Strong global convergence properties of algorithms for nonlinear symmetric cone programming*

R. Andreani† G. Haeser‡ A. Ramos§ D.O. Santos¶ L.D. Secchin‖
A. Serranoni**

January 6, 2024 (updated Jun 5, 2024)

Abstract

Sequential optimality conditions have played a major role in proving strong global convergence properties of numerical algorithms for many classes of optimization problems. In particular, the way complementarity is handled defines different optimality conditions, and is fundamental to achieving a strong condition. Typically, one uses the inner product structure to measure complementarity, which gives a very general approach to a general conic optimization problem, even in the infinite dimensional case. In this paper we exploit the Jordan algebraic structure of symmetric cones in order to measure complementarity, which gives rise to a stronger sequential optimality condition related to the well-known complementary approximate Karush-Kuhn-Tucker conditions in standard nonlinear programming. Our results improve some known results in the setting of semidefinite programming and second-order cone programming in a unified manner. In particular, we obtain global convergence results which are stronger than the ones known for augmented Lagrangian algorithms and interior point methods for general symmetric cones.

Key words: Nonlinear symmetric cone optimization, sequential optimality conditions, numerical algorithms, global convergence.

1 Introduction

The concept of a symmetric cone appeared for the first time in the context of quantum mechanics [27], in the quest for the adequate mathematical object describing an observable quantity. Although physicists have abandoned the idea early on, symmetric cones turned out to be extremely relevant in several fields of mathematics. In particular, the relation between symmetric cones and Jordan algebras has fueled many contributions from algebraists. More recently, the concept has proved to be in the center of the most relevant conic optimization problems.

Any symmetric cone is associated with an Euclidean Jordan algebra; if the algebra is simple then the symmetric cone is also called simple. Jordan, von Neumann, and Wigner presented a complete classification of simple Euclidean Jordan algebras in [27] up to isomorphisms; the same classification can be seen in [20] in a more modern exposition. In terms of such classification, the existing symmetric cones are: the *positive semidefinite cone* defined over the symmetric matrices with real, complex or quaternionic entries, the so-called *Lorentz* cone or *second-order cone*, and the *exceptional* or *Albert* cone, which is defined as a 27-dimensional cone of positive semidefinite 3×3 matrices with octonion entries. All other symmetric cones are obtained by taking the Cartesian product of the simple symmetric cones above; for instance, the nonnegative orthant in $\mathbb{R}^m$ is simply the Cartesian product of positive semidefinite one dimensional cones.

*This work has been partially supported by CEPID-CeMEAI (FAPESP 2013/07375-0), FAPESP (grants 2023/01655-2, 2018/24293-0 and 2017/18308-2), CNPq (grants 306988/2021-6, 302000/2022-4, 309136/2021-0 and 407147/2023-3), PRONEX - CNPq/FAPERJ (grant E-26/010.001247/2016) and ANID (Fondecyt grant 1231188).

†Department of Applied Mathematics, University of Campinas, Campinas, SP, Brazil. E-mail: andreani@ime.unicamp.br

‡Department of Applied Mathematics, University of São Paulo, São Paulo, SP, Brazil. E-mail: ghaeser@ime.usp.br

§Departamento de Matemática, Universidad de Tarapacá, Arica, Chile. Email: aramosf@academicos.uta.cl

¶Paulista School of Politics, Economics and Business, Federal University of São Paulo, Osasco-SP, Brazil. Email: daiana.santos@unifesp.br

‖Department of Applied Mathematics, Federal University of Espírito Santo, São Mateus, ES, Brazil. E-mail: leonardo.secchin@ufes.br

**Department of Computer Science, University of São Paulo, São Paulo, SP, Brazil. E-mail: ariel@ime.usp.br

In optimization, the importance of symmetric cones is clear. The positive semidefinite cone and the second-order cone allows defining the *nonlinear semidefinite programming* (NSDP) problem and the *nonlinear second-order cone programming* (NSOCP) problem. Each of these problems are well studied due to their presence in many applications. When the problem has more than one type of conic constraint, say, semidefinite and second-order cone ones, formulation as a symmetric cone may be beneficial. Also, by exploiting the Jordan algebraic structure of symmetric cones, one is able to highlight the most relevant features of the problem. In general, the study of optimization problems over general symmetric cones appears as an unifying tool of the theory developed for each of these problems mentioned, see [20, 33, 34]. See also [40] for strategies for dealing with mixed constraints by means of symmetric cones; in particular, even though a second-order cone constraint may be formulated as a semidefinite constraint, this reformulation usually introduces some undesired structure to the problem, at least when one is interested in working with weak constraint qualifications.

In this work, we propose a strong global convergence theory for an augmented Lagrangian method and an interior point method formulated for a general nonlinear symmetric cone problem defined on a real vector space. This is obtained by means of so-called sequential optimality conditions, which, by means of a penalization technique, builds a primal-dual sequence nearby a local minimizer of the problem. This gives a new necessary optimality condition, without constraint qualifications, which establishes the existence of a particular primal-dual sequence in order for it to be checked. We then show that the algorithms we study generate precisely this type of sequence; therefore, the primal feasible limit points of the sequences generated by the algorithms satisfy the new necessary (sequential) optimality condition. This gives a (stronger) confidence with respect to the optimality of the limit point. The strength of this confidence is measured by comparing our results with the classical one stating that under a (weak) constraint qualification the said limit point satisfies the standard Karush-Kuhn-Tucker conditions.

Sequential optimality conditions emerged as a theoretical tool for this type of analysis more than a decade ago [5] and they have been widely exploited in the most diverse contexts such as cardinality constrains [28], complementarity constraints [8], geometric constraints [18], infinite dimensional spaces [16], sequential quadratic programming for NSDP [36], bound-constrained minimization and complexity analysis [37, 26], Nash equilibrium [30, 17], constraint qualifications for NSDP [6] and NSOCP [7], among several other contexts. In the case of symmetric cones, its eigenvalue structure [9] or the inherent inner product [4] have been used in order to define a notion of approximate complementarity. However, although the use of the Jordan product provides stronger results [3, 2], this has only been partially explored in the literature. In this paper we present a novel way of exploiting the Jordan product in defining approximate complementarity, which allows us to present a proof that algorithms for general symmetric cones generate this sequence, instead of resorting to a case-by-case analysis such as in [3, 2]. The global convergence result for an augmented Lagrangian method obtained in this way is stronger than the previous ones for all classes of symmetric cones, in particular, the augmented Lagrangian method studied in [33] requires that the set of Lagrange multipliers be a singleton, while this set may even be unbounded in our analysis. We refer also to [39] for primal-dual global convergence of augmented Lagrangian methods under second-order sufficient conditions in a framework that includes conic constraints. See also [22] in the context of nonlinear programming. We also expand our studies to consider an interior point method which also generates this type of sequence, hence enjoying the same strong global convergence result. Interestingly, the interior point method we consider is inherently formulated in the context of symmetric cones, as it heavily makes use of the spectral decomposition in its formulation.

This article has the following structure: the basic theoretical tools we use are presented in Section 2. Specifically, we review the fundamental results from Euclidean Jordan algebras and symmetric cones, together with some fundamentals of convex analysis. In Section 3, we discuss optimality conditions. In particular, we define the *Approximate Gradient Projection* (AGP) optimality condition and its stronger counterpart, the *Complementary Approximate KKT* (CAKKT) condition and variations. Sections 4 and 5 are devoted to the algorithms. In Section 4 we present an augmented Lagrangian method and in Section 5 an interior point method for nonlinear symmetric cone programming. In both cases, we are able to show that such algorithms generate CAKKT sequences. Finally, in Section 6, we measure the strength of our results by means of defining suitable constraint qualifications. Section 7 presents some concluding remarks.

Summary of contributions:

- For the general nonlinear programming problem with a symmetric cone constraint, we introduce a new sequential optimality condition, which we call *Split-Complementary-Approximate-KKT* condition, which is stronger than similar notions defined previously.

- We prove a global convergence result for an augmented Lagrangian method in terms of the new optimality condition. This gives, in particular, that feasible limit points of the algorithm satisfy the Complementary-Approximate-KKT (CAKKT) condition, which was not previously known. This improves on the quality of the limit point found by this algorithm.

- We extend our results to prove that an interior point framework also satisfy the CAKKT sequential optimality condition.
- We introduce a new constraint qualification called CAKKT-regularity which is stronger than Robinson's condition such that under this regularity condition a standard KKT point is found by the algorithms above.

Notation: We represent by $\mathbb{S}^m$ the set of $m \times m$ real symmetric matrices. The positive semidefinite cone will be denoted by $\mathbb{S}^m_+ = \{A \in \mathbb{S}^m : x^T A x \geq 0, \forall x \in \mathbb{R}^n\}$ and a shorthand for a positive semidefinite matrix $A \in \mathbb{S}^m_+$ will be $A \succeq 0$. The m-dimensional second-order cone will be denoted by $\mathcal{Q} = \{(x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1} : x_0 \geq \|\bar{x}\|\}$ for $m \geq 2$, where $\|\cdot\|$ denotes the Euclidean norm. Given a finite dimensional vector space $\mathcal{V}$ equipped with an inner product, an operator $J\colon \mathbb{R}^n \to \mathcal{V}$ and a point $x \in \mathbb{R}^n$, we denote by $DJ(x)$ the derivative of J at x and its adjoint with respect to the inner product of $\mathcal{V}$ and the standard inner product in $\mathbb{R}^n$ by $DJ(x)^T\colon \mathcal{V} \to \mathbb{R}^n$. The identity matrix of appropriate dimensions will be denoted by I and the Hadamard product of $u, v \in \mathbb{R}^\ell$ with $u = (u_1, \dots, u_\ell)$ and $v = (v_1, \dots, v_\ell)$ is denoted by $u \bullet v = (u_1 v_1, \dots, u_\ell v_\ell)$. For a point-to-set mapping $C\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^\ell$ we denote the outer limit of $C(z)$ as $z \to \bar{z}$ by $\limsup_{z \to \bar{z}} C(z) = \{\bar{w} \in \mathbb{R}^\ell : \exists (z^k, w^k) \to (\bar{z}, \bar{w}), w^k \in C(z^k)\}$.

