On exact copositive representation of simplicial quadratic optimization problems, their strong conic duality and a new proof of the Frank-Wolfe theorem

Immanuel M. Bomze* E. Alper Yıldırım†

3 June 2026

Abstract

We are interested in exactness, strong conic duality and dual attainability in copositive relaxations of quadratic optimization problems (QPs) of a special form, in which any (feasible) QP can be recast. By using our results, the celebrated Frank-Wolfe theorem on the attainability of any bounded QP even over unbounded polyhedra, regardless of whether the objective function is convex or not, follows very easily.

Keywords — Attainability, conic duality, copositive relaxation, Frank-Wolfe theorem, quadratic optimization problems

AMS Subject Classification: 90C20, 90C25, 90C26

*Faculty of Mathematics University of Vienna, Oskar-Morgenstern-Platz 1 1090 Wien, Austria. ORCID ID: 0000-0002-6288-9226. E-mail: `immanuel.bomze@univie.ac.at`

†School of Mathematics and Maxwell Institute of Mathematical Sciences, The University of Edinburgh, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, United Kingdom. ORCID ID: 0000-0003-4141-3189. E-mail: `E.A.Yildirim@ed.ac.uk`

1 Introduction

Consider the following optimization problem:

$$p^* = \inf \{f(\mathbf{x}) : \mathbf{x} \in M\} \ , \tag{P}$$

where (abbreviating the index set $[1\!:\!m] = \{1, \ldots, m\}$)

$$M = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \leq 0 \quad \text{for all } j \in [1\!:\!m]\}$$

and $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function, $g_j : \mathbb{R}^n \to \mathbb{R}$ for each $j \in [1\!:\!m]$, $M \subseteq \mathbb{R}^n$ is the feasible region, and $\mathbf{x} \in \mathbb{R}^n$. If the optimal value denoted by p^* is finite, we say that it is *attained* if there exists $\mathbf{x}^* \in M$ such that $f(\mathbf{x}^*) = p^*$. Any such $\mathbf{x}^* \in M$ is called *an optimal solution.*

The well-known Weierstrass theorem states that the optimal value is attained whenever f is a continuous function and M is a nonempty compact set. On the other hand, it is easy to construct simple examples for which attainability fails if M is an unbounded set, even if f is a convex function and $M \subseteq \mathbb{R}^n$ is a convex set (see Section 4.2).

In optimization problems, the attainability question is an important one for the following reasons. First, attainability implies that an optimal solution of an optimization problem arising from a real-life application, such as radiation therapy or engineering design, can, in fact, be implemented. Second, if the optimal value is finite but not attainable, then there exists a sequence $\mathbf{x}_k \in M$, $k \in \mathbb{N}$ (where $\mathbb{N} = \{1, 2, \ldots\}$ denotes the set of positive integers), such that $f(\mathbf{x}_k) \to p^*$ and $\|\mathbf{x}_k\| \to \infty$ as $k \to \infty$. Therefore, solvers must handle arbitrarily large numbers, which may lead to serious numerical instabilities and unreliable results. Therefore, the understanding of classes of optimization problems (P) for which attainability holds is crucial (see, e.g., [2]).

If f is a linear function and g_j is an affine function for each $j \in [1\!:\!m]$, then M is a polyhedron and (P) is a linear optimization problem. The fundamental theorem of linear optimization states that the optimal value is always attained (see, e.g., [13]). The attainability property has been extended to various more general classes of optimization problems. In a seminal paper, Frank and Wolfe [17] established attainability for quadratic optimization problems (QPs), i.e., instances of (P) where f is a (possibly nonconvex) quadratic function and M is a (possibly unbounded) polyhedron. This attainability result is referred to as the Frank-Wolfe theorem. For a detailed literature review, we refer the reader to Section 4.

If, on the other hand, the problem (P) admits a representation as a conic optimization problem (i.e., minimizing a linear function over the intersection of an affine subset and a closed convex cone – see Section 1.2), then there exist several sufficient conditions in conic duality theory that ensure strong duality and attainability, such as the Slater condition [25] or the closedness condition [4, 33, 38]. Sometimes, these conditions are not needed to establish strong conic duality, including attainability, see, e.g., [19, 36]. In a similar vein, we will employ here a formulation which does not need any additional assumptions apart from mere feasibility.

1.1 Our contribution; organization of the paper

By using only elementary arguments and establishing strong conic duality for a particular family of optimization problems, we lead readers from a more elementary existence result for an optimal solution to a famous one in the context of nonconvex QP. Note that duality theory not only depends on the optimization problem itself but rather on its mathematical description by means of constraints and/or elementary building blocks like extremal elements used in the Minkowski-Weyl theorem (see Proposition 5.1). We will deal with this, and related aspects, in due course, delegating to an appendix some more fundamental and also elementary results, most of which are important on their own right. This should benefit a smooth flow of arguments. By this design, we try to be as modular as possible to enable the use of our approach in different educational contexts.

The paper is organized as follows: in Section 2 we provide some necessary results on irreducibility for matrices on the boundary of the copositive cone, building upon and extending previous ones by Hildebrand and coauthors [14, 18]. Section 3 discusses consequences for simplicial QPs, which consist of optimizing a quadratic function subject to a single linear equality constraint plus a cone constraint. Without any further assumption (in particular without assuming attainability as in the famous paper by Burer [12]), we will prove exactness of the copositive reformulation and strong conic duality. Furthermore, if the cone is polyhedral (i.e., finitely generated), we also prove (not assume) primal attainability for both the original problem and the conic reformulation.

Based upon this chain of arguments, Section 4 presents a new proof of the celebrated Frank-Wolfe theorem on attainability in QPs and also collects some counterexamples. In an appendix in Section 5, we provide all elementary yet fundamental results with their proofs needed in our chain of arguments. By modularity of design, parts of this appendix may be just recalled or cited or skipped, depending on the educational necessities.

1.2 Notation; background on lifting for conic relaxation of QPs

Let $\mathbb{R}^n, \mathbb{R}^n_+$, $\mathbb{R}^{m\times n}$, and $\mathcal{S}^n$ denote the n-dimensional Euclidean space, the nonnegative orthant, the set of $m\times n$ real matrices, and the space of $n\times n$ real symmetric matrices, respectively. We reserve $\mathbf{e}_j,\ j\in[1\!:\!n]$, to denote the unit vectors in $\mathbb{R}^n$. The vector of all ones and the identity matrix are denoted by $\mathbf{e}=\sum_{j=1}^n \mathbf{e}_j$ and $\mathsf{I}_n=\sum_{j=1}^n \mathbf{e}_j\,\mathbf{e}_j^\top \in \mathcal{S}^n$, respectively. We use 0 to denote the real number 0, $\mathbf{o}$ for the vector of all zeroes, as well as O for the matrix of all zeroes, whose dimensions will always be clear from the context. For a vector $\mathbf{x}\in\mathbb{R}^n$, denote by $\mathbf{x}^\top$ its transpose and by $\|\mathbf{x}\|=\sqrt{\mathbf{x}^\top\mathbf{x}}=\sqrt{\sum_{i=1}^n x_i^2}$ its Euclidean norm. Then, the Euclidean distance between $\mathbf{x}\in\mathbb{R}^n$ and $\mathbf{y}\in\mathbb{R}^n$ is given by $d(\mathbf{x},\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|$. For $\mathbf{x}\in\mathbb{R}^n_+$, we define the following index set:

$$\operatorname{supp}(\mathbf{x})=\{j\in[1\!:\!n] : x_j>0\}\ . \tag{1}$$

We reserve boldface lowercase letters, boldface uppercase letters, and uppercase calligraphic letters to denote vectors, matrices, and subsets of $\mathcal{S}^n$, respectively.

Recall that a set $C\subseteq\mathbb{R}^n$ is said to be *convex* if $\mathbf{x}\in C$ and $\mathbf{y}\in C$ imply $(1-\alpha)\,\mathbf{x}+\alpha\,\mathbf{y}=\mathbf{x}+\alpha\,(\mathbf{y}-\mathbf{x})\in C$ whenever $0\le\alpha\le 1$, i.e., C contains all convex combinations of any two elements in C. $C\subseteq\mathbb{R}^n$ is a cone if $\alpha\,\mathbf{x}\in C$ for any $\alpha\ge 0$. C is called *closed* if $\mathbf{z}_k\in C$ for all $k\in\mathbb{N}$ and $\lim\limits_{k\to\infty}\mathbf{z}_k=\bar{\mathbf{z}}$ (or, equivalently, $\mathbf{z}_k\to\bar{\mathbf{z}}$, both meaning $d(\mathbf{z}_k,\bar{\mathbf{z}})\to 0$) implies $\bar{\mathbf{z}}\in C$. A direction $\mathbf{d}\in\mathbb{R}^n$ is called a recession direction of a convex set C if, for each $\mathbf{x}\in C$ and each $\alpha\ge 0$, $\mathbf{x}+\alpha\,\mathbf{d}\in C$. The set of all recession directions is called the recession cone of C. For a given set $D\subseteq\mathbb{R}^n$, we denote by $\operatorname{cone}(D)$ and $\operatorname{conv}(D)$ the set of all nonnegative combinations of the elements in D (i.e., the smallest convex cone with respect to set inclusion that contains D) and the set of all convex combinations of D (i.e., the smallest convex set with respect to set inclusion that contains D), respectively.

For a nonempty closed convex cone $\mathbb{K}\subseteq\mathbb{R}^n$ (resp., a nonempty closed convex set $C\subseteq\mathbb{R}^n$), a nonempty convex cone $\mathbb{F}\subseteq\mathbb{K}$ (resp., a nonempty convex set $F\subseteq C$) is called a *face* of $\mathbb{K}$ (resp., a face of C) if, for all $\mathbf{x}\in\mathbb{K}$ and $\mathbf{y}\in\mathbb{K}$, $\mathbf{x}+\mathbf{y}\in\mathbb{F}$ implies $\mathbf{x}\in\mathbb{F}$ and $\mathbf{y}\in\mathbb{F}$ (resp., for all $\mathbf{x}\in C$ and $\mathbf{y}\in C$, if there exists $\alpha\in(0,1)$ such that $(1-\alpha)\,\mathbf{x}+\alpha\,\mathbf{y}\in F$, then $\mathbf{x}\in F$ and $\mathbf{y}\in F$). A one-dimensional face of a cone $\mathbb{K}$ is called an *extreme ray* and a zero-dimensional face of a convex set C is an *extreme point*. The dual cone of $\mathbb{K}$ is given by

$$\mathbb{K}^* = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y}^\top \mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \mathbb{K}\} . \tag{2}$$

It is obvious that $\mathbb{K}^*$ is again a nonempty closed convex cone.

