First- and second-order optimality conditions for second-order cone and semidefinite programming under a constant rank condition

R. Andreani*   G. Haeser†   L. M. Mito†   H. Ramírez C.‡   T. P. Silveira†

July 9, 2021

Abstract

The well known constant rank constraint qualification [Math. Program. Study 21:110–126, 1984] introduced by Janin for nonlinear programming has been recently extended to a conic context by exploiting the eigenvector structure of the problem. In this paper we propose a more general and geometric approach for defining a new extension of this condition to the conic context. The main advantage of our approach is that we are able to recast the strong second-order properties of the constant rank condition in a conic context. In particular, we obtain a second-order necessary optimality condition that is stronger than the classical one obtained under Robinson's constraint qualification, in the sense that it holds for every Lagrange multiplier, even though our condition is independent of Robinson's condition.

Keywords: Constraint qualifications; Constant rank; Second-order optimality conditions; Second-order cone programming; Semidefinite programming.

1   Introduction

In the classical *nonlinear programming* (NLP) context, the so-called *constant rank constraint qualification* (CRCQ) [37] was first presented as a tool for stability analysis, which stood out for being independent of the usual *Mangasarian-Fromovitz constraint qualification* (MFCQ) and strictly weaker than the *linear independence constraint qualification* (LICQ). For instance, it has been applied with this purpose in NLP [29, 37, 47, 48, 50], *mathematical programs with equilibrium constraints* (MPEC) [34], *generalized equations* [35], and *bilevel optimization* [45, 60]. Also, it is the origin of several other constant rank-type conditions, such as the *constant positive linear dependence* [10, 12, 52] and the *constant rank of the subspace component* [11], which have been successfully applied in the convergence analysis of iterative algorithms. To name a few algorithms whose convergence theory relies on CRCQ and its variants, we point out: an augmented Lagrangian method [3, 13], a regularized interior point method [53], sequential quadratic programming methods for NLP [42, 52, 59] and MPEC [39], and some relaxation schemes for MPEC [36, 58]. In fact, a particularly interesting aspect of CRCQ that makes it suitable for supporting practical algorithms is the fact it can be roughly interpreted as a relaxation of LICQ that is able to separate the core information of the problem, ignoring redundant constraints. Moreover, all linear programming problems satisfy CRCQ, in contrast with LICQ and MFCQ.

Besides convergence of algorithms and stability analysis, CRCQ was used in several contexts, such as NLP [4, 13, 46], MPEC [33], vector optimization [44], and continuous-time NLP [49], for studying necessary second-order optimality conditions. One of the main goals of this paper is to bring such results to more general conic programming contexts, namely *nonlinear second-order cone programming* (NSOCP) and *nonlinear semidefinite programming* (NSDP). As far as we know, the best second-order results for these problems have appeared in a well-known paper by Bonnans, Cominetti, and Shapiro [21]. Basically, they derived no-gap second-order optimality conditions for problems over *second-order regular cones* [21, Definition 3], such as

---
*Department of Applied Mathematics, University of Campinas, Campinas-SP, Brazil. Email: andreani@ime.unicamp.br

†Department of Applied Mathematics, University of São Paulo, São Paulo-SP, Brazil.   Email: {ghaeser,leokoto,thiagops}@ime.usp.br

‡Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático (CNRS UMI 2807), Universidad de Chile, Santiago, Chile. Email: hramirez@dim.uchile.cl

NSDP and NSOCP, under the well-known *Robinson's CQ* (see (7) on page 5, or [54]), which is the natural extension of MFCQ to conic programming. In particular, their second-order necessary condition states that every local solution that satisfies Robinson's CQ must also satisfy the following: for every critical direction, there exists a Lagrange multiplier (possibly depending on this direction), such that a certain quadratic form is nonnegative with respect to such direction and multiplier. However, the second-order condition that is obtained under CRCQ in NLP replaces "there exists a Lagrange multiplier" with "for every Lagrange multiplier", which is stronger than the one of [21]. Although this stronger condition can be obtained from [21] after assuming that the Lagrange multiplier is unique, which is ensured by stronger constraint qualifications such as the *nondegeneracy* condition (see (8) on page 5), this assumption is often regarded as too stringent. To the best of our knowledge, no second-order result concerning every Lagrange multiplier, without assuming its uniqueness, has been presented so far in the literature of nonlinear conic programming. Moreover, no extension of CRCQ has been proposed for nonlinear conic programming until very recently.

In 2019, Zhang and Zhang [61] proposed an extension of CRCQ and its relaxed version [47] for NSOCP, but it was later discovered that their results were incorrect [5]. This event has motivated us to investigate other possible extensions of CRCQ to conic problems, and their properties. The first step in this direction was made in [6], for NSOCP and NSDP problems with multiple constraints. The idea of [6] is to rewrite some of the conic constraints as locally equivalent NLP constraints, whenever possible, and then jointly applying nondegeneracy and the NLP version of CRCQ to the resulting problem. Later, in [8], based on the ideas from [7], we improved this strategy by exploiting the eigenvector structure of the semidefinite cone to deal with the conic constraints that could not be rewritten as NLP constraints. This approach was also extended to NSOCP problems in [9]. In simple terms, the condition of [8, 9] demands the rank of some families of functions to remain constant along every sequence converging to the point of interest – roughly speaking, a constant rank "by paths" – therefore, this extension is highly specialized to deal with sequences generated by iterative algorithms, but since this rank may vary between paths, it is likely unsuitable for other purposes. Indeed, the focus of [8, 9] was the global convergence of a large class of algorithms to first-order stationary points, and no second-order results were provided in it. Nevertheless, it is reasonable to expect that CRCQ may have multiple independent and correct extensions, each one of them generalizing at least one important aspect of it, but perhaps not all of them.

A common feature of all previous attempts of extending CRCQ to a conic context is an approach based on re-characterizing the conic program and the nondegeneracy condition, trying to make them as similar to NLP and LICQ as possible, so the extension of CRCQ would come out straightforwardly. This is somehow understandable because, even in NLP, the CRCQ condition has never received a geometrical interpretation before. In this paper, we present a new geometrical characterization of CRCQ for NLP in terms of the faces of the nonnegative orthant, which suggests a natural extension of it to NSOCP and NSDP. A point that we should stress is that contrary to our previously mentioned works, the definition of CRCQ that we present here is very simple. We prove that this extension is a constraint qualification strictly weaker than nondegeneracy and independent of Robinson's CQ, as it should be, and we also compare it with the condition of [8, 9]. Then, as an application, we show that every local solution of the problem satisfies the strong second order optimality condition, provided our extension of CRCQ holds. Moreover, just as it happens in NLP, our result does not demand *a priori* any specific condition over the Lagrange multiplier set, besides nonemptiness.

The structure of this paper is as follows: Section 2 consists of a nonlinear conic programming review emphasizing some aspects of the theory that are not commonly discussed in the literature; in Section 3, we analyze CRCQ for NLP and we show how it can be interpreted in terms of the faces of the nonnegative orthant. In Sections 4 and 5, we propose extensions of CRCQ for NSOCP and NSDP, respectively, and we prove some of its properties. Finally, in Section 6, we conclude this paper with a short discussion and some ideas of prospective work.

We end this section by introducing some of our basic notation: throughout this paper, $\mathbb{E}$ will denote a finite-dimensional linear space equipped with the inner product $\langle \cdot, \cdot \rangle$; and for a given set $S \subseteq \mathbb{E}$, we will denote the *polar* of S by

$$S^\circ := \{z \in \mathbb{E} \mid \langle z, y \rangle \leq 0, \ \forall y \in S\}$$

and the *orthogonal complement* of S will be denoted by $S^\perp$. The notations $\mathrm{cl}(S)$, $\mathrm{int}(S)$, $\mathrm{bd}(S)$, and $\mathrm{bd}^+(S)$ stand for the topological closure, interior, boundary, and boundary excluding the origin of S in $\mathbb{E}$, respectively. The *smallest cone* that contains S will be denoted by $\mathrm{cone}(S)$, and the *smallest linear space* that contains S will be denoted by $\mathrm{span}(S)$. Moreover, we denote by $Dg(x)$ the derivative of a twice continuously differentiable function $g : \mathbb{R}^n \to \mathbb{E}$ at a given point $x \in \mathbb{R}^n$, and by $Dg(x)^T$ the adjoint of $Dg(x)$, which by definition satisfies $\langle Dg(x)d, z \rangle = \langle d, Dg(x)^T z \rangle$ for all $d \in \mathbb{R}^n$ and $z \in \mathbb{E}$. Similarly, $D^2g(x)$ denotes the second-order derivative of g at x, and the action of $D^2g(x)$ over $d_1, d_2 \in \mathbb{R}^n$ will be denoted by $D^2g(x)[d_1, d_2]$.

2 Common framework: nonlinear conic programming

In this section, we will review some classical results of convex analysis, and first- and second-order optimality conditions and constraint qualifications for NSOCP and NSDP. These problems are the cornerstones of two independent research fields, but they can also be seen as particular cases of a *nonlinear conic programming* (NCP) problem, given by

$$
\begin{aligned}
\text{Minimize} \quad & f(x), \\
\text{s.t.} \quad & g(x) \in \mathcal{K},
\end{aligned}
\tag{NCP}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{E}$ are twice continuously differentiable, and $\mathcal{K} \subseteq \mathbb{E}$ is a closed convex pointed cone that is assumed to be nonempty. We will use (NCP) as a framework to discuss the common traits of NSOCP and NSDP simultaneously, before moving to specific traits. Throughout the whole paper, we will denote the feasible set of (NCP) by $\Omega := \{ x \in \mathbb{R}^n \mid g(x) \in \mathcal{K} \}$.

Let us begin with two key ideas that underlie all the results of this paper: *reducibility* and *faces*. Recall from [24, Definition 3.135] that for any given linear spaces $\mathbb{E}$ and $\mathbb{F}$, a cone $\mathcal{K} \subseteq \mathbb{E}$ is said to be *reducible* (more precisely, C^2-*reducible*) at a point $y \in \mathcal{K}$, to a closed convex pointed cone $\mathcal{C} \subseteq \mathbb{F}$, if there exists a neighborhood $\mathcal{N}$ of y and a twice continuously differentiable reduction function $\Xi \colon \mathcal{N} \to \mathbb{F}$ (possibly depending on y) such that $\Xi(y) = 0$, $D\Xi(y)$ is surjective, and

$$
\mathcal{K} \cap \mathcal{N} = \{ z \in \mathcal{N} \mid \Xi(z) \in \mathcal{C} \}.
$$

In general, reductions are meant to be used as a simplification tool that allows one to interpret any point of $\mathcal{K}$ as a vertex of some other cone $\mathcal{C}$, and then extend the results obtained at $\mathcal{C}$ to $\mathcal{K}$ in a smooth way. In this work, we are also interested in the geometrical properties of the reduced cone $\mathcal{C}$ as well; in particular, in its faces.

To make a brief revision, we recall that F is a *face* of $\mathcal{C}$ if every open line segment that contains a point of F also has its extrema in F; that is, if for every $y \in F$ and every $z, w \in \mathcal{C}$ such that $y = \alpha z + (1 - \alpha)w$ for some $\alpha \in (0, 1)$, we have that $z, w \in F$. Further, when there exists some $\eta \in \mathcal{C}^\circ$ such that

$$
F = \mathcal{C} \cap \{\eta\}^\perp,
$$

that is, when F is the intersection between $\mathcal{C}$ and one of its supporting hyperplanes, we say that F is an *exposed face* of $\mathcal{C}$. Some cones, like the nonnegative orthant, the semidefinite cone, and the second-order cone, are *facially exposed*, meaning all of their faces are exposed. We use the notation $F \trianglelefteq \mathcal{C}$ to say that F is a face of $\mathcal{C}$.

Now, to contextualize our results, we will revisit the classical theory of NCP in the next section, with a special emphasis in the work of Guignard [32], and Bonnans, Cominetti, and Shapiro [21]. In particular, we stress some aspects of the NCP theory that are often disregarded in the literature.

2.1 Review of first-order optimality conditions

For any set $S \subseteq \mathbb{E}$ and any $z \in S$, recall the (Bouligand) *tangent cone* to S at z, defined as

$$
\mathcal{T}_S(z) := \left\{ y \in \mathbb{E} \mid \exists \{t_k\}_{k \in \mathbb{N}} \to 0^+, \exists \{y^k\}_{k \in \mathbb{N}} \to y \text{ such that } z + t_k y^k \in S \text{ for all } k \in \mathbb{N} \right\}.
$$

Our review of first-order constraint qualifications for (NCP) revolves around two particular cones: the tangent cone $\mathcal{T}_\Omega(\bar{x})$ to Ω at a feasible point $\bar{x} \in \Omega$, and the *linearized tangent cone*

$$
\mathcal{L}_\Omega(\bar{x}) := \{ d \in \mathbb{R}^n \mid Dg(\bar{x})d \in \mathcal{T}_\mathcal{K}(g(\bar{x})) \},
$$

where $\mathcal{T}_\mathcal{K}(g(\bar{x}))$ is the tangent cone to $\mathcal{K}$ at $g(\bar{x})$. The importance of these cones for our analyses lies on the necessary optimality conditions for (NCP) associated with them. Namely, given any local minimizer $\bar{x} \in \Omega$ of (NCP), it is easy to see that $\langle \nabla f(\bar{x}), d \rangle \geq 0$ for all $d \in \mathcal{T}_\Omega(\bar{x})$; that is,

$$
-\nabla f(\bar{x}) \in \mathcal{T}_\Omega(\bar{x})^\circ.
\tag{1}
$$

This is one of the simplest necessary optimality conditions, sometimes called the *first-order geometric necessary condition* for the optimality of $\bar{x}$. However, it may be difficult to use (1) when Ω does not admit an explicit characterization since $\mathcal{T}_\Omega(\bar{x})^\circ$ may not be easily computable in this case. The polar of $\mathcal{L}_\Omega(\bar{x})$, on the other hand, admits a practical description, as it is shown in the following lemma, extracted from the proof of [32, Theorem 2] by Guignard:

3

Lemma 2.1. *Let $\bar{x} \in \Omega$. Then, $\mathcal{L}_\Omega(\bar{x})^\circ = \mathrm{cl}(H(\bar{x}))$, where*

$$H(\bar{x}) := Dg(\bar{x})^T \mathcal{N}_\mathcal{K}(g(\bar{x})) = \left\{ Dg(\bar{x})^T z \mid z \in \mathcal{N}_\mathcal{K}(g(\bar{x})) \right\}, \tag{2}$$

and $\mathcal{N}_\mathcal{K}(g(\bar{x})) := \mathcal{T}_\mathcal{K}(g(\bar{x}))^\circ$ is the normal cone *to $\mathcal{K}$ at $g(\bar{x})$.*

Proof. By the bipolar theorem (see e.g. [24, Proposition 2.40]), it suffices to prove that $\mathcal{L}_\Omega(\bar{x}) = H(\bar{x})^\circ$. Take any direction $d \in \mathcal{L}_\Omega(\bar{x})$ and let $z \in \mathcal{T}_\mathcal{K}(g(\bar{x}))^\circ$. By definition, $Dg(\bar{x})d \in \mathcal{T}_\mathcal{K}(g(\bar{x}))$ and then

$$0 \geq \langle Dg(\bar{x})d, z \rangle = \langle d, Dg(\bar{x})^T z \rangle.$$

Thus, since z is arbitrary, we obtain that $d \in H(\bar{x})^\circ$; and since d is also arbitrary, it follows that $\mathcal{L}_\Omega(\bar{x}) \subseteq (H(\bar{x}))^\circ$. Conversely, assume that there exists a vector $v \in H(\bar{x})^\circ$ such that $v \notin \mathcal{L}_\Omega(\bar{x})$, that is, $Dg(\bar{x})v \notin \mathcal{T}_\mathcal{K}(g(\bar{x}))$. By the strong separation theorem (see e.g. [24, Theorem 2.14]), there exists a vector y such that $\langle y, Dg(\bar{x})v \rangle > 0$ and $\langle y, z \rangle < 0$, for all $z \in \mathcal{T}_\mathcal{K}(g(\bar{x}))$, that is, $y \in \mathcal{N}_\mathcal{K}(g(\bar{x}))$. Therefore, $Dg(\bar{x})^T y \in H(\bar{x})$, which is a contradiction with $\langle Dg(\bar{x})^T y, v \rangle > 0$, since $v \in H(\bar{x})^\circ$.
$\qquad\square$

Recall that since $\mathcal{K}$ is a closed convex cone, we have

$$\mathcal{N}_\mathcal{K}(g(\bar{x})) = \{ z \in \mathcal{K}^\circ \mid \langle g(\bar{x}), z \rangle = 0 \}.$$

Then, combining the first-order geometric necessary condition and Lemma 2.1 yields the following theorem, also by Guignard:

Theorem 2.1 (Theorem 2 of [32]). *Let $\bar{x} \in \Omega$ be a local minimizer of* (NCP). *If $\mathcal{T}_\Omega(\bar{x})^\circ = \mathcal{L}_\Omega(\bar{x})^\circ$ and $H(\bar{x})$ is closed, then there exists some $\bar{\lambda} \in \mathcal{K}^\circ$ such that*

$$\nabla f(\bar{x}) + Dg(\bar{x})^T \bar{\lambda} = 0 \quad \text{and} \quad \langle g(\bar{x}), \bar{\lambda} \rangle = 0. \tag{3}$$

Theorem 2.1 can be seen as the "dual form" of the first-order geometric condition (1), and any vector $\bar{\lambda} \in \mathcal{K}^\circ$ that satisfies the *Karush-Kuhn-Tucker conditions* (3) is called a *Lagrange multiplier* associated with $\bar{x}$. Moreover, the collection of all Lagrange multipliers associated with $\bar{x}$ will be denoted by $\Lambda(\bar{x})$, and when $\Lambda(\bar{x}) \neq \emptyset$ we say that $\bar{x}$ is a *KKT point* of (NCP).