2 Preliminaries

In this section we present the essential concepts that form the core of our exposition, namely, Euclidean Jordan algebras and symmetric cones. We restrict our analysis to the case of a symmetric cone with real entries, however, one could easily extend our approach to deal with more general symmetric cones. We also state some well known concepts about projections that are used in our results. Most of the content presented here can be seen in more detail in [20]. See also [21, 33].

Definition 2.1. *Let $\mathcal{V}$ be a real finite-dimensional vector space. We say that $\mathcal{V}$ is a* Jordan algebra *when it is equipped with a* Jordan product*, that is, a bilinear commutative operator $\circ\colon \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ satisfying the Jordan property*

$$u \circ (u^2 \circ w) = u^2 \circ (u \circ w) \text{ for each } u, w \in \mathcal{V}, \text{ where } u^2 = u \circ u.$$

In addition, when $\mathcal{V}$ is equipped with a compatible inner product, that is, such that $x \mapsto u \circ x$ is self-adjoint with respect to this inner product for each $u \in \mathcal{V}$, we say that $\mathcal{V}$ is an Euclidean Jordan algebra*. In other words, we require the inner product $\langle \cdot, \cdot \rangle$ and the Jordan product $\circ$ to be compatible by satisfying*

$$\langle u \circ w, z \rangle = \langle w, u \circ z \rangle \text{ for each } u, w, z \in \mathcal{V}.$$

We assume throughout the paper the existence of a unique multiplicative identity e satisfying $u \circ e = u$ for each $u \in \mathcal{V}$.

Definition 2.2. *Let $\mathcal{V}$ be a Jordan algebra. The* rank *of $u \in \mathcal{V}$ is the smallest natural number r such that the elements*

$$e, u^1, \dots, u^r$$

are linearly dependent, where $u^r = u \circ u^{r-1}$ for $r > 1$ and $u^1 = u$. The rank of $\mathcal{V}$ is the maximum rank among its elements.

We adopt the letter r to represent the rank of a Jordan algebra $\mathcal{V}$, which will be clear from the context. An element $c \in \mathcal{V}$ is called an *idempotent* if $c \circ c = c$, while a *primitive idempotent* is an idempotent which is not the sum of two nonzero idempotents. Furthermore, two idempotents c_1 and c_2 are said to be orthogonal if $c_1 \circ c_2 = 0$.

Definition 2.3. *We say that $\{c_1, \dots, c_r\} \subset \mathcal{V}$ is a* complete system of orthogonal idempotents*, where r is the rank of the Jordan algebra $\mathcal{V}$, if the following conditions hold:*

(i) $c_i \circ c_j = 0$ *for all* $i \neq j$*;*

(ii) $c_i \circ c_i = c_i$ *for all* i*;*

(iii) $c_1 + \dots + c_r = e$.

If in addition each idempotent c_i is primitive, then we say that $\{c_1, \ldots, c_r\}$ is a Jordan frame *of $\mathcal{V}$.*

An essential tool for developing the theory of nonlinear symmetric cone programming is the spectral decomposition, which we present next.

Theorem 2.1. *([20, Theorem III.1.2]) Let $\mathcal{V}$ be an Euclidean Jordan algebra and $u \in \mathcal{V}$. Then there exist real numbers $\lambda_1(u), \ldots, \lambda_r(u)$ and a Jordan frame $\{c_1(u), \ldots, c_r(u)\}$ such that*

$$u = \sum_{i=1}^{r} \lambda_i(u) c_i(u).$$

The numbers $\lambda_1(u), \ldots, \lambda_r(u)$ are the eigenvalues *of u.*

Next, we state the notion of a symmetric cone and its relations with Euclidean Jordan algebras.

Definition 2.4. *Let $\mathcal{V}$ be a finite-dimensional vector space equipped with an inner product $\langle \cdot, \cdot \rangle$. A non-empty cone $\mathcal{K} \subseteq \mathcal{V}$ is a* symmetric cone *if its interior, denoted by $int(\mathcal{K})$, is non-empty,*

(i) $\mathcal{K}$ is self-dual*, that is, $\mathcal{K} = \mathcal{K}^* := \{u : \langle u, v \rangle \geq 0, \forall v \in \mathcal{K}\}$, and*

(ii) $\mathcal{K}$ is homogenous*, that is, for each $u, v \in int(\mathcal{K})$ there is a linear bijection T such that $T(u) = v$ and $T(\mathcal{K}) = \mathcal{K}$.*

The simplest examples of symmetric cones are the cone of $m \times m$ positive semidefinite matrices with real entries, the second-order cone in $\mathbb{R}^m$ and the non-negative orthant of $\mathbb{R}^m$. Next, we will present a very useful characterization of a symmetric cone in terms of Euclidean Jordan algebras.

Theorem 2.2. *([20, Theorem III.2.1]) Let $\mathcal{V}$ be a finite-dimensional inner product space and $\emptyset \neq \mathcal{K} \subseteq \mathcal{V}$. We have that $\mathcal{K}$ is a symmetric cone if, and only if, $\mathcal{V}$ can be equipped with a Jordan product $\circ$ compatible with the inner product of $\mathcal{V}$ so that $\mathcal{K} = \{u \circ u : u \in \mathcal{V}\}$.*

In summary, every symmetric cone can be seen as a cone of squares of a convenient Jordan product. For instance, a symmetric matrix A is positive semidefinite if, and only if, A has a square root, that is, a symmetric matrix B such that $A = B^2$. This can be seen from the previous result by endowing $\mathbb{S}^m$ with the compatible Jordan product:

$$X \circ Y := \frac{XY + YX}{2}$$

for $X, Y \in \mathbb{S}^m$. Similarly, the second-order cone $\mathcal{Q}$ is the cone of squares of $\mathbb{R}^m$ when equipped with the canonical inned product and the Jordan product

$$x \circ y := (\langle x, y \rangle, x_0 \bar{y} + y_0 \bar{x})$$

for $x = (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1}$ and $y = (y_0, \bar{y}) \in \mathbb{R} \times \mathbb{R}^{m-1}$. In turn, the non-negative orthant of $\mathbb{R}^m$ is the cone of squares of $\mathbb{R}^m$ equipped with the Jordan product $x \circ y := x \bullet y$ for $x, y \in \mathbb{R}^m$. The following propositions summarize some basic properties of symmetric cones which will be used in the sequel.

Proposition 2.1 ([20])**.** *Let $\mathcal{V}$ be an Euclidean Jordan algebra and $\emptyset \neq \mathcal{K} \subseteq \mathcal{V}$ its associated symmetric cone. Let $u, v \in \mathcal{V}$.*

(i) $u \in \mathcal{K}$ if, and only if, the eigenvalues of u are non-negative.

(ii) When $u, v \in \mathcal{K}$ we have $u \circ v = 0$ if, and only if, $\langle u, v \rangle = 0$.

Proposition 2.2 ([25])**.** *Let $\mathcal{V}$ be an Euclidean Jordan algebra.*

(i) We have that

$$|\langle u, v \rangle| \leq \|u \circ v\|_1 \leq \|u\|_2 \|v\|_2 \text{ for every } u, v \in \mathcal{V},$$

where $\| \cdot \|_1$ and $\| \cdot \|_2$ are the ℓ^1 and ℓ^2 spectral norms of $\mathcal{V}$, respectively.

(ii) If $\{u_k\}_{k \in \mathbb{N}}$ is a sequence in $\mathcal{V}$, then

$$u_k \circ u_k \to 0 \text{ if, and only if } \|u_k\| \to 0.$$

.

We end this introductory part with some elementary facts about the projection of an element $u \in \mathcal{V}$ onto a closed convex cone $\mathcal{K} \subseteq \mathcal{V}$, which will be denoted by $\Pi_{\mathcal{K}}(u)$. When $\mathcal{K}$ is a symmetric cone and $\mathcal{V}$ is endowed with the corresponding Jordan product, we can write the spectral decomposition $u = \sum_{i=1}^{r} \lambda_i c_i$; in this case, it is well-known that $\Pi_{\mathcal{K}}(u) = \sum_{i=1}^{r} \max\{0, \lambda_i\} c_i$. Notice that since $\mathcal{K}$ is self-dual, the polar of $\mathcal{K}$ is $-\mathcal{K}$. The next lemma brings useful properties of projections.

Lemma 2.1. *([4, Lemma 2.1]) Let $\mathcal{V}$ be an inner product space, $\mathcal{K} \subseteq \mathcal{V}$ a closed convex self-dual cone, and $u \in \mathcal{V}$. Then,*

(i) $\Pi_{\mathcal{K}}(\alpha u) = \alpha \Pi_{\mathcal{K}}(u)$ *for all* $\alpha \geq 0$*;*

(ii) $\Pi_{\mathcal{K}}(-u) = -\Pi_{-\mathcal{K}}(u)$*;*

(iii) (Moreau's decomposition) $u = \Pi_{\mathcal{K}}(u) + \Pi_{-\mathcal{K}}(u)$ *and* $\langle \Pi_{\mathcal{K}}(u), \Pi_{-\mathcal{K}}(u) \rangle = 0$.

3 Optimality conditions

The problem we are dealing with is the general nonlinear symmetric cone programming (NSCP) problem stated as

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{Minimize}} \quad & f(x), \\ \text{subject to} \quad & G(x) \in \mathcal{K}, \\ & h(x) = 0, \end{aligned} \tag{NSCP}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$, $h\colon \mathbb{R}^n \to \mathbb{R}^\ell$ and $G\colon \mathbb{R}^n \to \mathcal{V}$ are continuously differentiable mappings with $\mathcal{V}$ a finite dimensional inner-product space, and $\mathcal{K} \subseteq \mathcal{V}$ a symmetric cone. We denote by $\circ$ the compatible Jordan product of $\mathcal{V}$ given by Theorem 2.2.