We are interested in quadratic forms (or functions) over the cone $\mathbb{K}$. A common linearization approach consists of *(Shor) lifting*, which uses the Euclidean geometry on matrix spaces generated by the trace (Frobenius) inner product, denoted by

$$\langle \mathsf{U}, \mathsf{V}\rangle := \text{trace } (\mathsf{U}^\top \mathsf{V}) = \sum_{i=1}^{m} \sum_{j=1}^{n} U_{ij} V_{ij}$$

for any $\mathsf{U} \in \mathbb{R}^{m\times n}$ and $\mathsf{V} \in \mathbb{R}^{m\times n}$, yielding the Frobenius norm $\|\mathsf{U}\|_F = \sqrt{\langle \mathsf{U}, \mathsf{U}\rangle}$ of $\mathsf{U} \in \mathbb{R}^{m\times n}$. This structure is naturally inherited by the linear subspace $\mathcal{S}^n$ of $\mathbb{R}^{n\times n}$. Now, for any $\mathbf{x} \in \mathbb{R}^n$ and $\mathsf{Q} \in \mathcal{S}^n$, note that by circular symmetry of the trace function, we have

$$\mathbf{x}^\top \mathsf{Q}\, \mathbf{x} = \text{trace } (\mathbf{x}^\top \mathsf{Q}\, \mathbf{x}) = \text{trace } (\mathsf{Q}\, \mathbf{x}\, \mathbf{x}^\top) = \langle \mathsf{Q}, \mathbf{x}\, \mathbf{x}^\top\rangle .$$

However, the set $\{\mathbf{y}\, \mathbf{y}^\top : \mathbf{y} \in \mathbb{K}\}$ is typically nonconvex. By convexification, this suggests the consideration of the following matrix cone:

$$\mathcal{CP}(\mathbb{K}) := \text{conv}\left(\{\mathbf{y}\, \mathbf{y}^\top : \mathbf{y} \in \mathbb{K}\}\right) = \left\{ \mathsf{Y} \in \mathcal{S}^n : \mathsf{Y} = \sum_{j=1}^{k} \mathbf{y}_j\, \mathbf{y}_j^\top , \mathbf{y}_j \in \mathbb{K} \text{ for all } j \in [1\!:\!k] \right\} . \tag{3}$$

Note that it suffices to choose $k \leq n(n+1)/2$ by Carathéodory's theorem, cf. Lemma 5.1. In particular, for $\mathbb{K} = \mathbb{R}^n$ and $\mathbb{K} = \mathbb{R}^n_+$, this cone $\mathcal{CP}(\mathbb{K})$ is the cone of positive-semidefinite matrices and (classical) completely positive matrices, respectively. The dual cone of $\mathcal{CP}(\mathbb{K}) \subseteq \mathcal{S}^n$ with respect to (w.r.t.) the Frobenius inner product $\langle\cdot,\cdot\rangle$ is given by

$$\mathcal{COP}(\mathbb{K}) := \{\mathsf{Z} \in \mathcal{S}^n : \langle \mathsf{Z}, \mathsf{Y}\rangle \geq 0 \text{ for all } \mathsf{Y} \in \mathcal{CP}(\mathbb{K})\} , \tag{4}$$

which coincides the set of copositive matrices over $\mathbb{K}$, generating quadratic forms taking no negative values over $\mathbb{K}$. Indeed, it is easy to see by linearity that

$$\mathcal{COP}(\mathbb{K}) = \left\{ \mathsf{Z} \in \mathcal{S}^n : \mathbf{y}^\top \mathsf{Z}\, \mathbf{y} \geq 0 \text{ for all } \mathbf{y} \in \mathbb{K} \right\},$$

and therefore $\mathcal{COP}(\mathbb{K})$ is always a closed[1] convex cone. For $\mathbb{K} = \mathbb{R}^n$ and $\mathbb{K} = \mathbb{R}^n_+$, the cone $\mathcal{COP}(\mathbb{K})$ is the cone of positive-semidefinite matrices and (classical) copositive matrices, respectively. We will also need a subcone of $\mathcal{COP}(\mathbb{K})$, the cone of *copositive-plus* matrices, which is

[1] in the sequel, we will not need the fact [39, Proposition 1, Lemma 1] that also $\mathcal{CP}(\mathbb{K})$ is the dual cone of $\mathcal{COP}(\mathbb{K})$ and therefore as well closed (and convex)

well known in the classical case $\mathbb{K} = \mathbb{R}^n_+$:

$$\mathcal{COP}_+(\mathbb{K}) = \left\{ \mathsf{Y} \in \mathcal{COP}(\mathbb{K}) : \mathsf{Y}\mathbf{u} \in \mathbb{K}^\perp \text{ if } \mathbf{u}^\top \mathsf{Y}\, \mathbf{u} = 0 \text{ and } \mathbf{u} \in \mathbb{K} \right\} , \tag{5}$$

where $\mathbb{K}^\perp := \{\mathbf{v} \in \mathbb{R}^n : \mathbf{u}^\top \mathbf{v} = 0 \text{ for all } \mathbf{u} \in \mathbb{K}\}$ denotes the orthogonal complement of $\mathbb{K}$. Note, in particular, that if $\mathbb{K}$ is a solid cone (i.e., it has a nonempty interior), then $\mathbb{K}^\perp = \{\mathbf{o}\}$. The classical and more familiar case $\mathbb{K} = \mathbb{R}^n_+$ is retrieved by observing $(\mathbb{R}^n_+)^\perp = \{\mathbf{o}\}$ since $\mathbb{R}^n_+$ is a solid cone. For more recent results on $\mathcal{COP}_+(\mathbb{R}^n_+)$, see [10].

2 Irreducibility results

In this section, we focus on some technical results for matrices A located at the boundary of the copositive cone $\mathcal{COP}(\mathbb{K})$ given by (4). Recall that such matrices A satisfy $\mathbf{x}^\top \mathsf{A}\, \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{K}$ with equality $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$ for some $\mathbf{u} \in \mathbb{K} \setminus \{\mathbf{o}\}$, which are then nontrivial ($\mathbf{u} \neq \mathbf{o}$) global minimizers of the quadratic optimization problem $\min\{\mathbf{x}^\top \mathsf{A}\, \mathbf{x} : \mathbf{x} \in \mathbb{K}\}$.

Indeed, any $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$ satisfying the above property must be on the boundary, looking at $\mathsf{A} - \varepsilon\, \mathsf{I}_n$ as $\varepsilon \searrow 0$. The converse argument uses the compact base $C = \{\mathbf{x} \in \mathbb{K} : \|\mathbf{x}\| = 1\}$ for $\mathbb{K}$. If $\rho := \min\{\mathbf{x}^\top \mathsf{A} \mathbf{x} : \mathbf{x} \in C\} > 0$, then the Frobenius ball $\{\mathsf{B} \in \mathcal{S}^n : \|\mathsf{B} - \mathsf{A}\|_F \leq \rho\} \subset \mathcal{COP}(\mathbb{K})$, as is easily seen applying the Cauchy-Schwarz inequality and the identity $\|\mathbf{x}\,\mathbf{x}^\top\|_F = \|\mathbf{x}\|^2 = 1$ on C:

$$\mathbf{x}^\top \mathsf{B}\, \mathbf{x} = \mathbf{x}^\top (\mathsf{B} - \mathsf{A})\, \mathbf{x} + \mathbf{x}^\top \mathsf{A}\, \mathbf{x} \geq \langle \mathsf{B} - \mathsf{A}, \mathbf{x}\,\mathbf{x}^\top \rangle + \rho \geq \rho - \|\mathsf{B} - \mathsf{A}\|_F \|\mathbf{x}\,\mathbf{x}^\top\|_F \geq 0 \text{ for all } \mathbf{x} \in C\,.$$

Now positive-homogeneity of the quadratic forms implies $\mathsf{B} \in \mathcal{COP}(\mathbb{K})$.

The following results are an extension of findings by Hildebrand [18], which will play a key role in establishing our main result. First, for a nonempty closed convex cone $\mathbb{K} \subseteq \mathbb{R}^n$, we establish a useful property of $\mathcal{COP}(\mathbb{K})$ given by (4).

Lemma 2.1. *Let $\mathbb{K} \subseteq \mathbb{R}^n$ be a nonempty closed convex cone. If $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$ and $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$ for some $\mathbf{u} \in \mathbb{K}$, then $\mathsf{A}\, \mathbf{u} \in \mathbb{K}^*$, where $\mathbb{K}^*$ is defined as in* (2)*.*

Proof. Let $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$ and $\mathbf{u} \in \mathbb{K}$ be such that $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$. For an arbitrary $\mathbf{v} \in \mathbb{K}$, let $\phi : \mathbb{R} \to \mathbb{R}$ be given by

$$\phi(\alpha) := \tfrac{1}{2} \left(\mathbf{u} + \alpha\,(\mathbf{v} - \mathbf{u})\right)^\top \mathsf{A} \left(\mathbf{u} + \alpha\,(\mathbf{v} - \mathbf{u})\right) .$$

Note that ϕ is a quadratic function with its first derivative given by

$$\phi'(\alpha) = \alpha\,(\mathbf{v}-\mathbf{u})^\top \mathsf{A}\,(\mathbf{v}-\mathbf{u}) + \,(\mathbf{v}-\mathbf{u})^\top \mathsf{A}\,\mathbf{u}\,.$$

Since $\mathbb{K}$ is convex and $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$, we have $\phi(\alpha) \geq 0$ for each $\alpha \in [0,1]$. By the hypothesis, $\phi(0) = 0$. It follows that

$$\phi'(0) = (\mathbf{v}-\mathbf{u})^\top \mathsf{A}\,\mathbf{u} \geq 0,$$

which, together with $\mathbf{u}^\top \mathsf{A}\,\mathbf{u} = 0$, implies that $\mathbf{v}^\top \mathsf{A}\,\mathbf{u} \geq 0$. Since $\mathbf{v} \in \mathbb{K}$ is arbitrary, the assertion follows from (2). □

Our next result is an immediate corollary of Lemma 2.1, which seems to be common knowledge by now, see, e.g. [14].

Corollary 2.1. *Let $\mathbb{K}$ be either of the choices in (i) or (ii) below. Suppose that $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$ and that $\mathbf{u}^\top \mathsf{A}\,\mathbf{u} = 0$ for some $\mathbf{u} \in \mathbb{K}$.*

(i) If $\mathbb{K} = \mathbb{R}^n$, then $\mathsf{A}\,\mathbf{u} = \mathbf{o}$.

(ii) If $\mathbb{K} = \mathbb{R}^n_+$, then

$$\begin{aligned} \mathbf{e}_j^\top \mathsf{A}\,\mathbf{u} &= 0, \quad \textit{if } j \in \operatorname{supp}(\mathbf{u})\,, \\ \mathbf{e}_j^\top \mathsf{A}\,\mathbf{u} &\geq 0, \quad \textit{if } j \notin \operatorname{supp}(\mathbf{u})\,. \end{aligned}$$

Proof. (i) If $\mathbb{K} = \mathbb{R}^n$, then $\mathbb{K}^* = \{\mathbf{o}\}$. The assertion follows from Lemma 2.1.

(ii) If $\mathbb{K} = \mathbb{R}^n_+$, then $\mathbb{K}^* = \mathbb{K} = \mathbb{R}^n_+$. By Lemma 2.1, $\mathsf{A}\,\mathbf{u} \in \mathbb{R}^n_+$, which implies that $\mathbf{e}_j^\top \mathsf{A}\,\mathbf{u} \geq 0$ for each $j \in [1\!:\!n]$. The assertion follows by observing that

$$\mathbf{u}^\top \mathsf{A}\,\mathbf{u} = \sum_{j=1}^{n} u_j\, \mathbf{e}_j^\top \mathsf{A}\,\mathbf{u} = \sum_{j: u_j > 0} u_j\, \mathbf{e}_j^\top \mathsf{A}\,\mathbf{u} = 0\,.$$

□

We next present an extension of a helpful result, which generalizes a lemma first established in [18]. This result relates irreducibility of a matrix (in the sense described in the following lemma) on the boundary of $\mathcal{COP}(\mathbb{K})$ with respect to another matrix in $\mathcal{COP}_+(\mathbb{K})$, for finitely generated cones $\mathbb{K}$, which are defined as

$$\mathbb{K} = \operatorname{cone}(\{\mathbf{d}_1, \ldots, \mathbf{d}_r\}) = \{\mathsf{D}\,\mathbf{x} : \mathbf{x} \in \mathbb{R}^r_+\} \text{ where } \mathsf{D} = \begin{bmatrix}\mathbf{d}_1 & \cdots & \mathbf{d}_r\end{bmatrix} \in \mathbb{R}^{n\times r}\,, \tag{6}$$

for some $\{\mathbf{d}_1, \ldots, \mathbf{d}_r\} \subset \mathbb{R}^n$. Note that by the Minkowski-Weyl Theorem, any polyhedral cone $\mathbb{K} = \{\mathbf{u} \in \mathbb{R}^n : \mathsf{A}\,\mathbf{u} \in \mathbb{R}^m_+\}$ can be represented in the form (6), see Proposition 5.1 in the appendix.