The hypothesis of Theorem 2.1,

$$\mathcal{T}_\Omega(\bar{x})^\circ = \mathcal{L}_\Omega(\bar{x})^\circ \quad \text{and} \quad H(\bar{x}) \text{ is closed}, \tag{4}$$

is known in the literature as *Guignard's CQ*, and it is the weakest assumption that makes the KKT conditions necessary for the local optimality of $\bar{x}$, in the sense of: if $\Lambda(\bar{x}) \neq \emptyset$ for every continuously differentiable function f that has a local minimizer constrained to Ω at $\bar{x}$, then Guignard's CQ must also hold at $\bar{x}$ [31, Corollary 3.4]. Börgens et al. [25, Definition 5.11] defined Guignard's CQ for optimization problems in Banach spaces as a single equality

$$\mathcal{T}_\Omega(\bar{x})^\circ = H(\bar{x}),$$

which is equivalent to (4) due to Lemma 2.1. In NLP, Guignard's CQ is usually stated in the form $\mathcal{T}_\Omega(\bar{x})^\circ = \mathcal{L}_\Omega(\bar{x})^\circ$, since the closedness of $H(\bar{x})$ follows from the polyhedricity of $\mathbb{R}^m_+$. However, as it can be seen in the following example, the equality $\mathcal{T}_\Omega(\bar{x})^\circ = \mathcal{L}_\Omega(\bar{x})^\circ$ on its own may not ensure that $\Lambda(\bar{x}) \neq \emptyset$ when $H(\bar{x})$ is not closed.

Example 2.1. *Consider the following problem, presented in [2, Subsection 2.1]:*

$$\begin{aligned} \text{Minimize} \quad & f(x) := -x_2, \\ s.t. \quad & g(x) := (x_1, x_1, x_2) \in K_3, \end{aligned}$$

where K_3 is the three-dimensional second-order cone, given by

$$K_3 = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 \geq \sqrt{x_2^2 + x_3^2} \right\}.$$

Note that its feasible set is given by $\Omega = \{x \in \mathbb{R}^2 \mid x_1 \geq 0 \text{ and } x_2 = 0\}$, and that the point $\bar{x} = (0,0) \in \mathbb{R}^2$ is a local minimizer of it. Any Lagrange multiplier $\lambda := (\lambda_1, \lambda_2, \lambda_3) \in K_3^\circ$ associated with $\bar{x}$ must satisfy

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{5}$$

which implies that $\lambda_3 = 1$ and $\lambda_1 = -\lambda_2$. But since $\lambda \in K_3^\circ = -K_3$, then this vector must also satisfy $-\lambda_1 \geq \sqrt{\lambda_1^2 + 1}$, which does not have a solution with $\lambda_3 = 1$ and $\lambda_1 = -\lambda_2$. Therefore, $\bar{x}$ does not satisfy the KKT conditions. However, note that $\mathcal{T}_\Omega(\bar{x}) = \Omega = \mathcal{L}_\Omega(\bar{x})$ and consequently, $\mathcal{T}_\Omega(\bar{x})^\circ = \mathcal{L}_\Omega(\bar{x})^\circ$. Additionally, note that

$$H(\bar{x}) = \{(y_1 + y_2, y_3) \in \mathbb{R}^2 \mid (y_1, y_2, y_3) \in K_3^\circ\}$$

is not closed, because the sequence $\left\{\left(-\frac{1}{k}, -1\right)\right\}_{k \in \mathbb{N}}$ is contained in $H(\bar{x})$ since $\left(-\frac{1}{k} - k, k, -1\right) \in K_3^\circ$, $\forall k \in \mathbb{N}$, but its limit point $(0, -1)$ does not belong to $H(\bar{x})$.

The condition

$$\mathcal{T}_\Omega(\bar{x}) = \mathcal{L}_\Omega(\bar{x}) \quad \text{and} \quad H(\bar{x}) \text{ is closed,} \tag{6}$$

which implies Guignard's CQ, is known as *Abadie's CQ* (see also Börgens et al. [25, Definition 5.5]), and Example 2.1 tells us that the closedness of $H(\bar{x})$ cannot be omitted in this case, either. The reason why we emphasize this point is that, as far as we know, it appears that Abadie's CQ and Guignard's CQ are rarely seen in the literature of finite-dimensional conic programming problems other than NLP, and the closedness of $H(\bar{x})$ is rarely regarded in the study of constraint qualifications. In contrast, $H(\bar{x})$ plays an important role in our results.

In finite-dimensional conic contexts, the focus is usually on constraint qualifications that already imply $H(\bar{x})$ is closed without requiring it explicitly, such as Robinson's CQ, that holds at a given point $\bar{x} \in \Omega$ when

$$0 \in \text{int}(\text{Im}(Dg(\bar{x})) - \mathcal{K} + g(\bar{x})). \tag{7}$$

In particular, if $\mathcal{K}$ has nonempty interior, then Robinson's CQ holds at $\bar{x}$ if, and only if, there exists some $d \in \mathbb{R}^n$ such that

$$g(\bar{x}) + Dg(\bar{x})d \in \text{int}(\mathcal{K}).$$

Robinson's CQ is stronger than Abadie's CQ, and it implies that $\Lambda(\bar{x})$, besides being closed and convex, is also nonempty and bounded [24, Theorem 3.9] when $\bar{x}$ is a local minimizer of (NCP). If $\mathcal{K}$ is reducible at the point $g(\bar{x})$ to a cone $\mathcal{C}$ by the reduction function Ξ, then the constraint $g(x) \in \mathcal{K}$ is locally equivalent to the reduced constraint $\mathcal{G}(x) \in \mathcal{C}$, where $\mathcal{G} := \Xi \circ g$. In this case, Robinson's CQ holds at $\bar{x}$ for the original constraint if, and only if, it holds for the reduced constraint at the same point.

Another well-known constraint qualification in the context of conic programming is the *nondegeneracy* condition, which holds at $\bar{x}$ when

$$\text{Im}(Dg(\bar{x})) + \text{lin}(\mathcal{T}_\mathcal{K}(g(\bar{x}))) = \mathbb{E}, \tag{8}$$

where $\text{lin}(\mathcal{T}_\mathcal{K}(g(\bar{x}))) = \mathcal{T}_\mathcal{K}(g(\bar{x})) \cap -\mathcal{T}_\mathcal{K}(g(\bar{x}))$ denotes the largest linear space contained in $\mathcal{T}_\mathcal{K}(g(\bar{x}))$; that is, its *lineality space*. This CQ has first appeared in Shapiro and Fan's article [57] for NSDP, by the name *transversality*, and then it was generalized to NCP by Shapiro, in [56]. Nondegeneracy is strictly stronger than Robinson's CQ and it is known that if $\bar{x}$ is a local minimizer of (NCP) that satisfies nondegeneracy, then $\Lambda(\bar{x})$ is a singleton (see, for instance, [24, Proposition 4.75]). Moreover, if $\mathcal{K}$ is reducible, nondegeneracy is equivalent to the surjectivity of $D\mathcal{G}(\bar{x})$, as it can be easily deduced from the equality $\text{lin}(\mathcal{T}_\mathcal{K}(g(\bar{x}))) = \text{Ker}(D\Xi(g(\bar{x})))$; see [24, Section 4.6.1].

Due to their implications over the Lagrange multiplier set, nondegeneracy and Robinson's CQ are currently the most important CQs in the study of second-order optimality conditions for (NCP), which will be reviewed in the next subsection.

2.2   Second-order optimality conditions

Before starting, recall that the (inner) second-order tangent set to a nonempty set $S \subseteq \mathbb{E}$, at a point $z \in S$, in a direction $y \in \mathcal{T}_S(z)$, is defined by

$$\mathcal{T}_S^2(z, y) := \left\{ w \in \mathbb{E} \;\middle|\; z + ty + \frac{t^2}{2}w + o(t^2) \in S, \; \forall t > 0 \right\}, \tag{9}$$

which is closed for all such z, y, and S. In addition, if S is convex, then $\mathcal{T}_S^2(z, y)$ is also convex [24, Page 163]; and if S is second-order regular, as it is the case of the semidefinite cone and the second-order cone, then $\mathcal{T}_S^2(z, y)$ is nonempty [24, Page 202].

The role of second-order necessary optimality conditions is to provide additional information when first-order conditions are not meaningful enough; that is, along the directions in the cone

$$C(\bar{x}) := \{d \in \mathbb{R}^n \mid d \in \mathcal{T}_\Omega(\bar{x}), \; \langle \nabla f(\bar{x}), d \rangle = 0\},$$

which is often called the *cone of critical directions*, or simply, the *critical cone* of (NCP) at $\bar{x}$. Ben-Tal and Zowe [19] presented a *geometric second-order necessary optimality condition* for (NCP), stating that if $\bar{x}$ is a local minimizer of the problem, then

$$\langle \nabla f(\bar{x}), s \rangle + \langle \nabla^2 f(\bar{x})d, d \rangle \geq 0 \tag{10}$$

for every $d \in C(\bar{x})$ and every $s \in \mathcal{T}_\Omega^2(\bar{x}, d)$. Then, Kawasaki [41, Theorem 5.1] made the first advances to derive a "dual form" of (10) under Robinson's CQ assuming that $\mathcal{K}$ is a closed convex cone with nonempty interior. This result was later generalized and refined by Cominetti [26, Theorem 4.2] to the case where $\mathcal{K}$ is assumed to be a closed convex set. An important improvement was made afterwards by Bonnans, Cominetti, and Shapiro [21], who clarified several key points of the previous works, and obtained no-gap[1] second-order conditions, in particular, for second-order regular cones [21, Section 4]. Let us recall Bonnans, Cominetti, and Shapiro's necessary condition in the context of second-order regular cones:

Theorem 2.2 (Theorem 3.1 of [21]). *Let $\bar{x} \in \Omega$ be a local minimizer of* (NCP) *that satisfies Robinson's CQ. Then, for every direction $d \in C(\bar{x})$, there exists some $\bar{\lambda}_d \in \Lambda(\bar{x})$, such that*

$$d^T \nabla^2 f(\bar{x})d + \langle D^2 g(\bar{x})[d, d], \bar{\lambda}_d \rangle - \sigma(d, \bar{x}, \bar{\lambda}_d) \geq 0, \tag{11}$$

where

$$\sigma(d, \bar{x}, \bar{\lambda}_d) := \sup \left\{ \langle w, \bar{\lambda}_d \rangle \mid w \in \mathcal{T}_\mathcal{K}^2(g(\bar{x}), Dg(\bar{x})d) \right\} \tag{12}$$

is the support function of $\mathcal{T}_\mathcal{K}^2(g(\bar{x}), Dg(\bar{x})d)$ with respect to $\bar{\lambda}_d$.

The term $\sigma(d, \bar{x}, \bar{\lambda}_d)$ characterizes a possible curvature of the set $\mathcal{K}$ at $g(\bar{x})$ along $Dg(\bar{x})d$, and it is often called the "sigma-term" in the classical literature (for instance, in the book [24]). Since $\bar{\lambda}_d \in \Lambda(\bar{x})$ and $\mathcal{K}$ is convex, $\sigma(d, \bar{x}, \bar{\lambda}_d)$ is always nonnegative; and if $\mathcal{K}$ is polyhedral, as in NLP, then the sigma-term is zero everywhere. See also the discussion on polyhedricity and extended polyhedricity in [24, Section 3.2.3]. It is also worth mentioning that the second-order optimality condition of Theorem 2.2 can be derived without constraint qualifications, using Fritz John (generalized) multipliers [24, Theorem 3.50].

Although the condition of Theorem 2.2 is generally considered very natural and useful in the conic programming context and in NLP, a stronger condition where the Lagrange multiplier $\bar{\lambda}$ does not depend on d has several potential uses, in view of the NLP literature. This motivates the following definition:

Definition 2.1. *Let $\bar{x} \in \Omega$ be a KKT point and let $\bar{\lambda} \in \Lambda(\bar{x})$ be given. We say that the pair $(\bar{x}, \bar{\lambda})$ satisfies the* second-order condition *(SOC) when*

$$d^T \nabla^2 f(\bar{x})d + \langle D^2 g(\bar{x})[d, d], \bar{\lambda} \rangle - \sigma(d, \bar{x}, \bar{\lambda}) \geq 0, \tag{13}$$

for every $d \in C(\bar{x})$.

In NLP, the existence of some $\bar{\lambda} \in \Lambda(\bar{x})$ such that SOC holds for the pair $(\bar{x}, \bar{\lambda})$ is known as the *semi-strong second-order necessary optimality condition* [20]. Moreover, when SOC holds for every $\bar{\lambda} \in \Lambda(\bar{x})$, then we obtain what is known as the *strong second-order necessary optimality condition* [4]. However, while the condition of Theorem 2.2 is necessary for optimality under Robinson's CQ, this is not true, in general, for the strong and semi-strong conditions. In fact, there is a counterexample published by Baccari [17, Section 3] (see also Anitescu [14] and Arutyunov [15]), that shows that Robinson's CQ does not guarantee the existence of a $\bar{\lambda} \in \Lambda(\bar{x})$ such that the pair $(\bar{x}, \bar{\lambda})$ satisfies SOC (see also the extended version of [18] for details). Under nondegeneracy, the set $\Lambda(\bar{x})$ is a singleton and, in this case, the semi-strong and the strong second-order conditions both coincide with the condition of Theorem 2.2.

As far as we know, there is no result concerning the semi-strong and strong second-order conditions without assuming uniqueness of Lagrange multipliers in the literature of conic programming, except for NLP. In NLP, this has been addressed by means of constant rank-type constraint qualifications, which is also the path we will follow in this paper.

---

[1]The term "zero gap", or "no gap", is often used in NLP to refer to a second-order condition that does not require constraint qualifications to be necessary (using Fritz John/generalized Lagrange multipliers), and that becomes sufficient after replacing an inequality by a strict inequality. However, in this paper, we say that a condition has zero gap when it satisfies the latter, possibly subject to a constraint qualification, in the same way as [21].

3 Revisiting constant rank CQs in NLP

In this section we will revisit some constant rank-type conditions for NLP from a geometrical point of view, in order to extend it to a more general conic context later on. Consider the standard NLP problem

$$
\begin{array}{lll}
\text{Minimize} & f(x), & \\
\text{s.t.} & g_j(x) \geq 0, & j = 1, \ldots, m, \\
& g_j(x) = 0, & j = m+1, \ldots, m+p,
\end{array}
\tag{NLP}
$$

which is a particular case of (NCP) with $\mathbb{E} = \mathbb{R}^{m+p}$, $\mathcal{K} = \mathbb{R}_+^m \times \{0\}^p$, and $g(x) := (g_1(x), \ldots, g_{m+p}(x))$. As usual in NLP, given a feasible point $\bar{x}$ of (NLP), we will denote the set of active inequality constraints at $\bar{x}$ as $\mathcal{A}(\bar{x}) := \{j \in \{1, \ldots, m\} \mid g_j(\bar{x}) = 0\}$.

Now, let us recall Janin's constant rank constraint qualification as it was first presented in [37].

Definition 3.1 (CRCQ [37])**.** *Let $\bar{x}$ be a feasible point of (NLP). We say that the* constant rank constraint qualification for NLP *(CRCQ) holds at $\bar{x}$ if there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ such that, for every subset $J \subseteq \mathcal{A}(\bar{x}) \cup \{m+1, \ldots, m+p\}$, the rank of the family $\{\nabla g_j(x)\}_{j \in J}$ remains constant for all $x \in \mathcal{V}$.*

To prove that CRCQ is a constraint qualification, Janin proved that it implies $\mathcal{L}_\Omega(\bar{x}) \subseteq \mathcal{T}_\Omega(\bar{x})$, which in turn implies Abadie's CQ in NLP. His proof is what motivates the requirement to consider every subset J of $\mathcal{A}(\bar{x}) \cup \{m+1, \ldots, m+p\}$ in Definition 3.1; indeed, after picking a direction

$$
d \in \mathcal{L}_\Omega(\bar{x}) = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \nabla g_j(\bar{x})^T d \geq 0, \; j \in \mathcal{A}(\bar{x}), \\ \nabla g_j(\bar{x})^T d = 0, \; j \in \{m+1, \ldots, m+p\} \end{array} \right\},
$$

in order to prove that $d \in \mathcal{T}_\Omega(\bar{x})$, it is sufficient to have the constant rank assumption for the constraints that correspond to the indices $j \in \mathcal{A}(\bar{x})$ such that $\nabla g_j(\bar{x})^T d = 0$. Since those indices depend on d, and they are not determined *a priori*, one considers all possibilities. However, as it was noted several years later by Minchenko and Stakhovski [47], taking subsets of the equality constraints is completely superfluous, even for Janin's proof. The "correct" definition of CRCQ was then presented in [47] as a relaxed version of CRCQ.

