_{exp}) is lower bounded by $\frac{1}{2}\|x_0\|^2$ when $\|d\| > 0$. For the case $\|d\| = 0$, we note that (P_{exp}) becomes:

$$\begin{aligned}
\min_{\theta_i,\eta_i,\omega_i,\rho_i} \quad & \theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \rho_1^\mathsf{T}\rho_2 \\
\text{s.t.} \quad & 0 \leq \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - \|\rho_i\|^2 \\
& \|x_0\|^2 \leq \theta_i \\
& 0 = -\theta_i - \eta_i\lambda^\mathsf{T} a.
\end{aligned} \tag{A.5}$$

Now we take any $\tilde{d} \neq 0$ such that $\|\tilde{d}\| < -\lambda^\mathsf{T} a$ and leveraging that (A.5) does not depend on ω, we can reformulate (A.5) as

$$\begin{aligned}
\min_{\theta_i,\eta_i,\omega_i,\rho_i} \quad & \theta_1\theta_2 + \eta_1\eta_2 + \theta_1\eta_2\lambda^\mathsf{T} a + \theta_2\eta_1\lambda^\mathsf{T} a - \omega_1\omega_2\|\tilde{d}\|^2 - \rho_1^\mathsf{T}\rho_2 \\
\text{s.t.} \quad & 0 \leq \theta_i^2 + \eta_i^2 + 2\theta_i\eta_i\lambda^\mathsf{T} a - \omega_i^2\|\tilde{d}\|^2 - \|\rho_i\|^2 \\
& \|x_0\|^2 \leq \theta_i \\
& \|\tilde{d}\|^2\omega_i = -\theta_i - \eta_i\lambda^\mathsf{T} a \\
& \omega_i = 0
\end{aligned}$$

Removing constraints $\omega_i = 0$ results in a problem of the type (P_{exp}) with a non-zero d vector. We can directly obtain the following result.

Corollary A.2 *Problem (P_{exp}) is lower bounded by $\frac{1}{2}\|x_0\|^2$.*

A.2 The set $C_\lambda \cap H - M$ is S^g-free

Recall that $C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} : \|y\| \leq \lambda^\mathsf{T} x\}$ and the apex of $C_\lambda \cap H$ is $(x_0, 0)$ with

$$(x_0, 0) := \left(\frac{-1}{1 - (\lambda^\mathsf{T} a)^2} a + \frac{\lambda^\mathsf{T} a}{1 - (\lambda^\mathsf{T} a)^2} \lambda, 0 \right).$$

Note that that $\|x_0\|^2 = \frac{1}{1 - \lambda^\mathsf{T} a^2}$. We begin with the following auxiliary result.

Lemma A.5 *The hyperplane $-x_0^\mathsf{T} x = 0$ defines a cross section of C_λ, i.e., every point of $(x, y) \in C_\lambda$ can be written as $(x, y) = (x_0, 0) + \tau((\bar{x}, \bar{y}) - (x_0, 0))$ where $(\bar{x}, \bar{y}) \in C_\lambda$ with $-x_0^\mathsf{T} \bar{x} = 0$ and $\tau \geq 0$.*

Proof Let $(x, y) \in C_\lambda$. If $(x, y) = (x_0, 0)$ the result clearly holds. For the other cases, we define $f(t) := (x_0, 0) + t((x, y) - (x_0, 0))$. We begin by noting that $f(t) \in C_\lambda$ for $t \geq 0$: indeed,

$$\lambda^\mathsf{T} f_x(t) = (1 - t)\lambda^\mathsf{T} x_0 + t\lambda^\mathsf{T} x \geq t\|y\| = \|f_y(t)\|$$

where we used that $(x, y) \in C_\lambda$ and that $\lambda^\mathsf{T} x_0 = 0$. Thus, if we are able to show that there exist a $t^* > 0$ such that $-(x_0, 0)f(t^*) = 0$ we are done, since in that case

$$(x, y) = (x_0, 0) + \frac{1}{t^*}(f(t^*) - (x_0, 0))$$

and we can define $\tau = 1/t^*$ and $(\bar{x}, \bar{y}) = f(t^*)$. Note that

$$-x_0 f_x(t) = -\|x_0\|^2 + t(\|x_0\|^2 - x_0^\mathsf{T} x),$$

therefore such $t^* > 0$ exists if and only if $\|x_0\|^2 - x_0^\mathsf{T} x > 0$. Replacing x_0 by its definition and using that $1 - \lambda^\mathsf{T} a^2 > 0$ yields

$$\|x_0\|^2 - x_0^\mathsf{T} x > 0 \Leftrightarrow a^\mathsf{T} x - \lambda^\mathsf{T} a \lambda^\mathsf{T} x > -1.$$

Recall that $\|d\| < -\lambda^\mathsf{T} a$. Since $(x, y) \in C_\lambda$ by assumption, we have $\|y\| \leq \lambda^\mathsf{T} x$, thus $-\lambda^\mathsf{T} a \lambda^\mathsf{T} x > \|d\|\|y\|$. Furthermore, we can rewrite $a^\mathsf{T} x = -1 - d^\mathsf{T} y$. Then,

$$a^\mathsf{T} x - \lambda^\mathsf{T} a \lambda^\mathsf{T} x > -1 - d^\mathsf{T} y + \|d\|\|y\| \geq -1,$$

where the latter inequality follows from Cauchy-Schwarz. This concludes the proof.

We are now ready to state the proof of Theorem 4.2, i.e., that $C_\lambda \cap H - M$ is S^g-free. Much like in the proof of Theorem 4.1, there will be an intermediate claim that we leave for a subsequent section.

Proof (Theorem 4.2) To show that $C_\lambda \cap H - M$ is S^g-free, we show that the translation $C_\lambda \cap H + m$ is S^g-free for all $m \in -M$. Since this is clearly true for $m = 0$, we assume that $m \neq 0$. From the definition of M, recall that $m \in \langle \{\lambda, a\} \rangle \times \mathbb{R}^m$. Lemma A.5 shows that every point of C_λ can be written as $(x_0, 0) + \tau((x, y) - (x_0, 0))$ where $(x, y) \in C_\lambda$ and $-x_0^\mathsf{T} x = 0$. Using this result, we show that if $(m_x, m_y) \in -M$, then the quadratic function

$\|x\|^2 - \|y\|^2$ restricted to the line $(m_x, m_y) + (x_0, 0) + \tau((x, y) - (x_0, 0))$ has all its roots in $(-\infty, 0]$. This implies that $C_\lambda - M$ is S^g-free.