Let us recall the well-known *Karush-Kuhn-Tucker* (KKT) conditions for NSCP. Let $L\colon \mathbb{R}^n \times \mathbb{R}^\ell \times \mathcal{V} \to \mathbb{R}$ be the so-called *Lagrangian function* associated with our problem, that is, $L(x, \lambda, \sigma) := f(x) + \langle \lambda, h(x) \rangle - \langle G(x), \sigma \rangle$. A triple $(\bar{x}, \lambda, \sigma) \in \mathbb{R}^n \times \mathbb{R}^\ell \times \mathcal{K}$ is a *KKT triple* for (NSCP) if

$$\begin{aligned} \nabla_x L(\bar{x}, \lambda, \sigma) &= 0, \\ \sigma \circ G(\bar{x}) &= 0, \end{aligned}$$

with $h(\bar{x}) = 0$ and $G(\bar{x}) \in \mathcal{K}$, where $\nabla_x L(\bar{x}, \lambda, \sigma) = \nabla f(\bar{x}) + Dh(\bar{x})^T \lambda - DG(\bar{x})^T \sigma$. The condition $\sigma \circ G(\bar{x}) = 0$ is called the *complementarity condition* and it can be replaced by $\langle \sigma, G(\bar{x}) \rangle = 0$ or by $\lambda_i(\sigma) \lambda_i(G(\bar{x})) = 0$ for all $i = 1, \ldots, r$, where $G(\bar{x})$ and σ share a common Jordan frame. We say that $\bar{x}$ satisfies *strict complementarity* if $\operatorname{rank}(\sigma) + \operatorname{rank}(G(\bar{x})) = r$. It is well-known that when we have some *constraint qualification* (CQ), the KKT conditions provide necessary optimality conditions for (NSCP). In order to obtain global convergence results for algorithms, some specific constraint qualifications are usually needed in order to guarantee that the sequence generated by the algorithm can be proven to give rise to a KKT triple. In [33], the authors propose an augmented Lagrangian algorithm for nonlinear programming problems on symmetric cones where their convergence results are obtained requiring the so-called *nondegeneracy* condition defined below.

Definition 3.1. *We say that a feasible point $\bar{x}$ satisfies the nondegeneracy condition if*

$$\mathbb{R}^\ell = Im(Dh(\bar{x})),$$

$$\mathcal{V} = lin\ T_{\mathcal{K}}(G(\bar{x})) + Im(DG(\bar{x})),$$

where $Im(\cdot)$ denotes the image space of the underlying operator and $lin\ T_{\mathcal{K}}(G(\bar{x}))$ denotes the lineality space of the tangent cone of $\mathcal{K}$ at $G(\bar{x})$.

Although nondegeneracy is a very widespread condition used in many conic contexts, it is considered too stringent as it guarantees, for instance, uniqueness of the Lagrange multiplier. In order to avoid this assumption, we will adopt a different approach that does not require a constraint qualification. This type of condition is called *sequential optimality condition* and the idea is to adopt an approximate version of the KKT conditions. In this version, the KKT conditions are approximately satisfied by requiring the existence of a sequence of primal points, not necessarily feasible, that converges to the feasible point of interest, alongside a sequence of (possibly unbounded) Lagrange multipliers that in some sense approximately satisfy the KKT conditions. Let us begin by explicitly defining the simplest of these conditions, named Approximate-KKT (AKKT) condition, initially introduced for NSDP in [9] and extended to symmetric cones in [3].

Definition 3.2 (AKKT for symmetric cones). *Let $\bar{x}$ be a feasible point for* (NSCP). *We say that $\bar{x}$ satisfies the* Approximate-KKT (AKKT) *conditions if there exist sequences $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$, $\{\lambda^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^\ell$, and $\{\sigma^k\}_{k\in\mathbb{N}} \subset \mathcal{K}$ such that*

$$\nabla f(x^k) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k \to 0, \tag{1}$$

$$\text{If } \lambda_i(G(\bar{x})) > 0 \text{ then } \lambda_i(\sigma^k) = 0 \text{ for each } i \in \{1,\dots,r\} \text{ and sufficiently large } k, \tag{2}$$

$$c_i(\sigma^k) \to c_i(G(\bar{x})) \text{ for each } i \in \{1,\dots,r\},$$

where we use the terminology introduced in Theorem 2.1 for the spectral decompositions of $G(\bar{x})$ and $\sigma^k, k \in \mathbb{N}$.

The correspondence between the Jordan frames of $G(\bar{x})$ and $\sigma^k, k \in \mathbb{N}$, given in (2) allows looking at strictly positive eigenvalues of $G(\bar{x})$ as "inactive" constraints, with the "corresponding" Lagrange multiplier eigenvalue being eventually equal to zero. This enables a more direct analogy with the concept of active constraints from classic nonlinear programming. As shown in [3], any local minimizer satisfies the AKKT conditions without a CQ. Also, an AKKT point satisfying Robinson's CQ, a condition strictly weaker than nondegeneracy (see Definition 6.1), is also a KKT point. Constraint qualifications that guarantee that any AKKT point is actually a KKT point are often called *strict* CQs [11]. Observe that any strict CQ is a constraint qualification by default. In later sections we will go further into the specifics of algorithms and strict constraint qualifications for (NSCP).

Over the last decade, new sequential optimality conditions have emerged in nonlinear optimization in order to present stronger global convergence results of algorithms. In particular, we are interested in the so-called *Complementarity-AKKT* (CAKKT) condition introduced in [10] in the nonlinear programming context. This concept was extended to the conic context in [3, 2, 4], however the analysis of algorithms for general symmetric cones under this condition is not complete. For instance, in [3, 2], an augmented Lagrangian method for specific classes of symmetric cones (semidefinite cone and second-order cone) have been shown to generate a CAKKT sequence, while in [4] a different extension of CAKKT has been proposed which does not make use of the Jordan algebraic structure. Both definitions generalize the notion from [10] for the case of nonlinear programming. We will bridge this gap by defining a new sequential optimality condition which we call *split-CAKKT*, which will imply both definitions of CAKKT in the conic context, and we will show that this condition is satisfied at feasible limit points of an augmented Lagrangian method. We will also show an interior point method in a general symmetric cone context which generates CAKKT sequences. Next, we define the CAKKT conditions for (NSCP) as in [3].

Definition 3.3 (CAKKT for symmetric cones). *Let $\bar{x}$ be a feasible point for* (NSCP). *We say that $\bar{x}$ satisfies the* Complementarity-AKKT condition *(CAKKT) if there exist sequences $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$, $\{\lambda^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^\ell$, and $\{\sigma^k\}_{k\in\mathbb{N}} \subset \mathcal{K}$ satisfying*

$$\begin{aligned} \nabla f(x^k) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k &\to 0, \\ \sigma^k \circ G(x^k) &\to 0, \\ \text{and} \quad \lambda^k \bullet h(x^k) &\to 0. \end{aligned}$$

Once an algorithm is known to generate a CAKKT sequence, one can also use this fact to propose a suitable stopping criterion [10] or a criterion for complexity analysis [26, 37]. Namely, the following characterization of CAKKT will be useful later on:

Proposition 3.1. *A point $\bar{x}$ satisfies the CAKKT condition if, and only if, for all $\varepsilon > 0$, there exists x with $\|x - \bar{x}\| \leq \varepsilon$ such that x is an ε-CAKKT point, that is, there exist $\sigma \in \mathcal{K}$ and $\lambda \in \mathbb{R}^\ell$ such that*

$$\|\nabla f(x) + Dh(x)^T\lambda - DG(x)^T\sigma\| \leq \varepsilon, \ \|h(x)\| \leq \varepsilon,$$

and

$$dist(G(x), \mathcal{K}) \leq \varepsilon, \ \|\sigma \circ G(x)\| \leq \varepsilon, \ \|\lambda \bullet h(x)\| \leq \varepsilon,$$

where $dist(G(x), \mathcal{K}) = \|G(x) - \Pi_{\mathcal{K}}(G(x))\|$ is the distance of $G(x)$ to $\mathcal{K}$.

Proof. This is straightforward by noticing that from continuity one has $\|h(x^k)\| \to 0$ and $\|\mathrm{dist}(G(x^k), \mathcal{K})\| \to 0$ for some $x^k \to \bar{x}$ if, and only if, $\bar{x}$ is feasible. ■

Finally, we propose next our extension of the CAKKT condition from [10] to the context of symmetric cones. We exploit further the splitting technique using Moreau's decomposition introduced in [4], however we use the Jordan product instead of the inner product. Thus, clearly, our condition implies both previously defined notions of the CAKKT concept in the context of symmetric cones [4, 3].

Definition 3.4 (**Split-CAKKT/AGP for symmetric cones**). *Let $\bar{x}$ be a feasible point of* (NSCP). *Suppose that there exist sequences $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$, $\{\lambda^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^\ell$, and $\{\sigma^k\}_{k\in\mathbb{N}} \subset \mathcal{K}$ such that (1) is satisfied.*

(i) We say that $\bar{x}$ satisfies the Split-CAKKT *condition if, in addition to satisfying* (1), *the sequences satisfy*

$$\lambda^k \bullet h(x^k) \to 0, \tag{3}$$

$$\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k)) \to 0, \tag{4}$$

and

$$\sigma^k \circ \Pi_{-\mathcal{K}}(G(x^k)) \to 0. \tag{5}$$

(ii) We say that $\bar{x}$ satisfies the Approximate Gradient Projection (AGP) *condition if, in addition to satisfying* (1), *the sequences satisfy* (4).

To see that Split-CAKKT implies CAKKT as in [3], it is enough to use Lemma 2.1 to write $G(x^k) = \Pi_{\mathcal{K}}(G(x^k)) + \Pi_{-\mathcal{K}}(G(x^k))$ and use the fact that the Jordan product is bilinear to conclude that $\sigma^k \circ G(x^k) \to 0$. Also, by Proposition 2.2, (4) and (5) imply $\langle \sigma^k, \Pi_{\mathcal{K}}(G(x^k))\rangle \to 0$ and $\langle \sigma^k, \Pi_{-\mathcal{K}}(G(x^k))\rangle \to 0$, which is the definition from [4]. In the previous definition we also propose a stronger notion of the Approximate Gradient Projection condition which makes use of the Jordan product (in place of the inner product as in [4]). The idea is that for the augmented Lagrangian method, we will be able to show that an AGP sequence is generated without further smoothness assumptions, while a Split-CAKKT sequence will be generated under an additional smoothness assumption of the functions $G(\cdot)$ and $h(\cdot)$. Next we prove that Split-CAKKT (and AGP) are necessary optimality conditions for NSCP without the need of a constraint qualification to be present. The idea is based on the classical external penalty method in a very similar way to what is done in [3, Theorem 3.1].

