SIAM J. OPTIM.
Vol. 26, No. 1, pp. 96–110

A CONE-CONTINUITY CONSTRAINT QUALIFICATION AND ALGORITHMIC CONSEQUENCES*

ROBERTO ANDREANI†, JOSÉ MÁRIO MARTÍNEZ†, ALBERTO RAMOS‡, AND PAULO J. S. SILVA†

Abstract. Every local minimizer of a smooth constrained optimization problem satisfies the sequential approximate Karush–Kuhn–Tucker (AKKT) condition. This optimality condition is used to define the stopping criteria of many practical nonlinear programming algorithms. It is natural to ask for conditions on the constraints under which AKKT implies KKT. These conditions will be called strict constraint qualifications (SCQs). In this paper we define a cone-continuity property (CCP) that will be shown to be the weakest possible SCQ. Its relation to other constraint qualifications will also be clarified. In particular, it will be proved that CCP is strictly weaker than the constant positive generator constraint qualification.

Key words. constrained optimization, optimality conditions, constraint qualifications, KKT conditions, approximate KKT conditions

AMS subject classifications. 90C46, 90C30

DOI. 10.1137/15M1008488

1. Introduction.

We will consider constrained optimization problems defined by

$$\text{minimize } f(x) \text{ subject to } h(x) = 0, g(x) \le 0, \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}^m$, and $g : \mathbb{R}^n \to \mathbb{R}^p$ admit continuous first derivatives onto $\mathbb{R}^n$.

Many scientific and technological problems require the solution of problems of this form. Finding global solutions of (1.1) is possible only when the problem is small or has some special structure. Even the verification that a given feasible point is a solution may be very hard. For this reason one relies on necessary optimality conditions. By this we mean computable conditions that must be verified by the minimizers of (1.1) and whose fulfillment indicates that, very likely, the point under consideration is an acceptable (perhaps approximate) solution of (1.1).

A point $x \in \mathbb{R}^n$ is said to satisfy the KKT conditions related to (1.1) if there exist $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{i=1}^{p} \mu_i \nabla g_i(x) = 0, \tag{1.2}$$

$$h_i(x) = 0 \ \forall i = 1, \ldots, m, \tag{1.3}$$

*Received by the editors February 17, 2015; accepted for publication (in revised form) October 20, 2015; published electronically January 7, 2016. This work was supported by PRONEX-Optimization (PRONEX-CNPq/FAPERJ E-26/111.449/2010-APQ1), CEPID-CeMEAI (FAPESP 2013/07375-0), CNPq (grants 304618/2013-6, 482549/2013-0, 303750/2014-6 and 303013/2013-3), FAPESP (grants 2012/20339-0 and 2013/05475-7), CAPES, and National Project on Industrial Mathematics.
http://www.siam.org/journals/siopt/26-1/M100848.html

†Department of Applied Mathematics, Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas, SP, Brazil (andreani@ime.unicamp.br, martinez@ime.unicamp.br, pjssilva@ime.unicamp.br).
‡Institute of Mathematics and Statistics, University of São Paulo, São Paulo, SP, Brazil (aramos27@gmail.com).

and

$$\min\{\mu_i, -g_i(x)\} = 0 \ \forall i = 1, \dots, p. \tag{1.4}$$

The condition (1.4) implies that $g(x) \le 0$, $\mu \ge 0$, and $\mu_i = 0$ for all i such that $g_i(x) < 0$.

Given $x \in \mathbb{R}^n$, it is easy to check the existence of $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ satisfying (1.2), (1.3), and (1.4). Unfortunately, the KKT conditions are not necessarily satisfied by minimizers of (1.1). For example $x^* = 0$ is a global minimizer of x subject to $x^2 = 0$ but there are no multipliers λ, μ that fulfill the KKT conditions for $x = x^*$. The properties of the constraints that guarantee that minimizers of the constrained optimization problem satisfy the KKT conditions are called *constraint qualifications* (CQs): If x is a local minimizer of (1.1) and some CQ is fulfilled at x, then the KKT conditions are satisfied for appropriate multipliers $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$. In other words, the property

$$\text{KKT or Not-CQ} \tag{1.5}$$

is fulfilled at every local minimizer of (1.1).