Let us introduce $(\bar{x}, \bar{y}) := (m_x + x_0, m_y)$. Note that since $m \neq 0$, $(\bar{x}, \bar{y})$ belongs to the set $L = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m \mid \|x\| \geq \|y\|, -x_0^\mathsf{T} x \geq 0, a^\mathsf{T} x + d^\mathsf{T} y = -1\}$. The quadratic restricted to the line is given by $\|\bar{x} + \tau(x - x_0)\|^2 - \|\bar{y} + \tau y\|^2$. Expanding this yields,

$$\begin{aligned}
&\|\bar{x} + \tau(x - x_0)\|^2 - \|\bar{y} + \tau y\|^2 \\
&= \|\bar{x}\|^2 + \tau^2\|x - x_0\|^2 + 2\tau\bar{x}^\mathsf{T}(x - x_0) - \|\bar{y}\|^2 - \tau^2\|y\|^2 - 2\tau\bar{y}^\mathsf{T} y \\
&= \tau^2(\|x - x_0\|^2 - \|y\|^2) + 2\tau(\bar{x}^\mathsf{T}(x - x_0) - \bar{y}^\mathsf{T} y) + \|\bar{x}\|^2 - \|\bar{y}\|^2 \\
&\geq \tau^2(\|x - x_0\|^2 - \|y\|^2) + 2\tau(\bar{x}^\mathsf{T}(x - x_0) - \bar{y}^\mathsf{T} y). \qquad \text{(A.6)}
\end{aligned}$$

The inequality follows from $\|\bar{x}\| \geq \|\bar{y}\|$ since $(\bar{x}, \bar{y}) \in L$.

Note that $\|x - x_0\| - \|y\| \geq \lambda^\mathsf{T}(x - x_0) - \|y\| = \lambda^\mathsf{T} x - \|y\| \geq 0$, and thus $\|x - x_0\|^2 - \|y\|^2 \geq 0$. Due to the this, the roots of (A.6) enclose the roots of $\|\bar{x} + \tau(x - x_0)\|^2 - \|\bar{y} + \tau y\|^2$. Therefore, it is enough to prove that the roots of (A.6) are in $(-\infty, 0]$.

One of the roots is clearly 0, and the other is

$$-2\frac{\bar{x}^\mathsf{T}(x - x_0) - \bar{y}^\mathsf{T} y}{\|x - x_0\|^2 - \|y\|^2}.$$

Hence, to show that this root is negative, we need to prove that $\bar{x}^\mathsf{T}(x - x_0) - \bar{y}^\mathsf{T} y \geq 0$. This translates to showing that the optimization problem

$$\begin{aligned}
\min_{\bar{x}, \bar{y}, x, y} \quad & \bar{x}^\mathsf{T}(x - x_0) - \bar{y}^\mathsf{T} y \\
\text{s.t.} \quad & \|\bar{x}\|^2 - \|\bar{y}\|^2 \geq 0 \\
& a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} = -1 \\
& -x_0^\mathsf{T}\bar{x} \geq 0 \qquad\qquad (P_{root}) \\
& (\lambda^\mathsf{T} x)^2 - \|y\|^2 \geq 0 \\
& a^\mathsf{T} x + d^\mathsf{T} y = -1 \\
& \lambda^\mathsf{T} x \geq 0
\end{aligned}$$

is non-negative.

Recall that we can restrict to vectors that live in a subspace orthogonal to the lineality space of $C_\lambda \cap H$. This implies that we can write $\bar{x} = \bar{\theta} a + \bar{\eta}\lambda$, $x = \theta a + \eta\lambda$. We can also write $\bar{y} = \bar{\omega} d + \bar{\rho}$ and $y = \omega d + \rho$.

Since $x_0^\mathsf{T} x = 0$ together with the fact that $x_0^\mathsf{T} a = -1$ and $x_0^\mathsf{T}\lambda = 0$, we get $x = \eta\lambda$. Replacing the variables with its decompositions in the objective as well as in the constraints yields the equivalent expanded problem

$$\begin{aligned}
\min_{\bar{\theta}, \bar{\eta}, \bar{\omega}, \bar{\rho}, \eta, \omega, \rho} \quad & \bar{\eta}\eta - \bar{\omega}\omega\|d\|^2 - \bar{\rho}\rho + \bar{\theta}\eta\lambda^\mathsf{T} a + \bar{\theta} \\
\text{s.t.} \quad & \bar{\theta}^2 + \bar{\eta}^2 + 2\bar{\theta}\bar{\eta}\lambda^\mathsf{T} a - \bar{\omega}^2\|d\|^2 - \|\bar{\rho}\|^2 \geq 0 \\
& \bar{\theta} + \bar{\eta}\lambda^\mathsf{T} a + \bar{\omega}\|d\|^2 = -1 \\
& \bar{\theta} \geq 0 \qquad\qquad (P_{root}^{exp}) \\
& \eta^2 - \omega^2\|d\|^2 - \|\rho\|^2 \geq 0 \\
& \eta\lambda^\mathsf{T} a + \omega\|d\|^2 = -1 \\
& \eta \geq 0.
\end{aligned}$$

The rest of the proof relies on the following claim, whose correctness is shown in Corollary A.3.

Claim Problem (P^{exp}_{root}) is bounded. Moreover, we can assume that an optimal solution satisfies $\bar{\theta} = 0$.

We now show that (P^{exp}_{root}) ≥ 0 by building a relaxation such that, when requiring $\bar{\theta} = 0$, has a non-negative optimal solution. Let us build a penalized objective using multipliers $-\frac{1}{2}$ for each nonlinear constraint of (P^{exp}_{root}):

$$\begin{aligned}
&\bar{\eta}\eta - \bar{\omega}\omega\|d\|^2 - \bar{\rho}\rho + \bar{\theta}\eta\lambda^\mathsf{T} a + \bar{\theta} \\
&- \frac{1}{2}\left(\bar{\theta}^2 + \bar{\eta}^2 + 2\bar{\theta}\bar{\eta}\lambda^\mathsf{T} a - \bar{\omega}^2\|d\|^2 - \|\bar{\rho}\|^2\right) \\
&- \frac{1}{2}\left(\eta^2 - \omega^2\|d\|^2 - \|\rho\|^2\right) \\
=& - \frac{1}{2}(\eta - \bar{\eta})^2 + \frac{1}{2}\|d\|^2(\omega - \bar{\omega})^2 + \frac{1}{2}\|\rho - \bar{\rho}\|^2 - \frac{\bar{\theta}^2}{2} + \bar{\theta}\eta\lambda^\mathsf{T} a - \bar{\theta}\bar{\eta}\lambda^\mathsf{T} a + \bar{\theta}
\end{aligned}$$

We replace the objective function in (P^{exp}_{root}) by the above aggregation, drop the nonlinear constraints and replace $\bar{\theta} = 0$. This yields

$$\begin{aligned}
\min_{\bar{\eta},\bar{\omega},\bar{\rho},\eta,\omega,\rho} \quad & - \frac{1}{2}(\eta - \bar{\eta})^2 + \frac{1}{2}\|d\|^2(\omega - \bar{\omega})^2 + \frac{1}{2}\|\rho - \bar{\rho}\|^2 \\
\text{s.t.} \quad & \bar{\eta}\lambda^\mathsf{T} a + \bar{\omega}\|d\|^2 = -1 \\
& \eta\lambda^\mathsf{T} a + \omega\|d\|^2 = -1 \\
& \eta \geq 0.
\end{aligned} \tag{P^{rel}_{root}}$$

As mentioned above, since we can assume $\bar{\theta} = 0$ in an optimal solution of (P^{exp}_{root}), showing that (P^{rel}_{root}) ≥ 0 suffices.