Theorem 3.1. *Let $\bar{x}$ be a local minimizer for* (NSCP). *Then, $\bar{x}$ satisfies the Split-CAKKT condition.*

Proof. Let us build an auxiliary problem

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{Minimize}} & f(x) + \frac{1}{2}\|x-\bar{x}\|^2, \\ \text{subject to} & G(x) \in \mathcal{K}, \\ & h(x) = 0, \\ & \|x-\bar{x}\| \leq \delta, \end{array} \tag{6}$$

for some $\delta > 0$ small enough. Thus, $\bar{x}$ is the unique global minimizer of (6). In fact, it is enough to consider $\{x : \|x-\bar{x}\| \leq \delta\}$ as the neighborhood that attests that $\bar{x}$ is a local minimizer in order to make it a global one, while the regularization term $\|x-\bar{x}\|^2$ asserts uniqueness. Now, let us consider the penalized problem

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{Minimize}} & f(x) + \frac{1}{2}\|x-\bar{x}\|^2 + \frac{\rho_k}{2}\left(\|h(x^k)\|^2 + \|\Pi_{\mathcal{K}}(-G(x))\|^2\right), \\ \text{subject to} & \|x-\bar{x}\| \leq \delta, \end{array} \tag{7}$$

where $\{\rho_k\}_{k\in\mathbb{N}} \to \infty$ for each $k \in \mathbb{N}$. Notice that by Lemma 2.1, $\|\Pi_{\mathcal{K}}(-G(x))\| = \|\Pi_{-\mathcal{K}}(G(x))\| = \|G(x) - \Pi_{\mathcal{K}}(G(x))\| = \operatorname{dist}(G(x), \mathcal{K})$, therefore problem (7) penalizes infeasibility. Now, considering x^k the solution of (7), we can prove by usual external penalty method arguments that $x^k \to \bar{x}$ and therefore $\|x^k - \bar{x}\| < \delta$ for k large enough. Thus,

$$\nabla f(x^k) + (x^k - \bar{x}) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k = 0,$$

where $\lambda^k = \rho_k h(x^k)$, $\sigma^k = \rho_k \Pi_{\mathcal{K}}(-G(x^k))$, which implies (1). Since $\{\sigma^k\}_{k\in\mathbb{N}} \subseteq \mathcal{K}$ and $\Pi_{\mathcal{K}}(-G(x^k)) \in \mathcal{K}$, we have by Lemma 2.1 that

$$\left\langle \sigma^k, \Pi_{\mathcal{K}}(G(x^k))\right\rangle = -\rho_k \left\langle \Pi_{-\mathcal{K}}(G(x^k)), \Pi_{\mathcal{K}}(G(x^k))\right\rangle = 0,$$

therefore, $\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k)) = 0$, which implies (4). Now, to prove (3) and (5) note that by Lemma 2.1,

$$\begin{array}{rcl} \sigma^k \circ \Pi_{-\mathcal{K}}(G(x^k)) & = & \rho_k \Pi_{\mathcal{K}}(-G(x^k)) \circ \Pi_{-\mathcal{K}}(G(x^k)) \\ & = & -\sqrt{\rho_k}\Pi_{\mathcal{K}}(-G(x^k)) \circ \sqrt{\rho_k}\Pi_{\mathcal{K}}(-G(x^k)). \end{array}$$

Note also that $\lambda^k \bullet h(x^k) \to 0$ if, and only if $\rho_k \|h(x^k)\|^2$. Then, since x^k is a solution of (7), we have that

$$f(x^k) + \frac{1}{2}\|x^k - \bar{x}\|^2 + \frac{\rho_k}{2}(\|h(x^k)\|^2 + \|\Pi_{\mathcal{K}}(-G(x^k))\|^2) \leq f(\bar{x}).$$

Thus,

$$\frac{\rho_k}{2}(\|h(x^k)\|^2 + \|\Pi_{\mathcal{K}}(-G(x^k))\|^2) \leq f(\bar{x}) - f(x^k) - \frac{1}{2}\|x^k - \bar{x}\|^2.$$

Since $x^k \to \bar{x}$ we have by the above inequality that $\rho_k\|h(x^k)\|^2 \to 0$ and $\|\sqrt{\rho_k}\Pi_{\mathcal{K}}(-G(x^k))\|^2 \to 0$. Thus, by the above calculation and using Proposition 2.2 item ii) we conclude that $\lambda^k \bullet h(x^k) \to 0$ and $\sigma^k \circ \Pi_{-\mathcal{K}}(G(x^k)) \to 0$. Hence, $\bar{x}$ satisfies the Split-CAKKT condition. ■

In the next sections, we will discuss the algorithmic implications of the conditions we defined. One of the reasons that justifies the interest in optimization over symmetric cones is the fact that in practical problems one is often dealing with problems over mixed conic constraints. Thus, an algorithm and its convergence theory can be elegantly and concisely formulated, being specified only when implementing. In the next sections we will present two well-known algorithms in this context, the augmented Lagrangian algorithm and the interior point method. Our goal is to show that the augmented Lagrangian algorithm generates Split-CAKKT sequences (consequently, CAKKT sequences). For the interior point method we will prove that it generates CAKKT sequences. The reason for using CAKKT sequences and not Split-CAKKT sequences for the interior point method is that in this way we can consider a standard interior point method without needing to redefine the algorithm. In addition, Split-CAKKT is not natural in this context due to the explicit need of considering projections onto the cone.

4 Global convergence of an augmented Lagrangian method

We are interested in the augmented Lagrangian algorithm for NSOCP and NSDP discussed in [3] and [2, 9], respectively. In [9] it is shown that the algorithm generates AKKT sequences when applied to NSDPs, while in [2] the results are extended to CAKKT sequences. In [3], similar results appear for NSOCP. We will show that the result holds for a general symmetric cone by means of showing that the algorithm in fact generates Split-CAKKT sequences, which in particular implies CAKKT. Notice also that equality constraints were not considered in these previous works. Our results are also stronger than the ones obtained in [4, 16, 33], which can be applied to our context. Let us introduce the augmented Lagrangian algorithm.

Given $\rho > 0$, let $L_\rho : \mathbb{R}^n \times \mathbb{R}^\ell \times \mathcal{K} \to \mathbb{R}$ be the *augmented Lagrangian function* of (NSCP), defined as

$$L_\rho(x, \lambda, \sigma) := f(x) + \frac{\rho}{2}\left(\left\|h(x) + \frac{\lambda}{\rho}\right\|^2 + \left\|\Pi_{\mathcal{K}}\left(-G(x) + \frac{\sigma}{\rho}\right)\right\|^2\right),$$

whose partial derivative with respect to x is given by

$$\nabla_x L_\rho(x, \lambda, \sigma) = \nabla f(x) + Dh(x)^T(\lambda + \rho h(x)) - DG(x)^T \Pi_{\mathcal{K}}(\sigma - \rho G(x)). \tag{8}$$

To see this, it is enough to note that $\|\Pi_{\mathcal{K}}(-u)\|^2 = \text{dist}(u, \mathcal{K})^2$ is continuously differential with derivative with respect to u given by $2(u - \Pi_{\mathcal{K}}(u))$. See [23].

We assume now that a sequence $\{x^k\}_{k\in\mathbb{N}}$ can be generated by the previous algorithm and that it admits a limit point $\bar{x}$. In practice [14], this is usually done by considering additional box constraints that are carried over to the subproblems. This additional constraint does not hinder the theory developed here at all and assures that the augmented Lagrangian subproblems admit a global minimizer, and that any sequence generated by approximately solving the augmented Lagrangian subproblems will be bounded. We also assume that the limit point is feasible. We do this in view of [4, Proposition 1], since the limit points tend to be feasible as they are stationary for an infeasibility measure. We start by showing that, without additional smoothness assumptions, the algorithm generates an AGP sequence.

Theorem 4.1. *Let $\bar{x}$ be a feasible limit point of a sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by the augmented Lagrangian method. Then, $\bar{x}$ satisfies the AGP condition.*

Proof. From (9) and (8), we get

$$\|\nabla f(x^k) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k\| \leq \varepsilon_k,$$

Algorithm 1 General framework: augmented Lagrangian

Inputs: A sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ of positive scalars such that $\varepsilon_k \to 0$; a non-empty compact set $\mathcal{B} \subset \mathbb{R}^\ell \times \mathcal{K}$; real parameters $\tau > 1$, $\sigma \in (0,1)$, and $\rho_0 > 0$; and initial points $(\hat{\lambda}^0, \hat{\sigma}^0) \in \mathcal{B}$.

For every $k \in \mathbb{N}$:

1. Compute x^k such that
$$\|\nabla_x L_{\rho_k}(x^k, \hat{\lambda}^k, \hat{\sigma}^k)\| \leq \varepsilon_k; \tag{9}$$

2. Update the multipliers
$$\lambda^k := \hat{\lambda}^k + \rho_k h(x^k), \qquad \sigma^k := \Pi_{\mathcal{K}}\left(\hat{\sigma}^k - \rho_k G(x^k)\right),$$
and compute some $(\hat{\lambda}^{k+1}, \hat{\sigma}^{k+1}) \in \mathcal{B}$ (typically, the projection of (λ^k, σ^k) onto $\mathcal{B}$);

3. Define
$$V^k := \left(h(x^k), \Pi_{\mathcal{K}}\left(\frac{\hat{\sigma}^k}{\rho_k} - G(x^k)\right) - \frac{\hat{\sigma}^k}{\rho_k}\right);$$

4. If $k = 0$ or $\|V^k\| \leq \sigma\|V^{k-1}\|$, set $\rho_{k+1} := \rho_k$. Otherwise, choose some $\rho_{k+1} \geq \tau\rho_k$.