Lemma 2.2. *Let* $\mathbb{K}$ *be given as in* (6) *and* $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$. *Further, let* $\mathsf{B} \in \mathcal{COP}_+(\mathbb{K})$. *Then,* $\mathsf{A} - \varepsilon\,\mathsf{B} \notin \mathcal{COP}(\mathbb{K})$ *for each* $\varepsilon > 0$ *if and only if there exists* $\mathbf{u} \in \mathbb{K}$ *be such that* $\mathbf{u}^\top \mathsf{A}\,\mathbf{u} = 0$ *and* $\mathsf{B}\,\mathbf{u} \notin \mathbb{K}^\perp$.

Proof. $\Longleftarrow$: Suppose that there exists $\mathbf{u} \in \mathbb{K}$ be such that $\mathbf{u}^\top \mathsf{A}\,\mathbf{u} = 0$ and $\mathsf{B}\,\mathbf{u} \notin \mathbb{K}^\perp$. Then, for each $\varepsilon > 0$, we have

$$\mathbf{u}^\top \left(\mathsf{A} - \varepsilon\,\mathsf{B}\right)\mathbf{u} = -\varepsilon\,\mathbf{u}^\top \mathsf{B}\,\mathbf{u} < 0\,,$$

where the inequality is inferred from $\mathsf{B} \in \mathcal{COP}_+(\mathbb{K})$ and $\mathsf{B}\,\mathbf{u} \notin \mathbb{K}^\perp$.
$\Longrightarrow$: Let $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$ and $\mathsf{B} \in \mathcal{COP}_+(\mathbb{K})$ be such that $\mathsf{A} - \varepsilon\,\mathsf{B} \notin \mathcal{COP}(\mathbb{K})$ for each $\varepsilon > 0$. Let D be as in (6) and define the sequence of Standard Quadratic optimization problems

$$\mu_k := \min\left\{\tfrac{1}{2}\mathbf{x}^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right)\mathsf{D}\,\mathbf{x} : \mathbf{e}^\top \mathbf{x} = 1, \quad \mathbf{x} \in \mathbb{R}^r_+\right\}, \quad k \in \mathbb{N}\,. \tag{P_k}$$

Since the feasible region is compact and the objective function is continuous, the optimal value is attained. Let $\mathbf{x}_k \in \mathbb{R}^r_+$ denote an optimal solution of (P_k). By the hypothesis,

$$\mathbf{x}_k^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right)\mathsf{D}\,\mathbf{x}_k = 2\,\mu_k < 0, \quad k = 1, 2, \ldots. \tag{7}$$

By the compactness of the feasible region, and taking a subsequence if necessary, we can assume that $\mathbf{x}_k \to \mathbf{x}$. It follows from (7), putting $\mathbf{u} := \mathsf{D}\,\mathbf{x}$ and $\mathbf{u}_k := \mathsf{D}\,\mathbf{x}_k$ (so that $\{\mathbf{u}, \mathbf{u}_k\} \subset \mathbb{K}$ and $\mathbf{u}_k \to \mathbf{u}$), that

$$0 \geq \lim_{k\to\infty} \mathbf{u}_k^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right)\,\mathbf{u}_k = \mathbf{u}^\top \mathsf{A}\,\mathbf{u}\,.$$

Since $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$ and $\mathbf{u} \in \mathbb{K}$, we conclude that

$$\mathbf{u}^\top \mathsf{A}\,\mathbf{u} = 0\,. \tag{8}$$

Since $\mathbf{x}_k \to \mathbf{x}$, we have $\operatorname{supp}(\mathbf{x}) \subseteq \operatorname{supp}(\mathbf{x}_k)$ for all sufficiently large $k \in \mathbb{N}$. Let us fix such an index k. Consider the following one-dimensional quadratic function $\rho : \mathbb{R} \to \mathbb{R}$ given by

$$\rho(\alpha) = \tfrac{1}{2}\left[\mathbf{x}_k + \alpha(\mathbf{x} - \mathbf{x}_k)\right]^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\mathsf{B}\right)\mathsf{D}\,\left[\mathbf{x}_k + \alpha(\mathbf{x} - \mathbf{x}_k)\right]\,.$$

Since each of $\mathbf{x}$ and $\mathbf{x}_k$ is a feasible solution of (P_k), it follows from the convexity of the feasible region that $\alpha^* = 0$ is a global optimal solution of ρ over the interval $[0, 1]$. Therefore

$$\mu_k = \rho(0) \leq \rho(1) = -\tfrac{1}{2k}\,\mathbf{u}^\top \mathsf{B}\,\mathbf{u}\,,$$

where we used (8). If $\rho(0) = \rho(1)$, then $0 > \mu_k = -\frac{1}{2k}\,\mathbf{u}^\top \mathsf{B}\,\mathbf{u}$, which implies that $\mathsf{B}\,\mathbf{u} \notin \mathbb{K}^\perp$.

Let us suppose now that $\rho(0) < \rho(1)$. The first and second derivatives of ρ are given by

$$\begin{aligned}\rho'(\alpha) &= \alpha\,(\mathbf{x} - \mathbf{x}_k)^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,(\mathbf{x} - \mathbf{x}_k) + (\mathbf{x} - \mathbf{x}_k)^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,\mathbf{x}_k,\\ \rho''(\alpha) &= (\mathbf{x} - \mathbf{x}_k)^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,(\mathbf{x} - \mathbf{x}_k)\,.\end{aligned}$$

By Proposition 3.1 below or [9, Theorem 7], we have $\mathsf{D}^\top \left(\mathsf{A} - \frac{1}{k}\,\mathsf{B}\right) \mathsf{D} - 2\,\mu_k\,\mathbf{e}\,\mathbf{e}^\top \in \mathcal{COP}(\mathbb{R}^r_+)$. Since

$$\mathbf{x}_k^\top \left(\mathsf{D}^\top \left(\mathsf{A} - \frac{1}{k}\,\mathsf{B}\right) \mathsf{D} - 2\,\mu_k\,\mathbf{e}\,\mathbf{e}^\top\right) \mathbf{x}_k = 0\,,$$

Corollary 2.1(ii) and $\operatorname{supp}(\mathbf{x}) \subseteq \operatorname{supp}(\mathbf{x}_k)$ ensure that, for all $j \in [1\!:\!n]$, the following implication holds:

$$\mathbf{e}_j^\top \mathbf{x} > 0 \Longrightarrow \mathbf{e}_j^\top \mathbf{x}_k > 0 \Longrightarrow \mathbf{e}_j^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,\mathbf{x}_k = 2\,\mu_k\,. \tag{9}$$

Therefore,

$$\begin{aligned}\rho'(0) &= (\mathbf{x} - \mathbf{x}_k)^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,\mathbf{x}_k\\ &= \mathbf{x}^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,\mathbf{x}_k - 2\,\mu_k\\ &= \sum_{j:x_j>0} x_j\,\mathbf{e}_j^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,\mathbf{x}_k - 2\,\mu_k\\ &= 0\,,\end{aligned}$$

where we used (7) in the second line and (9) together with $\mathbf{e}^\top \mathbf{x} = 1$ in the third line. Since $\rho(0) < \rho(1)$, $\rho'(0) = 0$, and ρ is a quadratic function, we conclude that $\rho''(0) > 0$. Then,

$$\begin{aligned}0 &< \rho''(0)\\ &= (\mathbf{x} - \mathbf{x}_k)^\top \mathsf{D}^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathsf{D}\,(\mathbf{x} - \mathbf{x}_k)\\ &= (\mathbf{u} - \mathbf{u}_k)^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) (\mathbf{u} - \mathbf{u}_k)\\ &= (\mathbf{u} - \mathbf{u}_k)^\top \left(\mathsf{A} - \tfrac{1}{k}\,\mathsf{B}\right) \mathbf{u}\\ &= \mathbf{u}^\top \mathsf{A}\,\mathbf{u} - \tfrac{1}{k}\,\mathbf{u}^\top \mathsf{B}\,\mathbf{u} - (\mathbf{u}_k)^\top \mathsf{A}\mathbf{u} + \tfrac{1}{k}\,\mathbf{u}_k^\top \mathsf{B}\,\mathbf{u}\\ &= -\tfrac{1}{k}\,\mathbf{u}^\top \mathsf{B}\,\mathbf{u} - (\mathbf{u}_k)^\top \mathsf{A}\mathbf{u} + \tfrac{1}{k}\,\mathbf{u}_k^\top \mathsf{B}\,\mathbf{u}\,,\end{aligned}$$

where we used $(\mathbf{u} - \mathbf{u}_k)^\top \left(\mathsf{A} - \frac{1}{k}\,\mathsf{B}\right) \mathbf{u}_k = \rho'(0) = 0$ in the fourth line and (8) in the last line. Note that $\mathsf{A}\,\mathbf{u} \in \mathbb{K}^*$ by (8) and Lemma 2.1. Since $\mathbf{u} \in \mathbb{K}$, $\mathsf{B} \in \mathcal{COP}_+(\mathbb{K})$, and $\mathbf{u}_k \in \mathbb{K}$, we conclude that the first two terms in the last line are nonpositive. Therefore, we obtain $(\mathbf{u}_k)^\top \mathsf{B}\,\mathbf{u} > 0$, which once again implies that $\mathsf{B}\,\mathbf{u} \notin \mathbb{K}^\perp$. Together with (8), this completes the proof. □

We now specialize Lemma 2.2 to a case needed later on:

Corollary 2.2. *Let* $\mathbb{K} \subseteq \mathbb{R}^n$ *be given by* (6), $\mathsf{A} \in \mathcal{S}^n$, *and* $\mathbf{a} \in \mathbb{R}^n$. *Then,* $\mathsf{A} - \varepsilon\, \mathbf{a}\,\mathbf{a}^\top \notin \mathcal{COP}(\mathbb{K})$ *for each* $\varepsilon > 0$ *if and only if there exists* $\mathbf{u} \in \mathbb{K}$ *be such that* $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$ *and* $\mathbf{a}^\top \mathbf{u} \neq 0$.