By solving the linear constraints for $\bar{\eta} = -\frac{1}{\lambda^\mathsf{T} a}(\bar{\omega}\|d\|^2 + 1)$ and $\eta = -\frac{1}{\lambda^\mathsf{T} a}(\omega\|d\|^2 + 1)$ and replacing the variables in the objective of (P^{rel}_{root}) gives

$$\begin{aligned}
&- \frac{1}{2}(\eta - \bar{\eta})^2 + \frac{1}{2}\|d\|^2(\omega - \bar{\omega})^2 + \frac{1}{2}\|\rho - \bar{\rho}\|^2 \\
=& - \frac{\|d\|^4}{2\lambda^\mathsf{T} a^2}(\omega - \bar{\omega})^2 + \frac{1}{2}\|d\|^2(\omega - \bar{\omega})^2 + \frac{1}{2}\|\rho - \bar{\rho}\|^2 \\
=& \frac{\|d\|^2}{2}(1 - \frac{\|d\|^2}{\lambda^\mathsf{T} a^2})(\omega - \bar{\omega})^2 + \frac{1}{2}\|\rho - \bar{\rho}\|^2.
\end{aligned}$$

Since $\frac{\|d\|^2}{\lambda^\mathsf{T} a^2} < 1$, we have that $\frac{\|d\|^2}{2}(1 - \frac{\|d\|^2}{\lambda^\mathsf{T} a^2})(\omega - \bar{\omega})^2 + \frac{1}{2}\|\rho - \bar{\rho}\|^2 \geq 0$, proving that ($P^{rel}_{root}$) ≥ 0 and therefore also (P^{exp}_{root}) ≥ 0 This concludes the proof that that the set $C_\lambda \cap H - M$ is S^g-free.

A.2.1 Auxiliary results

The purpose of the following results is to show that Problem (P^{exp}_{root}) is bounded and that we can assume that an optimal solution satisfies $\bar{\theta} = 0$, which was used in the proof in Appendix A.2.

Lemma A.6 *The optimization problem* (P^{exp}_{root}) *is bounded.*

Proof First, consider the change of variables $\hat{\theta} = \bar{\theta} + \frac{1}{1-\lambda^\mathsf{T} a^2}$ and $\hat{\eta} = \bar{\eta} - \frac{\lambda^\mathsf{T} a}{1-\lambda^\mathsf{T} a^2}$. Replacing these in transforms (P^{exp}_{root}) into the equivalent problem

$$\begin{aligned}
\min_{\hat{\theta},\hat{\eta},\bar{\omega},\bar{\rho},\eta,\omega,\rho} \quad & \hat{\eta}\eta - \bar{\omega}\omega\|d\|^2 - \bar{\rho}\rho + \hat{\theta}\eta\lambda^\mathsf{T} a + \hat{\theta} - \frac{1}{1-\lambda^\mathsf{T} a} \\
\text{s.t.} \quad & \hat{\theta}^2 + \hat{\eta}^2 - \bar{\omega}^2\|d\|^2 - \|\bar{\rho}\|^2 + 2\lambda^\mathsf{T} a\hat{\theta}\hat{\eta} - 2\hat{\theta} - \frac{\lambda^\mathsf{T} a^2}{(1-\lambda^\mathsf{T} a^2)^2} \geq 0 \\
& \hat{\theta} + \hat{\eta}\lambda^\mathsf{T} a + \bar{\omega}\|d\|^2 = 0 \\
& \hat{\theta} \geq \|x_0\|^2 \\
& \eta^2 - \omega^2\|d\|^2 - \|\rho\|^2 \geq 0 \\
& \eta\lambda^\mathsf{T} a + \omega\|d\|^2 = -1 \\
& \eta \geq 0.
\end{aligned}$$

Since $\hat{\theta} \geq \|x_0\|^2$, we have that $-2\hat{\theta} - \frac{\lambda^\mathsf{T} a^2}{(1-\lambda^\mathsf{T} a^2)^2} \leq 0$ and we can relax the nonlinear constraint in the previous optimization problem to obtain

$$\begin{aligned}
\min_{\hat{\theta},\hat{\eta},\bar{\omega},\bar{\rho},\eta,\omega,\rho} \quad & \hat{\eta}\eta - \bar{\omega}\omega\|d\|^2 - \bar{\rho}\rho + \hat{\theta}\eta\lambda^\mathsf{T} a + \hat{\theta} - \frac{1}{1-\lambda^\mathsf{T} a} \\
\text{s.t.} \quad & \hat{\theta}^2 + \hat{\eta}^2 + 2\lambda^\mathsf{T} a\hat{\theta}\hat{\eta} - \bar{\omega}^2\|d\|^2 - \|\bar{\rho}\|^2 \geq 0 \\
& \hat{\theta} + \hat{\eta}\lambda^\mathsf{T} a + \bar{\omega}\|d\|^2 = 0 \\
& \hat{\theta} \geq \|x_0\|^2 \\
& \eta^2 - \omega^2\|d\|^2 - \|\rho\|^2 \geq 0 \\
& \eta\lambda^\mathsf{T} a + \omega\|d\|^2 = -1 \\
& \eta \geq 0.
\end{aligned} \tag{P_{conv}}$$

In (P_{conv}) we can analyze the feasible region of two sets of variables separately. The constraints involving the variables (η, ω, ρ) clearly define a convex region: since $\eta \geq 0$, the nonlinear constraint defines a second order cone. On the other hand, the convexity of the constraints involving variables $(\hat{\theta}, \hat{\eta}, \bar{\omega}, \bar{\rho})$ follows from the proof of Lemma A.2. Thus, the feasible region of (P_{conv}) is convex.

In particular, we obtain that $\hat{\eta} \geq 0$ and the first nonlinear constraint can be reformulated as

$$\sqrt{\bar{\omega}^2\|d\|^2(1-\|d\|^2) + \|\bar{\rho}_i\|^2} \leq \hat{\eta}\sqrt{(1-\lambda^\mathsf{T} a^2)}. \tag{A.7}$$

Now, replacing $\hat{\theta} = -\hat{\eta}\lambda^\mathsf{T} a - \bar{\omega}\|d\|^2$ in the objective function yields

$$\begin{aligned}
&\hat{\eta}\eta - \bar{\omega}\omega\|d\|^2 - \bar{\rho}\rho + \hat{\theta}\eta\lambda^\mathsf{T} a + \hat{\theta} - \frac{1}{1-\lambda^\mathsf{T} a} \\
=&\hat{\eta}\eta(1-\lambda^\mathsf{T} a^2) - \bar{\omega}\eta\|d\|^2\lambda^\mathsf{T} a - \hat{\eta}\lambda^\mathsf{T} a - \bar{\omega}\|d\|^2 - \bar{\omega}\omega\|d\|^2 - \bar{\rho}\rho + \frac{1}{1-\lambda^\mathsf{T} a} \\
=&\hat{\eta}\eta(1-\lambda^\mathsf{T} a^2) - \bar{\omega}\omega\|d\|^2(1-\|d\|^2) - \bar{\rho}\rho - \hat{\eta}\lambda^\mathsf{T} a + \frac{1}{1-\lambda^\mathsf{T} a}.
\end{aligned}$$

The last equality is obtained by replacing the constraint $\eta = -\frac{1}{\lambda^\mathsf{T} a}(1+\omega\|d\|^2)$ in $\bar{\omega}\eta\|d\|^2\lambda^\mathsf{T} a$.

