On constraint qualifications for second-order optimality conditions depending on a single Lagrange multiplier *

Gabriel Haeser † Alberto Ramos ‡

November 29, 2019

Abstract

Second-order optimality conditions play an important role in continuous optimization. In this paper, we present and discuss new constraint qualifications to ensure the validity of some well-known second-order optimality conditions. Our main interest is on second-order conditions that can be associated with numerical methods for solving constrained optimization problems. Such conditions depend on a single Lagrange multiplier, instead of the whole set of Lagrange multipliers, and they are consistent with second-order algorithms where, usually, at each iteration, one only has access to a single approximate Lagrange multiplier. For each condition, we characterize the weakest second-order constraint qualification that guarantees its fulfilment at local minimizers, while proposing new weak conditions implying them. Relations with other constraint qualifications stated in the literature are discussed.

Keywords: Nonlinear optimization, constraint qualifications, second-order optimality conditions.

Mathematics Subject Classification 90C30 · 90C46

1 Introduction

In this paper, we study second-order optimality conditions for the constrained optimization problem of the form

$$\begin{array}{ll} \text{minimize} & f(x), \\ \text{subject to} & h_i(x) = 0 \quad \forall i \in \mathcal{E} := \{1, \dots, m\}, \\ & g_j(x) \leq 0 \quad \forall j \in \mathcal{J} := \{1, \dots, p\}, \end{array} \tag{1}$$

*This work was supported by CNPq (Grants 438185/2018-8, 404656/2018-8 and 303427/2018-3) and FAPESP (Grants 2017/18308-2 and 2018/24293-0).

†Department of Applied Mathematics, University of São Paulo, São Paulo SP, Brazil. Email: ghaeser@ime.usp.br

‡Department of Mathematics, Federal University of Paraná, Curitiba, PR, Brazil. Email: albertoramos@ufpr.br

where we assume that all functions f, h_i, $g_j : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable. Furthermore, we denote by Ω the feasible set of (1).

Numerical methods for solving (1) are usually iterative, and thus, in their implementation suitable stopping criteria must be considered. Useful tools for designing stopping criteria are the necessary optimality conditions, that is, conditions satisfied at every local minimizer. However, not all necessary optimality conditions serve this purpose. Hence, we only deal with optimality conditions that can be (approximately) checked with the information readily available at the current iteration, that is, a primal-dual solution (see [2]). For first-order methods for solving constrained optimization problems, this role is usually played by the Karush-Kuhn-Tucker (KKT) conditions. For more details, see [23, Framework 7.13], [17, Chapter 12] and [1, 7].

When one seeks more reliable methods and possibly with a faster convergence rate, one usually relies on second-order methods, which are based on second-order necessary optimality conditions. Observe that second-order necessary optimality conditions allow ruling out possible non-minimizers accepted as solutions when only first-order information is considered. In contrast with the first-order case, where KKT conditions are dominant, several second-order optimality conditions have been proposed in the literature, both from a theoretical and a practical point of view, see for instance [23, 17, 10, 25, 12, 15, 13] and references therein.

Following our previous paper [19], where we focused on an optimality condition depending on the whole set of Lagrange multipliers, in this paper, we focus on the *weak second-order optimality condition* (WSOC) and the *strong second-order optimality condition* (SSOC) defined in Definition 2.1. Both conditions rely on the existence of an adequate Lagrange multiplier, and not on the whole set of Lagrange multipliers.

Similarly to the first-order case, for the validity of SSOC or WSOC at local minimizer, one needs a constraint qualification (CQ) to hold. For the KKT conditions, it is known that Guignard's CQ is the weakest possible assumption that yields the KKT conditions at local minimizers, independently of the objective function (see [18]). In other words, the fact that Guignard's CQ is a constraint qualification can be seen as the implication

$$\text{local minimizer} \Rightarrow \text{KKT conditions hold or Guignard's CQ fails,}$$

that is, "KKT or not-Guignard" is a genuine first-order necessary optimality condition. However, because of the minimality property of Guignard's CQ, this optimality condition is too strong for practical purposes, since no algorithm is known that fulfills it at limit points of its sequence of iterates. The assumptions yielding global convergence of algorithms, in the sense described above, are stronger, see [3, 4, 7, 6].

Given the role of WSOC and SSOC in numerical algorithms, in this paper we are interested in discussing the weakest possible assumptions guaranteeing their validity at local minimizers. See Theorems 3.1 and 3.3. Furthermore, new weak constraint qualifications are introduced and some relations among them are analyzed. To develop our new constraint qualifications that guarantee the fulfilment of SSOC or WSOC at local minimizer of (1), we rely on some theoretical tools defined in [19], which will be properly presented.

This paper is organized as follows. In Section 2 we present the formal definition of the second-order conditions. In Section 3, new CQs for WSOC and SSOC are introduced. In Section 4 we give some concluding remarks.

Notations: By $\mathbb{R}^n$, we denote the n-dimensional real Euclidean space, $n \in \mathbb{N}$, while $\mathbb{R}^n_+ \subseteq \mathbb{R}^n$ is the set of vectors with non-negative components. The set of real symmetric matrices of order n is denoted by $\mathrm{Sym}(n)$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product on $\mathbb{R}^n$ and on $\mathrm{Sym}(n)$, with $\| \cdot \|$ the associated norm. Given $A \in \mathrm{Sym}(n)$ and $v \in \mathbb{R}^n$, we use Av^2 to denote $\langle v, Av \rangle$ and $\|A\|_2$ the operator-norm induced by $\| \cdot \|$. For a set $S \subseteq \mathbb{R}^n$, the symbol $z \xrightarrow{S} z^*$ means that $z \to z^*$ with $z \in S$. For a given cone $\mathcal{K} \subseteq \mathbb{R}^s$, its polar is denoted by $\mathcal{K}^\circ := \{v \in \mathbb{R}^s : \langle v, k \rangle \leq 0 \text{ for all } k \in \mathcal{K}\}$. For every set $C \subseteq \mathbb{R}^n$, we use $\mathrm{lin}(C)$ to denote its lineality space, that is, the largest subspace included in C. When C is a closed convex cone, $\mathrm{lin}(C) = C \cap -C$. For $A, B \in \mathrm{Sym}(n)$, and $\mathcal{V}$ a subset of $\mathbb{R}^n$, we use $A \preceq B$ on $\mathcal{V}$ to say that $\langle v, Av \rangle \leq \langle v, Bv \rangle$ for every $v \in \mathcal{V}$, and when the previous relation holds with A replaced by a matrix with all entries equal to zero, we say that B is positive semi-definite on $\mathcal{V}$. Given a set-valued mapping $\varGamma : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the *sequential Painlevé-Kuratowski outer limit* of $\varGamma(z)$ as $z \to z^*$ is defined by

$$\limsup_{z \to z^*} \varGamma(z) := \{w^* \in \mathbb{R}^d : \exists \ (z^k, w^k) \to (z^*, w^*) \text{ with } w^k \in \varGamma(z^k)\}.$$

We say that $\varGamma$ is *outer semicontinuous* at z^* if $\limsup_{z \to z^*} \varGamma(z) \subseteq \varGamma(z^*)$.

Finally, the regular normal cone to Ω at $\bar{z}$ is defined as:

$$N_1(\bar{z}) := \left\{ w : \operatorname*{limsup}_{z \in \Omega, z \to \bar{z}} \frac{\langle w, z - \bar{z} \rangle}{\|z - \bar{z}\|^2} \leq 0 \right\}.$$

2 Second-order Conditions for Nonlinear Optimization

Let us describe the second-order conditions that we will deal with in this paper. Let us start by the following standard notations and definitions. The Lagrangian function for problem (1) is defined as

$$L(x, \lambda, \mu) := f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x),$$

where $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p_+$. We use $\nabla_x L(x, \lambda, \mu)$ and $\nabla^2_{xx} L(x, \lambda, \mu)$ for the gradient and Hessian of $L(x, \lambda, \mu)$ with respect to x, respectively.