Proof. This is a special case of Lemma 2.2 for $\mathsf{B} = \mathbf{a}\,\mathbf{a}^\top$, which is positive-semidefinite and hence in $\mathcal{COP}_+(\mathbb{K})$ by Corollary 2.1(i). Indeed, if $\mathbf{u}^\top \mathsf{B}\, \mathbf{u} = (\mathbf{a}^\top \mathbf{u})^2 = 0$, then $\mathbf{a}^\top \mathbf{u} = 0$ and therefore $\mathsf{B}\, \mathbf{u} = (\mathbf{a}^\top \mathbf{u})\, \mathbf{a} = \mathbf{o} \in \mathbb{K}^\perp$. Likewise we conclude from $\mathsf{B}\, \mathbf{u} \notin \mathbb{K}^\perp$ that necessarily ($\mathbf{a} \neq \mathbf{o}$ and) $\mathbf{a}^\top \mathbf{u} \neq 0$. Hence the result. $\square$

A further specialization yields Hildebrand's result:

Corollary 2.3. *[18, Lemma 4.3] Let* $\mathsf{A} \in \mathcal{COP}(\mathbb{R}^n_+)$, *and* $\mathbf{a} \in \mathbb{R}^n$. *Then,* $\mathsf{A} - \varepsilon\, \mathbf{a}\,\mathbf{a}^\top \notin \mathcal{COP}(\mathbb{R}^n_+)$ *for each* $\varepsilon > 0$ *if and only if there exists* $\mathbf{u} \in \mathbb{R}^n_+$ *be such that* $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$ *and* $\mathbf{a}^\top \mathbf{u} \neq 0$.

Proof. Immediate from Corollary 2.2 for the choice $\mathbb{K} = \mathbb{R}^n_+ = \operatorname{cone}(\{\mathbf{e}_1, \ldots, \mathbf{e}_n\})$. $\square$

Corollary 2.2 extends Corollary 2.3 from the nonnegative orthant $\mathbb{R}^n_+$ to any finitely generated cone $\mathbb{K}$. For instance, since $\mathbb{R}^n = \operatorname{cone}\left(\left\{\mathbf{e}_1, \ldots, \mathbf{e}_n, -\sum\limits_{j=1}^{n} \mathbf{e}_j\right\}\right)$ and $\mathcal{COP}(\mathbb{R}^n) = \mathcal{COP}_+(\mathbb{R}^n)$ — see Corollary 2.1(i) — is the cone of positive-semidefinite matrices, Lemma 2.2 and thus also Corollary 2.2 provides a result on irreducibility for the semidefinite cone.

Remark 1. *An interesting question is whether Lemma 2.2 can be extended to any nonempty closed convex cone* $\mathbb{K}$. *We illustrate by an example that the answer, in general, is negative. Let* $\mathbb{K} \subseteq \mathbb{R}^3$ *denote the second-order cone given by*

$$\mathbb{K} = \left\{ [x_1, x_2, x_3]^\top \in \mathbb{R}^3 : x_1 \geq \sqrt{x_2^2 + x_3^2} \right\}. \tag{10}$$

The faces of $\mathbb{K}$ *are given by* $\{\mathbf{o}\} \subseteq \mathbb{R}^3$, *its extreme rays generated by the elements on the boundary (i.e.,* $\mathbf{x} = [x_1, x_2, x_3]^\top \in \mathbb{R}^3$ *such that* $x_1 = \sqrt{x_2^2 + x_3^2} > 0$*), and* $\mathbb{K}$ *itself (see, e.g., [32]), which implies that* $\mathbb{K}$ *is not a polyhedral cone since it is not finitely generated. Let* $\mathsf{A} \in \mathcal{S}^3$ *and* $\mathbf{a} \in \mathbb{R}^3$ *be given by*

$$\mathsf{A} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{11}$$

First, we argue that $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$. *Indeed, for any* $\mathbf{x} \in \mathbb{K}$, *we have*

$$\mathbf{x}^\top \mathsf{A}\, \mathbf{x} = 2\, x_1^2 - x_2^2 + 2\, x_1\, x_3 = (x_1 + x_3)^2 + x_1^2 - x_2^2 - x_3^2 \geq 0,$$

where we used $\mathbf{x} \in \mathbb{K}$ *to derive the inequality. It follows that* $\mathsf{A} \in \mathcal{COP}(\mathbb{K})$*. Furthermore,* $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$ *and* $\mathbf{u} \in \mathbb{K}$ *if and only if* $\mathbf{u} \in \{[\alpha, 0, -\alpha]^\top : \alpha \geq 0\}$*. Next, consider* $\mathsf{A} - \varepsilon\, \mathbf{a}\, \mathbf{a}^\top$*, where* $\varepsilon > 0$*. Let* $\mathbf{u}(\varepsilon) := [\sqrt{1+\varepsilon}, \sqrt{\varepsilon}, -1]^\top \in \mathbb{R}^3$*. For any* $\varepsilon > 0$*, we have* $\mathbf{u}(\varepsilon) \in \mathbb{K}$*. Furthermore, let* $\phi : \mathbb{R} \to \mathbb{R}$ *be given by*

$$\phi(\varepsilon) = \mathbf{u}(\varepsilon)^\top \left(\mathsf{A} - \varepsilon\, \mathbf{a}\, \mathbf{a}^\top\right) \mathbf{u}(\varepsilon) = 2 + \varepsilon - \varepsilon^2 - 2\sqrt{1+\varepsilon}.$$

We claim that $\phi(\varepsilon) < 0$ *for each* $\varepsilon > 0$*. To that end, we have*

$$\begin{aligned} \phi'(\varepsilon) &= 1 - 2\varepsilon - \frac{1}{\sqrt{1+\varepsilon}}, \\ \phi''(\varepsilon) &= -2 + \frac{1}{2(1+\varepsilon)^{3/2}}. \end{aligned}$$

We have $\phi(0) = 0$*,* $\phi'(0) = 0$*, and* $\phi''(\varepsilon) < 0$ *for each* $\varepsilon > 0$*, which implies that* ϕ *is strictly concave on* $(0, +\infty)$*. We therefore conclude that* $\phi(\varepsilon) < 0$ *for each* $\varepsilon > 0$*. Since* $\mathbf{u}(\varepsilon) \in \mathbb{K}$*, it follows that* $\mathsf{A} - \varepsilon\, \mathbf{a}\, \mathbf{a}^\top \notin \mathcal{COP}(\mathbb{K})$ *for each* $\varepsilon > 0$*. On the other hand, note that* $\mathbf{a}^\top \mathbf{u} = 0$ *for any* $\mathbf{u} \in \mathbb{K}$ *such that* $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$*. Therefore, we conclude that Lemma 2.2 with* $\mathsf{B} := \mathbf{a}\, \mathbf{a}^\top$ *(cf. Corollary 2.2) does not hold in this setting.*

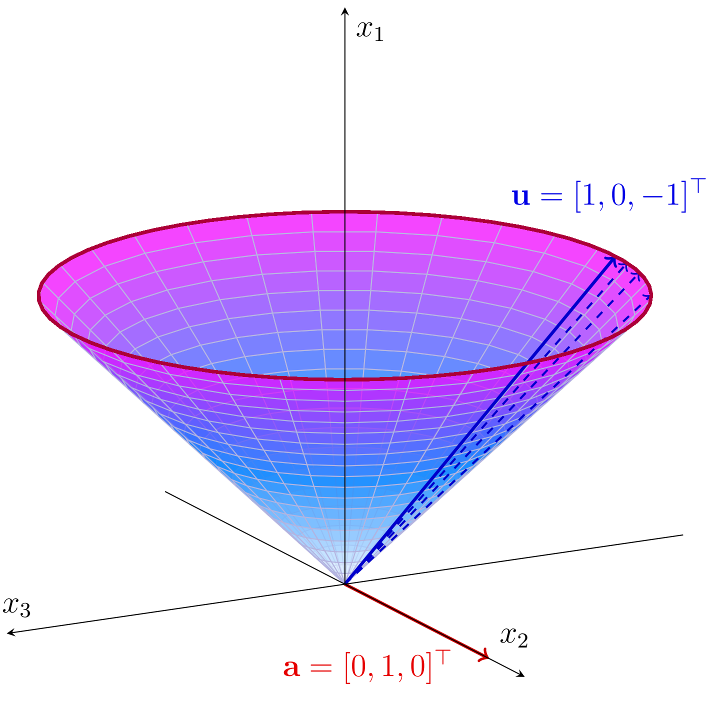

Figure 1: The three-dimensional second-order cone and the vectors $\mathbf{a}$ (in red) and $\mathbf{u}$ (in blue) in the example in Remark 1 (the dashed arrows illustrate $\mathbf{u}(\varepsilon)$ as $\varepsilon \to 0$)

It is insightful to highlight where the proof of Lemma 2.2 breaks down in this example. The proof exploits the special structure of the nonnegative orthant and relies on the following crucial observation: whenever $\mathbf{x}_k \in \mathbb{R}^n_+$ *for each* $k \in \mathbb{N}$ *and* $\mathbf{x}_k \to \mathbf{x}$*, then, for all sufficiently large* k*, the smallest (w.r.t set inclusion) face of the nonnegative orthant that contains* $\mathbf{x}$ *is contained in the smallest face that contains* $\mathbf{x}_k$*. In this example, note that* $\mathbf{u}(\varepsilon)$ *is an extreme ray for each* $\varepsilon > 0$*,* $\mathbf{u}(\varepsilon) \to \mathbf{u} = [1, 0, -1]^\top$ *as* $\varepsilon \to 0$*, and* $\mathbf{u}$ *is also an extreme ray. On the other hand, since* $\mathbf{u}(\varepsilon)$ *is not a multiple of* $\mathbf{u}$ *for any* $\varepsilon > 0$*, it follows that* $\mathbf{u}$ *is not contained in the smallest face of* $\mathbb{K}$ *that contains* $\mathbf{u}(\varepsilon)$ *(please see Figure 1). The reader is referred to [29] for facial structures of* $\mathcal{COP}(\mathbb{K})$ *and* $\mathcal{CP}(\mathbb{K})$*.*

3 Implications on quadratic optimization problems

3.1 Exact copositive reformulation and strong conic duality – simplicial case

In a famous paper, Burer [12] proved that all quadratic optimization problems admit an exact copositive reformulation. However, in that paper, it is assumed that the primal conic problem has an optimal solution, and this may not be automatically granted as the conic dual need not be strictly feasible, i.e., the Slater condition may not hold. We will proceed without this *a priori* assumption (actually the main focus of our paper is to prove attainability, not assuming it from the outset). Note that this strong conic duality result was proved recently in [19] as well, with a slightly different scope.

In [12], Burer built upon (and cited) an earlier result [11] on Standard Quadratic Optimization. The arguments used in [11] can be easily and directly adapted to the following class of optimization problems, the so-called *simplicial conic quadratic optimization problems,* which obviously are always feasible:[2]

$$p^* := \inf \left\{ f(\mathbf{y}) := \tfrac{1}{2}\, \mathbf{y}^\top \mathsf{F}\, \mathbf{y} + \mathbf{f}^\top \mathbf{y} : \mathbf{a}^\top \mathbf{y} = 1\,,\, \mathbf{y} \in \mathbb{K} \right\}, \tag{SCQP}$$

where $\mathbb{K} \subseteq \mathbb{R}^n$ is a closed convex cone, $\mathsf{F} \in \mathcal{S}^n$, $\mathbf{f} \in \mathbb{R}^n$, and $\mathbf{a} \in \mathbb{K}^* \backslash \mathbb{K}^\perp$.