Obviously, necessary optimality conditions should be as strong as possible. Moreover, the strength of (1.5) is linked to the weakness of the CQ. The most popular CQ is the linear independence of the gradients of active constraints (LICQ). Its attractiveness is due to two independent properties: On the one hand, LICQ is easily verifiable and, on the other hand, it can be associated with many practical optimization algorithms, for which it can be proved that convergence occurs to points that satisfy "KKT or Not-LICQ." The Mangasarian–Fromovitz CQ (MFCQ), which states that the gradients of active constraints are "positively linearly independent" at the feasible point under consideration, is obviously weaker than LICQ [18, 22]. Qi and Wei [21] introduced the "constant positive linear dependence" (CPLD) condition, which says that if some gradients of active constraints are positively linearly dependent at a point x, then the same gradients are linearly dependent in a neighborhood of x. They also showed that a particular sequential quadratic programming algorithm converges to points that satisfy "KKT or not-CPLD." Curiously, Qi and Wei did not prove that CPLD was a CQ. This property of CPLD was proved by Andreani, Martínez, and Schuverdt [8], who also described the status of CPLD with respect to other CQs proving that the new condition implies quasi-normality. CPLD is weaker than MFCQ and is necessarily satisfied if the constraints of the problem are linear (a property that is not shared by MFCQ). This motivated a sequence of papers in which weaker CQs were introduced, with proved association with practical algorithms. See [4, 5] and references therein. This effort seemed to come to an end with the introduction of the constant positive generator (CPG) CQ in [5]. The definition of the CPG CQ is the following.

DEFINITION 1.1. *Assume that $h(x^*) = 0$ and $g(x^*) \le 0$. Define $I = \{1, \dots, m\}$. Let $J(x^*) \subset \{1, \dots, p\}$ be the indices of the active inequality constraints at x^*. Let J_- be the set of indices $\ell \in J(x^*)$ such that, for all $\ell \in J_-$, there exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ and $\mu_j \in \mathbb{R}_+$ for all $j \in J(x^*)$, such that*

$$-\nabla g_\ell(x^*) = \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x^*). \tag{1.6}$$

Define $J_+ = J(x^*) \setminus J_-$. We say that the CPG condition holds at x^* if there exist (maybe empty) sets $I' \subset I$ and $J' \subset J_-$ such that

(i) *the gradients* $\nabla h_i(x^*)$ *and* $\nabla g_j(x^*)$ *indexed by* $i \in I'$ *and* $j \in J'$ *are linearly independent;*

(ii) *for all* x *in a neighborhood of* x^**, if*

$$z = \sum_{i=1}^{m} \lambda'_i \nabla h_i(x) + \sum_{j \in J(x^*)} \mu'_j \nabla g_j(x)$$

with $\mu'_j \geq 0$ *for all* $j \in J(x^*)$*, then for all* $i \in I'$*,* $\ell \in J'$*, and* $j \in J_+$*, there exist* $\lambda''_i \in \mathbb{R}$*,* $\lambda'''_\ell \in \mathbb{R}$*, and* $\mu''_j \in \mathbb{R}_+$ *such that*

$$z = \sum_{i \in I'} \lambda''_i \nabla h_i(x) + \sum_{\ell \in J'} \lambda'''_\ell \nabla g_\ell(x) + \sum_{j \in J_+} \mu''_j \nabla g_j(x).$$

Remark 1. The item (i) of Definition 1.1 above is equivalent to stating that the gradients $\{\nabla h_i(x^*), i \in I\}$, $\{\nabla g_l(x^*), l \in J'\}$, and $\{\nabla g_j(x^*), j \in J_+\}$ form a positive linear independent set, that is, the only solution of

$$\sum_{i \in I'} \lambda_i \nabla h_i(x) + \sum_{\ell \in J'} \gamma_\ell \nabla g_\ell(x) + \sum_{j \in J_+} \mu_j \nabla g_j(x) = 0$$

with $\lambda_i \in \mathbb{R}, i \in I'$, $\gamma_\ell \in \mathbb{R}, \ell \in J'$, and $\mu_j \geq 0, j \in J_+$, is the trivial one.