For a feasible point $\bar{x}$, we denote by $A(\bar{x}) := \{j \in \mathcal{J} : g_j(\bar{x}) = 0\}$, the set of indexes of active inequality constraints. Furthermore, we denote by $\Lambda(\bar{x})$

the set of Lagrange multipliers associated with problem (1) at $\bar{x}$. That is, $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ belongs to $\Lambda(\bar{x})$ if, and only if

$$\nabla_x L(x, \lambda, \mu) = 0, \quad \text{and} \quad \mu_j g_j(\bar{x}) = 0, \;\; \forall j \in \mathcal{J}. \tag{2}$$

Now, we consider the cone of critical directions (critical cone) defined as follows:

$$C(\bar{x}) := \left\{ d \in \mathbb{R}^n : \begin{array}{l} \langle \nabla f(\bar{x}), d \rangle = 0; \;\; \langle \nabla h_i(\bar{x}), d \rangle = 0, i \in \mathcal{E} \\ \langle \nabla g_j(\bar{x}), d \rangle \leq 0, j \in A(\bar{x}) \end{array} \right\}. \tag{3}$$

Obviously, $C(\bar{x})$ is a non-empty closed convex cone. From the algorithmic point of view, an important set is the critical subspace, given by:

$$S(\bar{x}) := \{ d \in \mathbb{R}^n : \langle \nabla h_i(\bar{x}), d \rangle = 0, i \in \mathcal{E}; \langle \nabla g_j(\bar{x}), d \rangle = 0, j \in A(\bar{x}) \}. \tag{4}$$

Now, we are able to define the classical second-order conditions.

Definition 2.1. *Let $\bar{x}$ be a feasible point of* (1) *with $\Lambda(\bar{x}) \neq \emptyset$. We say that*

1. *the strong second-order optimality condition (SSOC) holds at $\bar{x}$ if there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ such that $\nabla^2_{xx} L(\bar{x}, \lambda, \mu) d^2 \geq 0$ for every $d \in C(\bar{x})$;*

2. *the weak second-order optimality condition (WSOC) holds at $\bar{x}$ if there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ such that $\nabla^2_{xx} L(\bar{x}, \lambda, \mu) d^2 \geq 0$ for every $d \in S(\bar{x})$.*

Note that both conditions depend only on the existence of a suitable Lagrange multiplier, while other second-oder conditions consider all the set of Lagrange multipliers, for instance, in [19] we did a similar discussion for the *basic second-order optimality condition* (BSOC), which requires that for every $d \in C(\bar{x})$, there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ (which depends on d) such that $\nabla^2_{xx} L(\bar{x}, \lambda, \mu) d^2 \geq 0$. See [14].

We refer the reader to the extended version of [11] for a discussion on the pivotal role of WSOC on global convergence of numerical algorithms.

2.1 Preliminaries and Basic Results

We proceed to present some results from [19] which will be used in our analysis. Given a set-valued mapping $B(\cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \times \mathbb{R}_+, d \mapsto B(d)$ and a set $A \subseteq \mathbb{R}^n$, we define the *polar pairing* of A and $B(\cdot)$, denoted by $[A, B(\cdot)]$, as the closed convex set defined as

$$[A, B(\cdot)] := \left\{ (w, W) \in \mathbb{R}^n \times \mathrm{Sym}(n) : \begin{array}{l} \langle w, d \rangle \leq 0, \quad \text{for all } d \in A \\ \langle w, z \rangle + \langle W, rdd^T \rangle \leq 0, \text{ for all} \\ (d, (z, r)) \in K_A(w) \times B(d) \end{array} \right\}, \tag{5}$$

where $K_A(v) := A \cap v^{\perp}$ for $v \in \mathbb{R}^n$. Note that for a given cone $K \subseteq \mathbb{R}^n$, we can recover the polar K°, using the polar pairing, in two distinct ways. By defining $K(\cdot) := K \times \{0\}$ we get $[\{0\}, K(\cdot)] = [K, \emptyset] = K^{\circ} \times \mathrm{Sym}(n)$.

The second-order normal cone $N_2(\bar{x}) \subseteq \mathbb{R}^n \times \mathrm{Sym}(n)$ to $\Omega \subseteq \mathbb{R}^n$ at $\bar{x} \in \Omega$ is defined as follows:

$$N_2(\bar{x}) := \left\{ (w, W) : \limsup_{x \in \Omega, x \to \bar{x}} \frac{\langle w, x - \bar{x} \rangle + (1/2) W (x - \bar{x})^2}{\|x - \bar{x}\|^2} \leq 0 \right\}. \tag{6}$$

From [19], we have the following characterization of $N_2(\bar{x})$:

Theorem 2.1. *Let $\bar{x} \in \Omega$, $w \in \mathbb{R}^n$ and $W \in \mathit{Sym}(n)$. Then, $(w, W) \in N_2(\bar{x})$ iff there exists a twice continuously differentiable function f that attains its global minimum relative to Ω at $\bar{x}$ such that $-\nabla f(\bar{x}) = w$ and $-\nabla^2 f(\bar{x}) = W$.*

In [19], the following inclusion is proved:

Theorem 2.2. *Let $\bar{x} \in \Omega$. Then, we have that $N_2(\bar{x}) \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$, where $T_1(\bar{x})$ is the (Bouligand) tangent cone to Ω at $\bar{x}$, $T_2^{proj}(x, d)$ denotes the projective second-order cone [24] to Ω at $\bar{x}$ and with direction d and $T_2^{proj}(\bar{x}, \cdot)$ is the set-valued mapping $d \rightrightarrows T_2^{proj}(x, d)$. See* (10) *and* (12) *below for the proper definition.*

Theorem 2.1 is a second-order analogue of the well known fact that the regular normal cone $N_1(\bar{x})$ coincides with $-\nabla f(\bar{x})$ for every continuously differentiable function f that attains its global minimum relative to Ω at $\bar{x}$, while Theorem 2.2 is a second-order version of the well known inclusion $N_1(\bar{x}) \subseteq T_1(\bar{x})^\circ$ (see [25, Theorem 6.28]), which together are knwon as the *geometric* necessary optimality condition. Hence, Theorems 2.1 and 2.2 yield a second-order geometric optimality condition. In the first-order case, it actually holds that $N_1(\bar{x}) = T_1(\bar{x})^\circ$, which simplifies the definition of the weakest CQ implying the KKT conditions (that is, Guignard's CQ), whereas in the second-order case, we do not have a characterization of $N_2(\bar{x})$.

For passing from the geometric necessary optimality condition to WSOC or SSOC, one must rely on a CQ. Besides the linear independence constraint qualification (LICQ), few CQs guarantee the existence of a Lagrange multipliers satisfying SSOC or WSOC. Not even the Mangasarian-Fromovitz CQ (MFCQ), which is a very useful condition for stability analysis, convergence analysis and many other applications, serves this purpose. We say that MFCQ holds at $\bar{x}$ if there is a nonzero vector $d \in \mathbb{R}^n$ such that

$$\langle \nabla h_i(\bar{x}), d \rangle = 0, \quad i \in \mathcal{E} \quad \text{and} \quad \langle \nabla g_j(\bar{x}), d \rangle < 0, \quad j \in A(\bar{x}). \tag{7}$$

It is well-known that MFCQ alone is not enough to imply WSOC at a local minimizer (see [8, 9]). By adding to MFCQ a constant rank assumption, WSOC is guaranteed to hold, see [5]. The constant rank assumption needed for this task is the so-called *weak constant rank* condition (WCR), which holds at a feasible point $\bar{x}$ if the set $\{\nabla h_i(x), \nabla g_j(x) : i \in \mathcal{E}, j \in A(\bar{x})\}$ has constant rank for every x near $\bar{x}$.

We end this section by a version of the well known Constant Rank Theorem [21], which will be needed in our analysis. Here, we prove an uniform bound on the second-order derivative of the curve α guaranteed to exist. The proof will be given in the appendix.