As in the proof of Lemma A.2, consider the vectors $\nu = (\rho, \|d\|\sqrt{1-\|d\|^2}w)$ and $\bar{\nu} = (\bar{\rho}, \|d\|\sqrt{1-\|d\|^2}\bar{\omega})$. By Cauchy-Schwarz, we have

$$\begin{aligned}
\bar{\nu}^\mathsf{T}\nu &= \bar{\omega}\omega\|d\|^2(1-\|d\|^2) + \bar{\rho}\rho \\
&\le \sqrt{\bar{\omega}^2\|d\|^2(1-\|d\|^2) + \|\bar{\rho}\|^2}\sqrt{\omega^2\|d\|^2(1-\|d\|^2) + \|\rho\|^2} \\
&\le \hat{\eta}\sqrt{1-\lambda^\mathsf{T}a^2}\sqrt{\omega^2\|d\|^2(1-\|d\|^2) + \|\rho\|^2} \qquad \text{(A.8)}
\end{aligned}$$

where the last inequality follows from (A.7). Since $\omega^2\|d\|^2 = \omega^2\|d\|^2(1-\|d\|^2) + \omega^2\|d\|^4$ together with the fact that $\eta\lambda^\mathsf{T}a + \omega\|d\|^2 = -1 \Leftrightarrow \omega^2\|d\|^4 = (\eta\lambda^\mathsf{T}a+1)^2$, the nonlinear constraint $\eta^2 - \omega^2\|d\|^2 - \|\rho\|^2 \ge 0$ can be reformulated to

$$\sqrt{\omega^2\|d\|^2(1-\|d\|^2) + \|\rho\|^2} \le \sqrt{\eta^2 - (\eta\lambda^\mathsf{T}a+1)^2}.$$

Since $\hat{\eta} \ge 0$, we can use this in (A.8) to obtain

$$\bar{\nu}^\mathsf{T}\nu \le \hat{\eta}\sqrt{1-\lambda^\mathsf{T}a^2}\sqrt{\eta^2 - (\eta\lambda^\mathsf{T}a+1)^2},$$

Hence, the objective function of (P_{conv}) can be lower bounded by

$$\begin{aligned}
&\hat{\eta}\eta(1-\lambda^\mathsf{T}a^2) - \bar{\omega}\omega\|d\|^2(1-\|d\|^2) - \bar{\rho}\rho - \hat{\eta}\lambda^\mathsf{T}a + \frac{1}{1-\lambda^\mathsf{T}a} \\
&\ge \hat{\eta}\eta(1-\lambda^\mathsf{T}a^2) - \hat{\eta}\sqrt{1-\lambda^\mathsf{T}a^2}\sqrt{\eta^2 - \omega^2\|d\|^4} - \hat{\eta}\lambda^\mathsf{T}a + \frac{1}{1-\lambda^\mathsf{T}a} \\
&= \hat{\eta}\left(\eta(1-\lambda^\mathsf{T}a^2) - \lambda^\mathsf{T}a - \sqrt{1-\lambda^\mathsf{T}a^2}\sqrt{\eta^2 - (\eta\lambda^\mathsf{T}a+1)^2}\right) + \frac{1}{1-\lambda^\mathsf{T}a}.
\end{aligned}$$

If we show that $\eta(1-\lambda^\mathsf{T}a^2) - \lambda^\mathsf{T}a - \sqrt{1-\lambda^\mathsf{T}a^2}\sqrt{\eta^2 - (\eta\lambda^\mathsf{T}a+1)^2} \ge 0$, we obtain that (P_{conv}) is lower-bounded by 0. This condition is equivalent to

$$\begin{aligned}
&\eta(1-\lambda^\mathsf{T}a^2) - \lambda^\mathsf{T}a \overset{?}{\ge} \sqrt{1-\lambda^\mathsf{T}a^2}\sqrt{\eta^2 - (\eta\lambda^\mathsf{T}a+1)^2} \\
\Leftrightarrow\quad &(\eta(1-\lambda^\mathsf{T}a^2) - \lambda^\mathsf{T}a)^2 \overset{?}{\ge} (1-\lambda^\mathsf{T}a^2)(\eta^2 - (\eta\lambda^\mathsf{T}a+1)^2) \\
\Leftrightarrow\quad &\eta^2(1-\lambda^\mathsf{T}a^2)^2 - 2\eta\lambda^\mathsf{T}a(1-\lambda^\mathsf{T}a^2) + \lambda^\mathsf{T}a^2 \overset{?}{\ge} (1-\lambda^\mathsf{T}a^2)(\eta^2(1-\lambda^\mathsf{T}a^2) - 2\eta\lambda^\mathsf{T}a - 1) \\
\Leftrightarrow\quad &\lambda^\mathsf{T}a^2 \overset{?}{\ge} -1 + \lambda^\mathsf{T}a^2 \\
\Leftrightarrow\quad &0 \overset{?}{\ge} -1.
\end{aligned}$$

This shows that $(P_{root}^{exp}) \ge (P_{conv}) \ge 0$, concluding the proof.

Corollary A.3 *Problem (P_{root}^{exp}) is bounded. Moreover, we can assume that an optimal solution satisfies $\bar{\theta} = 0$.*

Proof Lemma A.6 shows that the problem is bounded thus an optimum exists. Furthermore, by Lemma A.4 we know that the optimum must be attained at an extreme point of the feasible region.

To show that an extreme point must satisfy the constraint $\bar{\theta} \ge 0$ with equality, we note that the following set of inequalities in (P_{root}^{exp})

$$\begin{aligned}
&\bar{\theta}^2 + \bar{\eta}^2 + 2\bar{\theta}\bar{\eta}\lambda^\mathsf{T}a - \bar{\omega}^2\|d\|^2 - \|\bar{\rho}\|^2 \ge 0 \\
&\bar{\theta} + \bar{\eta}\lambda^\mathsf{T}a + \bar{\omega}\|d\|^2 = -1 \\
&\bar{\theta} \ge 0
\end{aligned}$$

have the same structure of the set in Lemma A.3. Using this result, we can easily see that $\bar{\theta} = 0$ in an extreme point.

A.3 Proof of Proposition 5.2

Proposition 5.2 *Let ϕ be the minimal representation of $C - f$ given in* (5.1)*. Then $\{z : \phi(z) \leq \tau\} = C - \nu - \tau(f - \nu)$, where ν is defined in* (4.3) *(the apex of $C_\lambda \cap H$ in the space $\langle\{\lambda, a\}\rangle \times \mathbb{R}^m$).*