In [5] it was proved that CPG is a CQ, strictly weaker than CPLD and some of its relaxations, and that it is useful in the context of several algorithms, for which it can be proved that limit points that satisfy CPG are KKT points.

The question that arose so far is, is CPG the weakest CQ that satisfies the above properties? Before going to the answer of this question we need to formulate it more precisely.

The second "nice property" of a CQ is its association with practical algorithms. This association consists of the possibility of proving that accumulation points of the sequences generated by an algorithm must necessarily be KKT points under such a CQ. Some CQs have this property and others do not. This leads us to discuss the notion of *sequential optimality condition* [19, 3, 9]. To fix ideas we will consider the most popular of these conditions, called *approximate KKT* (AKKT) [21, 3, 10].

DEFINITION 1.2. *Assume that* $h(x^*) = 0$ *and* $g(x^*) \leq 0$*. We say that* x^* *satisfies AKKT if there exist sequences* $\{x^k\} \subset \mathbb{R}^n$ *(*$\{x^k\}$ *is called an AKKT sequence),* $\{\lambda^k\} \subset \mathbb{R}^m$*, and* $\{\mu^k\} \subset \mathbb{R}^p$ *such that* $\lim_{k\to\infty} x^k = x^*$*,*

$$\lim_{k\to\infty} \nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla g_i(x^k) = 0, \tag{1.7}$$

and

$$\lim_{k\to\infty} \min\{\mu_i^k, -g_i(x^k)\} = 0 \quad \forall i = 1, \ldots, p. \tag{1.8}$$

AKKT, as do other sequential optimality conditions, has two main properties. The first is that it is a genuine necessary optimality condition, independently of the fulfillment of CQs [3, 10]. The second is that many optimization algorithms (but not all, see [7]) generate primal-dual $\{x^k, \lambda^k, \mu^k\}$ sequences for which (1.7) and (1.8) are fulfilled. These properties motivate the definition of *strict constraint qualifications* (SCQs). An SCQ is a property of feasible points of constrained optimization problem that, when satisfied by an AKKT point, guarantees that such points satisfy the KKT conditions [10]. In other words, SCQs are characterized by the property

$$AKKT + SCQ \Rightarrow KKT. \tag{1.9}$$

Since all local minimizers satisfy AKKT, the property (1.9) implies that SCQs are, in fact, CQs. The reciprocal is not true. For instance, Abadie's CQ [1] or quasi-normality [11] are CQs that are not SCQs.

Now we are able to define precisely what we mean by "CQs associated with algorithms." Essentially, those CQs are the SCQs. The question about the weakness of CPG can be formulated as, is CPG the weakest SCQ?

An attentive reader could argue in the following way: Being AKKT is a genuine necessary optimality condition obviously associated with the usual stopping criteria of practical algorithms, why should one worry about the CQs under which AKKT implies KKT (which, in fact, is not a genuine optimality condition)? The reason is that, since AKKT is a necessary optimality condition, it should be as strong as possible. Its strength can be analyzed in several ways. One of these ways is to analyze the strength of the propositions that are implied by AKKT. The typical form of such propositions is [KKT or not-CQ]. Clearly, this proposition is stronger, the weaker CQ is. Therefore, the analysis of the CQ such that AKKT implies [KKT or not-CQ] leads us to practical conjectures on the strength of AKKT.

Note that AKKT, under several assumptions, can also be a sufficient optimality condition. For instance, in the convex case, AKKT with the hypothesis $\lim_{k\to\infty}\sum_{i=1}^{m} |\lambda_i^k h_i(x^k)| + \sum_{i=1}^{p} |\mu_i^k g_i(x^k)| = 0$ implies optimality. See [9, Theorem 4.2].

In this paper we will prove that CPG is not the weakest SCQ. The weakest SCQ will be completely characterized as being the continuity of the cone generated by the gradients of active constraints (cone continuity property (CCP)) and we will prove that CPG is strictly stronger than CCP.[1]

As a consequence of these results we are able to present an updated landscape of CQs, SCQs, and sequential optimality conditions; see Figure 1. Open questions remain that will be probably the subject of future research.