Lemma 2.1. *Let $\{F_1, \ldots, F_N\}$ be a family of C^2 real valued functions and $x \in \mathbb{R}^n$ be a point such that $F_j(x) = 0, \forall j = 1, \ldots, N$. Assume that there is an open neighbourhood $\mathcal{O}$ of x such that $\{\nabla F_1(y), \ldots, \nabla F_N(y)\}$ has constant rank for every $y \in \mathcal{O}$. Take $d \in \mathbb{R}^n$ such that $\langle \nabla F_j(x), d\rangle = 0$, for $j = 1, \ldots, N$.*

Then, there exists a C^2 curve $\alpha :]-t_0, t_0[\to \mathbb{R}^n$ $(t_0 > 0)$ such that

$$\alpha(0) = x, \quad \alpha'(0) = d, \text{ and } F_j(\alpha(t)) = 0, \forall t \in]-t_0, t_0[, \forall j = 1, \ldots, N. \tag{8}$$

Furthermore, we can choose $\alpha(t)$ such that

$$\|\alpha''(0)\| \leq K_F \|d\|^2, \tag{9}$$

where K_F is independent of d and depends only on the functions $F_1, \ldots, F_N$ and their derivatives up to order two evaluated at x.

3 New Constraint Qualifications for Second-Order Optimality Conditions

A constraint qualification is needed to relate some geometric objects (as the tangent cones) with some analytic objects (as the linearized tangent cones). In the first-order case, we define the *tangent cone* to Ω at $\bar{x}$ as follows:

$$T_1(\bar{x}) := \{d \in \mathbb{R}^n : \exists \ \ d^k \to d, t_k \downarrow 0, \text{ such that } \bar{x} + t_k d^k \in \Omega, \quad k \in \mathbb{N}\}. \tag{10}$$

Since $T_1(\bar{x})$ is a geometric object that can be difficult to compute, we consider its first-order approximation, which is known as the *linearized tangent cone* to Ω at $\bar{x}$ and is defined as follows:

$$L_1(\bar{x}) := \{d \in \mathbb{R}^n : \langle \nabla h_i(\bar{x}), d\rangle = 0, i \in \mathcal{E}; \quad \langle \nabla g_j(\bar{x}), d\rangle \leq 0, j \in \mathcal{A}(\bar{x})\}. \tag{11}$$

Using such objects, several important CQs were introduced in the literature. Guignard's CQ states that $T_1(\bar{x})^\circ = L_1(\bar{x})^\circ$, which gives rise to the more well-known (and stronger) Abadie's CQ, namely $T_1(\bar{x}) = L_1(\bar{x})$. For the second-order analysis, we consider the *projective second-order tangent cone* to Ω at $\bar{x}$, [24], defined as

$$T_2^{proj}(\bar{x}, d) := \limsup_{(t,r,t/r)\downarrow 0} \frac{\Omega - td - \bar{x}}{\frac{t^2}{2r}}. \tag{12}$$

As the tangent cone, the set $T_2^{proj}(\bar{x}, d)$ can be a difficult object to deal with, hence, we define the *projective linearized second-order tangent cone* of Ω at $\bar{x}$ as

$$L_2^{proj}(\bar{x}, d) := \left\{ (z, r) \in \mathbb{R}^n \times \mathbb{R}_+ : \begin{array}{l} \langle \nabla h_i(\bar{x}), z\rangle + r\nabla^2 h_i(\bar{x}) d^2 = 0, i \in \mathcal{E}; \\ \langle \nabla g_j(\bar{x}), z\rangle + r\nabla^2 g_j(\bar{x}) d^2 \leq 0, j \in A(\bar{x}, d) \end{array} \right\}, \tag{13}$$

where $A(\bar{x}, d) := \{j \in A(\bar{x}) : \langle \nabla g_j(\bar{x}), d\rangle = 0\}$. Clearly, $T_2^{proj}(\bar{x}, d) \subseteq L_2^{proj}(\bar{x}, d)$, $\forall d \in T_1(\bar{x})$. In fact, the relations among $L_1(\bar{x})$, $T_1(\bar{x})$, $L_2^{proj}(\bar{x}, d)$ and $T_2^{proj}(\bar{x}, d)$

is given by the use of the polar pairing. By [19, Proposition 4.1], the following inclusion holds:

$$[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)] \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]. \tag{14}$$

Such inclusion can be considered as a second-order version of the first-order inclusion $L_1(\bar{x})^\circ \subseteq T_1(\bar{x})^\circ$. Other useful cones are the non-projective versions of L_2 and T_2 as follows:

$$L_2(\bar{x}, d) := \left\{ z \in \mathbb{R}^n : \begin{array}{l} \langle \nabla h_i(\bar{x}), z \rangle + \nabla^2 h_i(\bar{x}) d^2 = 0, i \in \mathcal{E}; \\ \langle \nabla g_j(\bar{x}), z \rangle + \nabla^2 g_j(\bar{x}) d^2 \leq 0, j \in A(\bar{x}, d) \end{array} \right\},$$

which is a first-order approximation of the *second-order tangent cone* to Ω at $\bar{x}$, [14], defined as

$$T_2(\bar{x}, d) := \limsup_{t \downarrow 0} \frac{\Omega - td - \bar{x}}{t^2/2}.$$

In general, $T_2(\bar{x}, d) \subseteq L_2(\bar{x}, d)$. It is known that under MFCQ at $\bar{x}$, the equality $T_2(\bar{x}, d) = L_2(\bar{x}, d)$ holds, for every $d \in T_1(\bar{x})$ such that $T_2(\bar{x}, d) \neq \emptyset$, see [14]. The definition of the projective objects is motivated by the fact that $T_2(\bar{x}, d)$ may be empty. See [24]. In fact, note that $z \in L_2(\bar{x}, d)$ implies that $(z, 1) \in L_2^{proj}(\bar{x}, d)$, and if $(z, r) \in L_2^{proj}(\bar{x}, d)$ with $r > 0$, then $z/r \in L_2(\bar{x}, d)$.

As $L_2^{proj}(\bar{x}, d)$ is defined by linear equalities and inequalities, its polar can be easily calculated. Indeed, $(w, \eta) \in L_2^{proj}(\bar{x}, d)^\circ$ if, and only if, there exist multipliers $\lambda_i \in \mathbb{R}$, $i \in \mathcal{E}$; $\mu_j \geq 0$, $j \in A(\bar{x}, d)$ and $\beta \geq 0$, such that

$$\begin{pmatrix} w \\ \eta \end{pmatrix} = \sum_{i \in \mathcal{E}} \lambda_i \begin{pmatrix} \nabla h_i(\bar{x}) \\ \nabla^2 h_i(\bar{x}) d^2 \end{pmatrix} + \sum_{j \in A(\bar{x}, d)} \mu_j \begin{pmatrix} \nabla g_j(\bar{x}) \\ \nabla^2 g_j(\bar{x}) d^2 \end{pmatrix} - \begin{pmatrix} 0 \\ \beta \end{pmatrix}. \tag{15}$$

Finally, some remarks about the nomeclature used in this paper. When we say that certain CQ is of Abadie-type, we refer to a CQ that in its definition relates some notion of tangent cone and its linearized counter-part. Similarly, we say that a CQ is of Guignard-type if its definition consider relations of certain tangent cones and linerized cones but via the polar pairing or the polar operation.