Proof We first show that $\{z : \phi(z) \leq \tau\} \subseteq C - \nu - \tau(f - \nu)$. Let z be such that $\phi(z) \leq \tau$. We have to show that there exists a $\bar{z} \in C$ with $z = \bar{z} - \nu - \tau(f - \nu)$. Choose $\bar{z} = z + \nu + \tau(f - \nu)$. It is left to prove that $\bar{z} \in C$, that is, $\phi(\bar{z} - f) \leq 1$:

$$\begin{aligned}
\phi(\bar{z} - f) &= \phi(z + (1 - \tau)(\nu - f)) \\
&\leq \phi(z) + \phi((1 - \tau)(\nu - f)) \\
&\leq \tau + \phi((1 - \tau)(\nu - f)),
\end{aligned}$$

where the first inequality follows from the sublineality of ϕ as shown in Proposition 5.1 and the second inequality follows from the hypothesis $\phi(z) \leq \tau$. If $\phi(\alpha(\nu - f)) = \alpha\phi(\nu - f)$ for arbitrary $\alpha \in \mathbb{R}$ (not only $\alpha > 0$), then $\phi(\bar{z} - f) \leq 1$ follows immediately since the apex ν is on the boundary of C and therefore $\phi(\nu - f) = 1$. We have

$$\begin{aligned}
\phi(\alpha(\nu - f)) &= \sup_{\|\beta\|=1} \alpha \frac{-\beta^\mathsf{T} f_y - \lambda^\mathsf{T}(x_0 - f_x)}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} \\
&= \sup_{\|\beta\|=1} \alpha \left(\frac{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} + \frac{-\lambda^\mathsf{T} x_0}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} \right) \\
&= \alpha\phi(\nu - f).
\end{aligned}$$

The latter equation follows from $\lambda^\mathsf{T} x_0 = 0$. Thus, $\phi(\bar{z} - f) \leq 1$ and so $\bar{z} \in C$.

Now, we show the other inclusion, i.e., $C - \nu - \tau(f - \nu) \subseteq \{z : \phi(z) \leq \tau\}$. Let $z \in C - \nu - \tau(f - \nu)$, then there exists a $\bar{z} \in C$ such that $z = \bar{z} - \nu - \tau(f - \nu)$. Since

$$\begin{aligned}
\phi(z) &= \phi(\bar{z} - \nu - \tau(f - \nu)) \\
&= \phi(\bar{z} - f + (\tau - 1)(\nu - f)) \\
&\leq \phi(\bar{z} - f) + (\tau - 1)\phi(\nu - f) \\
&\leq 1 + \tau - 1 = \tau,
\end{aligned}$$

it holds that $z \in \{z : \phi(z) \leq \tau\}$. This proves the proposition.

A.4 Proof of $C - M = L \cup C$

Proposition A.1 *It holds that $C - M = L \cup C$.*

Proof Recall that

$$L = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m : a^\mathsf{T} x + d^\mathsf{T} y = -1, \|x\| \geq \|y\|, (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x \geq 0\}.$$

First, we show that $C - M \subseteq L \cup C$. Let $(\bar{x}, \bar{y}) \in C - M$. There exists $(x, y) \in C$ and $m \in M$ such that $(\bar{x}, \bar{y}) = (x, y) - m$. We need to show that $(\bar{x}, \bar{y}) \in L \cup C$.

Observe that if $m = 0$, then $(\bar{x}, \bar{y}) \in C$ and we are done. Therefore, we assume $m \neq 0$.

We now show that $(\bar{x}, \bar{y}) \in L$. It trivially holds that $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} = -1$. Since $C - M$ is S^g-free by Theorem 4.2, we have that $\|\bar{x}\| \geq \|\bar{y}\|$. Lastly, it holds that

$$\begin{aligned}(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T}\bar{x} &= (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x - (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} m_x \\ &\geq (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x + 1.\end{aligned}$$

As proven in the proof of Lemma A.5, $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x \geq -1$ for all $x \in C$. This shows that the constraint $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T}\bar{x} \geq 0$ is satisfied and therefore $(\bar{x}, \bar{y}) \in L$.

Now, we show that $L \cup C \subseteq C - M$. Let $(\bar{x}, \bar{y}) \in L$, since the case $(\bar{x}, \bar{y}) \in C$ is trivial. To show that $(\bar{x}, \bar{y}) \in C - M$, we have to show that there exist $(x, y) \in C$ and $m \in -M$ with $(\bar{x}, \bar{y}) = (x, y) + m$. We choose $m := (\bar{x}, \bar{y}) - \nu \in L - \nu = -M$ and $(x, y) := \nu \in C$. This concludes the proof.

A.5 Proof of Proposition 5.3

Proposition 5.3 *Let $\bar{\tau}$ be the largest root of the univariate quadratic equation $\|l_x(\tau)\|^2 = \|l_y(\tau)\|^2$. If the root exists and $l(\bar{\tau}) \in L$, then $\psi(r) = \bar{\tau}$. Otherwise, $\psi(r) = \phi(r)$.*

Proof Recall that $\psi(r) = \min\{\inf\{\tau : l(\tau) \in L\}, \phi(r)\}$. In order to prove the proposition, we need to see how $\inf\{\tau : l(\tau) \in L\}$ compares with $\phi(r)$. To do this, we study the problem $\inf\{\tau : l(\tau) \in L\}$.

First, notice that if the infimum exists, then it is in the relative boundary of L. From the definition of L (see (4.2)) it is easy to see that

$$L \subseteq R := \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m : a^\mathsf{T} x + d^\mathsf{T} y = -1, (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x \geq 0\}.$$

Thus, $\inf\{\tau : l(\tau) \in L\} \geq \inf\{\tau : l(\tau) \in R\}$ and it is enough to prove that $\inf\{\tau : l(\tau) \in R\}$ exists. We have that

$$\begin{aligned}(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} l_x(\tau) &= (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T}(r_x + x_0 + \tau(f_x - x_0)) \\ &= (a - \lambda^\mathsf{T} a\lambda)^\mathsf{T}(r_x + x_0) + \tau,\end{aligned}$$

where the last equality follow from $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} f_x = 0$ and $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x_0 = -1$. This shows that there exists τ_0 such that if $\tau \leq \tau_0$ then $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} l_x(\tau) < 0$. That is, $l(\tau) \notin R$ for all $\tau \leq \tau_0$ and so $\inf\{\tau : l(\tau) \in R\}$ exists.

Let $\tau_1 = \inf\{\tau : l(\tau) \in L\}$. There are two cases, depending on which part of the relative boundary of L $l(\tau_1)$ is located.

Case 1 $\|l_x(\tau_1)\| = \|l_y(\tau_1)\|$ and $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} l_x(\tau_1) \geq 0$.
We proceed to show that τ_1 corresponds to the largest root of $\|l_x(\tau)\|^2 = \|l_y(\tau)\|^2$ and that $\psi(r) = \tau_1$. This proves the first statement of the proposition.

We start by showing that $\tau_1 \leq \phi(r)$. Assume, by contradiction, that $\phi(r) < \tau_1$. Then, $\tau_2 = \phi(r)$ is such that $l(\tau_2) \in C$. However, since $f - \nu \in \mathrm{ri}(\mathrm{rec}(C))$, we have that for every

$\tau > \tau_2$, $l(\tau) \in \mathrm{ri}(C)$. To see this, recall that $l(\tau) = r + \nu + \tau(f - \nu)$. If $\tau > \tau_2$, then $l(\tau) = l(\tau_2) + (\tau - \tau_2)(f - \nu) \in C + \mathrm{ri}(\mathrm{rec}(C)) \subseteq \mathrm{ri}(C)$. In particular, $l(\tau_1) \in \mathrm{ri}(C)$, which is a contradiction with the fact that $l(\tau_1) \in S^g$ and C is S^g-free.