Definition 3.2 (RCRCQ [47])**.** *Let $\bar{x}$ be a feasible point of (NLP). We say that* relaxed constant rank constraint qualification for NLP *(RCRCQ) holds at $\bar{x}$ if there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ such that, for every subset $J \subseteq \mathcal{A}(\bar{x})$, the rank of the family $\{\nabla g_j(x)\}_{j \in J \cup \{m+1, \ldots, m+p\}}$ remains constant for all $x \in \mathcal{V}$.*

In order to bring these CQs to the conic setting, our approach in this manuscript consists first in generalizing two key ideas of NLP: the notion of "active constraints" and the notion of "subsets of indices of active constraints". The former can be interpreted in the general context as a consequence of reducibility. Indeed, for any given $\bar{x} \in \Omega$, let $s := |\mathcal{A}(\bar{x})|$ and note that $\mathbb{R}_+^m \times \{0\}^p$ is reducible at $g(\bar{x})$ to the cone

$$
\mathcal{C} := \mathbb{R}_+^s \times \{0\}^p
$$

in a neighborhood $\mathcal{N}$ of $g(\bar{x})$ by the mapping $\Xi \colon \mathcal{N} \to \mathbb{R}^{s+p}$ such that

$$
\Xi(y) := (y_j)_{j \in \mathcal{A}(\bar{x}) \cup \{m+1, \ldots, m+p\}}
$$

for every $y \in \mathcal{N}$, and in this case the reduced constraint function of (NLP) at $\bar{x}$ takes the form

$$
\mathcal{G}(x) := \Xi(g(x)) = (g_j(x))_{j \in \mathcal{A}(\bar{x}) \cup \{m+1, \ldots, m+p\}}.
\tag{14}
$$

Therefore, in NLP, reducing the problem is essentially the same as simply disregarding inactive constraints around the point $\bar{x}$. The notion of "subsets of indices of the active constraints", on the other hand, can be interpreted in terms of faces.

It is easy to see that every face of $\mathbb{R}_+^s$ can be written in terms of a unique subset of the canonical vectors of $\mathbb{R}^s$, which we will denote by $c_1, \ldots, c_s$. That is, $F \trianglelefteq \mathbb{R}_+^s$ if, and only if, there exists some $J \subseteq \{1, \ldots, s\}$ such that

$$
F = \mathbb{R}_+^s \bigcap_{j \in J} \{c_i\}^\perp,
\tag{15}
$$

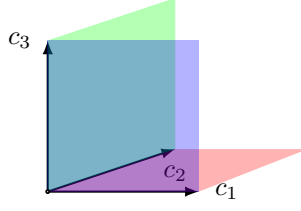where F and J are clearly in a one-to-one correspondence.

Figure 1: Faces of $\mathbb{R}^3_+$

For example, in Figure 1, the vertex of $\mathbb{R}^3_+$ corresponds to $J = \{1, 2, 3\}$; the one-dimensional faces cone(c_1), cone(c_2), and cone(c_3) correspond to $J = \{2, 3\}$, $J = \{1, 3\}$, and $J = \{1, 2\}$, respectively; the left, front, and bottom two-dimensional faces correspond to $J = \{1\}$, $J = \{2\}$, and $J = \{3\}$, respectively; and $\mathbb{R}^3_+$ itself corresponds to $J = \emptyset$.

Thus, considering all subsets of active constraints at $\bar{x}$ is the same as considering all faces of the reduced cone $\mathcal{C} = \mathbb{R}^s_+ \times \{0\}^p$. This discussion suggests a natural characterization of RCRCQ in terms of the faces of the reduced cone, as follows:

Proposition 3.1. *Let $\bar{x}$ be a feasible point of (NLP). Then, RCRCQ holds at $\bar{x}$ if, and only if, there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ such that, for each $F \trianglelefteq \mathbb{R}^{|\mathcal{A}(\bar{x})|}_+ \times \{0\}^p$, the dimension of*

$$D\mathcal{G}(x)^T[F^\perp]$$

remains constant for every $x \in \mathcal{V}$, where $\mathcal{G}$ is as defined in (14).

Proof. Let $s := |\mathcal{A}(\bar{x})|$ and, without loss of generality, let us assume that $\mathcal{A}(\bar{x}) = \{1, \ldots, s\}$. Moreover, let $c_1, \ldots, c_{s+p}$ be the canonical basis of $\mathbb{R}^{s+p}$, and let $F \trianglelefteq \mathbb{R}^s_+ \times \{0\}^p$. Note that $F = R \times \{0\}^p$, where $R \trianglelefteq \mathbb{R}^s_+$. Then, there exists some $J \subseteq \{1, \ldots, s\}$ such that

$$F = \left(\mathbb{R}^s_+ \bigcap_{j \in J} \{c_i\}^\perp \right) \times \{0\}^p,$$

which implies

$$F^\perp = R^\perp \times \mathbb{R}^p = \mathrm{span}\left(\{c_j \mid j \in J \cup \{s+1, \ldots, s+p\}\} \right),$$

so

$$D\mathcal{G}(x)^T[F^\perp] = \mathrm{span}(\{D\mathcal{G}(x)^T c_j\}_{j \in J \cup \{s+1, \ldots, s+p\}}) = \mathrm{span}(\{\nabla g_j(x)\}_{j \in J \cup \{m+1, \ldots, m+p\}}). \quad (16)$$

Consequently,

$$\dim(Dg(x)^T[F^\perp]) = \mathrm{rank}(\{\nabla g_j(x)\}_{j \in J \cup \{m+1, \ldots, m+p\}}).$$

The conclusion follows from the one-to-one correspondence between F and J.

$\square$

The equivalent form of RCRCQ presented in Proposition 3.1 allows us to visualize what it actually describes, geometrically. Indeed, recall that $\mathbb{R}^n = D\mathcal{G}(x)^{-1}(\mathrm{span}(F)) + (D\mathcal{G}(x)^{-1}(\mathrm{span}(F)))^\perp$ and it is elementary to see that

$$(D\mathcal{G}(x)^{-1}(\mathrm{span}(F)))^\perp = D\mathcal{G}(x)^T[F^\perp].$$

This implies the following relation:

$$\dim(D\mathcal{G}(x)^{-1}(\mathrm{span}(F))) + \dim(D\mathcal{G}(x)^T[F^\perp]) = n.$$

Thus, RCRCQ can be equivalently stated as the constant dimension of $D\mathcal{G}(x)^{-1}(\mathrm{span}(F))$ for every $x \in \mathcal{V}$ at each $F \trianglelefteq \mathcal{C} = \mathbb{R}^{|\mathcal{A}(\bar{x})|}_+ \times \{0\}^p$. The set $D\mathcal{G}(x)^{-1}(\mathrm{span}(F))$, on the other hand, can be regarded as a "linear approximation" of $\mathcal{G}^{-1}(\mathcal{C})$ around $\bar{x}$, since $D\mathcal{G}(x)$ is the best linear approximation of $\mathcal{G}$ at $x \in \mathcal{V}$ and, similarly, the faces of $\mathcal{C}$ can also be seen as "linear approximations" of it at $\mathcal{G}(\bar{x})$. In fact, each face induces a potentially different linear approximation of $\mathcal{G}^{-1}(\mathcal{C})$, which in turn coincides with Ω around $\bar{x}$. So roughly speaking: RCRCQ holds at $\bar{x}$ when the dimension of every linear approximation of the feasible set Ω at $\bar{x}$ is invariant to small perturbations. In particular, defining $g_J(x) := (g_j(x))_{j \in J \cup \{m+1, \ldots, m+p\}}$ for every $J \subseteq \mathcal{A}(\bar{x})$, this characterization is equivalent to the constant dimension of $\mathrm{Ker}(Dg_J(x))$ for all x in a neihborhood of $\bar{x}$ at every $J \subseteq \mathcal{A}(\bar{x})$, which can also be trivially seen from the original definition of RCRCQ.

Note that the characterization of RCRCQ from Proposition 3.1 and the discussion above do not appear to be limited to the context of NLP, contrary to its original definition. In the next two sections, we will prove that the same idea can be applied to NSOCP and NSDP, respectively, giving rise to new constraint qualifications.

Remark 3.1. *It is possible to obtain a characterization of CRCQ in the same style of Proposition 3.1. To do this, it suffices to reformulate the equality constraints $g_j(x) = 0$ as a pair of inequality constraints $g_j(x) \geq 0$ and $-g_j(x) \geq 0$, for $j \in \{m+1, \ldots, m+p\}$. That is, consider $\mathcal{K} := \mathbb{R}^m_+ \times \mathbb{R}^p_+ \times \mathbb{R}^p_+$ and $g(x) := (g_1(x), \ldots, g_{m+p}(x), -g_{m+1}(x), \ldots, -g_{m+p}(x))$ in Proposition 3.1.*

In view of Remark 3.1, we see that there are multiple ways of dealing with equality constraints in our approach, and they are not all equivalent. The suitability of each approach may depend on the application, but we highlight that our approach is able to deal with equality constraints regardless of how they are modelled. For simplicity, equality constraints are omitted in our exposition. See also Remarks 4.2 and 5.3. In the following two sections, we extend the ideas of this section to NSOCP and NSDP.

4   Nonlinear second-order cone programming

In this section, we consider the following problem:

$$\begin{aligned} \text{Minimize} \quad & f(x), \\ \text{s.t.} \quad & g_j(x) \in K_{m_j}, \quad j = 1, \ldots, q, \end{aligned} \qquad \text{(NSOCP)}$$

where $K_{m_j} := \{(z_0, \widehat{z}) \in \mathbb{R} \times \mathbb{R}^{m_j - 1} \mid z_0 \geq \|\widehat{z}\|\}$ when $m_j > 1$ and $K_1 = \{x \in \mathbb{R} \mid x \geq 0\}$. Since K_{m_j} is self-dual, we have that $z \in K_{m_j}^\circ$ if, and only if, $-z \in K_{m_j}$, for any $j = 1, \ldots, q$. Also, note that (NSOCP) can be seen as a particular case of (NCP) with

$$\mathcal{K} := K_{m_1} \times \ldots \times K_{m_q} \quad \text{and} \quad g(x) := (g_1(x), \ldots, g_q(x)).$$

Given a feasible point $\bar{x} \in \Omega$, let us define the following index sets:

$$\begin{aligned} I_{\text{int}}(\bar{x}) &:= \{j \in \{1, \ldots, q\} \mid g_j(\bar{x}) \in \text{int}(K_{m_j})\}, \\ I_B(\bar{x}) &:= \{j \in \{1, \ldots, q\} \mid g_j(\bar{x}) \in \text{bd}^+(K_{m_j})\}, \\ I_0(\bar{x}) &:= \{j \in \{1, \ldots, q\} \mid g_j(\bar{x}) = 0\}, \end{aligned}$$

which consist of the indices of the constraints that hit the interior, the boundary excluding zero, and the vertex of their respective cones. For simplicity, we will omit equality constraints; we should mention, nevertheless, that our results can be easily adapted to deal with equality constraints — see Remark 4.2 for details. As another measure to avoid cumbersome notation, we will assume that $I_B(\bar{x}) = \{1, \ldots, |I_B(\bar{x})|\}$; this assumption will often be recalled throughout this section.

Following Bonnans and Ramírez [22], for any given $\bar{x} \in \Omega$, we see that $\mathcal{K}$ is reducible to

$$\mathcal{C} := \prod_{j \in I_0(\bar{x})} K_{m_j} \times \mathbb{R}^{|I_B(\bar{x})|}_+ \tag{17}$$

in a neighborhood $\mathcal{N}_1 \times \ldots \times \mathcal{N}_q$ of $g(\bar{x})$ by the function $\Xi := (\Xi_j)_{j \in I_0(\bar{x}) \cup I_B(\bar{x})}$, where $\Xi_j : \mathcal{N}_j \to \mathbb{R}^{m_j}$ is the identity function for every $j \in I_0(\bar{x})$, and $\Xi_j : \mathcal{N}_j \to \mathbb{R}$ is given by

$$\Xi_j(y) := y_0 - \|\widehat{y}\| \tag{18}$$

for every $j \in I_B(\bar{x})$, and every $y \in \mathbb{R}^{m_j}$. This leaves us with the reduced constraint

$$\mathcal{G}(x) \in \mathcal{C},$$

where $\mathcal{G}(x) := \Xi(g(x)) = (\mathcal{G}_j(x))_{j \in I_0(\bar{x}) \cup I_B(\bar{x})}$,

$$\mathcal{G}_j(x) := \Xi_j(g_j(x)) = \begin{cases} g_j(x), & \text{if } j \in I_0(\bar{x}), \\ \phi_j(x), & \text{if } j \in I_B(\bar{x}), \end{cases} \tag{19}$$

and $\phi \colon \mathbb{R}^n \to \mathbb{R}^{|I_B(\bar{x})|}$ has its j-th component given by

$$\phi_j(\bar{x}) := [g_j(x)]_0 - \|\widehat{g_j(\bar{x})}\|. \tag{20}$$

Note that $g(x) \in \mathcal{K}$ if, and only if, $\mathcal{G}(x) \in \mathcal{C}$ for every x sufficiently close to $\bar{x}$.

By [22, Lemma 25], we see that the linearized cone of the original constraints of (NSOCP) at a given $\bar{x} \in \Omega$ can be computed as

$$\mathcal{L}_\Omega(\bar{x}) = \left\{ \ d \in \mathbb{R}^n \ \left| \ \begin{array}{cc} Dg_j(\bar{x})d \in K_{m_j}, & j \in I_0(\bar{x}) \\ D\phi(\bar{x})d \in \mathbb{R}_+^{|I_B(\bar{x})|} & \end{array} \right. \right\},$$

and that it coincides with the linearized cone of the reduced constraint at $\bar{x}$. Moreover, it follows from [1, Lemma 15] that for each $j = I_{\text{int}}(\bar{x}) \cup I_B(\bar{x})$, we have $\langle \bar{\lambda}_j, g_j(\bar{x}) \rangle = 0$, if, and only if,

$$\bar{\lambda}_j = \begin{cases} 0, & \text{if } j \in I_{\text{int}}(\bar{x}), \\ \frac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} R_{m_j} g_j(\bar{x}), & \text{if } j \in I_B(\bar{x}), \end{cases} \tag{21}$$

where R_{m_j} is a matrix defined as

$$R_{m_j} := \begin{bmatrix} 1 & 0 \\ 0 & -\mathbb{I}_{m_j-1} \end{bmatrix}, \tag{22}$$

and $\mathbb{I}_{m_j-1}$ is the $(m_j - 1) \times (m_j - 1)$ identity matrix. Therefore, still following [22], the point $\bar{x}$ satisfies the KKT conditions with respect to the constraint $g(x) \in \mathcal{K}$ if, and only if, there exist some vectors $\bar{\lambda}_j \in K_{m_j}^\circ$, $j \in I_0(\bar{x}) \cup I_B(\bar{x})$, such that:

$$\nabla f(\bar{x}) + \sum_{j \in I_0(\bar{x})} Dg_j(\bar{x})^T \bar{\lambda}_j + \sum_{j \in I_B(\bar{x})} \frac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} Dg_j(\bar{x})^T R_{m_j} g_j(\bar{x}) = 0, \tag{23}$$

which also coincides with the KKT conditions with respect to the reduced constraint $\mathcal{G}(x) \in \mathcal{C}$. In fact, note that for each $j \in I_B(\bar{x})$, the reduced Lagrange multiplier with respect to the reduced constraint $\phi_j(x) \geq 0$ is simply $[\bar{\lambda}_j]_0$.

With this in mind, we are ready to present our extension of CRCQ (and RCRCQ) to NSOCP inspired by the characterization of Proposition 3.1.

4.1 A facial constant rank constraint qualification for NSOCP

Recall that, for each $j = 1, \ldots, q$, the cone K_{m_j} is facially exposed, meaning every $F \trianglelefteq K_{m_j}$ can be written as the intersection of one of its supporting hyperplanes, say $\{\eta\}^\perp$ with $\eta \in K_{m_j}$. In fact, although K_{m_j} has infinitely many faces when $m_j > 2$, they are limited to only three types:

- The vertex, $\{0\}$, which can be characterized by any $\eta \in \text{int}(K_{m_j})$;

- The cone K_{m_j} itself, which is characterized by $\eta = 0$;

- A ray at the boundary of K_{m_j}, starting at the vertex and passing through a point $z \in \text{bd}^+(K_{m_j})$, which can be characterized by any $\eta \in \text{cone}(R_{m_j} z) \setminus \{0\}$.