The following conic primal-dual pair gives a relaxation of (SCQP):

$$\mu := \inf \left\{ \langle \mathsf{G}, \mathsf{Y} \rangle : \langle \mathbf{a}\,\mathbf{a}^\top, \mathsf{Y} \rangle = 1, \quad \mathsf{Y} \in \mathcal{CP}(\mathbb{K}) \right\}, \tag{CP}$$

[2] observe that the feasible set can be unbounded if the point $\mathbf{a}$ lies on the boundary of $\mathbb{K}^*$

where $\mathsf{G} := \frac{1}{2}\left(\mathsf{F} + \mathbf{a}\,\mathbf{f}^\top + \mathbf{f}\,\mathbf{a}^\top\right) \in \mathcal{S}^n$. Indeed, consider the rank-one matrices $\mathsf{Y} = \mathbf{y}\,\mathbf{y}^\top$ with (SCQP)-feasible $\mathbf{y}$; then $\langle \mathbf{a}\,\mathbf{a}^\top, \mathsf{Y}\rangle = (\mathbf{a}^\top\mathbf{y})^2 = 1$ and $\langle \mathsf{G}, \mathsf{Y}\rangle = \mathbf{y}^\top \mathsf{G}\,\mathbf{y} = f(\mathbf{y})$. Therefore also problem (CP) is always feasible.

Consider the conic dual[3] of (CP) given by

$$\nu := \sup\left\{\sigma : \mathsf{G} - \sigma\,\mathbf{a}\,\mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})\right\} . \tag{CD}$$

By feasibility, weak duality and relaxation, we would, by general duality theory, obtain the inequality chain

$$\nu \le \mu \le p^* < +\infty \,. \tag{12}$$

However, we will provide a much simpler direct argument now to establish this relation:

Proposition 3.1. *Consider p^* from* (SCQP)*, μ from* (CP) *and ν from* (CD)*. Then* (12) *holds. Moreover, if $p^* > -\infty$, then $\mathsf{G} - \sigma\,\mathbf{a}\,\mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})$ holds if and only if $\sigma \le p^*$.*

Proof. If $\mathsf{Z} := \mathsf{G} - \sigma\,\mathbf{a}\,\mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})$ and $\mathsf{Y} \in \mathcal{CP}(\mathbb{K})$ is feasible with $\langle \mathbf{a}\,\mathbf{a}^\top, \mathsf{Y}\rangle = 1$, we get by definition of dual cones

$$0 \le \langle \mathsf{Z}, \mathsf{Y}\rangle = \langle \mathsf{G}, \mathsf{Y}\rangle - \sigma\langle \mathbf{a}\,\mathbf{a}^\top, \mathsf{Y}\rangle = \langle \mathsf{G}, \mathsf{Y}\rangle - \sigma$$

or equivalently $\sigma \le \langle \mathsf{G}, \mathsf{Y}\rangle$, which shows $\nu \le \mu$. We already argued for $\mathsf{Y} = \mathbf{y}\,\mathbf{y}^\top$ with $\mathbf{y} \in \mathbb{K}$ and $\mathbf{a}^\top\mathbf{y} = 1$ that

$$\mu \le \langle \mathsf{G}, \mathsf{Y}\rangle = f(\mathbf{y})\,,$$

so $\mu \le p^*$ follows as well. So relation (12) holds.

Next, assume that $p^* \in \mathbb{R}$ is finite. We now show the only missing assertion, namely that $\sigma \le p^*$ implies $\mathsf{G} - \sigma\,\mathbf{a}\,\mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})$. Take any $\mathbf{x} \in \mathbb{K}$. If $\mathbf{a}^\top\mathbf{x} > 0$, then $\mathbf{y} := \frac{1}{\mathbf{a}^\top\mathbf{x}}\,\mathbf{x} \in \mathbb{K}$ is (SCQP)-feasible, and thus

$$\mathbf{y}^\top \mathsf{G}\,\mathbf{y} = f(\mathbf{y}) \ge p^* \ge \sigma = \sigma(\mathbf{a}^\top\mathbf{y})^2$$

which implies

$$\mathbf{x}^\top(\mathsf{G} - \sigma\,\mathbf{a}\,\mathbf{a}^\top)\,\mathbf{x} = (\mathbf{a}^\top\mathbf{x})^2\,\mathbf{y}^\top(\mathsf{G} - \sigma\,\mathbf{a}\,\mathbf{a}^\top)\,\mathbf{y} \ge 0\,.$$

If however $\mathbf{a}^\top\mathbf{x} = 0$, then feasibility and boundedness of (SCQP) allow us to argue as follows: take any (SCQP)-feasible $\mathbf{y}$ and form $\mathbf{y}(\alpha) = \mathbf{y} + \alpha\,\mathbf{x}$, which is (SCQP)-feasible as well for arbitrary $\alpha \ge 0$. We conclude

$$-\infty < p^* \le f(\mathbf{y}(\alpha)) = \langle \mathsf{G}, (\mathbf{y} + \alpha\,\mathbf{x})(\mathbf{y} + \alpha\,\mathbf{x})^\top\rangle = f(\mathbf{y}) + 2\,\alpha\,\mathbf{x}^\top\mathsf{G}\mathbf{y} + \alpha^2\,\mathbf{x}^\top\mathsf{G}\,\mathbf{x} \quad \text{for all } \alpha > 0\,,$$

[3]we do not need any Lagrangian duality here and rather start from scratch with the definition (CD).

which implies $\mathbf{x}^\top \mathsf{G}\,\mathbf{x} \geq 0$, or $\mathsf{G} \in \mathcal{COP}(\mathbb{K} \cap \mathbf{a}^\perp)$, and therefore also in this case

$$\mathbf{x}^\top (\mathsf{G} - \sigma\, \mathbf{a}\, \mathbf{a}^\top)\, \mathbf{x} = \mathbf{x}^\top \mathsf{G}\, \mathbf{x} \geq 0\,,$$

so that we arrive at $\mathsf{G} - \sigma\, \mathbf{a}\, \mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})$ for all $\sigma \leq p^*$, as claimed. □

We now prove that always $\nu = \mu = p^*$ holds, so that both the relaxation gap and the conic duality gap are zero, and furthermore that (conic dual) attainability holds in (CD). While we do not need any additional assumption like strict feasibility or Slater condition here, the result still does not settle the question of (primal) attainability, neither in problem (CP) nor in problem (SCQP). Remark that primal conic attainability, of problem (CP), is always assumed in the formulation of theorems in [12], but we will avoid this assumption for the specially structured QP here and in particular in the next Section 3.2, owing to the technique adapted from [11].

Theorem 3.1. *We always have $\nu = \mu = p^* < +\infty$, and ν in* (CD) *is attained by $\sigma = p^*$ if $p^* \in \mathbb{R}$. So, if p^* is finite (but not necessarily attained), then full strong duality holds (zero duality gap and dual attainability), and the conic relaxations* (CP) *and* (CD) *of the problem* (SCQP) *are exact.*

Proof. If $p^* = -\infty$, then both problems (SCQP) and (CP) are unbounded (and (CD) is infeasible), due to (12). Otherwise, for finite p^*, we have $\mathsf{G} - p^*\, \mathbf{a}\, \mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})$ according to Proposition 3.1, and therefore $p^* \leq \nu$ by definition (CD), which proves that equalities hold in (12) and that p^* is the optimal solution of (CD). □

3.2 Primal attainability for the polyhedral simplicial case

We now request the cone $\mathbb{K}$ to be polyhedral, specializing (SCQP) to (SPQP) below. With the help of the previous elementary results (which as mentioned are also shown in [19]), we can now prove primal attainability both in (SPQP) and in (CP):

Theorem 3.2. *Consider the following class of quadratic optimization problems:*

$$p^* := \inf\left\{ \tfrac{1}{2}\, \mathbf{y}^\top \mathsf{F}\, \mathbf{y} + \mathbf{f}^\top \mathbf{y} : \mathbf{a}^\top \mathbf{y} = 1, \quad \mathbf{y} \in \mathbb{K} \right\}, \tag{SPQP}$$

where $\mathsf{F} \in \mathcal{S}^n$, $\mathbf{f} \in \mathbb{R}^n$, $\mathbb{K} \subseteq \mathbb{R}^n$ is a polyhedral cone and $\mathbf{a} \in \mathbb{K}^*$. If p^* is finite (in particular, if (SPQP) *is feasible and therefore $\mathbf{a} \notin \mathbb{K}^\perp$), then* (SPQP) *has an optimal solution and its copositive relaxation is exact with full strong conic duality in vigour.*

Proof. By Theorem 3.1, we have $p^* = \mu = \nu$ and the optimal value ν is attained in (CD). Therefore, we conclude that $\mathsf{A} := \mathsf{G} - p^* \mathbf{a}\mathbf{a}^\top \in \mathcal{COP}(\mathbb{K})$ and $\mathsf{A} - \varepsilon\, \mathbf{a}\mathbf{a}^\top \notin \mathcal{COP}(\mathbb{K})$ for any $\varepsilon > 0$. By Corollary 2.2, there exists $\mathbf{u} \in \mathbb{K}$ such that $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$ and $\mathbf{a}^\top \mathbf{u} \neq 0$. Since $\mathbf{a} \in \mathbb{K}^* \backslash \mathbb{K}^\perp$, we get $\mathbf{a}^\top \mathbf{u} > 0$. Let us define $\mathbf{y} = \frac{1}{\mathbf{a}^\top \mathbf{u}}\, \mathbf{u} \in \mathbb{K}$. Clearly, $\mathbf{a}^\top \mathbf{y} = 1$ and hence $\mathbf{y}$ is a feasible solution of (SPQP). Furthermore, since $\mathbf{u}^\top \mathsf{A}\, \mathbf{u} = 0$, we have $\mathbf{y}^\top \mathsf{A}\, \mathbf{y} = 0$, which implies that

$$0 = \mathbf{y}^\top \mathsf{A}\, \mathbf{y} = \mathbf{y}^\top \left(\mathsf{G} - p^* \mathbf{a}\mathbf{a}^\top \right) \mathbf{y} = \tfrac{1}{2}\, \mathbf{y}^\top \mathsf{F}\, \mathbf{y} + \mathbf{f}^\top \mathbf{y} - p^* = f(\mathbf{y}) - p^* \,.$$

It follows that $\mathbf{y}$ is an optimal solution of (SPQP), and that $\mathbf{y}\,\mathbf{y}^\top$ is an optimal solution of (CP) with the same optimal value. This completes the proof. $\square$

Remark 2. *Theorem 3.2 establishes the attainability property for* (SPQP) *under the assumption that* $\mathbb{K} \subseteq \mathbb{R}^n$ *is a polyhedral cone. We now use the example in Remark 1 to illustrate that the attainability result may not necessarily hold for a more general convex cone. Consider an instance of* (SCQP)*, where* $\mathbf{f} = \mathbf{o} \in \mathbb{R}^3$*,* $(\mathsf{F}, \mathbf{a}) = (2\,\mathsf{A}, \mathbf{a}) \in \mathcal{S}^3 \times \mathbb{R}^3$*, where* A *and* $\mathbf{a}$ *are defined as in* (11)*, and* $\mathbb{K} \subseteq \mathbb{R}^3$ *is the second-order cone given by* (10)*. By the arguments in Remark 1,* $\mu \geq 0$*. Let* $\mathbf{x}_k := [\sqrt{k+1}, 1, -\sqrt{k}]^\top \in \mathbb{R}^3$*,* $k \in \mathbb{N}$*. It is easy to verify that* $\mathbf{x}_k$ *is a feasible solution of* (SCQP) *for each* $k \in \mathbb{N}$*, and that* $f(\mathbf{x}_k) = 1 + 2\,k - 2\sqrt{k(k+1)} \to 0$ *as* $k \to \infty$*. It follows that* $p^* = 0$*, however, the optimal value is not attained. In particular, this example clearly highlights the crucial role of Lemma 2.2 in establishing Theorem 3.2. Note, however, that* $\mathbb{K}^* = \mathbb{K}$ *in this example (see, e.g., [32]), whereas* $\mathbf{a} \notin \mathbb{K}^*$*(cf. Theorem 3.2). As such, this example does not necessarily constitute a counterexample to Theorem 3.2 (please see the discussion in Section 6).*

4 A new proof of the Frank-Wolfe theorem

4.1 The theorem and its extensions in the literature

The attainability result of the Frank-Wolfe theorem for quadratic (possibly nonconvex) objectives over (possibly unbounded) polyhedra (see also [16] and [8] for alternative proofs and further refinements), was extended to the case where the feasible set M is the Minkowski sum of a compact set and a polyhedral cone by Kummer [21] (see [23] for further extensions). We remark that any polyhedron can be written as the Minkowski sum of a bounded polyhedron (i.e., a polytope) and a polyhedral cone by the fundamental Minkowski-Weyl theorem (already contained in Motzkin's thesis [24]; a standard reference is [37]; see also Proposition 5.2 in the appendix for an elementary argument). Perold [34] extended the attainability result to a larger class of nonquadratic objective functions over a polyhedral set. In a study that has remained obscure for some time, Belousov [5] established attainability for the case in which each of f and g_j, $j \in [1\!:\!m]$, is a general convex polynomial function of arbitrary degree (see also [6] and [22, 40] for the all-quadratic case, and [15, 27, 30, 35] for extensions to other classes of polynomial functions). Andronov and coauthors [1] extended the Frank-Wolfe theorem to the problem of minimizing a cubic objective function over a polyhedron (see also [20]; and [26, 28] for extensions). The reader is referred to [3, 7, 31, 36] for other extensions.