We now show that τ_1 is the largest root of the quadratic equation $\|l_x(\tau)\|^2 = \|l_y(\tau)\|^2$. Since $l(\tau_1) \in L$ and $L \subseteq C - M$ (see Proposition A.1), we have that $l(\tau_1) \in C - M$. Similar to the argument above, $l(\tau_1 + \epsilon) = l(\tau_1) + \epsilon(f - \nu) \in C - M + \mathrm{ri}(\mathrm{rec}(C))$, for every $\epsilon > 0$. Observe that $C - M + \mathrm{ri}(\mathrm{rec}(C)) \subseteq \mathrm{ri}(C - M)$. To see this, let $(x, y) \in C, m \in M, (z, w) \in \mathrm{ri}(\mathrm{rec}(C))$ and let $\bar{\epsilon} > 0$ small enough such that $(z, w) + B_{\bar{\epsilon}}((0, 0)) \subseteq \mathrm{rec}(C)$. Then, $(x, y) - m + (z, w) + B_{\bar{\epsilon}}((0, 0)) \subseteq C - M$, i.e., $(x, y) - m + (z, w) \in \mathrm{ri}(C - M)$.

Therefore, $l(\tau_1 + \epsilon) \in \mathrm{ri}(C - M)$, for every $\epsilon > 0$. Thus, the equation (on ϵ) $\|l_x(\tau_1 + \epsilon)\|^2 = \|l_y(\tau_1 + \epsilon)\|^2$ cannot have a positive root, since otherwise $l(\tau_1 + \epsilon) \in S^g$ contradicting the fact that $C - M$ is S^g-free as shown in Theorem 4.2. This implies that there is no root larger than τ_1, i.e., τ_1 is the largest root.

Case 2 $\|l_x(\tau_1)\| > \|l_y(\tau_1)\|$ and $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} l_x(\tau_1) = 0$.
We proceed to show that $\tau_1 > \phi(r)$, which shows the second claim and proves the proposition. It is enough to show that $l(\tau_1) \in \mathrm{ri}\, C$, since then $\phi(r) = \inf\{\tau : l(\tau) \in C\} < \tau_1$.

Let $(x, y) = l(\tau_1)$. We have to show that $\|y\| < \lambda^\mathsf{T} x$. Given that $(x, y) \in L \subseteq \langle\{\lambda, a\}\rangle \times \mathbb{R}^m$, (x, y) can be rewritten as $x = \bar{\theta} a + \theta\lambda$ and $y = \omega d + \rho$ with $\bar{\theta}, \theta, \omega \in \mathbb{R}$ and $\rho^\mathsf{T} d = 0$. The condition $(a - \lambda^\mathsf{T} a\lambda)^\mathsf{T} x = 0$ then implies that $\bar{\theta} = 0$. We thus need to prove that $\|y\| < \theta$. Notice that the assumption $\|x\| > \|y\|$ is equivalent to $|\theta| > \|y\|$, so it suffices to show that $\theta \geq 0$.

Expanding x and y in the constraint $a^\mathsf{T} x + d^\mathsf{T} y = -1$, yields $\lambda^\mathsf{T} a\theta + \omega\|d\|^2 = -1$. Solving for θ yields $\theta = -\frac{1+\omega\|d\|^2}{\lambda^\mathsf{T} a}$. Note that θ can be described as an increasing, linear function of ω, i.e., $\theta(\omega) = -\frac{1+\omega\|d\|^2}{\lambda^\mathsf{T} a}$. Showing that $\theta(\omega) \geq 0$ for the smallest possible ω is then enough to prove $\theta \geq 0$.

Squaring the constraint $|\theta| > \|y\|$ and expanding the definitions we get the following constraint on ω:

$$\|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)\omega^2 + 2\|d\|^2\omega + (1 - \lambda^\mathsf{T} a^2\|\rho\|^2) > 0.$$

Note that $\|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2) < 0$, so the left-hand side is a concave quadratic function. The smallest value ω can attain is the smallest root ω_0 of the quadratic function, that is,

$$\omega_0 = \frac{-\|d\|^2 + \sqrt{\|d\|^4 - \|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)(1 - \lambda^\mathsf{T} a^2\|\rho\|^2)}}{\|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)}.$$

We have

$$\theta(w_0) = \frac{-1}{\lambda^\mathsf{T} a} + \frac{\|d\|^2 - \sqrt{\|d\|^4 - \|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)(1 - \|\rho\|^2\lambda^\mathsf{T} a^2)}}{\lambda^\mathsf{T} a(\|d\|^2 - \lambda^\mathsf{T} a^2)}$$

and therefore

$$\begin{aligned}
\theta(w_0) > 0 \iff & \frac{\|d\|^2 - \sqrt{\|d\|^4 - \|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)(1 - \|\rho\|^2 \lambda^\mathsf{T} a^2)}}{\|d\|^2 - \lambda^\mathsf{T} a^2} < 1 \\
\iff & \|d\|^2 - \sqrt{\|d\|^4 - \|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)(1 - \|\rho\|^2 \lambda^\mathsf{T} a^2)} > \|d\|^2 - \lambda^\mathsf{T} a^2 \\
\iff & \lambda^\mathsf{T} a^2 > \sqrt{\|d\|^4 - \|d\|^2(\|d\|^2 - \lambda^\mathsf{T} a^2)(1 - \|\rho\|^2 \lambda^\mathsf{T} a^2)} \\
\iff & \lambda^\mathsf{T} a^4 - \|d\|^4 > \|d\|^2(\lambda^\mathsf{T} a^2 - \|d\|^2)(1 - \|\rho\|^2 \lambda^\mathsf{T} a^2) \\
\iff & (\lambda^\mathsf{T} a^2 - \|d\|^2)(\lambda^\mathsf{T} a^2 + \|d\|^2) > \|d\|^2(\lambda^\mathsf{T} a^2 - \|d\|^2)(1 - \|\rho\|^2 \lambda^\mathsf{T} a^2) \\
\iff & \lambda^\mathsf{T} a^2 + \|d\|^2 > \|d\|^2 - \|d\|^2 \|\rho\|^2 \lambda^\mathsf{T} a^2 \\
\iff & 1 > -\|d\|^2 \|\rho\|^2.
\end{aligned}$$

The last equivalence is clearly satisfied, concluding the proof.

A.6 Proof of Proposition 7.1

To show Proposition 7.1, we are going to map C to $\hat{C} := T(C)$, where it is easier to find the apex $\hat{\nu}$. The apex of C is then given by $\nu = T^{-1}(\hat{\nu})$.