Moreover, every $F \trianglelefteq \mathcal{C}$ has the form

$$F = \left(\prod_{j \in I_0(\bar{x})} F_j \right) \times R,$$

where $F_j \trianglelefteq K_{m_j}$ for every $j \in I_0(\bar{x})$, and $R \trianglelefteq \mathbb{R}_+^{|I_B(\bar{x})|}$. Then, for every $x \in \mathbb{R}^n$, sufficiently close to $\bar{x}$, we have

$$D\mathcal{G}(x)^T[F^\perp] = \sum_{j \in I_0(\bar{x})} Dg_j(x)^T[F_j^\perp] + D\phi(x)^T[R^\perp],$$

where $\phi(x) := (\phi_j(x))_{j \in I_B(\bar{x})}$. This motivates the following definition:

Definition 4.1. *Let $\bar{x}$ be a feasible point of* (NSOCP)*. We say that the* facial constant rank property *holds at $\bar{x}$ if there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ such that for each $F \trianglelefteq \mathcal{C}$, the dimension of $D\mathcal{G}(x)^T[F^\perp]$ remains constant for all $x \in \mathcal{V}$, where $\mathcal{G}$ is given by* (19) *and $\mathcal{C}$ is given by* (17)*.*

Recall the discussion after Proposition 3.1 and note that Definition 4.1 can be equivalently stated in terms of the constant dimension of $D\mathcal{G}(x)^{-1}(\text{span}(F))$ for all $x \in \mathcal{V}$ and every $F \trianglelefteq \mathcal{C}$. That is, the facial constant rank property holds at $\bar{x}$ when the dimension of every linear approximation of the feasible set remains locally invariant around $\bar{x}$. Although this characterization is somewhat more intuitive than Definition 4.2, the latter is easier to use.

10

The facial constant rank condition is sufficient for the equality $\mathcal{T}_\Omega(\bar{x}) = \mathcal{L}_\Omega(\bar{x})$ to hold. To prove this, we employ the main result of Janin's paper [37], but the version we use is a slightly different characterization found in [4, Proposition 3.1]. Despite the fact we work in a context more general than NLP, we use the same result that was used in NLP.

Proposition 4.1. *([4, Proposition 3.1]) Let $\{\zeta_i(x)\}_{i\in\mathcal{I}}$ be a finite family of twice continuously differentiable functions $\zeta_i \colon \mathbb{R}^n \to \mathbb{R}$, $i \in \mathcal{I}$, such that the family of its gradients $\{\nabla\zeta_i(x)\}_{i\in\mathcal{I}}$ remains with constant rank in a neighborhood of $\bar{x}$, and consider the linear subspace*

$$\mathcal{S} := \{y \in \mathbb{R}^n \mid \langle \nabla\zeta_i(\bar{x}), y \rangle = 0, \ i \in \mathcal{I}\}.$$

Then, there exists some neighborhoods V_1 and V_2 of $\bar{x}$, and a diffeomorphism $\psi: V_1 \to V_2$, such that:

(i) $\psi(\bar{x}) = \bar{x}$;

(ii) $D\psi(\bar{x}) = \mathbb{I}_n$;

(iii) $\zeta_i(\psi^{-1}(\bar{x} + y)) = \zeta_i(\psi^{-1}(\bar{x}))$ for every $y \in \mathcal{S} \cap (V_2 - \bar{x})$ and every $i \in \mathcal{I}$.

Moreover, the degree of differentiability of ψ is the same as of ζ_i, for all $i \in \mathcal{I}$.

For the last part of the above proposition, about the degree of differentiability of ψ, we refer to Minchenko and Stakhovski [48, Page 328]. Now, we are able to prove the main result of this section:

Theorem 4.1. *Let $\bar{x}$ be a feasible point of (NSOCP). If the facial constant rank property holds at $\bar{x}$, then $\mathcal{T}_\Omega(\bar{x}) = \mathcal{L}_\Omega(\bar{x})$.*

Proof. It suffices to show that $\mathcal{L}_\Omega(\bar{x}) \subseteq \mathcal{T}_\Omega(\bar{x})$. Let $d \in \mathcal{L}_\Omega(\bar{x})$ and suppose that $\bar{x}$ satisfies the facial constant rank property. Let

$$F := \left(\prod_{j\in I_0(\bar{x})} F_j\right) \times R, \tag{24}$$

where $F_j \trianglelefteq K_{m_j}$, $j \in I_0(\bar{x})$, are defined as

$$F_j := \begin{cases} K_{m_j} & \text{if } Dg_j(\bar{x})d \in \text{int}(K_{m_j}), \\ \text{cone}(Dg_j(\bar{x})d), & \text{if } Dg_j(\bar{x})d \in \text{bd}^+(K_{m_j}), \\ \{0\}, & \text{if } Dg_j(\bar{x})d = 0. \end{cases} \tag{25}$$

and $R \trianglelefteq \mathbb{R}^{|I_B(\bar{x})|}$ is given by

$$R := \mathbb{R}_+^{|I_B(\bar{x})|} \bigcap_{j\in J}\{c_j\}^\perp, \tag{26}$$

where c_j is the j-th vector of the canonical basis of $\mathbb{R}^{|I_B(\bar{x})|}$, and $J := \{j \in I_B(\bar{x}) \mid \nabla\phi_j(\bar{x})^T d = 0\}$. Recall that we are assuming for simplicity that $I_B(\bar{x}) = \{1, \ldots, |I_B(\bar{x})|\}$, and note that $D\mathcal{G}(\bar{x})d \in F$.

Now, for every $j \in I_0(\bar{x})$ such that $Dg_j(\bar{x})d \in \text{bd}^+(K_{m_j})$, let $A_j \in \mathbb{R}^{m_j \times m_j - 1}$ be any matrix with full column rank such that $\text{Im}(A_j) = \{Dg_j(\bar{x})d\}^\perp$, and observe that

$$Dg_j(x)^T[F_j^\perp] = \text{span}\left(\left\{Dg_j(x)^T A_j^i\right\}_{i=1,\ldots,m_j-1}\right)$$

for every such j, where A_j^i denotes the i-th column of A_j. Similarly, for every $j \in I_0(\bar{x})$ such that $Dg_j(\bar{x})d = 0$, we have

$$Dg_j(x)^T[F_j^\perp] = \text{span}(\{\nabla g_{j,i}(x)\}_{i=0,\ldots,m_j-1}),$$

where $\nabla g_{j,i}(x)$ denotes the i-th column of $Dg_j(x)^T$. And for every j such that $Dg_j(\bar{x})d \in \text{int}(K_{m_j})$, we have $Dg_j(x)^T[F_j^\perp] = \{0\}$. Finally, observe that $R^\perp = \text{span}(\{c_j\}_{j\in J})$ and then

$$D\phi(x)^T[R^\perp] = \text{span}\left(\{\nabla\phi_j(x)\}_{j\in J}\right).$$

Therefore, for every $x \in \mathcal{V}$, where $\mathcal{V}$ is the neighborhood of $\bar{x}$ given by Definition 4.1, the linear space

$$D\mathcal{G}(x)^T[F^\perp] = \sum_{j\in I_0(\bar{x})} Dg_j(x)^T[F_j^\perp] + D\phi(x)^T[R^\perp] \tag{27}$$

11

is generated by the family of vectors:

$$\bigcup_{\substack{j\in I_0(\bar{x}) \\ Dg_j(\bar{x})d\in \mathrm{bd}^+(K_{m_j}) \\ i=1,\ldots,m_j-1}} \left\{ Dg_j(x)^T A_j^i \right\} \bigcup_{\substack{j\in I_0(\bar{x}) \\ Dg_j(\bar{x})d=0 \\ i=0,\ldots,m_j-1}} \{\nabla g_{j,i}(x)\} \bigcup_{j\in J} \{\nabla \phi_j(x)\}, \tag{28}$$

which implies that the dimension of (27) equals the rank of (28), for every $x \in \mathcal{V}$. Since this dimension remains constant in $\mathcal{V}$, so does the rank of (28). This means we can apply Proposition 4.1 to the family of functions

$$\zeta_{i,j}(x) := \begin{cases} \langle A_j^i, g_j(x) \rangle, & \text{if } j \in I_0(\bar{x}),\ Dg_j(\bar{x})d \in \mathrm{bd}^+(K_{m_j}),\ i=1,\ldots,m_j-1, \\ g_{j,i}(x), & \text{if } j \in I_0(\bar{x}),\ Dg_j(\bar{x})d=0,\ i=0,\ldots,m_j-1, \\ \phi_j(x), & \text{if } j \in J, \end{cases} \tag{29}$$

where $g_{j,i}(x)$ denotes the i-th entry of $g_j(x)$ for $j \in J$. Then, consider the following subspace:

$$\mathcal{S} := \left\{ y \in \mathbb{R}^n \ \middle| \ \begin{array}{ll} A_j^T Dg_j(\bar{x})y = 0, & \text{if } j \in I_0(\bar{x}),\ Dg_j(\bar{x})d \in \mathrm{bd}^+(K_{m_j}) \\ Dg_j(\bar{x})y = 0, & \text{if } j \in I_0(\bar{x}),\ Dg_j(\bar{x})d = 0 \\ \nabla\phi_j(\bar{x})^T y = 0, & \text{if } j \in J \ , \end{array} \right\},$$

and note that $d \in \mathcal{S}$, so it follows that there exists a local diffeomorphism ψ for which items $(i), (ii)$ and (iii) of Proposition 4.1 are satisfied. Now, define the arc $\xi(t)$ by

$$\xi(t) := \psi^{-1}(\bar{x} + td),$$

for $t \in \mathbb{R}$ small enough so that $\bar{x} + td \in V_2$, where V_2 is given by Proposition 4.1. Then, we obtain that

$$\lim_{t\to 0^+} \xi(t) = \bar{x}, \qquad \lim_{t\to 0^+} \frac{\xi(t) - \bar{x}}{t} = d.$$

To complete the proof, it suffices to show that $\xi(t)$ remains feasible for every sufficiently small $t \geq 0$, so this is our goal from this point onwards. Proposition 4.1 tells us that there exists some $\varepsilon > 0$ such that $\zeta_{i,j}(\xi(t)) = \zeta_{i,j}(\bar{x}) = 0$ for every $t \in (-\varepsilon, \varepsilon)$. In terms of F, this means that

$$\mathcal{G}(\xi(t)) \in \mathrm{span}(F)$$

for every such t, which follows directly from (29). Now, let us analyse each case separately:

1. For each index $j \in I_0(\bar{x})$, consider the Taylor expansion of $g_j(\xi(t))$ around $t = 0$, given by

$$\begin{aligned} g_j(\xi(t)) &= g_j(\xi(0)) + tDg_j(\xi(0))\xi'(0) + o(t) \\ &= g_j(\bar{x}) + tDg_j(\bar{x})D\psi^{-1}(\bar{x})d + o(t) \\ &= tDg_j(\bar{x})d + o(t) \end{aligned} \tag{30}$$

   Then, we split in three sub-cases:

   - If $Dg_j(\bar{x})d \in \mathrm{int}(K_{m_j})$, then it follows from (30) that $g_j(\xi(t)) \in K_{m_j}$ for every $t \in [0, \varepsilon)$, shrinking ε if necessary;
   - If $Dg_j(\bar{x})d \in \mathrm{bd}^+(K_{m_j})$, then $g_j(\xi(t)) \in \mathrm{span}(Dg_j(\bar{x})d)$ due to (25), and it follows from (30) that $g(\xi(t)) \in \mathrm{cone}(Dg_j(\bar{x})d)$ for every $t \in [0, \varepsilon)$, taking a smaller ε if needed;
   - If $Dg_j(\bar{x})d = 0$, then $g(\xi(t)) = 0$ for every $t \in [0, \varepsilon)$, due to (25).

2. Since $\phi(\xi(t)) \in R$ for every $t \in [0, \varepsilon)$, for each index $j \in J$, we have $\phi_j(\xi(t)) = 0$. On the other hand, for each $j \notin J$, consider the Taylor expansion of $\phi_j(\xi(t))$ around $t = 0$:

$$\phi_j(\xi(t)) = \phi_j(\xi(0)) + t\nabla\phi_j(\xi(0))^T \xi'(0) + o(t) = t\nabla\phi_j(\bar{x})^T d + o(t),$$

   and since $\nabla\phi_j(\bar{x})^T d > 0$ for every $j \notin J$, it also follows that $\phi_j(\xi(t)) > 0$ for every $t \in [0, \varepsilon)$, taking a smaller ε if necessary.

Thus, $\mathcal{G}(\xi(t)) \in F$ for every $t \in [0, \varepsilon)$, which also implies that $G(\xi(t)) \in \mathcal{K}$ for every such t, completing the proof.

$\square$

A useful information that can be extracted from the proof above is an equivalent characterization of the facial constant rank property (Definition 4.1) without faces:

Corollary 4.1. *Let $\bar{x} \in \Omega$. Then, the facial constant rank property holds at $\bar{x}$ if, and only if, there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ such that: for all subsets $J_1, J_2 \subseteq I_0(\bar{x})$, $J_3 \subseteq I_B(\bar{x})$, such that $m_j > 1$ for all $j \in J_1$, and for all $\eta_j \in \mathrm{bd}^+(K_{m_j})$, $j \in J_1$, the rank of the family*

$$\bigcup_{\substack{j \in J_1 \\ i=1,\ldots,m_j}} \left\{ Dg_j(x)^T A_j^i \right\} \bigcup_{\substack{j \in J_2 \\ i=0,\ldots,m_j-1}} \{\nabla g_{j,i}(x)\} \bigcup_{j \in J_3} \{\nabla \phi_j(x)\}$$

remains the same for all $x \in \mathcal{V}$, where $A_j \in \mathbb{R}^{m_j \times m_j - 1}$ can be any matrix with full column rank such that $\mathrm{Im}(A_j) = \{\eta_j\}^\perp$, for each $j \in J_1$, and A_j^i denotes the i-th column of A_j.

Before proceeding, we will make a short discussion about Theorem 4.1 and its implications:

Remark 4.1. *Note that if all constraints are affine, then every feasible point satisfies the facial constant rank property. Then, it follows from Theorem 4.2 that $\mathcal{T}_\Omega(\bar{x}) = \mathcal{L}_\Omega(\bar{x})$ in this case, for every $\bar{x} \in \Omega$. We highlight this fact because when it is paired with Example 2.1, two things can be concluded: first, the facial constant rank property alone is not a CQ for (NSOCP); second, the only reason why constraint linearity is not a CQ for NSOCP is that $H(\bar{x})$ may not be closed. When $H(\bar{x})$ is closed, facial constant rank is a CQ, and so is constraint linearity. In other words, the above discussion, in view of the minimality of Guignard's CQ, allows us to conclude that the closedness of $H(\bar{x})$ is the weakest CQ for linear second-order cone programming problems.*

The discussion of Remark 4.1, together with Theorem 4.2, motivates our extension of CRCQ (and RCRCQ) to NSOCP:

Definition 4.2. *Let $\bar{x}$ be a feasible point of* (NSOCP) *and let $H(\bar{x})$ be the set defined in (2). We say that the* constant rank constraint qualification for NSOCP *(CRCQ) holds at $\bar{x}$, if it satisfies the facial constant rank property and, in addition, the set $H(\bar{x})$ is closed.*

When $m_1 = m_2 = \ldots = m_q = 1$, problem (NSOCP) reduces to a NLP problem. Moreover, since the faces of K_1 are $\{0\}$ and $\mathbb{R}_+$, the facial constant rank property (Definition 4.1) reduces to CRCQ in this case, and so does Definition 4.2. Moreover, as mentioned before, it follows directly from Theorem 4.1, that:

Theorem 4.2. *The CRCQ condition of Definition 4.2 implies Abadie's CQ.*

Since the nondegeneracy condition for (NSOCP) holds at a given $\bar{x} \in \Omega$ if, and only if, $D\mathcal{G}(\bar{x})^T$ is injective, then by continuity of $D\mathcal{G}$, nondegeneracy implies that $D\mathcal{G}(x)^T$ remains injective for every x close enough to $\bar{x}$. Therefore, it follows that the nondegeneracy condition implies CRCQ as in Definition 4.2. However, the converse is not true, as it can be seen in the following example:

Example 4.1. *Consider the following constraint*

$$g(x) := (x, x) \in K_2,$$

at the feasible point $\bar{x} = 0$. Since K_2 is a polyhedral cone and g is linear, then CRCQ as in Definition 4.2 holds at $\bar{x}$. However, Robinson's CQ is not satisfied at $\bar{x} = 0$, since

$$Dg(\bar{x})d = d(1,1) \notin \mathrm{int}(K_2)$$

for every $d \in \mathbb{R}$. Consequently, nondegeneracy is not satisfied, either.

Observe that Example 4.1 also shows that CRCQ does not imply Robinson's CQ. Conversely, Robinson's CQ does not imply CRCQ either, meaning they are not related, just as it happens with CRCQ and MFCQ in NLP. Let us show this with an example:

Example 4.2. *Consider the constraint:*

$$g(x) := (x_2, x_1^2) \in K_2$$

at the point $\bar{x} = (0,0)$. Robinson's CQ holds at $\bar{x}$, since $d = (0,1)$ satisfies

$$g(\bar{x}) + Dg(\bar{x})d = (1,0) \in \mathrm{int}(K_2).$$

On the other hand, take the face $F = \{0\}$ and note that

$$Dg(x)^T[F^\perp] = \mathrm{span}\left(\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2x_1 \\ 0 \end{bmatrix} \right\} \right)$$

has dimension 2 for every x such that $x_1 \neq 0$, and dimension 1 at $\bar{x}$.