This paper is organized as follows. In section 2, we review basic definitions of optimization and variational analysis and introduce our principal object of study (2.11). In section 3, we introduce the CCP and we prove that CCP is the weakest SCQ. Section 4 shows the relationship between the CCP and other constraint qualifications as Abadie's CQ and quasi-normality. Concluding remarks are discussed in section 5.

2. Basic definitions and preliminaries.

Our notation is standard in optimization and variational analysis; cf. [23, 13, 20]. $\mathbb{N}$ denotes the set of natural numbers, $\mathbb{R}^n$ stands for the n-dimensional real Euclidean space, $n \in \mathbb{N}$. We denote by $\mathbb{B}$ the closed unit ball in $\mathbb{R}^n$, and $\mathbb{B}(x, \eta) := x + \eta\mathbb{B}$ the closed ball centered at x with radius $\eta > 0$. $\mathbb{R}_+$ is the set of positive scalars, $\mathbb{R}_-$ is the set of negative scalars, and $a^+ = \max\{0, a\}$, the positive part of a. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product, $\|\cdot\|$ the associated norm. Given a set-valued mapping (multifunction) $F : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the *sequential Painlevé–Kuratowski outer/upper limit* of $F(z)$ as $z \to z^*$ is denoted by

$$\limsup_{z\to z^*} F(z) := \{w^* \in \mathbb{R}^d : \exists \ (z^k, w^k) \to (z^*, w^*) \text{ with } w^k \in F(z^k)\} \tag{2.1}$$

[1] This result was announced in the Workshop of Continuous Optimization held in Florianópolis in February 2014 [6] and cited in the book [10].

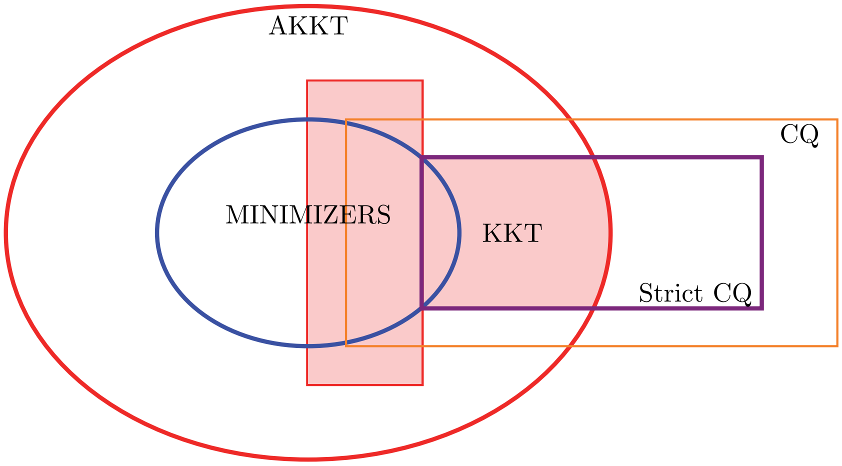

FIG. 1. *Optimality relations.* Small ellipse: *local minimizers.* Big ellipse: *feasible points that satisfy the AKKT condition.* Small rectangle: *feasible points that satisfy an SCQ.* Big rectangle: *feasible points that satisfy a (not necessarily strict) CQs.* Shadowed area: *KKT points.*

and the inner limit by

$$\liminf_{z \to z^*} F(z) := \{w^* \in \mathbb{R}^d : \forall z^k \to z^* \ \ \exists \ \ w^k \to w^* \text{ with } w^k \in F(z^k)\}. \tag{2.2}$$

We say that F is *outer semicontinuous* at z^* if

$$\limsup_{z \to z^*} F(z) \subset F(z^*). \tag{2.3}$$

It is *inner semicontinuous* at z^* if

$$F(z^*) \subset \liminf_{z \to z^*} F(z). \tag{2.4}$$

If F is both outer semicontinuous and inner semicontinuous at z^* we say that F is *continuous* at z^*.