---

where $\sigma^k = \rho_k \Pi_{\mathcal{K}}\left(\rho_k^{-1}\hat{\sigma}^k - G(x^k)\right)$ and $\lambda^k = \rho_k\left(h(x^k) + \frac{\hat{\lambda}^k}{\rho_k}\right)$. Thus, (1) holds with $\sigma^k \in \mathcal{K}$ for every k. Taking a subsequence if necessary, we can suppose that $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$. Our aim is to show that $\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k))$ converges to zero. Note that, by setting $u_k = \rho_k^{-1}\hat{\sigma}^k - G(x^k)$, we have that $\sigma^k = \rho_k\Pi_{\mathcal{K}}(u_k)$ and then $\langle \rho_k\Pi_{\mathcal{K}}(u_k), \Pi_{\mathcal{K}}(-u_k)\rangle = \langle \rho_k\Pi_{\mathcal{K}}(u_k), -\Pi_{-\mathcal{K}}(u_k)\rangle = 0$. Thus, $\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k) - \rho_k^{-1}\hat{\sigma}^k) = 0$, since $\langle \sigma^k, \Pi_{\mathcal{K}}(G(x^k) - \rho_k^{-1}\hat{\sigma}^k)\rangle = 0$. Thus, we have that

$$\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k)) = \sigma^k \circ \left(\Pi_{\mathcal{K}}(G(x^k)) - \Pi_{\mathcal{K}}(G(x^k) - \rho_k^{-1}\hat{\sigma}^k)\right). \tag{10}$$

We examine two scenarios based on whether the sequence $\{\rho_k\}_{k\in\mathbb{N}}$ is bounded or not:

1. Suppose that $\rho_k \to \infty$.

   By applying Proposition 2.2 item i) in (10) and the non-expansitivity of the projection, we get
$$\|\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k))\|_1 \leq \|\sigma^k\|_2 \left\|\frac{\hat{\sigma}^k}{\rho_k}\right\|_2 = \left\|\Pi_{\mathcal{K}}(\rho_k^{-1}\hat{\sigma}^k - G(x^k))\right\|_2 \|\hat{\sigma}\|_2 \to 0,$$
   due to the boundedness of $\{\hat{\sigma}^k\}_{k\in\mathbb{N}}$ and feasibility of $\bar{x}$.

2. If $\{\rho_k\}_{k\in\mathbb{N}}$ is a bounded sequence.

   By the Step 3 of the algorithm, we see that V^k must converge to zero, and hence $V_2^k \to 0$, where V_2^k is the second component of V^k. Using Moreau's decomposition and the definition of V_2^k, we see that
$$\begin{aligned}\Pi_{\mathcal{K}}(G(x^k)) - \Pi_{\mathcal{K}}(G(x^k) - \rho_k^{-1}\hat{\sigma}^k) &= \Pi_{\mathcal{K}}(G(x^k)) - G(x^k) + \rho_k^{-1}\hat{\sigma}^k - \Pi_{\mathcal{K}}(\rho_k^{-1}\hat{\sigma}^k - G(x^k)) \\ &= \Pi_{\mathcal{K}}(G(x^k)) - G(x^k) - V_2^k \to 0.\end{aligned}$$
   The result follows from (10) noting that $\{\sigma^k\}_{k\in\mathbb{N}}$ must be bounded when $\{\rho_k\}$ is bounded.

Therefore, we conclude that $\bar{x}$ satisfies the AGP condition. ■

Next, let us present the main convergence result associated with Algorithm 1, namely that it generates Split-CAKKT sequences. To do so, we will assume that the point $\bar{x}$ satisfies the so-called *generalized Lojasiewicz inequality.*

Assumption 4.1. *We say that a point $\bar{x}$ satisfies the generalized Lojasiewicz inequality when there exist some neighborhood $\mathcal{U}$ of $\bar{x}$ and a continuous function $\psi(x) : \mathcal{U} \to \mathbb{R}$ such that $\psi(x) \to 0$ when $x \to \bar{x}$, and*
$$|\Psi(x) - \Psi(\bar{x})| \leq \psi(x)\|\nabla\Psi(x)\| \text{ for every } x \in \mathcal{U},$$
where $\Psi(x) = (1/2)\left(\|\Pi_{\mathcal{K}}(-G(x))\|^2 + \|h(x)\|^2\right)$.

This assumption corresponds to the one introduced in [10] in the context of nonlinear programming. This type of assumption has been widely used in optimization and other research areas, primarily due to its non-restrictive nature. For a detailed understanding of this assumption, we refer to [15, 32] and the references therein. For instance, the condition is satisfied if the functions are analytic. The function $\Psi(x)$ is an infeasibility measure of the point x and every limit point of a sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by Algorithm 1 is a stationary point for the problem of minimizing $\Psi(x)$. See [3, Theorem 5.1] and [4, Proposition 1].

Theorem 4.2. *Let $\bar{x}$ be a feasible limit point of a sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by the augmented Lagrangian algorithm that satisfies Assumption 4.1. Then $\bar{x}$ satisfies the Split-CAKKT condition.*

Proof. Let us assume, for simplicity, that the sequence $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$. By Theorem 4.1 it is enough to prove that $\lambda^k \bullet h(x^k) \to 0$ and $\sigma^k \circ \Pi_{-\mathcal{K}}(G(x^k)) = -\sigma^k \circ \Pi_{\mathcal{K}}(-G(x^k))$ also converges to zero. We will make use of the fact that

$$\rho_k \left(\|\Pi_{\mathcal{K}}(-G(x^k))\|^2 + \|h(x^k)\|^2\right) \to 0, \tag{11}$$

as proved in Theorem 4 of [4] in the context of general cones. To account for equality constraints, we refer to Theorem 4.1 of [10]. Let us divide this proof in two cases based on $\{\rho_k\}_{k\in\mathbb{N}}$ being a bounded sequence or not, similarly to the proof of Theorem 4.1.

1. Suppose that $\{\rho_k\}_{k\in\mathbb{N}}$ is unbounded. Computing $\sigma^k \circ \Pi_{\mathcal{K}}(-G(x^k))$, we see that

$$\begin{aligned}\sigma^k \circ \Pi_{\mathcal{K}}(-G(x^k)) &= \rho_k \Pi_{\mathcal{K}}(\rho_k^{-1}\hat{\sigma}^k - G(x^k)) \circ \Pi_{\mathcal{K}}(-G(x^k)) \\ &= \left(\Pi_{\mathcal{K}}(\hat{\sigma}^k - \rho_k G(x^k)) - \Pi_{\mathcal{K}}(-\rho_k G(x^k))\right) \circ \Pi_{\mathcal{K}}(-G(x^k)) + \rho_k \Pi_{\mathcal{K}}(-G(x^k))^2.\end{aligned} \tag{12}$$

   Observe that, by (11), we have $\rho_k\|\Pi_{\mathcal{K}}(-G(x^k))\|^2 \to 0$ and then $\rho_k \Pi_{\mathcal{K}}(-G(x^k))^2 \to 0$. Now, we proceed by showing that the first expression of (12) goes to zero. By the non-expansitivity of the projection, the boundedness of $\{\hat{\sigma}^k\}_{k\in\mathbb{N}}$ and Proposition 2.2 item i), we get that:

$$\left\|\left(\Pi_{\mathcal{K}}(\hat{\sigma}^k - \rho_k G(x^k)) - \Pi_{\mathcal{K}}(-\rho_k G(x^k))\right) \circ \Pi_{\mathcal{K}}(-G(x^k))\right\|_1 \leq \|\hat{\sigma}^k\|_2 \|\Pi_{\mathcal{K}}(-G(x^k))\|_2 \to 0.$$

   Thus, $\sigma^k \circ \Pi_{\mathcal{K}}(-G(x^k)) \to 0$. It remains to show that $\lambda^k \bullet h(x^k) \to 0$. This follows from (11) and the fact that $\lambda^k \bullet h(x^k) = \hat{\lambda}^k \bullet h(x^k) + \rho_k h(x^k) \bullet h(x^k)$, noting that $\{\hat{\lambda}^k\}_{k\in\mathbb{N}}$ is bounded and $\bar{x}$ is feasible.

2. Suppose that $\{\rho_k\}_{k\in\mathbb{N}}$ is a bounded sequence. Similarly to the proof of Theorem 4.1, we have that $\{\sigma^k\}_{k\in\mathbb{N}}$ and $\{\lambda^k\}_{k\in\mathbb{N}}$ are bounded sequences, and thus it follows trivially that $\sigma^k \circ \Pi_{\mathcal{K}}(-G(x^k))$ and $\lambda^k \bullet h(x^k)$ converge to zero.

This concludes the proof that $\bar{x}$ satisfies the Split-CAKKT condition. ∎

We note that in [33] the global convergence of the augmented Lagrangian method was studied in the context of symmetric cones without assuming strict complementarity. They do so by employing the strategy of reformulating the problem by using squared slack variables, which allows them to avoid the conic constraints and treat them with standard nonlinear programming tools. However, it is well known that squared slack variables should be avoided due to the numerical instability it can introduce. In addition, as a general rule, reformulating the constraints instead of working with them in their original format introduces additional structure to the problem in such a way that a constraint qualification may be lost. The only exception in the case of the squared slack reformulation is the nondegeneracy condition, which was employed in [33]. That is, even though the approach of [33] is very interesting for dealing with symmetric cones, it is doomed to fail when the original problem does not satisfy nondegeneracy. For the use of squared slack variables in nonlinear programming, we refer to [19], and the references therein, where they report some second-order complexity results and competitive numerical results by means of a squared slack approach. In our approach, instead of reformulating the problem as a nonlinear programming problem and avoiding the conic structure, we do not reformulate the problem at all. Instead, we embrace the conic constraints and its rich mathematical structure, making use of the spectral decomposition given by the Jordan algebraic environment, allowing us to present the global convergence of the algorithm without strict complementarity and under a constraint qualification much weaker than nondegeneracy (where, for instance, the set of Lagrange multipliers may be unbounded), which we will present in Section 6.

5 Global convergence of an interior point method defined by spectral functions

Interior point methods have received considerable attention since the work of Karmarkar [31] due to its numerous advantages, especially in the linear case. Firstly, they are computationally efficient, especially for some structured large-scale problems [24]. Secondly, these methods excel in handling inequality constraints. Thirdly, they exhibit numerical stability [24], which makes them suitable for ill-conditioned problems. Lastly, their versatility allows for their application in various optimization problems, including linear programming, quadratic programming, and semidefinite programming, among others. In the context of linear semidefinite programming, Alizadeh introduced an interior point method with polynomial convergence in [1], while Nesterov and Nemirovskii obtained the same result in [35] in a more general setting. In [41], convergence rates are obtained in the context of symmetric cones. In this section, we will introduce an interior point method for (NSCP). In particular, we will exploit the Jordan algebraic structure introduced by the symmetric cone constraints and the spectral decomposition. To start, let us introduce some basic notions of spectral functions.