3.1 New Constraint Qualifications for WSOC

Let us define new CQs that guarantee that WSOC holds at every local minimizer. For this purpose, we consider yet another linearized cone $L_{2,W}(\bar{x}) \subseteq \mathbb{R}^n \times \mathrm{Sym}(n)$ related with WSOC, defined as follows:

$$L_{2,W}(\bar{x}) := \left\{ (z, Z) : \begin{array}{l} \langle \nabla h_i(\bar{x}), z \rangle + \langle \nabla^2 h_i(\bar{x}), Z \rangle = 0, \quad i \in \mathcal{E}, \\ \langle \nabla g_j(\bar{x}), z \rangle + \langle \nabla^2 g_j(\bar{x}), Z \rangle \leq 0, \quad j \in A(\bar{x}), \\ \text{and} \quad Z \in \mathrm{cl}(\mathrm{conv}(\{dd^T : d \in S(\bar{x})\})). \end{array} \right\},$$

where $S(\bar{x})$ is the critical subspace defined in (4). Note that if $z \in L_2(\bar{x}, d)$ for $d \in S(\bar{x})$, then $(z, Z) \in L_{2,W}(\bar{x})$ where $Z := dd^T$. Thus, we can think of $L_{2,W}(\bar{x})$ as a relaxation of $L_2(\bar{x}, d)$ where the constraint $Z = dd^T$ with $d \in S(\bar{x})$ is replaced by the relaxed constraint $Z \in \mathrm{cl}(\mathrm{conv}(\{dd^T : d \in S(\bar{x})\}))$. Note that we have $A(\bar{x}, d) = A(\bar{x})$ for every $d \in S(\bar{x})$. This cone allows us to characterize WSOC in the following way. WSOC holds at $\bar{x}$ for an objective function f if, and only if $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in L_{2,W}(\bar{x})^\circ$.

To see this characterization, let us compute the polar cone of $L_{2,W}(\bar{x})$. Indeed, $(w, W) \in L_{2,W}(\bar{x})^\circ$ if, and only if, there are multipliers $\lambda_i \in \mathbb{R}$, $i \in \mathcal{E}$; $\mu_j \geq 0$, $j \in A(\bar{x})$ and a matrix $D \in \mathbb{R}^{n \times n}$ satisfying $\langle D, dd^T \rangle \geq 0$, for every $d \in S(\bar{x})$ such that

$$\begin{pmatrix} w \\ W \end{pmatrix} = \sum_{i \in \mathcal{E}} \lambda_i \begin{pmatrix} \nabla h_i(\bar{x}) \\ \nabla^2 h_i(\bar{x}) \end{pmatrix} + \sum_{j \in A(\bar{x})} \mu_j \begin{pmatrix} \nabla g_j(\bar{x}) \\ \nabla^2 g_j(\bar{x}) \end{pmatrix} - \begin{pmatrix} 0 \\ D \end{pmatrix}. \tag{16}$$

Thus, using (16), we see that $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in L_{2,W}(\bar{x})^\circ$ if and only if there exist mulitpliers $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p_+$ and a matrix $D \in \mathbb{R}^{n \times n}$ such that $Dd^2 \geq 0$ for all $d \in S(\bar{x})$. Thus, for every $d \in S(\bar{x})$, we get that

$$\nabla f(\bar{x}) + \sum_{i \in \mathcal{E}} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla g_j(\bar{x}) = 0,$$
$$\nabla^2 f(\bar{x}) d^2 + \sum_{i \in \mathcal{E}} \lambda_i \nabla^d h_i(\bar{x}) d^2 + \sum_{j \in A(\bar{x})} \mu_j \nabla^2 g_j(\bar{x}) d^2 = Dd^2 \geq 0,$$

which is exactly the definition of WSOC. It is worth noting that the mulitpliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p_+$ do not depends on the vector $d \in S(\bar{x})$.

Using the cone $L_{2,W}(\bar{x})$, Theorem 2.1 and the characterization of WSOC above, we arrive at the following:

Theorem 3.1. *WSOC holds for every C^2 objective function f which has $\bar{x}$ as a local minimizer relative to Ω if, and only if, $N_2(\bar{x}) \subseteq L_{2,W}(\bar{x})^\circ$.*

Now, we proceed to define more computable CQs for WSOC. Having in mind that $L_1(\bar{x})$ is a first-oder approximation of $T_1(\bar{x})$, we continue by considering a new geometric object, that we call $T_{2,W}(\bar{x})$, in such a way that $L_{2,W}(\bar{x})$ can be interpreted as a first-order approximation of $T_{2,W}(\bar{x})$.

Thus, we consider the cone $T_{2,W}(\bar{x})$ defined as the closure of

$$\left\{ (\zeta, Z) : \quad \begin{array}{l} \exists \zeta_a \text{ such that } \zeta - \zeta_a \in T_1(\bar{x}) \\ (\zeta_a, Z) \in \mathrm{conv}\left\{ (z, r dd^T) : (z, r) \in T_2^{proj}(\bar{x}, d); d \in \mathrm{lin}(T_1(\bar{x})) \right\} \end{array} \right\}. \tag{17}$$

Clearly, $T_{2,W}(\bar{x})$ is a non-empty closed cone. We summarize some properties whose proofs are straightforward verifications:

- Since $\mathrm{lin}(T_1(\bar{x})) \subseteq S(\bar{x})$, it is easy to see that $T_{2,W}(\bar{x}) \subseteq L_{2,W}(\bar{x})$;
- From the definition of the polar pairing and Theorem 2.2, we get that:

$$N_2(\bar{x}) \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)] \subseteq T_{2,W}(\bar{x})^\circ.$$

Clearly, the equality $T_{2,W}(\bar{x})^\circ = L_{2,W}(\bar{x})^\circ$ is a Guignard-type weak CQ to guarantee WSOC at a local minimizer $\bar{x}$, while $[T_1(\bar{x}), T_2^{proj}(\bar{x},\cdot)] = L_{2,W}(\bar{x})^\circ$ is a weaker Guignard-type CQ. Also, $T_{2,W}(\bar{x}) = L_{2,W}(\bar{x})$ is an Abadie-type CQ for WSOC. It is known that the condition MFCQ+WCR implies that WSOC holds at local minimizers (see [5]). In fact, MFCQ+WCR implies a stronger condition:

Theorem 3.2. *Let $\bar{x}$ be a feasible point. If MFCQ and WCR hold at $\bar{x}$, then $T_{2,W}(\bar{x}) = L_{2,W}(\bar{x})$.*

Proof. Take $(z, Z) \in L_{2,W}(\bar{x})$. Since $Z \in \mathrm{cl}(\mathrm{conv}(\{dd^T : d \in S(\bar{x})\}))$, there is a sequence $\{Z^k\} \subseteq \mathrm{conv}(\{dd^T : d \in S(\bar{x})\})$ such that $Z^k \to Z$. Without loss of generality we can assume that $Z^k = \sum_{\ell \in I(k)} d_{\ell k} d_{\ell k}^T$, where $\{d_{\ell k} : \ell \in I(k)\} \subseteq S(\bar{x})$ and the cardinality of $I(k)$ is uniformly bounded, namely, by Carathéodory's Lemma we have that $|I(k)| \le n^2 + 1$, $\forall k \in \mathbb{N}$.

Since $\{\|Z^k\|\}$ is bounded by some constant M, we can show that the sequence $\{d_{\ell k} : \ell \in I(k), k \in \mathbb{N}\}$ is also bounded. Indeed, this is a consequence of the following inequality: for every $v \in \mathbb{R}^n$, we always have that

$$|\langle d_{\ell k}, v\rangle|^2 \le \sum_{\ell \in I(k)} |\langle d_{\ell k}, v\rangle|^2 = \langle \sum_{\ell \in I(k)} d_{\ell k} d_{\ell k}^T, vv^T \rangle = \langle Z^k, vv^T\rangle \le M\|vv^T\|,$$

where the last step follows from the Cauchy-Schwarz inequality.

To show that (z, Z) belongs to $T_{2,W}(\bar{x})$ (see (17)), it will be sufficient to find sequences $\{z_\ell^k : \ell \in I(k)\}$ and $\{\zeta^k\}$ such that $(z_\ell^k, 1) \in T_2^{proj}(\bar{x}, d_{\ell k})$, $\zeta^k - \sum_{\ell \in I(k)} z_\ell^k \in T_1(\bar{x})$ and (ζ^k, Z^k) converges to (z, Z).