Now we present an apparently novel argument following from above strong conic duality approach.

Theorem 4.1. *Consider a general quadratic optimization problem given by*

$$p^* := \min \left\{ \tfrac{1}{2}\, \mathbf{x}^\top \mathsf{Q}\, \mathbf{x} + \mathbf{c}^\top \mathbf{x} : \mathbf{x} \in P \right\}, \tag{QP}$$

where $\mathsf{Q} \in \mathcal{S}^n$, $\mathbf{c} \in \mathbb{R}^n$, *and* $P \subseteq \mathbb{R}^n$ *is a nonempty polyhedron. If* p^* *is finite, then* (QP) *has an optimal solution.*

Proof. Let $P \subseteq \mathbb{R}^n$ be a nonempty polyhedron. By the Minkowski-Weyl theorem (see Proposition 5.2 in the appendix for a concise proof), there exist $\{\mathbf{z}_1, \ldots, \mathbf{z}_k\} \subset \mathbb{R}^n$ and $\{\mathbf{d}_1, \ldots, \mathbf{d}_\ell\} \subset \mathbb{R}^n$ such that

$$P = \operatorname{conv}\left(\{\mathbf{z}_1, \ldots, \mathbf{z}_k\}\right) + \operatorname{cone}\left(\{\mathbf{d}_1, \ldots, \mathbf{d}_\ell\}\right).$$

We argue that $k \geq 1$. If $P = \mathbb{R}^n$, then we can choose $\mathbf{z}_1 = \mathbf{o}$ and $\mathbf{d}_j = \mathbf{e}_j$ for $j \in [1\!:\!n]$ and $\mathbf{d}_{n+1} = -\sum_{j=1}^n \mathbf{e}_j$. Otherwise, $\mathbf{z}_1, \ldots, \mathbf{z}_k$ can be chosen to be any arbitrary point on each of the finitely many minimal faces of P and $\mathbf{d}_1, \ldots, \mathbf{d}_\ell$ as the extreme rays of the polyhedral recession cone of P. Therefore, $\mathbf{x} \in P$ if and only if there exists $\mathbf{y} \in \mathbb{R}^{k+\ell}$ such that

$$\mathbf{x} = \sum_{i=1}^k y_i\, \mathbf{z}_i + \sum_{j=1}^\ell y_{k+j}\, \mathbf{d}_j, \tag{13}$$

$$\sum_{i=1}^k y_i = 1, \tag{14}$$

$$\mathbf{y} \in \mathbb{R}_+^{k+\ell}. \tag{15}$$

Let $\mathsf{Z} = \begin{bmatrix} \mathbf{z}_1 & \dots & \mathbf{z}_k \end{bmatrix} \in \mathbb{R}^{n\times k}$ and $\mathsf{D} = \begin{bmatrix} \mathbf{d}_1 & \dots & \mathbf{d}_\ell \end{bmatrix} \in \mathbb{R}^{n\times \ell}$. By defining

$$\mathsf{F} = \begin{bmatrix} \mathsf{Z}^\top \mathsf{Q}\mathsf{Z} & \mathsf{Z}^\top \mathsf{Q}\mathsf{D} \\ \mathsf{D}^\top \mathsf{Q}\mathsf{Z} & \mathsf{D}^\top \mathsf{Q}\mathsf{D} \end{bmatrix} \in \mathcal{S}^{k+\ell},$$

$$\mathbf{f} = \begin{bmatrix} \mathsf{Z}^\top \mathbf{c} \\ \mathsf{D}^\top \mathbf{c} \end{bmatrix} \in \mathbb{R}^{k+\ell},$$

$$\mathbf{a} = \begin{bmatrix} \mathbf{e} \\ \mathbf{o} \end{bmatrix} \in \mathbb{R}^{k+\ell},$$

(QP) can be equivalently formulated as an instance of (SCQP). By Theorem 3.2, the optimal value of (SCQP) is attained, from which an optimal solution of (QP) can be constructed using (13). The assertion follows. □

Remark 3. *The above reformulation (SQP) of (QP) is helpful to prove a mathematical existence result. It need not be minimal in terms of k and ℓ. Even if it were, $k+\ell$ can be exponential in n. In general, it is therefore not recommended to use this approach for practical implementation in optimization algorithms.*

4.2 Counterexamples

While there are extensions of Theorem 4.1 to cubic polynomial objectives and also for polynomial problems where all involved functions are convex (see Section 4.1), the following simple counterexamples illustrate that those assumptions, in general, cannot be relaxed further.

Example 0. *Let $n=1$, $P=\mathbb{R}_+$ and $f(x)=\frac{1}{x+1}$, a strictly convex yet non-polynomial objective. We have $p^*=0$ but it is not attained. Likewise, in the case of* quadratically constrained *quadratic optimization problems, the optimal value $p^* = \inf\left\{x_1^2 : \mathbf{x}\in\mathbb{R}^2, x_1x_2 \geq 1\right\} = 0$ cannot be attained.*

The next counterexample already appeared in [17].

Example 1. *The nonconvex polynomial $f(\mathbf{x}) = x_1^2 + (x_1x_2-1)^2$ of degree four (over the feasible set $P=\mathbb{R}^2$) has infimum zero, which cannot be attained.*

The next counterexamples appeared in [5, 6].

Example 2. *In $\mathbb{R}^3$, define $g_1(\mathbf{x}) := x_2^2 - x_1$ and $g_2(\mathbf{x}) := x_3^2 - x_1 - 1$ as well as $f(\mathbf{x}) := -g_1(\mathbf{x}) - g_2(\mathbf{x}) + (x_2-x_3)^2 = 2x_1 - 2x_2x_3 + 1$. Then the problem*

$$p^* = \inf\{f(\mathbf{x}) : g_1(\mathbf{x}) \leq 0\,,\, g_2(\mathbf{x}) \leq 0\}$$

suffers from inattainability, although all involved functions are quadratic, and all g_i are even convex constraints, whereas f is nonconvex. Indeed, for all $\mathbf{x} \in \mathbb{R}^3$ with $g_1(\mathbf{x}) \le 0$ and $g_2(\mathbf{x}) \le 0$, we have $f(\mathbf{x}) > 0$. On the other hand, for the sequence of feasible solutions $\mathbf{x}_k := [k, \sqrt{k}, \sqrt{k+1}]^\top$, $k \in \mathbb{N}$, we get $\inf_{k\in\mathbb{N}} f(\mathbf{x}_k) = 0 = p^*$. The Hessian of f has only one negative eigenvalue.

Example 3. *For the same constraints as in Example 2, consider now the objective*

$$f(\mathbf{x}) := -2g_1(\mathbf{x}) - 2g_2(\mathbf{x}) + (x_2 - x_3)^2 = -(x_2 + x_3)^2 + 4x_1 + 2\,.$$

which is a concave function strictly positive on the feasible set as argued in Example 2. But with the same sequence $\{\mathbf{x}_k : k \in \mathbb{N}\}$, we get $p^* = 0$. So again there is no attainability.

Finally, we present a counterexample with a cubic objective under a single convex quadratic constraint from [1]:

Example 4. *In $\mathbb{R}^3$, let $f(\mathbf{x}) := x_1^2 x_2 - 2x_1x_3$, and $g_1(\mathbf{x}) := (x_2 - 1)^2 + x_3^2 - 1$. Then, attainability fails for the problem*

$$p^* = \inf\left\{ f(\mathbf{x}) : g_1(\mathbf{x}) \le 0 \right\}\,.$$

Note that the feasible region is given by the Minkowski sum of a compact convex set and a polyhedral cone (i.e., an infinite cylinder of radius one with its axis given by the line $\{[\alpha, 1, 0]^\top : \alpha \in \mathbb{R}\}$). For any $\mathbf{x} \in \mathbb{R}^3$ such that $g_1(\mathbf{x}) \le 0$, we have $x_2 \ge 0$ and $x_3^2 \le -x_2^2 + 2x_2$. If $x_2 = 0$, then $x_3 = 0$, and $f([\alpha, 0, 0]^\top) = 0$ for each $\alpha \in \mathbb{R}$. On the other hand, if $x_2 > 0$, then

$$f(\mathbf{x}) = x_1^2 x_2 - 2x_1x_3 = x_2\left(x_1 - \frac{x_3}{x_2}\right)^2 - \frac{x_3^2}{x_2} \ge -\frac{x_3^2}{x_2} \ge -2 + x_2 > -2.$$

Therefore, $f(\mathbf{x}) > -2$ whenever $g_1(\mathbf{x}) \le 0$. On the other hand, for any positive integer k, it is easy to verify that $\mathbf{x}_k = \left[\sqrt{-1+2k}, \frac{1}{k}, \sqrt{\frac{2}{k} - \frac{1}{k^2}}\right]^\top$ satisfies $g_1(\mathbf{x}_k) = 0$ and $f(\mathbf{x}_k) = -2 + \frac{1}{k} \to -2$ as $k \to \infty$. Therefore, $p^* = -2$ but it is not attained.

5 Appendix: Minkowski–Weyl representation of polyhedra

Anything discussed below belongs to the canon of elementary convex analysis and optimization theory. However, we chose to repeat the simple arguments to illustrate that no advanced tools are necessary to establish them. Given their importance on their own, most of them would deserve mention as theorems; however, to highlight our focus, we formulate them as Lemmas or Propositions.

We start discussing a representation of polyhedral cones $\mathbb{K}$ contained in $\mathbb{R}^n_+$ of the form

$$\mathbb{K} = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathsf{A}\,\mathbf{x} = \mathbf{o}\}$$

where A is an $m \times n$ matrix with full row rank m. Note that any polyhedral cone, i.e., any solution set of a finite number of homogeneous linear inequalities in finitely many variables, can be written in the above form, if necessary, by doubling the number of variables and/or adding slack variables. It will turn out that these cones are exactly the finitely generated cones, i.e. $\mathbb{K} = \{\mathsf{D}\,\mathbf{u} : \mathbf{u} \in \mathbb{R}^k_+\}$ for some $\mathsf{D} \in \mathbb{R}^{n\times k}$.