Remark 4.2. *To consider* (NSOCP) *with an equality constraint in the form* $h(x) = 0$, *where* $h\colon \mathbb{R}^n \to \mathbb{R}^p$, *one should proceed as in Proposition 3.1. That is, consider*

$$g(x) := (g_1(x), \ldots, g_q(x), h(x))$$

and the cone

$$\mathcal{K} := K_{m_1} \times \ldots \times K_{m_q} \times \{0\}^p.$$

This will lead to an extension of RCRCQ. An extension of the original CRCQ condition can be obtained by writing the equality constraint as a pair of inequality constraints in the form $h(x) \in \mathbb{R}_+^p$ *and* $-h(x) \in \mathbb{R}_+^p$, *just as in Remark 3.1, then reducing, and applying Definition 4.2 to the new reduced cone.*

4.2   Strong second-order optimality conditions for NSOCP

In this subsection we will investigate second-order optimality conditions for (NSOCP) under the facial constant rank property; and, consequently, under CRCQ as well. Recall that the second-order condition of Definition 2.1 can be further specialized to the context of NSOCP by characterizing the sigma-term explicitly. Following Bonnans and Ramírez [22], we have for any $\bar{x} \in \Omega$ and any of its associate Lagrange multipliers $\bar{\lambda} := (\bar{\lambda}_1, \ldots, \bar{\lambda}_q) \in \Lambda(\bar{x})$, that

$$\sigma(d, \bar{x}, \bar{\lambda}) = \sum_{j=1}^{q} d^T \mathcal{H}_j(\bar{x}, \bar{\lambda}_j) d$$

for every $d \in C(\bar{x})$, where

$$\mathcal{H}_j(\bar{x}, \bar{\lambda}_j) := \begin{cases} -\dfrac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} Dg_j(\bar{x})^T R_{m_j} Dg_j(\bar{x}), & \text{if } j \in I_B(\bar{x}), \\[2ex] 0, & \text{otherwise.} \end{cases} \tag{31}$$

With this in mind, we can prove that SOC holds at $(\bar{x}, \bar{\lambda})$ under the facial constant rank property by means of analysing the problem along the curve $\xi(t)$ from the proof of Theorem 4.1.

Theorem 4.3. *Let* $\bar{x}$ *be a local minimizer of problem* (NSOCP) *that satisfies the facial constant rank property. Then, for any given Lagrange multiplier* $\bar{\lambda} \in \Lambda(\bar{x})$, *the pair* $(\bar{x}, \bar{\lambda})$ *satisfies SOC as in Definition 2.1; that is,*

$$d^T \nabla^2 f(\bar{x}) d + \sum_{j=1}^{q} \langle D^2 g_j(\bar{x})[d, d], \bar{\lambda}_j \rangle - \sigma(d, \bar{x}, \bar{\lambda}) \geq 0, \tag{32}$$

for every $d \in C(\bar{x}) = \mathcal{L}_\Omega(\bar{x}) \cap \{\nabla f(\bar{x})\}^\perp$.

Proof. If $\Lambda(\bar{x}) = \emptyset$, the result holds trivially. Otherwise, let $\bar{\lambda} := (\bar{\lambda}_1, \ldots, \bar{\lambda}_q) \in \Lambda(\bar{x})$ be arbitrary and fixed. Our aim is to prove that inequality (13) holds for the pair $(\bar{x}, \bar{\lambda})$, for every $d \in C(\bar{x})$. So let $d \in C(\bar{x})$ be also arbitrary, and let F be as in (24). Recall that, for the sake of simplicity and without loss of generality, we are assuming $I_B(\bar{x}) = \{1, \ldots, |I_B(\bar{x})|\}$.

Proceeding in the same way as in the proof of Theorem 4.1, since the facial constant rank property holds at $\bar{x}$ and $d \in \mathcal{L}_\Omega(\bar{x})$, we can construct a curve $\xi : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$, for some $\varepsilon > 0$, such that: $\xi(0) = \bar{x}$, $\xi'(0) = d$, and $\mathcal{G}(\xi(t)) \in \text{span}(F)$ for every $t \in (-\varepsilon, \varepsilon)$. In addition, $\mathcal{G}(\xi(t)) \in F$ for every $t \in [0, \varepsilon)$, meaning $\xi(t)$ is feasible for all such t. Since $\bar{x}$ is a local minimizer of (NSOCP), then $t = 0$ is a local minimizer of the function $\varphi(t) := f(\xi(t))$ subject to the constraint $t \geq 0$. Then, it is easy to see that

$$\varphi''(0) = d^T \nabla^2 f(\bar{x}) d + \nabla f(\bar{x})^T \xi''(0) \geq 0. \tag{33}$$

The rest of this proof consists of computing $\nabla f(\bar{x})^T \xi''(0)$. To do this, we will use an auxiliary complementarity function defined as

$$R(t) := \sum_{j \in I_0(\bar{x})} \langle g_j(\xi(t)), \bar{\lambda}_j \rangle + \sum_{j \in I_B(\bar{x})} [\bar{\lambda}_j]_0 \phi_j(\xi(t)).$$

First, we claim that $R(t) = 0$ for every $t \in (-\varepsilon, \varepsilon)$. Let us prove this, analysing each term separately:

1. For each $j \in I_0(\bar{x})$, it follows from the KKT conditions that

$$\langle Dg_j(\bar{x})d, \bar{\lambda}_j \rangle = \langle d, Dg_j(\bar{x})^T \bar{\lambda}_j \rangle = \langle d, -\nabla f(\bar{x}) \rangle = 0, \tag{34}$$

which implies the following:

- If $Dg_j(\bar{x})d \in \text{int}(K_{m_j})$, then $\bar{\lambda}_j = 0$, since $\bar{\lambda}_j \in K_{m_j}^\circ$;
- If $Dg_j(\bar{x})d \in \text{bd}^+(K_{m_j})$ we have $g_j(\xi(t)) \in \text{span}(Dg_j(\bar{x})d)$, and consequently, $\langle g_j(\xi(t)), \bar{\lambda}_j \rangle = 0$ for every $t \in (-\varepsilon, \varepsilon)$ due to (34);
- If $Dg_j(\bar{x})d = 0$, then $g(\xi(t)) = 0$ also for every $t \in (-\varepsilon, \varepsilon)$, due to (25).

The above reasoning implies that $\langle g_j(\xi(t)), \bar{\lambda}_j \rangle = 0$ for every $t \in [0, \varepsilon)$ and every $j \in I_0(\bar{x})$.

2. For each $j \in I_B(\bar{x})$, consider J as in (26) and it follows that if $\nabla \phi_j(\bar{x})^T d = 0$, then $\phi_j(\xi(t)) = 0$ for every $t \in (-\varepsilon, \varepsilon)$. On the other hand, recall that $\bar{\lambda}_j = \frac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} R_{m_j} g_j(\bar{x})$ for every $j \in I_B(\bar{x})$ due to complementarity, but using (21), we see that

$$[\bar{\lambda}_j]_0 \langle \nabla \phi_j(\bar{x}), d \rangle = \frac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} \langle Dg_j(\bar{x})^T R_{m_j} g_j(\bar{x}), d \rangle = \langle \bar{\lambda}_j, Dg_j(\bar{x})d \rangle = 0.$$

Therefore, if $\langle \nabla \phi_j(\bar{x}), d \rangle > 0$, then $[\bar{\lambda}_j]_0 = 0$.

Knowing that $R(t) = 0$ for every $t \in (-\varepsilon, \varepsilon)$, we obtain that the derivatives of $R(t)$ are also zero for all such t. Let us compute them: the first derivative of $R(t)$ is given by

$$R'(t) = \sum_{j \in I_0(\bar{x})} \langle Dg_j(\xi(t))\xi'(t), \bar{\lambda}_j \rangle + \sum_{j \in I_B(\bar{x})} [\bar{\lambda}_j]_0 \langle \nabla \phi_j(\xi(t)), \xi'(t) \rangle,$$

and the derivative of $R'(t)$ is

$$\begin{aligned}
R''(t) = & \sum_{j \in I_0(\bar{x})} \langle D^2 g_j(\xi(t))[\xi'(t), \xi'(t)], \bar{\lambda}_j \rangle + \sum_{j \in I_0(\bar{x})} \langle Dg_j(\xi(t))^T \bar{\lambda}_j, \xi''(t) \rangle \\
& + \sum_{j \in I_B(\bar{x})} [\bar{\lambda}_j]_0 \left(\langle D^2 \phi_j(\xi(t))\xi'(t), \xi'(t) \rangle + \langle \nabla \phi_j(\xi(t)), \xi''(t) \rangle \right).
\end{aligned}$$

Since $R''(t)$ is continuous, taking the limit $t \to 0$, we obtain

$$\begin{aligned}
R''(0) = \lim_{t \to 0^+} R''(t) = & \sum_{j \in I_0(\bar{x})} \langle D^2 g_j(\bar{x})[d, d], \bar{\lambda}_j \rangle + \sum_{j \in I_0(\bar{x})} \langle Dg_j(\bar{x})^T \bar{\lambda}_j, \xi''(0) \rangle \\
& + \sum_{j \in I_B(\bar{x})} [\bar{\lambda}_j]_0 \left(\langle D^2 \phi_j(\bar{x})d, d \rangle + \frac{1}{[g_j(\bar{x})]_0} \langle Dg_j(\bar{x})^T R_{m_j} g_j(\bar{x}), \xi''(0) \rangle \right).
\end{aligned}$$

The above expression can be simplified using the relation

$$\begin{aligned}
\langle D^2 \phi_j(\bar{x})d, d \rangle &= \frac{\langle \widehat{Dg_j(\bar{x})}d, \widehat{g_j(\bar{x})} \rangle^2}{\|\widehat{g_j(\bar{x})}\|^3} - \frac{\|\widehat{Dg_j(\bar{x})}d\|^2}{\|\widehat{g_j(\bar{x})}\|} + \langle D^2 g_j(\bar{x})[d, d], \nabla \phi_j(\bar{x}) \rangle \\
&= \frac{1}{[g_j(\bar{x})]_0} \left(\langle R_{m_j} Dg_j(\bar{x})d, \ Dg_j(\bar{x})d \rangle + \langle D^2 g_j(\bar{x})[d, d], R_{m_j} g_j(\bar{x}) \rangle \right),
\end{aligned}$$

which can be directly computed from the definition of ϕ_j, $j \in I_B(\bar{x})$, since in this case $[g_j(\bar{x})]_0 = \|\widehat{g_j(\bar{x})}\|$ and $\langle Dg_j(\bar{x})d, R_{m_j} g_j(\bar{x}) \rangle = 0$. Then, we get

$$\begin{aligned}
R''(0) = & \sum_{j \in I_0(\bar{x}) \cup I_B(\bar{x})} \langle D^2 g_j(\bar{x})[d, d], \bar{\lambda}_j \rangle + \sum_{j \in I_0(\bar{x}) \cup I_B(\bar{x})} \langle Dg_j(\bar{x})^T \bar{\lambda}_j, \xi''(0) \rangle \\
& + \sum_{j \in I_B(\bar{x})} \frac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} \langle R_{m_j} Dg_j(\bar{x})d, \ Dg_j(\bar{x})d \rangle = 0.
\end{aligned} \tag{35}$$

Moreover, by the KKT conditions, we have

$$\nabla f(\bar{x})^T \xi''(0) = - \sum_{j \in I_0(\bar{x}) \cup I_B(\bar{x})} \langle Dg_j(\bar{x})^T \bar{\lambda}_j, \xi''(0) \rangle,$$

which yields together with equation (35), the following:

$$\nabla f(\bar{x})^T \xi''(0) = \sum_{j \in I_0(\bar{x}) \cup I_B(\bar{x})} \langle D^2 g_j(\bar{x})[d, d], \bar{\lambda}_j \rangle + \sum_{j \in I_B(\bar{x})} \frac{[\bar{\lambda}_j]_0}{[g_j(\bar{x})]_0} d^T Dg_j(\bar{x})^T R_{m_j} Dg_j(\bar{x})d. \tag{36}$$

Therefore, since $\bar{\lambda}_j = 0$ for every $j \in I_{\text{int}}(\bar{x})$,

$$d^T \nabla^2 f(\bar{x})d + \sum_{j=1}^{q} \langle D^2 g_j(\bar{x})[d,d], \bar{\lambda}_j \rangle - \sigma(d, \bar{x}, \bar{\lambda}) \geq 0.$$

Since $d \in C(\bar{x})$ is arbitrary, we conclude that $\bar{x}$ satisfies SOC with respect to $\bar{\lambda}$, which was also chosen arbitrarily and remained fixed from the very beginning. Thus, the proof is complete.

$\square$

Observe that Theorem 4.3 implies that the facial constant rank property ensures the fulfilment of the strong second-order necessary condition at a given point $\bar{x}$, in the sense that for every $\bar{\lambda} \in \Lambda(\bar{x})$, and every $d \in C(\bar{x})$, inequality (13) holds true. If, in addition, $H(\bar{x})$ is closed (CRCQ), then $\Lambda(\bar{x}) \neq \emptyset$, and as consequence, we obtain that the strong second-order condition is satisfied in the presence of CRCQ. It is also worth mentioning that since the strong necessary condition of Theorem 4.3 implies the classical condition of Theorem 2.2, then it also induces a sufficient second-order optimality condition after replacing $\geq$ by $>$ in inequality (32).

Remark 4.3. *In contrast with the facial constant rank property, the condition presented in [61, Definition 2.1] fails to be a CQ even when $H(\bar{x})$ is closed. In fact, let us recall the counterexample presented in [5]:*

$$\begin{aligned} \text{Minimize} \quad & f(x) := -x, \\ s.t. \quad & g(x) := (x, x + x^2) \in K_2, \end{aligned}$$

The unique solution of this problem is $\bar{x} = 0$. For this particular example, [61, Definition 2.1] holds if, and only if, $\{1, 1 + 2x\}$ remain with constant rank in some neighborhood of $\bar{x}$ (one may consider also all of its subfamilies, see [5]). Of course, this is verified, and since K_2 is polyhedral, the set $H(\bar{x})$ is closed. However, $\bar{x}$ does not satisfy the KKT conditions.

On the other hand, to see that CRCQ as in Definition 4.2 does not hold at $\bar{x}$, take $F := \text{cone}((1,1)) \trianglelefteq K_2$ and note that

$$Dg(x)^T[F^\perp] = \text{span}(-2x)$$

has dimension 1 for every $x \neq 0$, but has dimension zero at $\bar{x}$. In particular, this example shows that CRCQ as in Definition 4.2 is not a mere correction of the condition presented in [61], and that the condition of [61] cannot be corrected by simply adding the closedness of $H(\bar{x})$ to its definition.

4.3   About the sequential constant rank CQ

In [9], we introduced an alternative extension of CRCQ for (NSOCP) that was based on a special re-characterization of the nondegeneracy condition [7] in terms of the eigenvectors of some perturbations of $g(\bar{x})$. Let us recall an equivalent characterization of it, which will be used here as a definition for simplicity.

Definition 4.3 (Seq-CRCQ for NSOCP). *Let $\bar{x} \in \Omega$. We say that the* Sequential-CRCQ *(Seq-CRCQ) condition holds at $\bar{x}$ if for every vector $\bar{w}_j \in \mathbb{R}^{m_j-1}$ with $\|\bar{w}_j\| = 1$, $j \in I_0(\bar{x})$, there is a neighborhood $\mathcal{V}$ of $(\bar{x}, \bar{w})$, $\bar{w} := (\bar{w}_j)_{j \in I_0(\bar{x})}$, such that: for all subsets $J_1, J_2 \subseteq I_0(\bar{x})$ and $J_3 \subseteq I_B(\bar{x})$, if the family*

$$\mathcal{D}(x, w) := \left\{ Dg_j(x)^T(1, -w_j) \right\}_{j \in J_1} \bigcup \left\{ Dg_j(x)^T(1, w_j) \right\}_{j \in J_2} \bigcup \left\{ Dg_j(x)^T \left(1, -\frac{\widehat{g_j(x)}}{\|\widehat{g_j(x)}\|} \right) \right\}_{j \in J_3}$$

is linearly dependent at $(x, w) := (\bar{x}, \bar{w})$, then $\mathcal{D}(x, w)$ remains linearly dependent for all $(x, w) \in \mathcal{V}$ such that $\|w_j\| = 1$, $j \in J_1 \cup J_2$, where $w := (w_j)_{j \in I_0(\bar{x})}$.