Consider the set $\mathcal{C}$ of convex, lower-semicontinuous, and proper functions defined on $\mathbb{R}$ with values in $\mathbb{R}\cup\{+\infty\}$ such that ϕ is twice continuously differentiable on $(0,\infty)$ and

$$(0,\infty)\subset \text{dom }\phi\subset[0,\infty),\quad \lim_{t\to 0^+}\phi'(t)=-\infty \text{ and } \phi''(t)>0, \forall t>0. \tag{13}$$

Now, in the context of (NSCP), for $\phi\in\mathcal{C}$, we consider the function $\Phi:\mathcal{V}\to\mathbb{R}\cup\{\infty\}$ as

$$\Phi(z):=\sum_{j=1}^{r}\phi(\lambda_j(z)),\quad z\in\mathcal{V}. \tag{14}$$

It is well known that Φ is well-defined and does not depend on the spectral decomposition of z. From [13] and by Proposition 2.3 of [38], we have that Φ is a *spectral function.* Furthermore, we have the following:

Proposition 5.1. *The statements below hold:*

- *int(dom* Φ*)* $=$ *int*$(\mathcal{K})$*;*
- Φ *is strictly convex and twice differentiable on int*$(\mathcal{K})$*;*
- *Let* $z\in$ *int*$(\mathcal{K})$ *with the spectral decomposition* $z=\sum_{j=1}^{r}\lambda_j(z)c_j(z)$*. Then,*

$$\nabla\Phi(z)=\sum_{j=1}^{r}\phi'(\lambda_j(z))c_j(z).$$

In addition, we will assume that $\phi\in\mathcal{C}$ satisfies

$$\phi'(s)s\geq M,\ \text{for every } s\in(0,\delta) \tag{15}$$

for some $\delta>0$ and a scalar $M\in\mathbb{R}$. We mention that classical barrier functions such as $\phi_1(s)=s\log s-s$, $\phi_2(s)=-\frac{1}{r}s^r$ $(r\in(0,1))$, $\phi_3(s)=-\log s$ (logarithmic barrier) and $\phi_4(s)=s-\log s$ (modified logarithm barrier) satisfy (15). For other classical barrier functions, as $\phi_5(s)=s^{-1}$, (15) may fail. By Proposition 2.2 of [38], any $\phi\in\mathcal{C}$ satisfies (15) if $0\in dom(\phi)$.

In order to introduce the interior point method, let us, as usual, introduce slack variables in (NSCP). That is, we replace the constraint $G(x)\in\mathcal{K}$ with $G(x)=z$, where $z\in\mathcal{K}$. Given a barrier function $\phi\in\mathcal{C}$ and a parameter $\mu>0$, we aim at solving the following subproblem

$$\begin{aligned}
\underset{(x,z)\in\mathbb{R}^n\times\mathcal{V}}{\text{Minimize}}\quad & f(x)+\mu\Phi(z),\\
\text{subject to}\quad & h(x)=0,\\
& G(x)=z,
\end{aligned} \tag{16}$$

where Φ is given by (14) and z is assumed to lie in $\text{int}(\mathcal{K})$. The interior point algorithm is defined as follows:

In Step 1 of Algorithm 2, we will assume that the subproblem is solved with an algorithm that generates CAKKT sequences. This can be done by simply considering Newton's method to a suitable KKT system that includes complementarity for equality constraints and taking into account that z must always remain in the interior of $\mathcal{K}$. In particular, by Proposition 3.1, we assume that an iterate $(x^k,z^k)\in\mathbb{R}^n\times\text{int}(\mathcal{K})$

Algorithm 2 A framework for interior-point methods based on a general spectral functions

Let $\theta \in (0,1)$ and $\{\varepsilon_k\} \subset \mathbb{R}_+$ be a sequence of positive scalars with $\varepsilon_k \to 0$.
For every $k \in \mathbb{N}$:

1. Find an ε_k approximate minimizer $(x^k, z^k) \in \mathbb{R}^n \times \operatorname{int}(\mathcal{K})$ of subproblem (16) with $\mu = \mu^k$.

2. Set $\mu_{k+1} := \theta\mu_k$.

---

is computed in Step 1 satisfying the ε_k-CAKKT condition, which can be stated to problem (16) in the following form: there exist $\lambda^k \in \mathbb{R}^\ell$ and $\Lambda^k \in \mathcal{V}$ such that

$$\|\nabla f(x^k) + Dh(x^k)^T\lambda^k + DG(x^k)^T\Lambda^k\| \le \varepsilon_k, \quad \|\mu_k\nabla\Phi(z^k) - \Lambda^k\| \le \varepsilon_k, \tag{17}$$

and

$$\max\{\|h(x^k)\|, \|G(x^k) - z^k\|\} \le \varepsilon_k, \quad \|\lambda^k \bullet h(x^k)\| \le \varepsilon_k, \quad \|\Lambda^k \circ (G(x^k) - z^k)\| \le \varepsilon_k. \tag{18}$$

Remark 1. *Notice that Proposition 3.1 applies to equality constraints of the form $h(x) = 0$ where $h : \mathbb{R}^n \to \mathbb{R}^\ell$, while the equality constraint of* (16) *is of the form $(h(x), G(x) - z) = (0, 0)$, which maps $\mathbb{R}^n \times \mathcal{V}$ to $\mathbb{R}^\ell \times \mathcal{V}$. The modification in the theory to account for this case is straightforward by considering the Hadamard product $\bullet$ in $\mathbb{R}^\ell$ and the Jordan product $\circ$ in $\mathcal{V}$, which we decided not to present in the preliminary material for clarity of the presentation.*

Theorem 5.1. *Consider a sequence $\{(x^k, z^k)\}_{k\in\mathbb{N}}$ generated by Algorithm 2 such that (x^k, z^k) is an ε_k-CAKKT point of subproblem* (16) *with $\mu = \mu^k$. Then, any limit point of $\{x^k\}$ is a feasible CAKKT point of* (NSCP).

Proof. Let $\bar{x}$ be any limit point of $\{x^k\}$. By continuity of G, $\{z^k\}$ has a limit point $\bar{z}$ such that $(\bar{x}, \bar{z})$ is a limit point of $\{(x^k, z^k)\}$. Thus, taking an adequate subsequence, we assume that $\{(x^k, z^k)\}$ converges to $(\bar{x}, \bar{z})$ where $\bar{x}$ is feasible to (NSCP).

Since for each $k \in \mathbb{N}$, (x^k, z^k) is an ε_k-CAKKT point of (16) with $\mu = \mu^k$, there exist sequences $\{\lambda^k\} \subset \mathbb{R}^\ell$ and $\{\Lambda^k\} \subset \mathcal{V}$ such that (17) and (18) hold. We will show that $\{x^k\}$ is a CAKKT sequence of $\bar{x}$ associated with (NSCP) having as approximate multiplier sequence $\{\lambda^k\} \subset \mathbb{R}^\ell$ and $\{\sigma^k := \Pi_{\mathcal{K}}(-\Lambda^k)\} \subset \mathcal{K}$.

Indeed, from (17) and non-expansivity of the projection, we get

$$\|\Pi_{-\mathcal{K}}(\Lambda^k) - \Pi_{-\mathcal{K}}(\mu_k\nabla\Phi(z^k))\| \le \varepsilon_k \text{ and } \|\Pi_{\mathcal{K}}(\Lambda^k) - \Pi_{\mathcal{K}}(\mu_k\nabla\Phi(z^k))\| \le \varepsilon_k. \tag{19}$$

We proceed to analyze $\Pi_{\mathcal{K}}(\mu_k\nabla\Phi(z^k))$. As $z^k \in \operatorname{int}\mathcal{K}$, we have $z^k = \sum_{j=1}^r \lambda_j(z^k)c_j(z^k)$ with $\lambda_j(z^k) > 0$, $\forall j$. Thus, by Proposition 5.1, we get $\mu_k\nabla\Phi(z^k) = \sum_{j=1}^r \mu_k\phi'(\lambda_j(z^k))c_j(z^k)$ and hence,

$$\Pi_{\mathcal{K}}(\mu_k\nabla\Phi(z^k)) = \sum_{j=1}^r \mu_k \max\{\phi'(\lambda_j(z^k)), 0\}c_j(z^k).$$

Furthermore, after taking a further subsequence, we assume that $\{\lambda_j(z^k)\} \subset \mathbb{R}_{++}$ and $\{c_j(z^k)\} \subset \mathcal{V}$ converge for every j. This can be done since an idempotent has unit norm. Now, if $\{\lambda_j(z^k)\}$ converges to a positive scalar, then $\phi'(\lambda_j(z^k))$ also converges. So using that $\mu_k \to 0$, we get $\mu_k \max\{\phi'(\lambda_j(z^k)), 0\}c_j(z^k) \to 0$. In the case that $\lambda_j(z^k) \to 0$, property (13) implies that $\phi'(\lambda_j(z^k)) \to -\infty$ and hence $\mu_k \max\{\phi'(\lambda_j(z^k)), 0\}c_j(z^k) = 0$ for sufficiently large k. From these observations, we get $\Pi_{\mathcal{K}}(\mu_k\nabla\Phi(z^k)) \to 0$ and by (19), as $\varepsilon_k \to 0$, we get $\Pi_{\mathcal{K}}(\Lambda^k) \to 0$.