Thus, for each $k \in \mathbb{N}$, consider the optimization problem

$$\begin{array}{ll} \text{Minimize} & \frac{1}{2}\|z - \zeta\|^2 \\ \text{s.t.} & \langle \nabla h_i(\bar{x}), \zeta\rangle + \langle \nabla^2 h_i(\bar{x}), Z^k\rangle = 0, \quad i \in \mathcal{E}, \\ & \langle \nabla g_j(\bar{x}), \zeta\rangle + \langle \nabla^2 g_j(\bar{x}), Z^k\rangle \le 0, \quad j \in A(\bar{x}). \end{array} \tag{18}$$

Let ζ^k be a solution of (18), which exists if the feasible set of (18) is non empty. Then, we proceed by showing that the feasible set of (18) is non empty. Since $Z^k = \sum_{\ell \in I(k)} d_{\ell k} d_{\ell k}^T$, for each $d_{\ell k} \in S(\bar{x})$, using WCR and Lemma 2.1, there is a C^2 curve $\alpha_{\ell k}(t)$ and a $t_0 > 0$ such that $\alpha_{\ell k}(0) = \bar{x}$, $\alpha'_{\ell k}(0) = d_{\ell k}$ and for every $t \in]-t_0, t_0[$: $h_i(\alpha_{\ell k}(t)) = 0$, $\forall i$ and $g_j(\alpha_{\ell k}(t)) = 0$, $\forall j \in A(\bar{x})$. Furthermore, $z_\ell^k := \alpha''_{\ell k}(0)$ is bounded by a constant that depends on the data and $d_{\ell k}$. Indeed, by (9) of Lemma 2.1, we see that $\|z_{\ell k}\| \le K\|d_{\ell k}\|^2$. Since $\{d_{\ell k} : \ell \in I(k), k \in \mathbb{N}\}$ is bounded and $I(k)$ is finite, $\{\sum_{\ell \in I(k)} z_{\ell k} : k \in \mathbb{N}\}$ is a bounded sequence. Deriving twice the equations $h_i(\alpha_{\ell k}(t)) = 0$, $\forall i$ and $g_j(\alpha_{\ell k}(t)) = 0$, $\forall j \in A(\bar{x})$ at $t = 0$, we obtain that for every $\ell \in I(k)$

$$\begin{array}{l} \langle \nabla h_i(\bar{x}), z_\ell^k\rangle + \nabla^2 h_i(\bar{x}) d_{\ell k}^2 = 0, \quad i \in \mathcal{E}, \\ \langle \nabla g_j(\bar{x}), z_\ell^k\rangle + \nabla^2 g_j(\bar{x}) d_{\ell k}^2 = 0, \quad j \in A(\bar{x}). \end{array} \tag{19}$$

Furthermore, from (19), $z_\ell^k \in L_2(\bar{x}, d_{\ell k})$, $\forall \ell \in I(k)$ and thus $z_\ell^k \in T_2(\bar{x}, d_{\ell k})$, $\forall \ell \in I(k)$ (here, we have used that MFCQ implies $L_2(\bar{x}, d_{\ell k}) = T_2(\bar{x}, d_{\ell k})$, see

[14, Example 3.39]). Therefore, $(z_\ell^k,1) \in T_2^{proj}(\bar{x}, d_{\ell k})$. Now, summing over $\ell \in I(k)$ all the expressions in (19), we get that

$$\begin{array}{l}\langle \nabla h_i(\bar{x}), \sum_{\ell\in I(k)} z_\ell^k \rangle + \langle \nabla^2 h_i(\bar{x}), \sum_{\ell\in I(k)} d_{\ell k} d_{\ell k}^T \rangle = 0, \quad i \in \mathcal{E}, \\ \langle \nabla g_j(\bar{x}), \sum_{\ell\in I(k)} z_\ell^k \rangle + \langle \nabla^2 g_j(\bar{x}), \sum_{\ell\in I(k)} d_{\ell k} d_{\ell k}^T \rangle = 0, \quad j \in A(\bar{x}).\end{array} \tag{20}$$

Thus, the feasible set of (18) is nonempty and hence, there exists a global minimizer ζ^k. Moreover, the sequence $\{\zeta^k\}$ is bounded, since by optimality, $\|z - \zeta^k\| \leq \|z - \sum_{\ell\in I(k)} z_\ell^k\|$ (let us recall that $\{\sum_{\ell\in I(k)} z_\ell^k : k \in \mathbb{N}\}$ is a bounded sequence). We continue by showing that $\zeta^k - \sum_{\ell\in I(k)} z_\ell^k \in T_1(\bar{x})$. Indeed, take $j \in A(\bar{x})$. Thus, using (20) and the feasibility of ζ^k, we see that

$$\begin{aligned}\langle \nabla g_j(\bar{x}), \zeta^k - \textstyle\sum_{\ell\in I(k)} z_\ell^k \rangle &= \langle \nabla g_j(\bar{x}), \zeta^k \rangle - \langle \nabla g_j(\bar{x}), \textstyle\sum_{\ell\in I(k)} z_\ell^k \rangle \\ &= \langle \nabla g_j(\bar{x}), \zeta^k \rangle + \langle \nabla^2 g_j(\bar{x}), \textstyle\sum_{\ell\in I(k)} d_{\ell k} d_{\ell k}^T \rangle \\ &= \langle \nabla g_j(\bar{x}), \zeta^k \rangle + \langle \nabla^2 g_j(\bar{x}), Z^k \rangle \leq 0.\end{aligned} \tag{21}$$

Similarly, for every $i \in \mathcal{E}$, we see that $\langle \nabla h_i(\bar{x}), \zeta^k - \sum_{\ell\in I(k)} z_\ell^k \rangle = 0$. Thus, $\zeta^k - \sum_{\ell\in I(k)} z_\ell^k \in L_1(\bar{x})$, and hence $\zeta^k - \sum_{\ell\in I(k)} z_\ell^k \in T_1(\bar{x})$, since MFCQ at $\bar{x}$ implies the equailty $L_1(\bar{x}) = T_1(\bar{x})$.

Now, we will show that $\{\zeta^k\}$ converges to z. Let $\bar{z}$ be a limit point of $\{\zeta^k\}$ (which exists since $\{\zeta^k\}$ is bounded). By the KKT conditions of (18) at ζ^k there are multipliers $(\lambda^k, \mu^k) \in \mathbb{R}^m \times \mathbb{R}_+^p$ with $\mu_j^k = 0$, for $j \notin A(\bar{x})$ such that

$$-\left(z - \eta^k\right) + \sum_{i\in\mathcal{E}} \lambda_i^k \nabla h_i(\bar{x}) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(\bar{x}) = 0, \tag{22}$$

and

$$\begin{array}{l}\lambda_i^k(\langle \nabla h_i(\bar{x}), \zeta^k \rangle + \langle \nabla^2 h_i(\bar{x}), Z^k \rangle) = 0, \quad i \in \mathcal{E}, \\ \mu_j^k(\langle \nabla g_j(\bar{x}), \zeta^k \rangle + \langle \nabla^2 g_j(\bar{x}), Z^k \rangle) = 0, \quad j \in A(\bar{x}).\end{array} \tag{23}$$

Since MFCQ holds at the feasible $\bar{x}$ for the original optimization problem, the sequence $\{(\lambda^k, \mu^k)\}$ is bounded. Otherwise dividing (22) by $\|(\lambda^k, \mu^k)\|$ and taking an adequate convergent subsequence we get a contradiction with MFCQ.