In line with our previous simplicial approaches, we will first restrict our attention to the *compact base* $C := \{\mathbf{x} \in \mathbb{K} : \mathbf{e}^\top\mathbf{x} = 1\}$ of $\mathbb{K}$ and observe that any extreme ray of $\mathbb{K}$ intersects C in one of its extreme points, and vice versa any such extreme point generates an extreme ray. This is immediate by scaling. So we next discuss characterizations of extreme points of compact convex sets (actually, bounded polyhedra also known as polytopes) of the form

$$C = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathsf{A}\,\mathbf{x} = \mathbf{o}\,,\, \mathbf{e}^\top\mathbf{x} = 1\}\,.$$

The next auxiliary results are seemingly obvious but still need somehow involved proofs. The first one is closely related to the celebrated theorem of Carathéodory.

Lemma 5.1. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix and* $\mathbf{b} = \mathsf{A}\,\widehat{\mathbf{x}}$ *for (at least one)* $\widehat{\mathbf{x}} \in \mathbb{R}^n_+$*. Then there is a (not necessarily unique)* minimal representation $\mathbf{b} = \sum\limits_{i\in I} z_i\,\mathbf{a}_i$ *with* $z_i \geq 0$ *for all* $i \in I$*, where* $I \subseteq [1\!:\!n]$ *and* $\{\mathbf{a}_i : i \in I\}$ *are* linearly independent *columns of* $\mathbf{A}$.

Proof. Suppose that there exists $\widehat{\mathbf{x}} \in \mathbb{R}^n_+$ such that $\mathbf{b} = \mathsf{A}\,\widehat{\mathbf{x}} = \sum_{i\in I} \widehat{x}_i\,\mathbf{a}_i$, where $I = \operatorname{supp}(\widehat{\mathbf{x}})$. If $\{\mathbf{a}_i : i \in I\}$ are linearly independent, then we are done. Otherwise, there exist $d_i \in \mathbb{R}$, $i \in I$, not all zero, such that $\sum\limits_{i\in I} d_i\,\mathbf{a}_i = \mathbf{o}$. Without loss of generality, we may and do assume that $d_i < 0$ for some $i \in I$ (otherwise replace each d_i with $-d_i$). Let $\widehat{\mathbf{d}} \in \mathbb{R}^n$ be such that $\widehat{d_i} = d_i$ for each $i \in I$ and $\widehat{d_j} = 0$ otherwise. Then obviously $\mathsf{A}\,\widehat{\mathbf{d}} = \mathbf{o}$. Let $\mathbf{x}(\alpha) := \widehat{\mathbf{x}} + \alpha\,\widehat{\mathbf{d}}$, where $\alpha \in \mathbb{R}$. Note that $\mathsf{A}\,\mathbf{x}(\alpha) = \mathbf{b}$ for any $\alpha \in \mathbb{R}$. Furthermore, since $x_i > 0$ for each $i \in I$, we have $\mathbf{x}(\alpha) := \widehat{\mathbf{x}} + \alpha\,\mathbf{d} \in \mathbb{R}^n_+$ for all $\alpha \in [0, \alpha^*]$, where

$$\alpha^* := \min_{i\in I: d_i<0} \frac{\widehat{x}_i}{|d_i|} > 0.$$

It follows that $\mathrm{supp}(\mathbf{x}(\alpha^*))$ is a strict subset of $\mathrm{supp}(\widehat{\mathbf{x}})$. Iterating the above reduction procedure and shrinking the index set I further if necessary yields the result after finitely many steps. $\square$

Remark 4. *If* $\mathsf{A} = [1, 1] \in \mathbb{R}^{1\times 2}$ *and* $\mathbf{b} = 1$, *then there are two minimal representations given by* $\mathbf{x}_1 = [1, 0]^\top$ *and* $\mathbf{x}_2 = [0, 1]^\top$.

Corollary 5.1. *Let* $C = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathsf{A}\,\mathbf{x} = \mathbf{b}\}$ *be a nonempty polyhedron where* $\mathsf{A} \in \mathbb{R}^{m\times n}$ *has full row rank* m. *Then there always exists an extreme point of* C, *which after suitable rearrangement of coordinates, is given as follows:*
there is a nonsingular $m \times m$ *submatrix* B *of* $\mathsf{A} = [\mathsf{B}|\mathsf{N}]$ *such that* $\mathsf{B}^{-1}\mathbf{b} \in \mathbb{R}^m_+$. *Then the point*

$$\mathbf{x} := \begin{bmatrix} \mathsf{B}^{-1}\mathbf{b} \\ \mathbf{o} \end{bmatrix} \in \mathbb{R}^n_+ \tag{16}$$

defines an extreme point of C.

Proof. First observe that the point $\mathbf{x}$ defined in (16) is indeed in C, as by assumption $\mathbf{x} \in \mathbb{R}^n_+$ and

$$\mathsf{A}\,\mathbf{x} = \mathsf{B}\,(\mathsf{B}^{-1}\,\mathbf{b}) + \mathsf{N}\,\mathbf{o} = \mathbf{b}\,.$$

Next, assume that $\mathbf{x} = \frac{1}{2}\mathbf{y} + \frac{1}{2}\mathbf{z}$ where $\{\mathbf{y}, \mathbf{z}\} \subset C \subseteq \mathbb{R}^n_+$ as well. Obviously, $\mathbf{y}$ and $\mathbf{z}$ inherit the zero pattern of $\mathbf{x}$ in the sense that $\mathrm{supp}(\mathbf{y}) \subseteq \mathrm{supp}(\mathbf{x})$ and as well $\mathrm{supp}(\mathbf{z}) \subseteq \mathrm{supp}(\mathbf{x})$, which implies after rearrangement of indices so that $I = [1\!:\!m]$, they can be written as $\mathbf{y}^\top = [\mathbf{u}^\top, \mathbf{o}^\top]$ and $\mathbf{z}^\top = [\mathbf{v}^\top, \mathbf{o}^\top]$ with $\{\mathbf{u}, \mathbf{v}\} \subset \mathbb{R}^m_+$. By feasibility we thus know that

$$\mathbf{b} = \mathsf{B}\,\mathbf{u} = \mathsf{B}\,\mathbf{v} = \mathsf{A}\,\mathbf{x}\,,$$

but B is nonsingular, hence $\mathbf{u} = \mathbf{v} = \mathsf{B}^{-1}\mathbf{b}$ and thus $\mathbf{y} = \mathbf{z} = \mathbf{x}$, proving extremality. Finally, existence follows by Lemma 5.1: since we know there is a $\widehat{\mathbf{x}} \in C$ by assumption, we can obtain a submatrix $\mathsf{B} \in \mathbb{R}^{m\times|I|}$ with full column rank, complementing the columns $\{\mathbf{a}_i : i \in I\}$ by other suitable $m - |I|$ linearly independent columns of A if necessary (recall that A has rank m). $\square$

Lemma 5.2. *Any extreme point of a nonempty polyhedron* C *can be written as in* (16) *from Corollary 5.1.*

Proof. Let $\mathbf{x} \in C$ be an extreme point. If $\mathbf{x} = \mathbf{o}$, then $\mathbf{b} = \mathbf{o}$, and we can choose any nonsingular B for (16). Otherwise, again denote by $I := \mathrm{supp}(\mathbf{x})$. Furthermore let $0 < x_r := \min\{x_i : i \in I\}$.

If the columns $\{\mathbf{a}_i : i \in I\}$ are linearly independent, we are done as in the proof of Corollary 5.1. Otherwise, assume by contraposition that they are not, so that there is $\mathbf{v} \in \mathbb{R}^{|I|} \setminus \{\mathbf{o}\}$ such that

$$\sum_{i \in I} v_i \, \mathbf{a}_i = \mathbf{o} \quad \text{with } \gamma := \max_{i \in I} |v_i| > 0 \,. \tag{17}$$

Now, for all $i \in I$, define $y_i = x_i - \frac{x_r}{\gamma} v_i \geq 0$ and $z_i = x_i + \frac{x_r}{\gamma} v_i \geq 0$ while $y_j = z_j = 0$ for all $j \in [1\!:\!n] \setminus I$. Then $\{\mathbf{y}, \mathbf{z}\} \subset \mathbb{R}^n_+$, and via (17) it is easy to verify $\mathsf{A}\,\mathbf{y} = \mathsf{A}\,\mathbf{z} = \mathsf{A}\,\mathbf{x} = \mathbf{b}$, so actually $\{\mathbf{y}, \mathbf{z}\} \subset C$. But $\mathbf{x} = \frac{1}{2}\,\mathbf{y} + \frac{1}{2}\,\mathbf{z}$ with $\mathbf{y} \neq \mathbf{x}$ and $\mathbf{z} \neq \mathbf{x}$, contradicting extremality of $\mathbf{x}$. Hence the result. $\square$

Remark 5. *There are only finitely many choices of extreme points* $\mathbf{x}$ *given as in* (16)*, given by selecting* m *linearly independent columns of* A*, defining* B *as the submatrix of* A *consisting of these columns, and* N *as the submatrix that consists of the remaining ones. So Corollary 5.1 and Lemma 5.2 together prove that there are only finitely many extreme points of* C*, namely at most* $\binom{n}{m}$*, but also at least one.*

Finally, we can prove that any compact (nonempty) polyhedron C is finitely generated in the sense that $C = \operatorname{conv}(\{\mathbf{v}_i : i \in [1\!:\!k]\})$:

Lemma 5.3. *Let* $C = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathsf{A}\mathbf{x} = \mathbf{b}\}$ *be a compact (nonempty) polyhedron and denote by* $\{\mathbf{v}_i : i \in [1\!:\!k]\} \neq \emptyset$ *its extreme points with finite* $k \geq 1$*. Then*

$$C = \operatorname{conv}(\{\mathbf{v}_i : i \in [1\!:\!k]\}) \,.$$

Proof. The proof is similar to, but different from, that of Lemma 5.1. Let $\mathbf{x} \in C$, hence $\sum_{i=1}^n x_i \, \mathbf{a}_i = \mathsf{A}\,\mathbf{x} = \mathbf{b}$. Again, let $I := \operatorname{supp}(\mathbf{x})$. If $\{\mathbf{a}_i : i \in I\}$ are linearly independent, $\mathbf{x}$ itself is extreme as shown in Corollary 5.1. Otherwise choose v_i as in the proof of Lemma 5.2 with $\sum_{i=1}^n v_i \, \mathbf{a}_i = \mathbf{o}$ and not all zero[4]. By the boundedness of C, we see that there must be a pair (i, j) such that $v_i < 0 < v_j$; indeed, if otherwise $\mathbf{v} \in \mathbb{R}^n_+ \setminus \{\mathbf{o}\}$, we would get an unbounded feasible ray of the form $\mathbf{x} + \alpha\,\mathbf{v} \in C$ as $\alpha \nearrow \infty$ (likewise for the case $-\mathbf{v} \in \mathbb{R}^n_+$, using $\alpha \searrow -\infty$). Now define

$$\lambda_1 = \min \left\{ \frac{x_j}{v_j} : j \in I, v_j > 0 \right\} > 0 \quad \text{and} \quad \lambda_2 = \min \left\{ \frac{x_i}{|v_i|} : i \in I, v_i < 0 \right\} > 0 \,,$$

where we used $\mathbf{x} \neq \mathbf{o}$ and the construction of $\mathbf{v}$ to derive strict positivity. So we generate two feasible points $\mathbf{y} = \mathbf{x} - \lambda_1 \, \mathbf{v} \in \mathbb{R}^n_+$ and $\mathbf{z} = \mathbf{x} + \lambda_2 \, \mathbf{v} \in \mathbb{R}^n_+$ (obvious by $\mathsf{A}\,\mathbf{v} = \mathbf{o}$) such that

$$\mathbf{x} = \frac{\lambda_2}{\lambda_1 + \lambda_2} \, \mathbf{y} + \frac{\lambda_1}{\lambda_1 + \lambda_2} \, \mathbf{z} \,,$$