To compute the apex of $\hat{C}$, we have to intersect it with $\operatorname{lin}(\hat{C})^\perp$, that is, the orthogonal of its lineality space. We have that

$$\hat{C} = \{(\hat{x}, \hat{y}, \hat{z}) \in \mathbb{R}^{|I_+|+|I_-|+|I_0|} \, : \, \|\hat{y}\| \leq \lambda(\hat{x}, 1)\}.$$

Since $\operatorname{lin}(\hat{C}) = \langle\{\lambda_x\}\rangle^\perp \times \{0\}^{|I_-|} \times \mathbb{R}^{|I_0|}$, we have $\operatorname{lin}(\hat{C})^\perp = \langle\{\lambda_x\}\rangle \times \mathbb{R}^{|I_-|} \times \{0\}^{|I_0|}$. Thus,

$$\hat{C} \cap \operatorname{lin}(\hat{C})^\perp = \{(\alpha\lambda_x, \hat{y}, 0) \, : \, \|\hat{y}\| \leq \alpha\|\lambda_x\|^2 + \lambda_{-1}, \alpha \in \mathbb{R}\}.$$

Let us define $\hat{\nu} := (\hat{x}_0, \hat{y}_0, 0)$. Since the apex must satisfy $\hat{y} = 0$ and $\alpha\|\lambda_x\|^2 + \lambda_{-1} = 0$, we get $\hat{x}_0 = -\frac{\lambda_{-1}}{\|\lambda_x\|^2}\lambda_x = -\frac{\hat{x}(\bar{s})}{\|\hat{x}(\bar{s})\|^2}$ and $\hat{y}_0 = 0$.

For the next step to compute ν, we need to find the inverse transformation T^{-1}. Notice that $T(s) = \frac{1}{\sqrt{\kappa}}\hat{\Theta}(V^\mathsf{T} s + \hat{b})$ where $\hat{\Theta} = \operatorname{diag}(\hat{\theta})$ and $\hat{b} := (\hat{b}_i)_{i\in[p]}$ with

$$\hat{\theta}_i = \begin{cases} \sqrt{|\theta_i|}, & \text{if } i \in I_+ \cup I_-, \\ 1, & \text{otherwise.} \end{cases}, \quad \hat{b}_i = \begin{cases} \frac{v_i^\mathsf{T} b}{2\theta_i}, & \text{if } i \in I_+ \cup I_-, \\ 0, & \text{otherwise.} \end{cases}.$$

Since V and $\hat{\Theta}$ are both invertible, the inverse of T is given by $T^{-1}(\hat{x}, \hat{y}, \hat{z}) = V(\kappa\hat{\Theta}^{-1}(\hat{x}, \hat{y}, \hat{z}) - \hat{b})$.

Finally, the apex of C is

$$\nu = T^{-1}(-\frac{\hat{x}(\bar{s})}{\|\hat{x}(\bar{s})\|^2}, 0, 0) = V(\kappa\hat{\Theta}^{-1}(-\frac{\hat{x}(\bar{s})}{\|\hat{x}(\bar{s})\|^2}, 0, 0) - \hat{b}),$$

which concludes the proof.

A.7 Minimal representation of $C \cap H - f$

Here we prove Proposition 5.1, which we repeat for convenience.

Proposition 5.1 *The minimal representation of $C - f$ is*

$$\phi(x, y) = \begin{cases} \sup_{\|\beta\|=1} \frac{\beta^\mathsf{T} y - \lambda^\mathsf{T} x}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} & \text{if } a^\mathsf{T} x + d^\mathsf{T} y = 0 \text{ and } x \in \langle \{\lambda, a\} \rangle \\ +\infty & \text{otherwise.} \end{cases} \tag{5.1}$$

In [4,9], the authors characterize the minimal representation of a *full-dimensional* convex set with the origin in its interior. In [22], the author characterizes minimal representations of arbitrary convex sets. In our setting, we can only apply the characterization of [22]. However, this requires the computation of polars, reverse polars, and co-kernels of $C \cap H - f$, see [22] for the definitions. We want to avoid this, so we present a result the builds on the result of [4] and is enough to prove Proposition 5.1.

Lemma A.7 *Let I be an arbitrary index set and let K be a convex set of the form*

$$K = \{x \in \mathbb{R}^n \,:\, a_i^\mathsf{T} x \leq 1, \forall i \in I, Ax = 0\},$$

such that for each $i \in I$ there exists an $x^i \in K$ with $a_i^\mathsf{T} x^i = 1$, and $\{a_i \,:\, i \in I\}$ is compact.

Then,

$$\phi(x) = \begin{cases} \sup_{i \in I} a_i^\mathsf{T} x & \text{if } x \in \ker(A) \\ +\infty & \text{otherwise} \end{cases}.$$

is the minimal representation of K.

Proof It is clear that $K = \{x \in \mathbb{R}^n \,:\, \phi(x) \leq 1\}$ and that ϕ is sublinear, so ϕ is a representation of K. Let ρ be any representation of K. It remains to show that $\phi(x) \leq \rho(x)$ for all $x \in \mathbb{R}^n$.

Let k be the dimension of $\ker(A)$, B be a basis of $\ker(A)$, and consider the embedding of K, $K_e = \{z \in \mathbb{R}^k \,:\, a_i^\mathsf{T} Bz \leq 1\}$. Note that if $K_e = \{z \in \mathbb{R}^k \,:\, \rho(Bz) \leq 1\}$ since ρ is a representation of K, i.e., $\rho \circ B$ is a representation of K_e.

We proceed to compute the minimal representation of K_e using [4, Theorem 1]. The above theorem says that the minimal representation of K_e is the support function of $\hat{K_e} = \{y \in K_e^* \,:\, z^\mathsf{T} y = 1 \text{ for some } z \in K_e\}$, where $K_e^* = \{y \in \mathbb{R}^k \,:\, z^\mathsf{T} y \leq 1 \text{ for all } z \in K_e\}$, is the polar of K_e.

Given that $\{a_i \,:\, i \in I\}$ is compact, the set $\{(B^\mathsf{T} a_i, 1) \,:\, i \in I\}$ is compact. By [19, Theorem 17.3], we conclude that if $\alpha^\mathsf{T} z \leq \beta$ is valid for K_e and $\alpha \neq 0$, then $\alpha = \sum_{i \in J} h_i B^\mathsf{T} a_i$ and $\beta \geq \sum_{i \in J} h_i$, where $J \subseteq I$ is finite. Since $0 \leq 1$ is also valid for K_e, we conclude that

$$K_e^* = \operatorname{conv}(\{0\} \cup \{B^\mathsf{T} a_i\}_{i \in I}).$$

Note that if $y \in K_e^*$ is such that $y = \sum_{i \in J} h_i B^\mathsf{T} a_i + h_0 0$ with $h_0 > 0$, then $y \notin \hat{K_e}$. Therefore, $\hat{K_e} \subseteq \operatorname{conv}(\{B^\mathsf{T} a_i\}_{i \in I})$. Furthermore, $B^\mathsf{T} a_i \in \hat{K_e}$ since there exists an z^i such that $Bz^i = x^i$ and so $a_i^\mathsf{T} Bz^i = 1$, by hypothesis. Thus, $\{B^\mathsf{T} a_i\}_{i \in I} \subseteq \hat{K_e} \subseteq \operatorname{conv}(\{B^\mathsf{T} a_i\}_{i \in I})$.

Since the support function of a set S is equal to the support function of $\operatorname{conv}(S)$. From the above we conclude that the minimal representation of K_e is

$$\sigma(z) = \sup_{i \in I} a_i^\mathsf{T} Bz.$$

Now we can show that $\phi(x) \leq \rho(x)$ for all $x \in \mathbb{R}^n$ where ρ be a representation of K. Let $x_0 \in \mathbb{R}^n$. If $x_0 \notin \ker(A)$, then $\phi(x_0) = \rho(x_0) = +\infty$. So, let us assume that $x_0 \in \ker(A)$ and let z_0 be such that $Bz_0 = x_0$. As we mentioned at the beginning of the proof, $\rho \circ B$ is a representation of K_e. Since σ is the minimal representation, $\sigma(z_0) \leq \rho(Bz_0)$. The above inequality is equivalent to $\phi(x_0) \leq \rho(x_0)$, which is what we wanted to prove.