Thus, we can assume that $(\lambda^k, \mu^k) \to (\lambda, \mu)$. Taking limit in (22), we get

$$\|z - \bar{z}\|^2 = \sum_{i\in\mathcal{E}} \lambda_i \langle \nabla h_i(\bar{x}), z - \bar{z} \rangle + \sum_{j=1}^{p} \mu_j \langle \nabla g_j(\bar{x}), z - \bar{z} \rangle. \tag{24}$$

Finally, we will show that $z = \bar{z}$. Thus, it will sufficient to see that the right-hand side of (24) is not positive. Indeed, taking an adequate limit in (23):

$$\begin{array}{l}\lambda_i(\langle \nabla h_i(\bar{x}), \bar{z} \rangle + \langle \nabla^2 h_i(\bar{x}), Z \rangle) = 0, \quad i \in \mathcal{E}, \\ \mu_j(\langle \nabla g_j(\bar{x}), \bar{z} \rangle + \langle \nabla^2 g_j(\bar{x}), Z \rangle) = 0, \quad j \in A(\bar{x}).\end{array} \tag{25}$$

With this information, we proceed to analyse (24). Take $j \in A(\bar{x})$. Then,

$$\begin{aligned}\mu_j \langle \nabla g_j(\bar{x}), z - \bar{z} \rangle &= \mu_j \langle \nabla g_j(\bar{x}), z \rangle - \mu_j \langle \nabla g_j(\bar{x}), \bar{z} \rangle \\ &= \mu_j \langle \nabla g_j(\bar{x}), z \rangle + \mu_j \langle \nabla^2 g_j(\bar{x}), Z \rangle \leq 0,\end{aligned} \tag{26}$$

Characterization of WSOC	$N_2(\bar{x}) \subseteq L_{2,W}(\bar{x})^\circ$
Weak Guignard-type CQ for WSOC	$[T_1(\bar{x}), T_2^{proj}(\bar{x},\cdot)] = L_{2,W}(\bar{x})^\circ$
Guignard-type CQ for WSOC	$T_{2,W}(\bar{x})^\circ = L_{2,W}(\bar{x})^\circ$
Abadie-type CQ for WSOC	$T_{2,W}(\bar{x}) = L_{2,W}(\bar{x})$

Table 1: New constraint qualifications for WSOC.

where, in the second equality we used (25) and in the last inequality, we used that $(z, Z) \in L_{2,W}(\bar{x})$. Similarly, for $i \in \mathcal{E}$, we get $\lambda_i \langle \nabla h_i(\bar{x}), z - \bar{z} \rangle = 0$. Using such inequalities into (24), we get that the right-hand side of (24) is not positive and thus $\bar{z} = z$. In this manner, we get that $\zeta^k \to z$ (since the limit point is unique) and (ζ^k, Z^k) converges to (z, Z) as we wanted to show. □

3.2 New Constraint Qualifications for SSOC

We proceed by defining new CQs that guarantee that SSOC holds at any local minimizer. For this purpose, we define a cone that characterizes SSOC in the following way:

$$K_2^S(\bar{x}) := \bigcup_{\substack{(\lambda,\mu)\in\mathbb{R}^m\times\mathbb{R}^p_+,\\ \mu_j=0 \text{ for } j\notin A(\bar{x})}} \left\{ \left(\sum_{i\in\mathcal{E}} \lambda_i \nabla h_i(\bar{x}) + \sum_{j\in A(\bar{x})} \mu_j \nabla g_j(\bar{x}), H \right) : \right.$$
$$\left. H \preceq \sum_{i\in\mathcal{E}} \lambda_i \nabla^2 h_i(\bar{x}) + \sum_{j\in A(\bar{x})} \mu_j \nabla^2 g_j(\bar{x}) \text{ on } \mathcal{C}(\bar{x},\mu) \right\},$$

where $\mathcal{C}(\bar{x},\mu)$ is the *critical cone associated with the vector* μ given by

$$\mathcal{C}(\bar{x},\mu) := \left\{ d \in \mathbb{R}^n : \quad \begin{array}{l} \langle \nabla h_i(\bar{x}), d\rangle = 0, i \in \mathcal{E}, \\ \langle \nabla g_j(\bar{x}), d\rangle = 0, \text{ if } \mu_j > 0, \quad j \in A(\bar{x}), \\ \langle \nabla g_j(\bar{x}), d\rangle \leq 0, \text{ if } \mu_j = 0, \quad j \in A(\bar{x}). \end{array} \right\}.$$

When $\mu \in \mathbb{R}^p_+$ is such that there is a vector $\lambda \in \mathbb{R}^m$ such that (λ, μ) is a Lagrange multiplier at $\bar{x}$ associated with an objective function f, the cone $\mathcal{C}(\bar{x},\mu)$ coincides with the corresponding critical cone $C(\bar{x}) = K_{L_1(\bar{x})}(\nabla f(\bar{x})) = L_1(\bar{x}) \cap \nabla f(\bar{x})^\perp$.

The cone $K_2^S(\bar{x})$ allows writing SSOC in a geometric way. Indeed, the following equivalence is straightforward:

Proposition 3.1. *The strong second-order condition (SSOC) holds at $\bar{x}$ if, and only if $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in K_2^S(\bar{x})$.*

By using Theorem 2.1, we get the following characterization of SSOC:

Theorem 3.3. *SSOC holds for every C^2 objective function f which has $\bar{x}$ as a local minimizer relative to Ω if, and only if $N_2(\bar{x}) \subseteq K_2^S(\bar{x})$.*

Let us introduce a Guignard-type condition for SSOC. By a straightforward verification, we have:

$$K_2^S(\bar{x}) \subseteq [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)] \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)], \tag{27}$$

where the second inclusion follows from [19, Proposition 4.1].

Since $N_2(\bar{x}) \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$ always holds at every feasible point $\bar{x}$, the equality $K_2^S(\bar{x}) = [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$ is a Guignard-type CQ for SSOC. From (27), we clearly have that the equality $K_2^S(\bar{x}) = [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$ implies $[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)] = [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$. Thus, as expected, the Guignard-type condition for SSOC ($K_2^S(\bar{x}) = [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$) implies the equality $[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)] = [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$, which is the weak second-order Guignard-type condition for the basic second-order condition as presented in [19].

An important CQ that implies SSOC at local minimizers is the *constant rank* CQ. Such CQ was introduced in [20] which is weaker than LICQ and independent of MFCQ. Later, in [22], a relaxed form of CRCQ was provided, which is useful in the stability analysis of nonlinear programs and the convergence analysis of some first- and second-order numerical methods. Here, we say that the *relaxed constant rank* CQ (RCRCQ) holds at the feasible point $\bar{x}$, if, for every $\mathcal{I} \subseteq A(\bar{x})$ the set $\{\nabla h_i(x), \nabla g_j(x) : i \in \mathcal{E}, j \in \mathcal{I}\}$ has constant rank for every x near $\bar{x}$.

Proposition 3.2. *Assume that $\bar{x}$ is a feasible point satisfying RCRCQ. Then, $K_2^S(\bar{x}) = [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$. In particular, under RCRCQ, we have that SSOC holds at any local minimizer.*

Proof. From the expression (27), it will be sufficient to show the non trivial inclusion $[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)] \subseteq K_2^S(\bar{x})$.

Take $(w, W) \in [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$. By RCRCQ, there exists a multiplier $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ such that $w = \sum_{i \in \mathcal{E}} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla g_j(\bar{x})$ with $\mu_j = 0, j \notin A(\bar{x})$. Define the symmetric matrix

$$H(\lambda, \mu) := \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla^2 g_j(\bar{x}) - W. \tag{28}$$

To show that $(w, W) \in K_2^S(\bar{x})$, we will need to prove that $H(\lambda, \mu)$ is a symmetric positive semi-definite matrix on $\mathcal{C}(\bar{x}, \mu)$.