[4] we define $v_j = 0$ for all $j \in [1\!:\!n] \setminus \operatorname{supp}(\mathbf{x})$

each of which has fewer strictly positive coordinates than $\mathbf{x}$. In other words, the supports $\operatorname{supp}(\mathbf{y})$ and $\operatorname{supp}(\mathbf{z})$ are both strict subsets of I. If both $\mathbf{y}, \mathbf{z}$ are extreme points, we are done. Otherwise, we repeat the process with the non-extreme endpoints of the line segment replacing $\mathbf{x}$, strictly shrinking the supports further. Since this iterative process must stop with points satisfying (16), the claim is proved. $\square$

Proposition 5.1. *Any polyhedral cone $\mathbb{K} = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathsf{A}\mathbf{x} = \mathbf{o}\}$ is finitely generated by extreme ray generators $\mathbf{d}_r$ in the sense that*

$$\mathbb{K} = \operatorname{cone}(\{\mathbf{d}_r : r \in [1\!:\!k]\})\,.$$

Proof. Consider the compact base $C = \{\mathbf{x} \in \mathbb{K} : \mathbf{e}^\top \mathbf{x} = 1\}$. By Lemma 5.3, we have $C = \operatorname{conv}\{\mathbf{v}_i : i \in [1\!:\!k]\}$ for some finite k. The extreme points $\mathbf{v}_i$ generate extreme rays of $\mathbb{K}$, as discussed above. Hence the result. $\square$

In the last step, we transfer this result to general polyhedra $P = \{\mathbf{x} \in \mathbb{R}^n : \mathsf{A}\mathbf{x} \leq \mathbf{b}\}$ as follows:

Proposition 5.2. *Let $P = \{\mathbf{x} \in \mathbb{R}^n : \mathsf{A}\mathbf{x} \leq \mathbf{b}\}$ be an arbitrary polyhedron. Then there are finitely many generators $\{\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{x}_{m+1}, \ldots \mathbf{x}_k\}$ of P in the sense that $P = Q + K$ with*

$$Q = \operatorname{conv}(\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}) \quad \textit{and} \quad K = \operatorname{cone}(\{\mathbf{x}_{m+1}, \ldots, \mathbf{x}_k\})\,.$$

Proof. Given P, we form the polyhedral cone

$$\mathbb{K} = \left\{ \begin{bmatrix} \mathbf{x} \\ \alpha \end{bmatrix} \in \mathbb{R}^{n+1} : \alpha \geq 0\,, \mathsf{A}\mathbf{x} - \alpha\,\mathbf{b} \leq 0 \right\}.$$

By Proposition 5.1, and after proper rescaling, we may assume that this cone is generated as follows:

$$\mathbb{K} = \operatorname{cone}\left\{ \begin{bmatrix} \mathbf{x}_1 \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{x}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{x}_{m+1} \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{x}_k \\ 0 \end{bmatrix} \right\}.$$

If we now define $Q = \operatorname{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ and $K = \operatorname{cone}\{\mathbf{x}_{m+1}, \ldots, \mathbf{x}_k\}$, we see $P = Q + K$ as a projection of the following relation for $[\mathbf{x}^\top, 1]^\top \in \mathbb{K}$ for any $\mathbf{x} \in P$:

$$\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} = \sum_{i=1}^{m} \alpha_i \begin{bmatrix} \mathbf{x}_i \\ 1 \end{bmatrix} + \sum_{i=m+1}^{k} \beta_i \begin{bmatrix} \mathbf{x}_i \\ 0 \end{bmatrix}, \quad \text{with} \quad \sum_{i=1}^{m} \alpha_i = 1 \ \text{ and all } \ \alpha_i, \beta_i \geq 0\,.$$

Hence the result. $\square$

6 Conclusion and outlook

By the Minkowski-Weyl theorem, any QP, i.e., any optimization problem that involves minimizing a (possibly nonconvex) quadratic function over a polyhedron in $\mathbb{R}^n$, can be written in simplicial conic form with a single linear constraint over a closed convex cone (and in particular over the nonnegative orthant). These (possibly nonconvex) simplicial conic quadratic optimization problems admit exact reformulations as convex conic optimization problems over appropriately defined matrix cones. For general convex cones, we presented simple and straightforward arguments illustrating how this variant of popular lifting relaxations can be shown to be exact by studying strong duality and attainment properties of the primal-dual pair of such conic optimization problems. Specializing the cone to the nonnegative orthant, our investigations led to a simple, alternative proof of the Frank-Wolfe theorem, based on attainment properties of the primal-dual simplicial conic reformulation.

As illustrated by Remark 1, Corollary 2.2, which is the crucial ingredient in the proof of Theorem 4.1, cannot be extended to nonpolyhedral cones. It would be interesting to investigate whether Corollary 2.2 can be generalized to nonpolyhedral cones under the additional assumption that $\mathbf{a} \in \mathbb{K}^* \setminus \mathbb{K}^\perp$, which, in turn, would imply that the attainment result of Theorem 4.1 could then be extended to nonpolyhedral cones. We leave this as a future research direction.

References

[1] V. G. Andronov, E. G. Belousov, and V. M. Shironin. On solvability of the problem of polynomial programming. *Izvestija Akademii Nauk SSSR. Tekhnicheskaja Kibernetika*, 20(4):194–197, 1982. In Russian. English translation in: Engineering Cybernetics, vol. 20, no. 4, pp. 1–4, 1982.

[2] A. Auslender. How to deal with the unbounded in optimization: Theory and algorithms. *Mathematical Programming*, 79:3–18, 1997.

[3] B. Bank and R. Mandel. *Parametric Integer Optimization.* De Gruyter, Berlin, Boston, 1988.

[4] A. I. Barvinok. *A Course in Convexity*, volume 54 of *Graduate Studies in Mathematics.* American Mathematical Society, 2002.

[5] E. G. Belousov. *Introduction to Convex Analysis and Integer Programming.* Moscow University Publ.: Moscow, 1977.

[6] E. G. Belousov and D. Klatte. A Frank-Wolfe type theorem for convex polynomial programs. *Computational Optimization and Applications*, 22(1):37–48, 2002.

[7] D. P. Bertsekas and P. Tseng. Set intersection theorems and existence of optimal solutions. *Mathematical Programming*, 110(2):287–314, 2007.

[8] E. Blum and W. Oettli. Technical note - direct proof of the existence theorem for quadratic programming. *Operations Research*, 20(1):165–167, 1972.

[9] I. M. Bomze. Evolution towards the maximum clique. *Journal of Global Optimization*, 10(2):143–164, 1997.

[10] I. M. Bomze and M. Dür. A square departure from symmetry in matrix cones. *Vietnam Journal of Mathematics*, 53:757—-771, 2025.

[11] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. J. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 18(4):301–320, 2000.

[12] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, 120(2):479–495, 2009.

[13] V. Chvátal. *Linear Programming.* A series of books in the mathematical sciences. W. H. Freeman, New York, 1983.

[14] P. J. Dickinson, M. Dür, L. Gijben, and R. Hildebrand. Irreducible elements of the copositive cone. *Linear Algebra and its Applications*, 439(6):1605–1626, 2013.

[15] S. T. Dinh, H. V. Ha, and T. S. Pham. A Frank-Wolfe type theorem for nondegenerate polynomial programs. *Mathematical Programming*, 147(1-2):519–538, 2014.

[16] B. C. Eaves. On quadratic programming. *Management Science*, 17(11):698–711, 1971.

[17] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3(1-2):95–110, 1956.

[18] R. Hildebrand. Minimal zeros of copositive matrices. *Linear Algebra and its Applications*, 459:154–174, 2014.

[19] S. Kim and M. Kojima. Strong duality of a conic optimization problem with a single hyperplane and two cone constraints. *Optimization*, 74(1):33–53, 2025.

[20] D. Klatte. On a Frank-Wolfe type theorem in cubic optimization. *Optimization*, 68(2-3):539–547, 2019.

[21] B. Kummer. Globale Stabilität quadratischer Optimierungsprobleme. *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Reihe*, XXVI(5):565–569, 1977.

[22] Z. Luo and S. Zhang. On extensions of the Frank-Wolfe theorems. *Computational Optimization and Applications*, 13(1-3):87–110, 1999.

[23] J. E. Martínez-Legaz, D. Noll, and W. Sosa. Non-polyhedral extensions of the frank and wolfe theorem. In H. H. Bauschke, R. S. Burachik, and D. R. Luke, editors, *Splitting Algorithms, Modern Operator Theory, and Applications*, pages 309–329. Springer International Publishing, Cham, 2019.

[24] T. S. Motzkin. *Beiträge zur Theorie der Linearen Ungleichungen.* PhD thesis, Universität Basel, 1936.

[25] Y. E. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *Studies in Applied and Numerical Mathematics.* SIAM, 1994.

[26] T. V. Nghi and N. N. Tam. A Frank-Wolfe-type theorem for cubic programs and solvability for quadratic variational inequalities. *Journal of Optimization Theory and Applications*, 187(2):448–468, 2020.

[27] T. V. Nghi and N. N. Tam. A Frank-Wolfe-type theorem for constrained nonconvex polynomial programs. *Optimization*, online(since 29 Dec.):1–21, 2025.

[28] T. V. Nghi and N. N. Tam. A non-polyhedral extension of the Frank-Wolfe theorem to cubic optimization. *Optimization Letters*, 19(6):1157–1169, 2025.

[29] M. Nishijima and B. F. Lourenço. Facial structure of copositive and completely positive cones over a second-order cone. *Linear Algebra and its Applications*, 743:211–249, 2026.

[30] W. T. Obuchowska. On generalizations of the Frank-Wolfe theorem to convex and quasi-convex programmes. *Computational Optimization and Applications*, 33(2-3):349–364, 2006.

[31] A. E. Özdaglar and P. Tseng. Existence of global minima for constrained optimization. *Journal of Optimization Theory and Applications*, 128(3):523–546, 2006.

[32] G. Pataki. The geometry of semidefinite programming. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, pages 29–65. Springer US, Boston, MA, 2000.

[33] G. Pataki. On the closedness of the linear image of a closed convex cone. *Mathematics of Operations Research*, 32(2):395–412, 2007.

[34] A. F. Perold. A generalization of the Frank - Wolfe theorem. *Mathematical Programming*, 18(1):215–227, 1980.

[35] T. Pham. Optimality conditions for minimizers at infinity in polynomial programming. *Mathematics of Operations Research*, 44(4):1381–1395, 2019.

[36] W. Schachinger and I. Bomze. A conic duality Frank-Wolfe-type theorem via exact penalization in quadratic optimization. *Mathematics of Operations Research*, 34(1):83–91, 2009.

[37] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience, Chichester, UK, 1986.

[38] A. Shapiro. On duality theory of conic linear problems. In M. Á. Goberna and M. A. López, editors, *Semi-Infinite Programming: Recent Advances*, pages 135–165. Springer US, Boston, MA, 2001.

[39] J. F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28(2):246–267, 2003.

[40] T. Terlaky. On ℓ_p programming. *European Journal of Operational Research*, 22(1):70–100, 1985.