Proof (Proposition 5.1) We have that $C - f = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m : \|y + f_y\| \leq \lambda^\mathsf{T} x + f_y, a^\mathsf{T} x + d^\mathsf{T} y = 0\}$ or, equivalently, $C - f = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m : \beta^\mathsf{T}(y + f_y) \leq \lambda^\mathsf{T}(x + f_x)\ \forall \beta \in D^m, a^\mathsf{T} x + d^\mathsf{T} y = 0\}$. Note that

$$\beta^\mathsf{T}(y + f_y) \leq \lambda^\mathsf{T}(x + f_x) \iff \frac{\beta^\mathsf{T} y - \lambda^\mathsf{T} x}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} \leq 1.$$

The equivalence is correct given that $\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y > 0$ since f is in the relative interior of C. Therefore,

$$C - f = \{(x, y) \in \langle\{\lambda, a\}\rangle \times \mathbb{R}^m : \frac{\beta^\mathsf{T} y - \lambda^\mathsf{T} x}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} \leq 1\ \forall \beta \in D^m, a^\mathsf{T} x + d^\mathsf{T} y = 0\}.$$

We are going to show that $\nu - f \in C - f$, where ν is defined by (4.3), and that every inequality is achieved at equality at that point. Since $\nu \in C$, we have $\nu - f \in C - f$. Evaluating the inequality defined by $\beta \in D^m$ at $\nu - f$ yields, $\frac{\beta^\mathsf{T}(-f_y) - \lambda^\mathsf{T}(\nu - f_x)}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} = 1$, where the equality follow from $\lambda^\mathsf{T} \nu = 0$. It remains to prove that $\{\frac{\beta^\mathsf{T} y - \lambda^\mathsf{T} x}{\lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y} : \beta \in D^m\}$ is compact in order to be able to apply Lemma A.7. It is clear that the set is closed. The boundedness of the set follows from the fact that the denominator is away from 0. This last claim can be verified as follows, $\inf_{\beta \in D^m} \lambda^\mathsf{T} f_x - \beta^\mathsf{T} f_y = \lambda^\mathsf{T} f_x - \|f_y\|$ and, by construction, $\lambda = \frac{f_x}{\|f_x\|}$. Thus, $\lambda^\mathsf{T} f_x - \|f_y\| = \|f_x\| - \|f_y\| > 0$ since $f \notin S^g$.

Applying Lemma A.7 finally proves the result.

References

1. Egon Balas. Intersection cuts—a new type of cutting planes for integer programming. *Operations Research*, 19(1):19–39, feb 1971.
2. Egon Balas and Robert G. Jeroslow. Strengthening cuts for mixed integer programs. *European Journal of Operational Research*, 4(4):224–234, apr 1980.
3. Amitabh Basu, Manoel Campelo, Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. Unique lifting of integer variables in minimal inequalities. *Mathematical Programming*, 141(1-2):561–576, jun 2012.
4. Amitabh Basu, Gérard Cornuéjols, and Giacomo Zambelli. Convex sets and minimal sublinear functions. *Journal of Convex Analysis*, 18(2):427–432, 2011.
5. Amitabh Basu, Santanu S. Dey, and Joseph Paat. Nonunique lifting of integer variables in minimal inequalities. *SIAM Journal on Discrete Mathematics*, 33(2):755–783, jan 2019.

6. Ksenia Bestuzheva, Mathieu Besançon, Wei-Kun Chen, Antonia Chmiela, Tim Donkiewicz, Jasper van Doornmalen, Leon Eifler, Oliver Gaul, Gerald Gamrath, Ambros Gleixner, Leona Gottwald, Christoph Graczyk, Katrin Halbig, Alexander Hoen, Christopher Hojny, Rolf van der Hulst, Thorsten Koch, Marco Lübbecke, Stephen J. Maher, Frederic Matter, Erik Mühmer, Benjamin Müller, Marc E. Pfetsch, Daniel Rehfeldt, Steffan Schlein, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Boro Sofranac, Mark Turner, Stefan Vigerske, Fabian Wegscheider, Philipp Wellner, Dieter Weninger, and Jakob Witzig. The SCIP Optimization Suite 8.0. ZIB-Report 21-41, Zuse Institute Berlin, December 2021.
7. Daniel Bienstock, Chen Chen, and Gonzalo Munoz. Outer-product-free sets for polynomial optimization and oracle-based cuts. *Mathematical Programming*, pages 1–44, 2020.
8. Antonia Chmiela, Gonzalo Muñoz, and Felipe Serrano. On the implementation and strengthening of intersection cuts for qcqps. *Mathematical Programming*, pages 1–38, 2022.
9. Michele Conforti, Gérard Cornuéjols, Aris Daniilidis, Claude Lemaréchal, and Jérôme Malick. Cut-generating functions and S-free sets. *Mathematics of Operations Research*, 40(2):276–391, may 2015.
10. Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. A geometric perspective on lifting. *Operations Research*, 59(3):569–577, jun 2011.
11. Santanu S. Dey and Laurence A. Wolsey. Constrained infinite group relaxations of MIPs. *SIAM Journal on Optimization*, 20(6):2890–2912, jan 2010.
12. Santanu S. Dey and Laurence A. Wolsey. Two row mixed-integer cuts via lifting. *Mathematical Programming*, 124(1-2):143–174, may 2010.
13. Ricardo Fukasawa, Laurent Poirrier, and Álinson S. Xavier. The (not so) trivial lifting in two dimensions. *Mathematical Programming Computation*, 11(2):211–235, sep 2018.
14. Fabio Furini, Emiliano Traversi, Pietro Belotti, Antonio Frangioni, Ambros Gleixner, Nick Gould, Leo Liberti, Andrea Lodi, Ruth Misener, Hans Mittelmann, Nikolaos Sahinidis, Stefan Vigerske, and Angelika Wiegele. QPLIB: A library of quadratic programming instances. *Programming Computation*, 2018.
15. Fred Glover. Convexity cuts and cut search. *Operations Research*, 21(1):123–134, feb 1973.
16. Ralph Gomory. An algorithm for the mixed integer problem. Technical report, RAND CORP SANTA MONICA CA, 1960.
17. MINLP library. `http://www.minlplib.org/`, 2010.
18. Gonzalo Muñoz and Felipe Serrano. Maximal quadratic-free sets. *Mathematical Programming*, pages 1–42, 2021.
19. Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 1970.
20. Asteroide Santana and Santanu S. Dey. The convex hull of a quadratic constraint over a polytope. *SIAM Journal on Optimization*, 30(4):2983—-2997, 2020.
21. Hoàng Tuy. Concave programming with linear constraints. In *Doklady Akademii Nauk*, volume 159, pages 32–35. Russian Academy of Sciences, 1964.
22. Alberto Zaffaroni. Convex radiant costarshaped sets and the least sublinear gauge. *Journal of Convex Analysis*, 20(2):307–328, 2013.