By Moreau's decomposition, we get $\sigma^k := \Pi_{\mathcal{K}}(-\Lambda^k) = \Pi_{\mathcal{K}}(\Lambda^k) - \Lambda^k$ and thus we conclude that

$$\begin{aligned}\nabla f(x^k) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k = \\ \nabla f(x^k) + Dh(x^k)^T\lambda^k + DG(x^k)^T\Lambda^k - DG(x^k)^T\Pi_{\mathcal{K}}(\Lambda^k) \to 0,\end{aligned} \tag{20}$$

since $\Pi_{\mathcal{K}}(\Lambda^k) \to 0$, $\{DG(x^k)\}$ is bounded and (17) holds. Note that $\lambda^k \bullet h(x^k) \to 0$ follows trivially from (18). It remains to show that $\sigma^k \circ G(x^k) \to 0$. In order to do this, let $w_1^k := \Lambda^k \circ (G(x^k) - z^k)$ and $w_2^k := \Pi_{-\mathcal{K}}(\Lambda^k) - \Pi_{-\mathcal{K}}(\mu_k\nabla\Phi(z^k))$. By (18), $\|w_1^k\| \le \varepsilon_k$ and by (19), $\|w_2^k\| \le \varepsilon_k$. So, $w_1^k \to 0$ and $w_2^k \to 0$. Now, using Moreau's decomposition we get

$$\begin{aligned}\sigma^k \circ G(x^k) &= -\Lambda^k \circ G(x^k) + \Pi_{\mathcal{K}}(\Lambda^k) \circ G(x^k) \\ &= -\Lambda^k \circ z^k - w_1^k + \Pi_{\mathcal{K}}(\Lambda^k) \circ G(x^k) \\ &= -\Pi_{-\mathcal{K}}(\Lambda^k) \circ z^k - \Pi_{\mathcal{K}}(\Lambda^k) \circ z^k - w_1^k + \Pi_{\mathcal{K}}(\Lambda^k) \circ G(x^k) \\ &= -\Pi_{-\mathcal{K}}(\mu_k\nabla\Phi(z^k)) \circ z^k - w_2^k \circ z^k - \Pi_{\mathcal{K}}(\Lambda^k) \circ z^k - w_1^k + \Pi_{\mathcal{K}}(\Lambda^k) \circ G(x^k).\end{aligned} \tag{21}$$

Clearly, as $\Pi_{\mathcal{K}}(\Lambda^k) \to 0$, we have $\Pi_{\mathcal{K}}(\Lambda^k) \circ z^k$, $w_2^k \circ z^k$ and $\Pi_{\mathcal{K}}(\Lambda^k) \circ G(x^k)$ converge to zero. Thus, by (21), we only need to show that $\Pi_{-\mathcal{K}}(\mu_k \nabla\Phi(z^k)) \circ z^k \to 0$. Since $z^k \in \operatorname{int} \mathcal{K}$, we get $z^k = \sum_{j=1}^r \lambda_j(z^k)c_j(z^k)$ with $\lambda_j(z^k) > 0$, $\forall j$, and by Proposition 5.1, $\mu_k \nabla\Phi(z^k) = \sum_{j=1}^r \mu_k \phi'(\lambda_j(z^k))c_j(z^k)$ and hence,

$$\Pi_{-\mathcal{K}}(\mu_k \nabla\Phi(z^k)) = \sum_{j=1}^r \mu_k \min\{\phi'(\lambda_j(z^k)), 0\}c_j(z^k). \tag{22}$$

Since $\{c_1(z^k), \dots, c_r(z^k)\}$ is a Jordan frame, $c_j(z^k) \circ c_i(z^k) = 0$, $i \neq j$ and $c_i^2(z^k) = c_i(z^k)$, $\forall i$. By (22), we get

$$\Pi_{-\mathcal{K}}(\mu_k \nabla\Phi(z^k)) \circ z^k = \sum_{j=1}^r \mu_k \min\{\phi^{'}(\lambda_j(z^k)), 0\}\lambda_j(z^k)c_j(z^k).$$

As done previously, we assume that $\{\lambda_j(z^k)\} \subset \mathbb{R}_{++}$ and $\{c_j(z^k)\} \subset \mathcal{V}$ converge for every j. If $\lambda_j(z^k)$ converges to a positive scalar, then $\phi'(\lambda_j(z^k))$ also converges. So, $\mu_k \to 0$ implies that $\mu_k \min\{\phi'(\lambda_j(z^k)), 0\}c_j(z^k) \to 0$. Now, if $\lambda_j(z^k) \to 0$, by (15) we have $\phi'(\lambda_j(z^k))\lambda_j(z^k) \geq M$ for some $M \in \mathbb{R}$ and k large enough. Thus, $0 \geq \mu_k \min\{\phi'(\lambda_j(z^k)), 0\}\lambda_j(z^k) \geq \mu_k \min\{M, 0\}$. As $\mu_k \to 0$, we get $\mu_k \min\{\phi'(\lambda_j(z^k)), 0\}\lambda_j(z^k) \to 0$ and hence $\Pi_{-\mathcal{K}}(\mu_k \nabla\Phi(z^k)) \circ z^k \to 0$. Substituting this into (21), we get $\sigma^k \circ G(x^k) \to 0$. From this and by (20), we conclude that $\bar{x}$ is a CAKKT point to (NSCP). ■

Notice that although it is somewhat unusual to include a complementarity constraint for equality constraints, it has been shown in [10] that this occurs somewhat naturally in the context of augmented Lagrangian methods. Our result suggests that this should be considered when solving the equality constrained subproblems of interior point methods in order to achieve a stronger global convergence result. This was already suggested in other contexts in [12] and [29].

6 The strength of the optimality condition

Our main objective in the previous sections was to introduce an optimality condition that does not depend on the validity of constraint qualification and that is generated by the augmented Lagrangian algorithm and the interior point method. Now, we will compare the results we obtained with previous ones known in the literature. To achieve this, we will introduce so-called *strict constraint qualifications* (SCQ). These conditions are properties satisfied by the constraints of the problem at a point in order to guarantee that when a sequential optimality condition holds at this point, the KKT conditions are also fulfilled. For more details in the context of nonlinear programming see [11].

We start by showing that our result is better than the one obtained under Robinson's constraint qualification. This CQ is strictly weaker than nondegeneracy and it is equivalent to the boundedness of the set of Lagrange multipliers. The definition for problem (NSCP) is as follows:

Definition 6.1. *We say that a feasible point $\bar{x}$ satisfies Robinson's condition if $Dh(\bar{x})$ is surjective and there exists $d \in \mathbb{R}^n$ such that*

$$G(\bar{x}) + DG(\bar{x})d \in int(\mathcal{K}),$$

$$Dh(\bar{x})d = 0.$$

Let us begin by showing that Robinson's condition is an SCQ associated with the Split-CAKKT condition. This is the statement of the next Theorem.

Theorem 6.1. *Let $\bar{x} \in \mathbb{R}^n$ be a point such that Split-CAKKT and Robinson's conditions are fulfilled. Then, $\bar{x}$ satisfies the KKT conditions.*

Proof. Let sequences $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$, $\{\lambda^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^\ell$, and $\{\sigma^k\}_{k\in\mathbb{N}} \subset \mathcal{K}$ be such that

$$\nabla f(x^k) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k \to 0, \tag{23}$$

and

$$\sigma^k \circ \Pi_{\mathcal{K}}(G(x^k)) \to 0, \sigma^k \circ \Pi_{-\mathcal{K}}(G(x^k)) \to 0. \tag{24}$$

If the sequence of Lagrange multipliers $\{\lambda^k\}_{k\in\mathbb{N}}$ and $\{\sigma^k\}_{k\in\mathbb{N}} \subset \mathcal{K}$ are bounded, it is easy to see that the limits given in (23) and (24) imply that $\bar{x}$ satisfies the KKT conditions. Otherwise, we can consider

$\alpha_k = \max\{1, \|\lambda^k\|, \|\sigma^k\|\}$, divide (23) and (24) by α_k, and take the limit in an appropriate subsequence to obtain:

$$Dh(\bar{x})^T\lambda - DG(\bar{x})^T\sigma = 0 \tag{25}$$

and $\sigma \circ \Pi_{\mathcal{K}}(G(\bar{x})) = 0$, $\sigma \circ \Pi_{-\mathcal{K}}(G(\bar{x})) = 0$ which implies $\sigma \circ G(\bar{x}) = 0$ and, in particular, $\langle G(\bar{x}), \sigma\rangle = 0$ for some $0 \neq (\lambda, \sigma) \in \mathbb{R}^\ell \times \mathcal{K}$. From Definition 6.1, we must have $\sigma \neq 0$ since otherwise (25) with $\lambda \neq 0$ contradicts the surjectivity of $Dh(\bar{x})$. Take $d \in \mathbb{R}^n$ as in Definition 6.1. Thus, we have by (25) that

$$\begin{aligned} 0 = \langle Dh(\bar{x})^T\lambda - DG(\bar{x})^T\sigma, d\rangle - \langle G(\bar{x}), \sigma\rangle &= \langle Dh(\bar{x})^T\lambda, d\rangle - \langle DG(\bar{x})^T\sigma, d\rangle - \langle G(\bar{x}), \sigma\rangle \\ &= \langle \lambda, Dh(\bar{x})d\rangle - \langle \sigma, G(\bar{x}) + DG(\bar{x})d\rangle. \end{aligned}$$

Thus, $\langle \sigma, G(\bar{x}) + DG(\bar{x})d\rangle = 0$ which contradicts the fact that $G(\bar{x}) + DG(\bar{x})d \in \mathrm{int}(\mathcal{K})$ and $0 \neq \sigma \in \mathcal{K}$. ■

Now let us introduce a condition which is weaker than Robinson's condition in such a way that it is still the case that any point satisfying this condition is such that every Split-CAKKT point is actually a KKT point. In fact, the condition we propose will be the weakest one with such property.