Take $d \in \mathcal{C}(\bar{x}, \mu)$. From the definition of $\mathcal{C}(\bar{x}, \mu)$, we see that $d \in K_{T_1(\bar{x})}(w)$. Now, set $\mathcal{I} := A(\bar{x}, d)$. By RCRCQ and Lemma 2.1 applied to the family $\{\nabla h_i(x), \nabla g_j(x) : i \in \mathcal{E}, j \in \mathcal{I}\}$, there is a C^2 curve $\alpha(t)$, $t \in]-t_0, t_0[$ $(t_0 > 0)$ such that $\alpha(0) = \bar{x}$, $\alpha'(0) = d$ and for every $t \in]-t_0, t_0[$, for every $i \in \mathcal{E}$ and for all $j \in A(\bar{x}, d)$ we have $h_i(\alpha(t)) = 0$ and $g_j(\alpha(t)) = 0$. Deriving twice, $h_i(\alpha(t)) = 0$, $\forall \in \mathcal{E}$ and $g_j(\alpha(t)) = 0$, $j \in A(\bar{x}, d)$ at $t = 0$, we get that

$$\begin{array}{l} \langle \nabla h_i(\bar{x}), z\rangle + \nabla^2 h_i(\bar{x}) d^2 = 0, \quad i \in \mathcal{E}, \\ \langle \nabla g_j(\bar{x}), z\rangle + \nabla^2 g_j(\bar{x}) d^2 = 0, \quad j \in A(\bar{x}, d), \end{array} \tag{29}$$

where $z := \alpha''(0)$. Note that if $\mu_j > 0$ then $j \in A(\bar{x}, d)$ (since $d \in \mathcal{C}(\bar{x}, \mu)$). Thus, using (29) into (28), we get that

$$\begin{array}{ll} H(\lambda, \mu) d^2 & = \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 h_i(\bar{x}) d^2 + \sum_{j \in A(\bar{x})} \mu_j \nabla^2 g_j(\bar{x}) d^2 - W d^2 \\ & = -\sum_{i \in \mathcal{E}} \lambda_i \langle \nabla h_i(\bar{x}), z\rangle - \sum_{j \in A(\bar{x})} \mu_j \langle \nabla g_j(\bar{x}), z\rangle - W d^2 \\ & = -\langle \sum_{i \in \mathcal{E}} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla g_j(\bar{x}), z\rangle - W d^2 \\ & = -\langle w, z\rangle - W d^2. \end{array} \tag{30}$$

Now, since $(w, W) \in [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$, we see that $\langle w, \bar{z}\rangle + \bar{r} W d^2 \leq 0$ for every $(\bar{z}, \bar{r}) \in T_2^{proj}(\bar{x}, d)$. In particular, we can take $(\bar{z}, \bar{r}) := (z, 1)$, since for $t > 0$ small enough $\alpha(t)$ is a feasible curve (note that $\langle \nabla g_j(\bar{x}), d\rangle < 0$ for $j \notin A(\bar{x}, d)$). Thus, $\langle w, z\rangle + W d^2 \leq 0$ and hence $H(\lambda, \mu) d^2 \geq 0$. This complete the proof. $\square$

Characterization of SSOC	$N_2(\bar{x}) \subseteq K_2^S(\bar{x})$
Guignard-type CQ for SSOC	$K_2^S(\bar{x}) = [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$

Table 2: New constraint qualifications for SSOC.

It is easy to see that $K_2^S(\bar{x}) \subseteq L_{2,W}^\circ$, as we expected. Thus, the inclusion $N_2(\bar{x}) \subseteq K_2^S(\bar{x})$ characterizing SSOC implies the inclusion $N_2(\bar{x}) \subseteq L_{2,W}^\circ$ characterizing WSOC.

All the relations among the CQs introduced in this paper are listed in Figure 1.

4 Conclusions

It is well known that Fritz-John type second-order optimality conditions have been thoroughly studied in the past years (see, for instance, [15, Proposition 5.48]). These conditions include a Lagrange multiplier associated with the objective function which gets rid of the necessity of a constraint qualification at

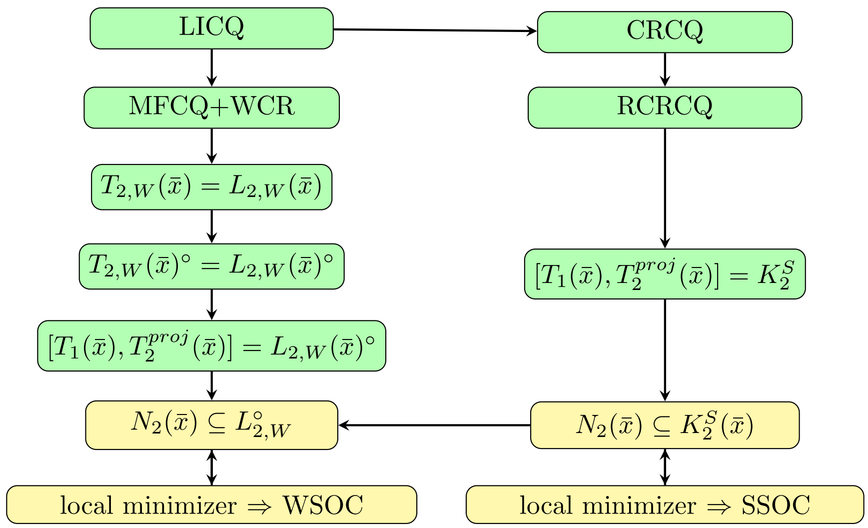

Figure 1: Relations among several CQs for second-order conditions.

all. They also provide no-gap necessary and sufficient optimality conditions, which is quite interesting. However, these type of analysis are present mostly in the most theoretical venues, given that they are difficult to be checked. This is mostly due to the necessity of knowing the full set of Fritz-John/ Lagrange multipliers, and also, the involvment of the true critical cone, where checking positive semidefiniteness whithin it is a hard computational problem. Hence, it is an interesting question to know which problems the Fritz-John multipliers can be replaced by Lagrange multipliers, besides the trivial case where all Fritz-John multipliers are Lagrange multipliers (that is, when MFCQ holds). Even though many research has appeared in the early days of optimization in second-order optimality conditions, minimal constraint qualifications, ensuring the validity of the conditions, have not been investigated thoroughly. In this paper, we reviewed the subject of second-order optimality conditions, with the view of defining minimal CQs with respect to two most well-known second-order optimality conditions under a CQ, namely, the weak and the strong second-order optimality conditions, while a previous analysis of the basic second-order condition has been done in [19]. The refined analysis of each condition is relevant due to their use in different contexts, namely, WSOC is mostly relevant in global convergence analysis, while SSOC is a refinement of the basic condition where only a single Lagrange multiplier is needed to check the condition. Considering gradients and Hessians of the constraints, we mainly investigated new Abadie- and Guignard-type constraint qualifications, where we mostly exploited the connections of geometric optimality conditions with a conic description of second-order optimality conditions.

References

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

[2] R. Andreani, G. Haeser, A. Ramos, and P.J.S. Silva. A second-order sequential optimality condition associated to the convergence of optimization algorithms. *IMA Journal of Numerical Analysis*, 37(4):1902–1929, 2017.

[3] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, 135:255–273, 2012.

[4] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. Two new weak constraint qualifications and applications. *SIAM Journal on Optimization*, 22:1109–1135, 2012.

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

[6] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. Strict constraint qualifications and sequential optimality conditions for constrained optimization. *Mathematics of Operations Research*, 43(3):693–717, 2008.

[7] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. A cone-continuity constraint qualification and algorithmic consequences. *SIAM Journal on Optimization*, 26(1):96–110, 2016.

[8] M. Anitescu. Degenerate nonlinear programming with a quadratic growth condition. *SIAM Journal on Optimization*, 10:4:1116–1135, 2000.

[9] A. V. Arutyunov. Perturbations of extremal problems with constraints and necessary optimality conditions. *Journal of Soviet Mathematics*, 54 (6):1342–1400, 1991.

[10] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Practical methods of Optimization: theory and algorithms*. John Wiley and Sons, 2006.

[11] 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: arXiv:1706.07833.

[12] D. P. Bertsekas. *Nonlinear programming*. Athenas Scientific, 1999.

[13] J. F. Bonnans, R. Cominetti, and A. Shapiro. Second-order optimality conditions based on parabolic second order tangent sets. *SIAM Journal of Optimization*, 9:466–492, 1999.

[14] J. F. Bonnans, A. Hermant, A. V. Arutyunov, and F. L. Pereira. No-gap second-order optimality conditions for optimal control problems with a single state constraint and control. *Mathematical Programming, Serie B.*, 117:21–50, 2009.

[15] J. F. Bonnans and A. Shapiro. *Pertubation Analysis of Optimization Problems*. Springer-Verlag, 2000.

[16] E. A. Coddington. *An Introduction to Ordinary Differential Equations*. Dover Publication, 1961.