This constraint qualification was used in [9] to achieve global convergence of a class of algorithms to KKT points, and some interesting properties were shown together with a weaker variant of Seq-CRCQ. Namely, it is also independent of Robinson's CQ, strictly weaker than nondegeneracy, and it implies the *metric subregularity CQ* (also known as *error bound CQ*). Moreover, note that if $I_0(\bar{x}) = \emptyset$, then Seq-CRCQ coincides with the facial constant rank property, which in turn coincides with CRCQ. However, this is not necessarily true otherwise. In the following example, we show that CRCQ according to Definition 4.2 does not imply Seq-CRCQ.

Example 4.3. *Consider the constraint:*

$$g(x) = (x, -x, 0) \in K_3,$$

and let $\bar{x} = 0$, a feasible point. Since g is affine, then the facial constant rank property holds at $\bar{x}$ (see Remark 4.1). Now, let us show that $H(\bar{x})$ is closed: since $g(\bar{x}) = 0$, it holds that

$$H(\bar{x}) = Dg(\bar{x})^T K_3 = \{v_1 - v_2 \mid (v_1, v_2, v_3) \in K_3\} = \mathbb{R}_+.$$

Therefore, $H(\bar{x})$ is a closed set, and CRCQ according to Definition 4.2 holds at $\bar{x}$.

On the other hand, Seq-CRCQ does not hold at $\bar{x}$, since for any $w = (w_1, w_2) \in \mathbb{R}^2$,

$$Dg(\bar{x})^T(1, w) = 1 - w_1 \quad \text{and} \quad Dg(\bar{x})^T(1, -w) = 1 + w_1;$$

then, take $\bar{w} = (1, 0)$ and any sequence $\{w^k\}_{k \in \mathbb{N}} \to \bar{w}$ such that $w_1^k \neq 1$ for all $k \in \mathbb{N}$, to see that $Dg(\bar{x})^T(1, w_1^k) \neq 0$ for every $k \in \mathbb{N}$, but $Dg(\bar{x})^T(1, \bar{w}) = 0$.

We were not able to prove nor find a counterexample for the converse statement. However, with only Example 4.3 at hand, we already know that CRCQ is in the worst case independent of Seq-CRCQ, and in the best case, strictly weaker than it, meaning the results of this paper either improve or are parallel to the results of [9].

5 Nonlinear semidefinite programming

In this section, we will study constant rank conditions for nonlinear semidefinite programming problems, which can be stated in standard form as follows:

$$\begin{aligned} \text{Minimize} \quad & f(x), \\ \text{s.t.} \quad & G(x) \succeq 0. \end{aligned} \tag{NSDP}$$

This problem can be seen as a particular case of (NCP), letting $\mathbb{E} = \mathbb{S}^m$ be the space of $m \times m$ symmetric matrices with real entries, and

$$\mathcal{K} = \mathbb{S}_+^m := \{A \in \mathbb{S}^m \mid z^T A z \geq 0, \ \forall z \in \mathbb{R}^m\}$$

be the cone of all $m \times m$ symmetric positive semidefinite matrices, with $G : \mathbb{R}^n \to \mathbb{E}$ being twice continuously differentiable. The symbol $\succeq$ denotes the partial order induced by $\mathbb{S}_+^m$, meaning that $A \succeq B$ if, and only if, $A - B \in \mathbb{S}_+^m$. In this section, for any given $A \in \mathbb{S}^m$ we will denote by $\mu_i(A)$ the i-th eigenvalue of A arranged in non-increasing order, and $u_i(A)$ will denote an associated unitary eigenvector.

Recall from Section 3 that the constant rank constraint qualification can be obtained in two steps: first, reduce the problem to consider only the locally relevant part of the constraint; then, analyse the image of the faces of the reduced cone by the derivative of the reduced constraint function. Therefore, we begin by recalling a matrix analysis lemma that will be useful for the reduction part.

Lemma 5.1. *Let $\bar{A} \succeq 0$ and denote by r the rank of $\bar{A}$. Also, let $\bar{E} \in \mathbb{R}^{m \times m - r}$ be a matrix whose columns form an orthonormal basis of $\mathrm{Ker}(\bar{A})$. Then, there exists an analytic matrix function $\mathcal{E} : \mathbb{S}^m \to \mathbb{R}^{m \times m - r}$ such that $\mathcal{E}(\bar{A}) = \bar{E}$ and, for all A close enough to $\bar{A}$, the columns of $\mathcal{E}(A)$ form an orthonormal basis of the space spanned by the eigenvectors associated with the $m - r$ smallest eigenvalues of A.*

Proof. Although this proof can be found in [24, Example 3.140] and [16, Section 2.3], we shall include it here for completeness purposes. For any given $A \in \mathbb{S}^m$, let $S(A)$ be the space spanned by the eigenvectors $u_{r+1}(A), ..., u_m(A)$, and let $\Pi(A)$ denote the orthogonal projection matrix onto $S(A)$. It is known that $\Pi(A)$ is an analytic function of A in a sufficiently small neighborhood of $\bar{A}$ (see, for example, [40, Theorem 1.8]). Consequently, the function $V(A) := \Pi(A)\bar{E}$ is also an analytic function of A in a neighborhood of $\bar{A}$, and moreover $V(\bar{A}) = \bar{E}$. It follows that for all A sufficiently close to $\bar{A}$, the rank of $V(A)$ is equal to the rank of $V(\bar{A}) = \bar{E}$, so the $m - r$ columns of $V(A)$ are linearly independent, and $\mathrm{Im}(V(A)) = S(A)$, when A is sufficiently close to $\bar{A}$. Now, let $\mathcal{E}(A)$ be a matrix whose columns are obtained by applying the Gram-Schmidt orthonormalization process to the columns of $V(A)$. The matrix $\mathcal{E}(A)$ is well-defined and also analytic in a neighborhood of $\bar{A}$. Moreover, it satisfies $\mathcal{E}(A)^T \mathcal{E}(A) = \mathbb{I}_{m-r}$ and $\mathrm{Im}(\mathcal{E}(A)) = S(A)$, for all A sufficiently close to $\bar{A}$, and $\mathcal{E}(\bar{A}) = V(\bar{A}) = \bar{E}$, which concludes the proof. Notice however that the columns of $\mathcal{E}(A)$ are not necessarily eigenvectors of A.

$\square$

Now, let $\bar{x} \in \Omega$, denote by r the rank of $G(\bar{x})$, and let $\bar{E} \in \mathbb{R}^{m \times m-r}$ be an arbitrary matrix with orthonormal columns that span $\text{Ker}(G(\bar{x}))$. Moreover, let $\mathcal{E}$ be the analytic matrix function with the properties described in Lemma 5.1, such that $\mathcal{E}(G(\bar{x})) = \bar{E}$. Observe that $\mathbb{S}_+^m$ is reducible to

$$\mathcal{C} := \mathbb{S}_+^{m-r}$$

in a neighborhood $\mathcal{N}$ of $G(\bar{x})$ by the mapping $\Xi \colon \mathcal{N} \to \mathbb{S}^{m-r}$ given by

$$\Xi(Y) := \mathcal{E}(Y)^T Y \mathcal{E}(Y),$$

for every $Y \in \mathcal{N}$ close enough to $G(\bar{x})$ so that $\mu_i(Y) > 0$ for every $i = 1, \ldots, r$. Then, define the function $E := \mathcal{E} \circ G$, consider the reduced constraint function

$$\mathcal{G}(x) := E(x)^T G(x) E(x),$$

and for every x sufficiently close to $\bar{x}$, we have that $G(x) \in \mathbb{S}_+^m$ if, and only if, $\mathcal{G}(x) \in \mathbb{S}_+^{m-r}$. Moreover, it is worth recalling that, since the function $\mathcal{E}$ is analytic, the degree of differentiability of $\mathcal{G}$ is the same as of G.

Following Bonnans and Shapiro [24, Equation 5.161], we see that the linearized cone of the original constraints of (NSDP) at $\bar{x} \in \Omega$ can be written as

$$\mathcal{L}_\Omega(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \bar{E}^T DG(\bar{x}) d\bar{E} \succeq 0 \right\},$$

which also coincides with the linearized cone of the reduced constraint at $\bar{x}$, because $E(\bar{x}) = \bar{E}$ and for each x close enough to $\bar{x}$, we have

$$D\mathcal{G}(x)[\,\cdot\,] = DE(x)[\,\cdot\,]^T G(x) E(x) + E(x)^T DG(x)[\,\cdot\,] E(x) + E(x)^T G(x) DE(x)[\,\cdot\,],$$

so $D\mathcal{G}(\bar{x})[\,\cdot\,] = \bar{E}^T DG(\bar{x})[\,\cdot\,]\bar{E}$. For more details on this reduction approach, see [21, 23].

In the next section, we will introduce a constant rank-type condition for NSDP by means of the faces of the reduced cone.

5.1 A facial constant rank constraint qualification for NSDP

Following the exposition of Pataki [51], the faces of $\mathcal{C} = \mathbb{S}_+^{m-r}$ can be represented in a very simple way: F is a face of $\mathbb{S}_+^{m-r}$ if, and only if, there exists an orthogonal matrix $U \in \mathbb{R}^{m-r \times m-r}$ and some $s \in \{1, \ldots, m-r\}$ such that

$$F = \left\{ U \begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix} U^T \;\middle|\; A_{11} \in \mathbb{S}_+^s \right\}.$$

With this in mind, let us define the analogue of Definition 4.2 for NSDP:

Definition 5.1. *Let $\bar{x} \in \Omega$ and let r be the rank of $G(\bar{x})$. We say that the* facial constant rank *property holds at $\bar{x}$ if there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ such that: for each $F \trianglelefteq \mathbb{S}_+^{m-r}$, the dimension of $D\mathcal{G}(x)^T[F^\perp]$ remains constant for every $x \in \mathcal{V}$.*

Following the discussion after Proposition 3.1 and also after Definition 4.1, to better visualize the meaning of Definition 5.1, recall that Ω is locally equivalent to $\mathcal{G}^{-1}(\mathbb{S}_+^{m-r})$ and that the faces of $\mathbb{S}_+^{m-r}$ can be regarded as linear approximations of $\mathbb{S}_+^{m-r}$, in some sense. Then, for every $F \trianglelefteq \mathbb{S}^{m-r}$, the set $D\mathcal{G}(x)^{-1}(\text{span}(F))$ defines a possible linear approximation of Ω around $\bar{x}$. The reasoning after Proposition 3.1 still holds in the context of NSDP and it follows that the facial constant rank property holds at $\bar{x} \in \Omega$ if, and only if, the dimension of $D\mathcal{G}(x)^{-1}(\text{span}(F))$ remains constant for all x in a neighborhood of $\bar{x}$, at every $F \trianglelefteq \mathbb{S}_+^{m-r}$. From this point of view, the facial constant rank property demands all linear approximations of the feasible set to remain with constant dimension in the vicinity of $\bar{x}$.

Now, we proceed to the main result of this section.

Theorem 5.1. *Let $\bar{x} \in \Omega$. If $\bar{x}$ satisfies the facial constant rank property, then $\mathcal{T}_\Omega(\bar{x}) = \mathcal{L}_\Omega(\bar{x})$.*

Proof. Let $d \in \mathcal{L}_\Omega(\bar{x})$, denote by s the rank of $\bar{E}^T DG(\bar{x}) d\bar{E}$, and let $\bar{Q} \in \mathbb{R}^{m-r \times m-r}$ be an orthogonal matrix such that

$$\bar{Q}^T \bar{E}^T DG(\bar{x}) d\bar{E} \bar{Q} = \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix},$$

where $R \succ 0$ is an $s \times s$ diagonal matrix. Let $\bar{W}$ be the matrix formed by the columns of $\bar{Q}$ corresponding to the positive eigenvalues of $\bar{E}^T DG(\bar{x})d\bar{E}$; that is, $\bar{W}^T \bar{E}^T DG(\bar{x})d\bar{E}\bar{W} = R$. Then, consider the face of $\mathbb{S}^{m-r}_+$ given by

$$F := \left\{ \bar{Q} \begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix} \bar{Q}^T \mid A_{11} \in \mathbb{S}^s_+ \right\} \tag{37}$$

and note that $\bar{E}^T DG(\bar{x})d\bar{E} \in F$. Let $\eta_1, \ldots, \eta_N$ be a basis of $F^\perp$, where $N := \dim(F^\perp)$, and note that

$$D\mathcal{G}(x)^T[F^\perp] = \operatorname{span}\left(\left\{ D\mathcal{G}(x)^T[\eta_i] \right\}_{i \in \{1,\ldots,N\}} \right). \tag{38}$$

Therefore, the facial constant rank property can be equivalently stated as the constant rank of the family

$$\left\{ D\mathcal{G}(x)^T[\eta_i] \right\}_{i \in \{1,\ldots,N\}}$$

in a neighborhood of $\bar{x}$. Furthermore, let $\zeta_i(x) := \langle \mathcal{G}(x), \eta_i \rangle$ and note that

$$\nabla \zeta_i(x) = D\mathcal{G}(x)^T[\eta_i]$$

for every $i \in \{1, \ldots, N\}$.

Then, by Proposition 4.1, there exist two neighborhoods V_1 and V_2 of $\bar{x}$, and a curve $\psi \colon V_1 \to V_2$ such that $\psi(\bar{x}) = \bar{x}$, $D\psi(\bar{x}) = \mathbb{I}_n$, and $\zeta_i(\psi^{-1}(\bar{x} + y)) = \zeta_i(\bar{x})$ for every $i \in \{1, \ldots, N\}$ and every y in the subspace

$$\mathcal{S} := \{ y \in \mathbb{R}^n \mid \langle \nabla \zeta_i(\bar{x}), y \rangle = 0, \ \forall i \in \{1, \ldots, N\} \}.$$

Since $D\mathcal{G}(\bar{x})d \in F$, we see that $\langle d, D\mathcal{G}(\bar{x})^T[\eta_i] \rangle = \langle D\mathcal{G}(\bar{x})d, \eta_i \rangle = 0$ for every $i \in \{1, \ldots, N\}$, so $d \in \mathcal{S}$. Then, let $\varepsilon > 0$ be such that $\bar{x} + td \in V_2$ for every $t \in (-\varepsilon, \varepsilon)$, and define $\xi(t) := \psi^{-1}(\bar{x} + td)$ for every such t. Moreover, note that $\xi'(t) = d$ and $\xi(0) = \bar{x}$.

Now, for every $t \in (-\varepsilon, \varepsilon)$, we have $\langle \mathcal{G}(\xi(t)), \eta_i \rangle = 0$ for every $i \in \{1, \ldots, N\}$, whence follows that

$$\mathcal{G}(\xi(t)) \in \operatorname{span}(F)$$

for every such t, meaning also

$$\bar{Q}^T \mathcal{G}(\xi(t))\bar{Q} = \begin{bmatrix} \bar{W}^T \mathcal{G}(\xi(t))\bar{W} & 0 \\ 0 & 0 \end{bmatrix}.$$

On the other hand, considering the Taylor expansion of $\mathcal{G}(\xi(t))$ around $t = 0$,

$$\mathcal{G}(\xi(t)) = \bar{E}^T G(\bar{x})\bar{E} + t\bar{E}^T DG(\bar{x})d\bar{E} + o(t) = t\bar{E}^T DG(\bar{x})d\bar{E} + o(t),$$

we observe that

$$\bar{W}^T \mathcal{G}(\xi(t))\bar{W} = t\bar{W}^T \bar{E}^T DG(\bar{x})d\bar{E}\bar{W} + o(t) = tR + o(t) \succ 0,$$

for $t \in [0, \varepsilon)$, shrinking ε if necessary. Thus,

$$\mathcal{G}(\xi(t)) \in F \subseteq \mathbb{S}^{m-r}_+$$

for every $t \in [0, \varepsilon)$, and then $G(\xi(t)) \succeq 0$ for all such t. Therefore, it follows that $d \in \mathcal{T}_\Omega(\bar{x})$. $\qquad \square$

Remark 5.1. *Similarly to Remark 4.1, we observe that if G is affine, then every $\bar{x} \in \Omega$ satisfies the facial constant rank property, which implies it is not a CQ on its own, unless $H(\bar{x})$ as defined in (2) is closed. We remark this fact because it implies that the weakest CQ that guarantees zero duality gap in linear SDP problems is the closedness of $H(\bar{x})$.*

With this in mind, we present our extension of CRCQ (and RCRCQ) for NSDP:

Definition 5.2 (CRCQ). *Let $\bar{x} \in \Omega$. We say that $\bar{x}$ satisfies the* constant rank constraint qualification *condition for NSDP (CRCQ) if it satisfies the facial constant rank property and, in addition, $H(\bar{x})$ is closed.*

And, as an immediate consequence of Theorem 5.1, we obtain the following:

Theorem 5.2. *Let $\bar{x} \in \Omega$. If $\bar{x}$ satisfies CRCQ, then it also satisfies Abadie's CQ.*

Now, we are led to compare our CRCQ condition with other CQs from the literature. First, let us show that it is, in general, independent of Robinson's CQ.