Definition 6.2. *We say that a feasible point $\bar{x}$ satisfies the Split-CAKKT-regularity condition if*

$$\limsup_{(x,\gamma)\to(\bar{x},0)} C(x,\gamma) \subset C(\bar{x}, 0),$$

that is, the set-valued function $(x, s) \in \mathbb{R}^n \times \mathbb{R} \rightrightarrows C(x, s)$ is outer semicontinuous at $(\bar{x}, 0)$, where

$$C(x,\gamma) := \left\{ Dh(x)^T\lambda - DG(x)^T\sigma \mid \max\{\|\sigma \circ \Pi_{\mathcal{K}}(G(x))\|, \|\sigma \circ \Pi_{-\mathcal{K}}(G(x))\|\} \leq \gamma, \lambda \in \mathbb{R}^\ell, \sigma \in \mathcal{K} \right\}.$$

Note that using a similar set, we can propose a similar constraint qualification (weaker than Robinson's condition) associated with the sequential optimality condition CAKKT. We omit the details. Next we state the minimality property of Split-CAKKT-regularity as follows:

Theorem 6.2. *A feasible point $\bar{x} \in \mathbb{R}^n$ satisfies Split-CAKKT-regularity if, and only if, for each continuously differentiable objective function f such that $\bar{x}$ satisfies the Split-CAKKT condition, $\bar{x}$ actually satisfies the KKT condition for* (NSCP) *with objective function f.*

Proof. Consider the feasible point $\bar{x} \in \mathbb{R}^n$ such that $\bar{x}$ satisfied Split-CAKKT-regularity and an objective function f such that $\bar{x}$ satisfies Split-CAKKT. Thus, there are sequences $\{x^k\}_{k\in\mathbb{N}} \to \bar{x}$, $\{\lambda^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^\ell$ and $\{\sigma^k\}_{k\in\mathbb{N}} \subset \mathcal{K}$ such that

$$\nabla f(x^k) + Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k \to 0, \ \sigma^k \circ \Pi_{\mathcal{K}}(G(x^k)) \to 0, \text{ and } \sigma^k \circ \Pi_{-\mathcal{K}}(G(x^k)) \to 0.$$

Then, since $Dh(x^k)^T\lambda^k - DG(x^k)^T\sigma^k \to -\nabla f(\bar{x})$ we have that $-\nabla f(\bar{x}) \in \limsup_{(x,\gamma)\to(\bar{x},0)} C(x,\gamma)$. Since $\bar{x}$ satisfies Split-CAKKT-regularity we have that $-\nabla f(\bar{x}) \in C(\bar{x}, 0)$. Thus, this clearly implies that $\bar{x}$ satisfies the KKT condition. Now we will prove the reciprocal implication. Take

$$w \in \limsup_{(x,s)\to(\bar{x},0)} C(x,s).$$

Then, there are sequences $x^k \to \bar{x}$ and $s^k \to 0$ such that $w^k \in C(x^k, s^k)$ for $w^k \to w$. Define $f(x) = -\langle w, x\rangle$ and we have that $\bar{x}$ satisfies the Split-CAKKT condition. Thus, $\bar{x}$ is a KKT point, which implies that $w = -\nabla f(\bar{x}) \in C(\bar{x}, 0)$ and the Split-CAKKT-regularity condition is satisfied. ■

While Theorem 6.2 gives that Robinson's condition implies Split-CAKKT-regularity, the simple example $G(x) := \begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix}$ with $\mathcal{K} = \mathbb{S}^2_+$ shows that the implication is strict.

7 Final Remarks

In this paper we provided a unified approach of the most relevant nonlinear conic programming problems that include second-order cones and semidefinite cones. We studied an augmented Lagrangian method and an interior point method where we showed a natural way to measure fulfillment of the complementarity condition by means of the underlying Jordan product. By means of introducing a new sequential optimality condition, defined in terms of the Jordan product, we show a strong global convergence property of the standard augmented Lagrangian method in the general context of symmetric cones. The condition introduced is natural enough such that a similar analysis is conducted in the context of an interior point method. We envision that by considering the Jordan algebraic structure of the problem one may obtain strong global convergence of other algorithms for symmetric cones such as the sequential quadratic programming algorithm studied in [36] and others.

References

[1] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5:13–51, 1995.

[2] R. Andreani, E. H. Fukuda, G. Haeser, D. O. Santos, and L. D. Secchin. On the use of Jordan Algebras for improving global convergence of an augmented Lagrangian method in nonlinear semidefinite programming. *Computational Optimization and Applications*, 79(3):633–648, 2021.

[3] R. Andreani, E. H. Fukuda, G. Haeser, D. O. Santos, and L. D. Secchin. Optimality conditions for nonlinear second-order cone programming and symmetric cone programming. *Journal of Optimization Theory and Applications*, 200(1):1–33, 2023.

[4] R. Andreani, W. Gómez, G. Haeser, L. M. Mito, and A. Ramos. On optimality conditions for nonlinear conic programming. *Mathematics of Operations Research*, 47(3):2160–2185, 2022.

[5] R. Andreani, G. Haeser, and J. Martínez. On sequential optimality conditions for smooth constrained optimization. *Optimization*, 60:627–641, 2011.

[6] R. Andreani, G. Haeser, L. Mito, and H. Ramírez. Weak notions of nondegeneracy in nonlinear semidefinite programming. *Mathematical Programming*, 2023. DOI: 10.1007/s10107-023-01970-4.

[7] R. Andreani, G. Haeser, L. M. Mito, and H. Ramírez. Sequential constant rank constraint qualifications for nonlinear semidefinite programming with algorithmic applications. *Set-Valued and Variational Analysis*, 31(1), Jan. 2023.

[8] R. Andreani, G. Haeser, L. D. Secchin, and P. J. S. Silva. New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences. *SIAM Journal on Optimization*, 29(4):3201–3230, 2019.

[9] R. Andreani, G. Haeser, and D. S. Viana. Optimality conditions and global convergence for nonlinear semidefinite programming. *Mathematical Programming*, 180(1-2):203–235, 2018.

[10] R. Andreani, J. M. Martínez, and B. F. Svaiter. A new sequential optimality condition for constrained optimization and algorithmic consequences. *SIAM Journal on Optimization*, 20(6):3533–3554, 2010.

[11] R. Andreani, J. M. Martínez, A. Ramos, and P. J. S. Silva. Strict constraint qualifications and sequential optimality conditions for constrained optimization. *Mathematics of Operations Research*, 43(3):693–717, 2018.

[12] R. Andreani, A. Ramos, A. A. Ribeiro, L. D. Secchin, and A. R. Velazco. On the convergence of augmented lagrangian strategies for nonlinear programming. *IMA Journal of Numerical Analysis*, 42:1733–1765, 2022.

[13] M. Baes. Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras. *Linear Algebra and its Applications*, 422:664–700, 2007.

[14] E. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization.* SIAM Publications, Philadelphia, PA, 2014.

[15] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Characterization of Lojasiewicz inequalities and applications: subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362:3319–3363, 2010.

[16] E. Börgens, C. Kanzow, P. Mehlitz, and G. Wachsmuth. New constraint qualifications for optimization problems in Banach spaces based on asymptotic KKT conditions. *SIAM Journal on Optimization*, 30(4):2956–2982, Jan. 2020.

[17] L. F. Bueno, G. Haeser, and F. N. Rojas. Optimality conditions and constraint qualifications for generalized Nash equilibrium problems and their practical implications. *SIAM Journal on Optimization*, 29(1):31–54, 2019.

[18] A. De Marchi, X. Jia, C. Kanzow, and P. Mehlitz. Constrained composite optimization and augmented Lagrangian methods. *Mathematical Programming*, 201(1-2):863–896, Sept. 2023.

[19] L. Ding and S. J. Wright. On squared-variable formulations. *arXiv:2310.01784*, 2023.

[20] J. Faraut and A. Korányi. *Analysis on Symmetric Cones.* Oxford University Press, New York, 1994.

[21] L. Faybusovich. Jordan-algebraic approach to convexity theorems for quadratic mappings. *SIAM Journal on Optimization*, 17(2):558–576, 2006.

[22] D. Fernandez and M. Solodov. Local convergence of exact and inexact augmented lagrangian methods under the second-order sufficient optimality condition. *SIAM Journal on Optimization*, 22:384–407, 2012.

[23] S. Fitzpatrick and R. R. Phelps. Differentiability of the metric projection in Hilbert space. *Transactions of the American Mathematical Society*, 270(2):483, 1982.

[24] J. Gondzio. Interior point methods 25 years later. *European Journal of Operational Research*, 218(3):587–601, 2012.

[25] M. S. Gowda. A Hölder type inequality and an interpolation theorem in Euclidean Jordan algebras. *Journal of Mathematical Analysis and Applications*, 474(1):248–263, 2019.

[26] G. Haeser, H. Liu, and Y. Ye. Optimality condition and complexity analysis for linearly-constrained optimization without differentiability on the boundary. *Mathematical Programming*, 178(1-2):263–299, Nov. 2019.

[27] P. Jordan, J. von Neumann, and E. Wigner. On an algebraic generalization of the quantum mechanical formalism. *Annals of Mathematics*, 35(1):26–64, 1934.

[28] C. Kanzow, A. B. Raharja, and A. Schwartz. Sequential optimality conditions for cardinality-constrained optimization problems with applications. *Computational Optimization and Applications*, 80:185–211, 2021.

[29] C. Kanzow and A. Schwartz. The price of inexactness: Convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited. *Mathematics of Operation Research*, 40:253–512, 2014.

[30] C. Kanzow and D. Steck. Augmented Lagrangian methods for the solution of generalized Nash equilibrium problems. *SIAM Journal on Optimization*, 26(4):2034–2058, Jan. 2016.

[31] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[32] S. Lojasiewicz. Sur les trajectories du gradient d'une fonction analytique. *in Seminari di Geometria 1982-1983, Universita di Bologna, Istituto di Geometria, Dipartamento di Matematica*, pages 115–117, 1984.

[33] B. Lourenço, E. Fukuda, and M. Fukushima. Optimality conditions for problems over symmetric cones and a simple augmented Lagrangian method. *Mathematics of Operations Research*, 43(4):1233–1251, 2018.

[34] Y. Lu and Y. Yuan. An interior-point trust-region algorithm for general symmetric cone programming. *SIAM Journal on Optimization*, 18(1):65–86, 2007.

[35] Y. Nesterov and A. Nemirovskii. Self-concordant functions and polynomial time methods in convex programming. *Central Economical and Mathematical Institute, USSR Academy of Science, Moscow,*, 1990.

[36] K. Okabe, Y. Yuya, and E. Fukuda. A revised sequential quadratic semidefinite programming method for nonlinear semidefinite optimization. *arXiv:2204.00369*, 2022.

[37] M. O'Neill and S. J. Wright. A log-barrier newton-CG method for bound constrained optimization with complexity guarantees. *IMA Journal of Numerical Analysis*, 41(1):84–121, Apr. 2020.

[38] H. Ramírez and D. Sossa. On the central paths in symmetric cone programming. *Journal of Optimization Theory and Applications*, 172:649–668, 2017.

[39] R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM Review*, 35:183–238, 1993.

[40] J. F. Sturm. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. *Optimization Methods and Software*, 17(6):1105–1154, 2002.

[41] T. Valkonen. Interior-proximal primal-dual methods. *Applied Analysis and Optimization*, 3:1–28, 2019.