[17] R. Fletcher. *Practical methods of Optimization: Constrained Optimization*. Wiley, 1981.

[18] F. J. Gould and J. W. Tolle. A necessary and sufficient qualification for constrained optimization. *SIAM Journal of Applied Mathematics*, 20:164–172, 1971.

[19] G. Haeser and A. Ramos. New constraint qualifications with second-order properties in nonlinear optimization. *To appear in Journal of Optimization Theory and Applications*, 2019. DOI: 10.1007/s10957-019-01603-x.

[20] R. Janin. Direction derivative of the marginal function in nonlinear programming. *Mathematical Programming Studies*, 21:127–138, 1984.

[21] J. M. Lee. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 1975.

[22] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. *Optimization*, 60:4:429–440, 2011.

[23] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, 2006.

[24] J. P. Penot. Optimality conditions in mathematical programming and composite optimization. *Mathematical Programming*, 67:225–245, 1994.

[25] R. T. Rockafellar and R. Wets. *Variational Analysis*. Series: Grundlehren der mathematischen Wissenschaften, Vol. 317, 2009.

A Proof of Lemma 2.1

First, by the constant rank assumption, we can assume, without loss of generality that $\{\nabla F_1(y), \ldots, \nabla F_M(y)\}$ $(M \leq N)$ is a linear basis for the subspace generated by $\{\nabla F_1(y), \ldots, \nabla F_N(y)\}$ for every y in $\mathcal{O}$. By the constant rank theorem, see [21, Theorem 7.1], there are C^2 mappings G_i, $i = M+1, \ldots, N$ such that $F_i(y) = G_i(F_1(y), \ldots, F_M(y))$ for y near x. Thus, if there is a C^2 curve $\alpha :]-t_0, t_0[\to \mathbb{R}^n$ $(t_0 > 0)$ such that $F_j(\alpha(t)) = 0$, $\forall t \in]-t_0, t_0[$, $\forall j = 1, \ldots, M$, we also get that $F_j(\alpha(t)) = 0$ for every $j = M+1, \ldots, N$.

Thereby, we proceed to find a C^2 curve $\alpha :]-t_0, t_0[\to \mathbb{R}^n$ ($t_0 > 0$) such that $F_j(\alpha(t)) = 0$, $\forall t \in]-t_0, t_0[$, $\forall j = 1, \ldots, M$. Similar to [10, Theorem 4.3.3], our approach consists in finding a C^2 curve α satisfying a differential equation with $\alpha(0) = x$, $\alpha'(0) = d$ and $\alpha''(0) = z$ for an adequate choice of z.

Let $A(y)$ be a matrix whose columns are $\{\nabla F_1(y), \ldots, \nabla F_M(y)\}$ and $P(y)$ be the orthogonal projection onto $ker A(y)^T$. Since $A(y)$ has linear independent columns, we have that $P(y) = [I - A(y)(A(y)^T A(y))^{-1} A(y)^T]$ and since the data is C^2, we see that $P(y)$ is a differentiable mapping for every y near x.

Take $d \in \mathbb{R}^n$ such that $\langle \nabla F_j(x), d \rangle = 0$, $\forall j = 1, \ldots, N$. Consider the minimization problem

$$\begin{array}{ll} \text{minimize} & (1/2)\|z\|^2, \\ \text{subject to} & \langle \nabla F_i(x), z \rangle + \nabla^2 F_i(x) d^2 = 0 \quad \forall i = 1, \ldots, M. \end{array} \tag{31}$$

Since $A(x)$ is an injective linear mapping, $A(x)^T$ is surjective and thus the feasible set of (31) is nonempty. Then, (31) admits a solution z. From the KKT conditions of (31), we see that $z \in im(A(x)) = ker(A(x)^T)^{\perp}$. Thus,

$$P(x)z = 0, \quad z = A(x)(A^T(x)A(x))^{-1}A^T(x)z, \quad A(x)^T z = -b, \tag{32}$$

where $b \in \mathbb{R}^M$ is a vector with coordinates $b_i = \nabla^2 F_i(x) d^2$, $\forall i = 1, \ldots, M$.

Consider the ordinary differential equation

$$\alpha'(t) = P(\alpha(t))(d + tz) \quad \text{and} \quad \alpha(0) = x. \tag{33}$$

Since $P(y)$ is differentiable at every y near x, from Picard's and uniqueness theorem, see [16], there is a C^2 curve $\alpha :]-t_0, t_0[\to \mathbb{R}^n$ such that $\alpha(0) = x$ and $\alpha'(0) = P(\alpha(0))(d) = P(x)d = d$ (since $d \in ker A(x)^T$ and thus $P(x)d = d$). Furthermore, for every $i = 1, \ldots, M$ and $\forall t \in]-t_0, t_0[$,

$$\frac{dF_i(\alpha(t))}{dt} = \langle \nabla F(\alpha(t)), \alpha'(t) \rangle = \langle \nabla F(\alpha(t)), P(\alpha(t))(d + tz) \rangle = 0.$$

Thus, $F_i(\alpha(t)) = 0$, $\forall i$ and (8) holds. To obtain the inequality in (9), we will show that $\alpha''(0) = z$ and $\|z\|$ is bounded by $K_F \|d\|^2$ for some constant K_F.

Computing $\alpha''(t)$ at $t = 0$:

$$\alpha''(t) = P'(\alpha(t))(d + tz) + P(\alpha(t))z. \tag{34}$$

Using (32), $A(x)^T d = 0$ and $b = (\nabla^2 F_1(x) d^2, \ldots, \nabla^2 F_M(x) d^2)$, we obtain that

$$\begin{aligned} P'(\alpha(t))|_{t=0} d &= -A(x)(A(x)^T A(x))^{-1} (A(\alpha(t))^T)'|_{t=0} d \\ &= -A(x)(A(x)^T A(x))^{-1} b \\ &= A(x)(A(x)^T A(x))^{-1} A(x)^T z = z. \end{aligned} \tag{35}$$

Substituting (35) into (34) at $t = 0$, we get that $\alpha''(0) = z$.

Finally, we proceed to bound $\|z\|$. If $z = 0$, there is noting to prove. If $z \neq 0$. Using the KKT conditions of (31) at z, there exist multiplier $\lambda \in \mathbb{R}^M$ such that

$$z + \sum_{i=1}^{M} \lambda_i \nabla F_i(x) = 0 \quad \lambda_i(\langle \nabla F_i(x), z \rangle + \nabla^2 F(x) d^2) = 0, \forall i = 1, \ldots, M. \tag{36}$$

Multiplying (36) by z and using $M \le N$, we get that

$$\begin{aligned} \|z\|^2 = \langle z, z\rangle = -\sum_{i=1}^{M} \lambda_i \langle \nabla F_i(x), z\rangle \quad &= \sum_{i=1}^{M} \lambda_i \nabla F_i^2(x) d^2 \\ &\le N\|\lambda\|_\infty \max\{|\nabla^2 F_i(x) d^2| : \forall i\} \\ &\le N\|\lambda\| \max\{\|\nabla^2 F_i(x)\|_2 : \forall i\}\|d\|^2. \end{aligned} \tag{37}$$

Observe that λ is bounded. Indeed, $\lambda = -(A^T(x)A(x))^{-1}A(x)^T z$, and thus $\|\lambda\| \le \|(A^T(x)A(x))^{-1}A(x)^T\|_2\|z\|$. Substituting this bound of λ into (37),

$$\begin{aligned} \|z\|^2 \quad &\le N\|\lambda\| \max\{\|\nabla^2 F_i(x)\|_2 : \forall i\}\|d\|^2 \\ &\le N\|(A^T(x)A(x))^{-1}A(x)^T\|_2 \max\{\|\nabla^2 F_i(x)\|_2 : \forall i\}\|d\|^2\|z\|. \end{aligned} \tag{38}$$

Thus, (9) holds with $K_F := N\|(A^T(x)A(x))^{-1}A(x)^T\|_2 \max\{\|\nabla^2 F_i(x)\|_2 : \forall i\}$.
□