Example 5.1. *Consider the following constraint:*

$$G(x) := \begin{bmatrix} -x & 0 \\ 0 & x \end{bmatrix}.$$

The only feasible point is $\bar{x} = 0$, for which one has $G(\bar{x}) = 0 \in \mathbb{S}_+^2$. Since G is linear it is enough to show that $H(\bar{x})$ is closed (see Remark 5.1). In this case,

$$H(\bar{x}) = \left\{ DG(\bar{x})^T A \ \middle| \ A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \in \mathbb{S}_+^2 \right\}.$$

Since $DG(\bar{x})^T A = \langle DG(\bar{x}), A \rangle = a_{22} - a_{11}$, but a_{11} and a_{22} are both nonnegative, it follows that $H(\bar{x}) = \mathbb{R}$, which is closed. On the other hand, Robinson's CQ does not hold at $\bar{x}$. In fact, given a real number $d \in \mathbb{R}$, we have that

$$G(\bar{x}) + DG(\bar{x})d = \begin{bmatrix} -d & 0 \\ 0 & d \end{bmatrix},$$

which is not in $\mathrm{int}(\mathbb{S}_+^2)$ regardless of $d \in \mathbb{R}$.

The above example shows that CRCQ does not imply Robinson's CQ. Conversely, we will show in the next example, that Robinson's CQ does not imply CRCQ either.

Example 5.2. *Consider the following constraint given by*

$$G(x) := \begin{bmatrix} x_2 & x_1^2 \\ x_1^2 & x_2 \end{bmatrix}$$

at the point $\bar{x} = (0,0)$. Then, for any direction $d = (d_1, d_2) \in \mathbb{R}^2$, it follows that

$$DG(x)d = \begin{bmatrix} 0 & 2x_1 \\ 2x_1 & 0 \end{bmatrix} d_1 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} d_2.$$

Now, consider $\overline{d} = (0,1)$. Then

$$G(\bar{x}) + DG(\bar{x})\overline{d} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathrm{int}(\mathbb{S}_+^2)$$

and Robinson's CQ holds at $\bar{x}$. In order to analyze whether CRCQ holds at $\bar{x}$ or not, let $\bar{E} \in \mathbb{R}^{2 \times 2}$ be a given orthogonal matrix and let us define $\mathcal{G}(x) := \bar{E}^T G(x) \bar{E}$. Then, take

$$F := \left\{ \bar{E}^T \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \bar{E} \ \middle| \ a \geq 0 \right\} \trianglelefteq \mathbb{S}_+^2$$

and note that

$$D\mathcal{G}(x)^T[F^\perp] = DG(x)^T[\bar{E}F^\perp \bar{E}^T] = \mathrm{span}\left(\{\nabla G_{12}(x), \nabla G_{22}(x)\}\right).$$

Then, since $\nabla G_{12}(x) = [2x_1, 0]^T$ and $\nabla G_{22}(x) = [0,1]^T$, for every $x \in \mathbb{R}^2$, the dimension of the subspace above is 1 at $\bar{x}$, but it is equal to 2 for every x close enough to $\bar{x}$ such that $x_1 \neq 0$. Therefore, CRCQ does not hold at $\bar{x}$.

Remark 5.2. *Note that the construction of Lemma 5.1 does not guarantee uniqueness of $\mathcal{E}$, and $\mathcal{G}$ depends on the choices of $\bar{E}$ and also of $\mathcal{E}$. Therefore, Definition 5.2 also depends on $\bar{E}$ and $\mathcal{E}$. In Example 5.2, we considered every $\bar{E}$, and the simplest choice of $\mathcal{E}$ (constantly equal to $\bar{E}$, which is admissible for Lemma 5.1 when $r = 0$), but a priori, different $\mathcal{E}$ may lead to different variants of CRCQ. Nevertheless, we should stress that they are all valid constraint qualifications and the results of this paper concerning CRCQ as in Definition 5.2 hold true for all choices of $\bar{E}$ and $\mathcal{E}$. Observe that the discussion of this remark, on the choice of the reduction, also applies for the NSOCP version of CRCQ (Definition 4.2), where a different choice of reduction function ϕ in (20) may lead to a different version of CRCQ.*

Recall that the nondegeneracy condition holds at $\bar{x}$ if, and only if, $D\mathcal{G}(\bar{x})^T$ is injective. By the continuity of $D\mathcal{G}$, we have that if $\bar{x}$ satisfies nondegeneracy, then $D\mathcal{G}(x)^T$ remains injective for every x sufficiently close to $\bar{x}$. Then, the dimension of $D\mathcal{G}(x)^T[F^\perp]$ remains constant for every such x, at every $F \trianglelefteq \mathbb{S}_+^{m-r}$, and it follows that nondegeneracy strictly implies CRCQ as in Definition 5.2.

Remark 5.3. *Note that our approach can be trivially extended to an NSDP problem with multiple constraints. Moreover, to deal with a separate equality constraint $h(x) = 0$, where $h\colon \mathbb{R}^n \to \mathbb{R}^p$, in the same spirit of Remark 4.2, one should consider a constraint in the form*

$$g(x) := (G(x), h(x)) \quad \text{and} \quad \mathcal{K} := \mathbb{S}_+^m \times \{0\}^p.$$

This yields an extension of RCRCQ after applying Definition 5.2 to the reduced form of this new problem, since $F \trianglelefteq \mathcal{K}$ if, and only if, $F = R \times \{0\}^p$, where $R \trianglelefteq \mathbb{S}_+^m$ in this case. To extend CRCQ one should write the equality constraint as a pair of inequality constraints in the form $h(x) \in \mathbb{R}_+^p$ and $-h(x) \in \mathbb{R}_+^p$, giving rise to a multifold NSDP problem where $\mathbb{R}_+$ is seen as a copy of $\mathbb{S}_+^1$.

For a last comparison, we should mention a constraint qualification presented in one of our previous works [8], which was called *Sequential CRCQ* (Seq-CRCQ) therein. As a matter of fact, Seq-CRCQ differs from Definition 5.2 in many aspects. For instance, Example 5.1 shows that Seq-CRCQ is not implied by CRCQ. This example has already appeared in [8, Example 4.1], where we show that Seq-CRCQ is not satisfied at $\bar{x} = 0$; on the other hand, we showed in Example 5.1, that CRCQ holds at $\bar{x}$. Thus, CRCQ is either strictly weaker than Seq-CRCQ, or completely independent of it. Despite our efforts to clarify the converse statement, we were not able to prove nor find a counterexample for it, so this is left as an open problem.

5.2 Strong second-order optimality conditions for NSDP

The earliest work that provides a practical characterization of the sigma-term in NSDP is Shapiro's [55], using second-order directional derivatives of the least eigenvalue function, $\mu_{\min}\colon \mathbb{S}^m \to \mathbb{R}$. Shapiro proved that

$$\sigma(d, \bar{x}, \bar{\lambda}) = d^T \mathcal{H}(\bar{x}, \bar{\lambda}) d,$$

for any $d \in C(\bar{x})$ and $\bar{\lambda} \in \Lambda(\bar{x})$, where

$$\mathcal{H}(\bar{x}, \bar{\lambda}) := \left[2 \left\langle D_{x_i} G(\bar{x}) G(\bar{x})^\dagger D_{x_j} G(\bar{x}), \ \bar{\lambda} \right\rangle \right]_{i,j=1,\dots,n}$$

and $G(\bar{x})^\dagger$ denotes the Moore-Penrose pseudoinverse of $G(\bar{x})$. Shapiro also proved that if a local minimizer $\bar{x} \in \Omega$ satisfies nondegeneracy and its associated Lagrange multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ is such that $\text{rank}(\bar{\lambda}) + \text{rank}(G(\bar{x})) = m$ (strict complementarity), then $\bar{x}$ satisfies SOC with respect to $\bar{\lambda}$.

Later, other authors provided creative ways of obtaining SOC via some local reformulation of (NSDP) with no curvature. For instance, Lourenço et al. [43] wrote $G(x) \succeq 0$ in the form $G(x) - Z^2 = 0$ with an additional variable $Z \in \mathbb{S}^m$, and then obtained SOC for NSDP out of SOC for NLP – under nondegeneracy and strict complementarity. Forsgren [27] rediscovered Shapiro's characterization of the sigma-term and obtained SOC (under nondegeneracy, but without strict complementarity) using a special reformulation of the problem. Jarre [38] provided an elementary construction of SOC via a certain Schur complement, under nondegeneracy, strict complementarity, and assuming that the tangent cone of the linearized constraint $G(\bar{x}) + DG(\bar{x})d \in \mathcal{K}$ coincides with $\mathcal{T}_\Omega(\bar{x})$. Fukuda et al. [28] used the characterization

$$\mathbb{S}_+^m = \{Z \in \mathbb{S}^m \mid \|\Pi_{\mathbb{S}_+^m}(-Z)\|^2 = 0\}$$

combined with an external penalty method and the Clarke subdifferential of $\Pi_{\mathbb{S}_+^m}$, to achieve a weaker second-order condition, which is stated only in terms of the lineality space of $C(\bar{x})$. However, their results were obtained assuming only Robinson's CQ together with the so-called *weak constant rank* (WCR) property, which is not a CQ on its own.

Following this line of research, the main contribution of this section consists of proving that every local minimizer $\bar{x}$ of (NSDP) satisfies SOC with respect to any Lagrange multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ under the facial constant rank property. In particular, when in addition $H(\bar{x})$ is closed (which leads to CRCQ), then the facial constant rank property implies $\Lambda(\bar{x}) \neq \emptyset$. A priori, we make no special requirement on $\Lambda(\bar{x})$.

Theorem 5.3. *Let $\bar{x} \in \Omega$ be a local minimizer of (NSDP) satisfying the facial constant rank property. Then, for every $\bar{\lambda} \in \Lambda(\bar{x})$, the pair $(\bar{x}, \bar{\lambda})$ satisfies SOC; that is,*

$$d^T \nabla^2 f(\bar{x}) d + \left\langle D^2 G(\bar{x})[d,d], \bar{\lambda} \right\rangle - \sigma(d, \bar{x}, \bar{\lambda}) \geq 0$$

holds for every $d \in C(\bar{x}) = \mathcal{L}_\Omega(\bar{x}) \cap \{\nabla f(\bar{x})\}^\perp$.

Proof. If $\Lambda(\bar{x}) = \emptyset$, then the result holds trivially; otherwise, let $\bar{\lambda} \in \Lambda(\bar{x})$ be arbitrary and fixed. Let $r := \text{rank}(G(\bar{x}))$, and let $\bar{E} \in \mathbb{R}^{m \times m-r}$ and $\bar{P} \in \mathbb{R}^{m \times r}$ be matrices with orthonormal eigenvector columns associated with the zero and positive eigenvalues of $G(\bar{x})$, respectively. Define $\bar{U} := [\bar{E}, \bar{P}]$. Moreover, let $\mathcal{E}$ be as in Lemma 5.1 and such that $\lim_{x \to \bar{x}} E(x) = \bar{E}$, where $E(x) := \mathcal{E}(G(x))$ for every $x \in \mathbb{R}^n$.

Now, let $d \in C(\bar{x})$ be arbitrary; so $\nabla f(\bar{x})^T d = 0$ and $\bar{E}^T DG(\bar{x}) d\bar{E} \succeq 0$. Following the proof of Theorem 5.1, let $\bar{Q} := [\bar{Z}, \bar{W}] \in \mathbb{R}^{m-r \times m-r}$ be an orthogonal matrix such that $\bar{Z}^T \bar{E}^T DG(\bar{x}) d\bar{E}\bar{Z} = 0$ and $\bar{W}^T \bar{E}^T DG(\bar{x}) d\bar{E}\bar{W} \succ 0$, and let s denote the rank of $\bar{E}^T DG(\bar{x}) d\bar{E}$. Moreover, let $F \trianglelefteq \mathbb{S}_+^{m-r}$ be defined as in (37); that is:

$$F := \left\{ \bar{Q} \begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix} \bar{Q}^T \; \middle| \; A_{11} \in \mathbb{S}_+^s \right\},$$

and note that $\bar{E}^T DG(\bar{x}) d\bar{E} \in F$. Similarly to the proof of Theorem 5.1, since the facial constant rank property holds at $\bar{x}$, there exists some $\varepsilon > 0$ and a twice continuously differentiable curve $\xi \colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ such that $\xi(0) = \bar{x}$, $\xi'(0) = d$, and

$$\mathcal{G}(\xi(t)) \in \text{span}(F)$$

for all $t \in (-\varepsilon, \varepsilon)$. Moreover, $\mathcal{G}(\xi(t)) \in F$ for every $t \in [0, \varepsilon)$. Since $\bar{x}$ is a local minimizer of (NSDP) and $\xi(t)$ is feasible for every small $t \geq 0$, then $t = 0$ is a local minimizer of the function $\phi(t) := f(\xi(t))$ subject to $t \geq 0$. Consequently, it is easy to see that

$$\phi''(0) = d^T \nabla^2 f(\bar{x}) d + \nabla f(\bar{x})^T \xi''(0) \geqslant 0. \tag{39}$$

The rest of the proof consists of computing the term $\nabla f(\bar{x})^T \xi''(0)$. By construction, we have $\mathcal{G}(\xi(t)) \in \text{span}(F)$ for every $t \in (-\varepsilon, \varepsilon)$, so $\bar{Z}^T \mathcal{G}(\xi(t))\bar{Z} = 0$ and the reduced complementarity function

$$R(t) := \left\langle \bar{Z}^T \mathcal{G}(\xi(t))\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle$$

has value zero, for all $t \in (-\varepsilon, \varepsilon)$. Therefore,

$$
\begin{aligned}
R'(t) &= \left\langle \bar{Z}^T (DE(\xi(t))\xi'(t))^T G(\xi(t)) E(\xi(t))\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle \\
&+ \left\langle \bar{Z}^T E(\xi(t))^T DG(\xi(t))\xi'(t) E(\xi(t))\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle \\
&+ \left\langle \bar{Z}^T E(\xi(t))^T G(\xi(t)) DE(\xi(t))\xi'(t)\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle
\end{aligned}
$$

also has value zero for every small t. Differentiating once more, and taking the limit $t \to 0$, we obtain

$$
\begin{aligned}
R''(0) &= \left\langle \bar{Z}^T \bar{E}^T D^2 G(\bar{x})[d,d]\bar{E}\bar{Z} + \bar{Z}^T \bar{E}^T DG(\bar{x})\xi''(0)\bar{E}\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle \\
&+ 2\left\langle \bar{Z}^T (DE(\bar{x})d)^T DG(\bar{x}) d\bar{E}\bar{Z} + \bar{Z}^T \bar{E}^T DG(\bar{x}) dDE(\bar{x})d\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle \\
&+ 2\left\langle \bar{Z}^T (DE(\bar{x})d)^T G(\bar{x}) DE(\bar{x}) d\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle = 0. \tag{40}
\end{aligned}
$$

However, following Shapiro and Fan [57, Equation 3.8], and Bonnans and Ramírez [23, Equation 67], we see that

$$DE(\bar{x})d = D\mathcal{E}(G(\bar{x})) DG(\bar{x}) d = -G(\bar{x})^\dagger DG(\bar{x}) d\bar{E}. \tag{41}$$

Substituting (41) into (40), the two last lines of expression (40) can be greatly simplified, which leads to the following:

$$
\begin{aligned}
R''(0) &= \left\langle \bar{Z}^T \bar{E}^T \left(D^2 G(\bar{x})[d,d] + DG(\bar{x})\xi''(0) \right) \bar{E}\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle \\
&- 2\left\langle \bar{Z}^T \bar{E}^T (DG(\bar{x})d)^T G(\bar{x})^\dagger DG(\bar{x}) d\bar{E}\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle = 0. \tag{42}
\end{aligned}
$$

However, by the complementarity condition we have $\bar{\lambda}\bar{P} = 0$, and using the KKT condition together with $\nabla f(\bar{x})^T d = 0$ we obtain

$$
\begin{aligned}
0 = \langle d, -\nabla f(\bar{x}) \rangle &= \langle d, DG(\bar{x})^T \bar{\lambda} \rangle = \langle DG(\bar{x}) d, \bar{\lambda} \rangle \\
&= \left\langle \bar{U}^T DG(\bar{x}) d\bar{U}, \; \bar{U}^T \bar{\lambda}\bar{U} \right\rangle = \left\langle \bar{E}^T DG(\bar{x}) d\bar{E}, \; \bar{E}^T \bar{\lambda}\bar{E} \right\rangle \\
&= \left\langle \bar{Q}^T \bar{E}^T DG(\bar{x}) d\bar{E}\bar{Q}, \; \bar{Q}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Q} \right\rangle \\
&= \left\langle \bar{Z}^T \bar{E}^T DG(\bar{x}) d\bar{E}\bar{Z}, \; \bar{Z}^T \bar{E}^T \bar{\lambda}\bar{E}\bar{Z} \right\rangle,
\end{aligned}
$$

but since $\bar{Z}^T \bar{E}^T DG(\bar{x})d\bar{E}\bar{Z} \succ 0$ and $\bar{W}^T \bar{E}^T \bar{\lambda} \bar{E}\bar{W} \preceq 0$, this implies

$$\bar{Z}^T \bar{E}^T \bar{\lambda} \bar{E}\bar{Z} = 0,$$

which in turn implies $\bar{\lambda}\bar{E}\bar{Z} = 0$. With this at hand, we obtain

$$R''(0) = \left\langle D^2 G(\bar{x})[d,d] + DG(\bar{x})\xi''(0) - 2(DG(\bar{x})d)^T G(\bar{x})^\dagger DG(\bar{x})d, \ \bar{\lambda} \right\rangle = 0, \tag{43}$$

and, by the KKT conditions, this leads to

$$\nabla f(\bar{x})^T \xi''(0) = -\langle DG(\bar{x})\xi''(0), \bar{\lambda}\rangle = \left\langle D^2 G(\bar{x})[d,d] - 2(DG(\bar{x})d)^T G(\bar{x})^\dagger DG(\bar{x})d, \ \bar{\lambda} \right\rangle. \tag{44}$$

Substituting (44) into (39) yields

$$d^T \nabla^2 f(\bar{x})d + \left\langle D^2 G(\bar{x})[d,d], \bar{\lambda} \right\rangle - d^T \mathcal{H}(\bar{x}, \bar{\lambda})d \geq 0.$$

Since $d \in C(\bar{x})$ was chosen arbitrarily, and $\bar{\lambda}$ is fixed from the beginning, the proof is complete.

$\square$

6   Final remarks

The constant rank constraint qualification (CRCQ) is one of the most important regularity conditions in nonlinear programming (NLP), with several relevant applications regarding global convergence of algorithms, second-order optimality conditions, and some topics of stability theory. However, one of the main reasons why most of these interesting results still remain exclusive to NLP is that CRCQ itself seems intrinsic to NLP. Until very recently, there was no extension or analogue of it in the conic programming context. In a recent pair of papers [8, 9], we presented an extension of CRCQ for nonlinear semidefinite and second-order cone programming using sequences and the eigenvector structure of their respective cones, which would allow us to adopt a strategy similar to the existing nonlinear programming literature. See also [7]. While this is interesting from the point of view of algorithms, it may not be an appropriate tool for other uses. Therefore, in this paper we adopted a more innovative approach: we first characterized CRCQ for NLP in a geometrical way, by means of the faces of a reduced cone, and then we showed this geometrical characterization could carry the essence of CRCQ to more general contexts. As far as we know, this is also the first time an intuitive interpretation of CRCQ was ever presented, and it is surprisingly simple: CRCQ describes the situations where every possible linear approximation of the feasible set (around a point of interest) preserves its dimension under small perturbations. As a side note, we should mention that this definition is either independent or strictly weaker than the ones presented in [8, 9].

As an application of our results, we obtained a strong second-order necessary optimality condition under CRCQ, in terms of any given Lagrange multiplier. This improves the classical result that is obtained under nondegeneracy, and serves as an alternative for the condition that can be obtained under Robinson's CQ, where for each direction in the critical cone, there is a Lagrange multiplier satisfying the second-order condition. We expect CRCQ to be an alternative to Robinson's CQ in other situations, especially those related with stability analysis of parametric nonlinear conic optimization programs, in view of the nonlinear programming literature – see, for instance, references [34, 37, 48]. Since CRCQ is independent of Robinson's CQ, we believe that this work allows the development of a new parallel strand in the study of stability. In a recent work, Gfrerer and Mordukhovich [29], fully characterized *tilt stable* local minimizers of NLP problems under the so-called *bounded extreme point property*, which is implied by CRCQ (improving a previous work that assumed CRCQ and MFCQ [50]), and one of the possibilities of future work mentioned by them is an extension for conic programs. In a related work, still for NLP, Gfrerer and Outrata [30] obtained similar results to [29] as an application of the generalized derivative of a particular set-valued mapping, which was computed assuming the *metric subregularity constraint qualification* (MSCQ) at the point of interest plus Robinson's CQ in its neighborhood. The CRCQ condition as presented in this paper may replace these assumptions in a possible extension of their results to NSOCP and NSDP. Moreover, we expect CRCQ to be useful for supporting the convergence theory of some iterative algorithms, and also to encourage the development of algorithms that rely on faces for solving nonlinear conic problems.

Regarding prospective work, the techniques employed in this paper strongly suggest that a further extension of CRCQ, for general reducible cones, is possible. In this paper we adopted a more pragmatic approach by working explicitly with NSOCP and NSDP for clarity, leaving the investigation of a more general result to future works. In fact, it would also be interesting to not rely on reducibility at all, which should be possible

by taking into account the faces of the tangent cone to $\mathcal{K}$ at $g(\bar{x})$, or perturbations of it, instead of the faces of the reduced cone $\mathcal{C}$. Furthermore, this work may inspire extensions of weaker constant rank-type conditions from NLP (together with their applications) to the conic environment, with emphasis on the well-established *constant positive linear dependence* condition [10, 12, 52] and the *constant rank of the subspace component* condition [11].

Funding

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming Series B*, 95:3–51, 2003. `doi:10.1007/s10107-002-0339-5`.

[2] E. D. Andersen, C. Roos, and T. Terlaky. Notes on duality in second order and p-order cone optimization. *Optimization*, 51(4):627–643, 2002. `doi:10.1080/0233193021000030751`.

[3] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On augmented lagrangian methods with general lower-level constraints. *SIAM Journal on Optimization*, 18(4):1286–1309, 2008. `doi:10.1137/060654797`.

[4] R. Andreani, C. E. Echagüe, and M. L. Schuverdt. Constant-rank condition and second-order constraint qualification. *Journal of Optimization theory and Applications*, 146(2):255–266, 2010. `doi:10.1007/s10957-010-9671-8`.

[5] R. Andreani, E. H. Fukuda, G. Haeser, H. Ramírez C., D. O. Santos, P. J. S. Silva, and T. P. Silveira. Erratum to: New constraint qualifications and optimality conditions for second order cone programs. *To appear in Set-Valued and Variational Analysis*, 2021. `doi:10.1007/s11228-021-00573-5`.

[6] R. Andreani, G. Haeser, L. M. Mito, H. Ramırez, D. O. Santos, and T. P. Silveira. Naive constant rank-type constraint qualifications for multifold second-order cone programming and semidefinite programming. *To appear in Optimization Letters*, 2020. `doi:10.1007/s11590-021-01737-w`.

[7] R. Andreani, G. Haeser, L. M. Mito, and H. Ramírez C. Weak notions of nondegeneracy in nonlinear semidefinite programming. Technical report, 2020. Available at ArXiv. URL: `https://arxiv.org/abs/2012.14810`.

[8] R. Andreani, G. Haeser, L. M. Mito, and H. Ramírez C. Sequential constant rank constraint qualifications for nonlinear semidefinite programming with applications. Technical report, 2021. Available at ArXiv. URL: `https://arxiv.org/abs/2106.00775v2`.

[9] R. Andreani, G. Haeser, L M. Mito, H. Ramírez C., and T. P. Silveira. Sequential constant rank for nonlinear second-order cone programming problems. Technical report, 2021.

[10] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, 135(1-2):255–273, 2012. `doi:10.1007/s10107-011-0456-0`.

[11] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. Two new weak constraint qualifications and applications. *SIAM Journal on Optimization*, 22(3):1109–1135, 2012. `doi:10.1137/110843939`.

[12] R. Andreani, J. M. Martínez, and M. L. Schuverdt. On the Relation between Constant Positive Linear Dependence Condition and Quasinormality Constraint Qualification. *Journal of Optimization Theory and Applications*, 125(2):473–483, 2005. `doi:10.1007/s10957-004-1861-9`.

[13] R. Andreani, J. M. Martínez, and M. L. Schuverdt. On second-order optimality conditions for nonlinear programming. *Optimization*, 56:529–542, 2007. `doi:10.1080/02331930701618617`.

[14] M. Anitescu. Degenerate nonlinear programming with a quadratic growth condition. *SIAM Journal on Optimization*, 10(4):1116–1135, 2000. `doi:10.1137/S1052623499359178`.

[15] A. Arutyunov. Second-order conditions in extremal problems. The abnormal points. *Transactions of the American Mathematical Society*, 350(11):4341–4365, 1998. `doi:10.1090/S0002-9947-98-01775-9`.

[16] A. Auslender and H. Ramírez. Penalty and barrier methods for convex semidefinite programming. *Mathematical Methods of Operations Research*, 63:Article number: 195, 2006. `doi:10.1007/s00186-005-0054-0`.

[17] A. Baccari. On the classical necessary second-order optimality conditions. *Journal of Optimization Theory and Applications*, 123(1):213–221, 2004. `doi:10.1023/B:JOTA.0000043998.04008.e6`.

[18] R. Behling, G. Haeser, A. Ramos, and D. S. Viana. On a conjecture in second-order optimality conditions. *Journal of Optimization Theory and Applications*, 176(3):625–633, 2018. Extended version at arXiv:1706.07833. `arXiv:arXiv:1706.07833v1`, `doi:10.1007/s10957-018-1229-1`.

[19] A. Ben-Tal and J. Zowe. A unified theory of first and second order conditions for extremum problems in topological vector spaces. In *Optimality and stability in mathematical programming (Mathematical Programming Studies)*, pages 39–76. Springer, 1982. `doi:10.1007/BFb0120982`.

[20] J. F. Bonnans. A semi-strong sufficiency condition for optimality in non convex programming and its connection to the perturbation problem. *Journal of Optimization Theory and Applications*, 60:7–18, 1989.

[21] J. F. Bonnans, R. Cominetti, and A. Shapiro. Second order optimality conditions based on parabolic second order tangent sets. *SIAM Journal on Optimization*, 9(2):466–492, 1999. `doi:10.1137/S1052623496306760`.

[22] J. F. Bonnans and H. Ramírez. Perturbation analysis of second-order cone programming problems. *Mathematical Programming*, 104(2):205–227, 2005. `doi:10.1007/s10107-005-0613-4`.

[23] J. F. Bonnans and H. Ramírez. Strong regularity of semidefinite programming problems. Technical report, 2005. DIM-CMM N° B-05/06-137.

[24] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Verlag, New York, 2000.

[25] 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, 2020. `doi:10.1137/19M1306804`.

[26] R. Cominetti. Metric regularity, tangent sets, and second-order optimality conditions. *Applied Mathematics and Optimization*, 21(1):265–287, 1990. `doi:10.1007/BF01445166`.

[27] A. Forsgren. Optimality conditions for nonconvex semidefinite programming. *Mathematical Programming*, 88(1):105–128, 2000. `doi:10.1007/PL00011370`.

[28] E. H. Fukuda, G. Haeser, and L. M. Mito. Second-order analysis for semidefinite and second-order cone programming via sequential optimality conditions. Technical report, 2020. Available at Optimization Online. URL: `http://www.optimization-online.org/DB_HTML/2020/08/7951.html`.

[29] H. Gfrerer and B. S. Mordukhovich. Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions. *SIAM Journal on Optimization*, 25(4):2081–2119, 2015. `doi:10.1137/15M1012608`.

[30] H. Gfrerer and J. V. Outrata. On computation of generalized derivatives of the normal-cone mapping and their applications. *Mathematics of Operations Research*, 41(4):1535–1556, 2016. `doi:10.1287/moor.2016.0789`.

[31] F. J. Gould and J. W. Tolle. Optimality conditions and constraint qualifications in banach space. *Journal of Optimization Theory and Applications*, 15(6):667–684, 1975. `doi:10.1007/BF00935506`.

[32] M. Guignard. Generalized Kunh-Tucker conditions for mathematical programming in a banach space. *SIAM Journal of Control*, 7:232–241, 1969. `doi:10.1137/0307016`.

[33] L. Guo, H.-H. Lin, and Ye. Second-order optimality conditions for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 158:33–64, 2013. `doi:10.1007/s10957-012-0228-x`.

[34] L. Guo, H.-H. Lin, J. J. Ye, and J. Zhang. Sensitivity analysis of the value function for parametric mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 24(3):1206–1237, 2014. `doi:10.1137/130929783`.

[35] R. Henrion, A. Y. Kruger, and J. V. Outrata. Some remarks on stability of generalized equations. *Journal of Optimization Theory and Applications*, 159:681–697, 2013. `doi:10.1007/s10957-012-0147-x`.

[36] T. Hoheisel, C. Kanzow, and A. Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Mathematical Programming*, 137:257–288, 2013. `doi:10.1007/s10107-011-0488-5`.

[37] R. Janin. Directional derivative of the marginal function in nonlinear programming. In *Sensitivity, Stability and Parametric Analysis (Mathematical Programming Studies)*, pages 110–126. Springer Berlin Heidelberg, 1984. `doi:10.1007/BFb0121214`.

[38] F. Jarre. Elementary optimality conditions for nonlinear SDPs. In *Handbook on Semidefinite, Conic and Polynomial Optimization*. International Series in Operations Research & Management Science, 2012. `doi:10.1007/978-1-4614-0769-0_16`.

[39] H. Jiang and D. Ralph. Smooth sqp methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, 10(3):779–808, 2000. `doi:10.1137/S1052623497332329`.

[40] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[41] H. Kawasaki. An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems. *Mathematical Programming*, 41(1):73–96, 1988. `doi:10.1007/BF01580754`.

[42] T. W. Liu. A reduced hessian sqp method for inequality constrained optimization. *Computational Optimization and Applications*, 49:31–59, 2011. `doi:10.1007/s10589-009-9285-y`.

[43] B. F. Lourenço, E. H. Fukuda, and M. Fukushima. Optimality conditions for nonlinear semidefinite programming via squared slack variables. *Mathematical Programming*, 166:1–24, 2016. `doi:10.1007/s10107-016-1040-4`.

[44] M. C. Maciel, S. A. Santos, and G. N. Sottosanto. On second-order optimality conditions for vector optimization. *Journal of Optimization Theory and Applications*, 149:332–351, 2011. `doi:10.1007/s10957-010-9793-z`.

[45] P. Mehlitz and L. I. Minchenko. R-regularity of set-valued mappings under the relaxed constant positive linear dependence constraint qualification with applications to parametric and bilevel optimization. *To appear in Set-Valued and Variational Analysis*, 2021. `doi:10.1007/s11228-021-00578-0`.

[46] L. Minchenko and A. Leschov. On strong and weak second-order necessary optimality conditions for nonlinear programming. *Optimization*, 65(9):1693–1702, 2016. `doi:10.1080/02331934.2016.1179300`.

[47] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. *Optimization*, 60(4):429–440, 2011. `doi:10.1080/02331930902971377`.

[48] L. Minchenko and S. Stakhovski. Parametric nonlinear programming problems under the relaxed constant rank condition. *SIAM Journal on Optimization*, 1(314–332), 2011. `doi:10.1137/090761318`.

[49] M. R. C. Monte and V. A. de Oliveira. A constant rank constraint qualification in continuous-time nonlinear programming. *Set-Valued and Variational Analysis*, 29:61–81, 2021. `doi:10.1007/s11228-020-00537-1`.

[50] B. S. Mordukhovich and T. T. A. Nghia. Second-order characterizations of tilt stability with applications to nonlinear programming. *Mathematical Programming*, 149:83–104, 2015. `doi:10.1007/s10107-013-0739-8`.

[51] G. Pataki. The geometry of semidefinite programming. In R. Saigal, L. Vandenberghe, and H. Wolkowicz, editors, *Handbook of Semidefinite Programming.*, pages 29–65. Kluwer Academic Publishers, Waterloo, Canada, 2000. `doi:10.1007/978-1-4615-4381-7_3`.

[52] L. Qi and Z. Wei. On the constant positive linear dependence conditions and its application to SQP methods. *SIAM Journal on Optimization*, 10(4):963–981, 2000. `doi:10.1137/S1052623497326629`.

[53] S. Qiu. A globally convergent regularized interior point method for constrained optimization. *To appear in Optimization Methods and Software*, 2021. `doi:10.1080/10556788.2021.1908283`.

[54] S. M. Robinson. First-order conditions for general nonlinear optimization. *SIAM Journal on Applied Mathematics*, 30(4):597–610, 1976. `doi:10.1137/0130053`.

[55] A. Shapiro. First and second order analysis of nonlinear semidefinite programs. *SIAM Journal on Optimization*, 77(1):301–320, 1997. `doi:10.1007/BF02614439`.

[56] A. Shapiro. On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints. *SIAM Journal on Optimization*, 7:508–518, 1997. `doi:10.1137/S1052623495279785`.

[57] A. Shapiro and M. K. H. Fan. On eigenvalue optimization. *SIAM Journal on Optimization*, 5(3):552–569, 1995. `doi:10.1137/0805028`.

[58] S. Steffensen and M. Ulbrich. A new relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 20(5):2504–2539, 2010. `doi:10.1137/090748883`.

[59] S. J. Wright. Modifying SQP for degenerate problems. *SIAM Journal on Optimization*, 13(2):470–497, 2002. `doi:10.1137/S1052623498333731`.

[60] M. Xu and J. J. Ye. Relaxed constant positive linear dependence constraint qualification and its application to bilevel programs. *Journal of Global Optimization*, 78:181–205, 2020. `doi:10.1007/s10898-020-00907-x`.

[61] Y. Zhang and L. Zhang. New constraint qualifications and optimality conditions for second order cone programs. *Set-Valued and Variational Analysis*, 27:693–712, 2019. `doi:10.1007/s11228-018-0487-2`.