Given the set S, the symbol $z \xrightarrow{S} z^*$ means that $z \to z^*$ with $z \in S$. For a cone $\mathcal{K} \subset \mathbb{R}^s$, its polar (negative dual) is $\mathcal{K}^\circ = \{v \in \mathbb{R}^s | \langle v, k\rangle \le 0 \text{ for all } k \in \mathcal{K}\}$. We use the notation $\phi(t) \le o(t)$ for any function $\phi : \mathbb{R}_+ \to \mathbb{R}^s$ such that $\limsup_{t \to 0_+} t^{-1}\phi(t) \le 0$.

Given $S \subset \mathbb{R}^n$ and $z^* \in S$, define the (Bouligand–Severi) *tangent/contingent cone* to S at z^* by

$$T_S(z^*) := \limsup_{t \downarrow 0} \frac{S - z^*}{t} = \{d \in \mathbb{R}^n : \exists \ \ t_k \downarrow 0, d^k \to d \text{ with } z^* + t_k d^k \in S\}. \tag{2.5}$$

The (Fréchet) *regular normal cone* to S at $z^* \in S$ is defined as

$$\widehat{N}_S(z^*) := \{w \in \mathbb{R}^n : \langle w, z - z^*\rangle \le o(|z - z^*|) \text{ for } z \in S\}. \tag{2.6}$$

The (Mordukhovich) *limiting normal cone* to S at $x^* \in S$ is

$$N_S(z^*) := \limsup_{z \xrightarrow{S} z^*} \widehat{N}_S(z). \tag{2.7}$$

When S is a convex set, both regular and limiting normal cones reduce to the classical normal cone of convex analysis and then the common notation $N_S(z^*)$ is used. For general sets we have the inclusion $\widehat{N}_S(z^*) \subset N_S(z^*)$ for all $z^* \in S$.

Denote by Ω the feasible set associated with (1.1),

$$\Omega := \{x \in \mathbb{R}^n \,|\, h(x) = 0, g(x) \leq 0\}.$$

Let $J(x^*)$ be the set of indices of active inequality constraints, and let $r = |J(x^*)|$ be the number of elements of $J(x^*)$. Let $x^* \in \Omega$ be a local minimizer of (1.1). The geometrical first-order necessary optimality condition states that $\langle \nabla f(x^*), d\rangle \geq 0$ for all $d \in T_\Omega(x^*)$. In other words,

$$-\nabla f(x^*) \in T_\Omega(x^*)^\circ. \tag{2.8}$$

Let us denote $I = \{1, \ldots, n\}$. The *linearized cone* $L_\Omega(x^*)$ is defined as follows:

$$L_\Omega(x^*) := \{d \in \mathbb{R}^n \mid \langle \nabla h_i(x^*), d\rangle = 0,\ \forall i \in I, \langle \nabla g_j(x^*), d\rangle \leq 0,\ \forall j \in J(x^*)\}\,. \tag{2.9}$$

By the geometric first-order necessary optimality condition (2.8), if x^* satisfies

$$T_\Omega(x^*)^\circ = L_\Omega(x^*)^\circ, \tag{2.10}$$

then the KKT conditions hold at x^*. The condition (2.10) was introduced by Guignard [15]. Gould and Tolle [16] proved that Guignard's condition (2.10) is the weakest CQ. Another CQ is Abadie's CQ, which reads $L_\Omega(x^*) = T_\Omega(x^*)$, which is stronger than Guignard's CQ.

Given $x^* \in \Omega$, we define

$$K(x) = \left\{ \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x) : \quad \mu_j \in \mathbb{R}_+, \lambda_i \in \mathbb{R} \right\}. \tag{2.11}$$

Clearly, $K(x)$ is a closed convex cone and coincides with $L_\Omega(x^*)^\circ$ at x^*. Using this cone, the KKT condition can now be written as $-\nabla f(x^*) \in K(x^*)$. The cone (2.11) also allows us to rewrite the CPG condition in a geometric setting. For this purpose, define the cone:

$$K_{I',J'}(x) = \left\{ \sum_{i \in I'} \lambda_i \nabla h_i(x) + \sum_{\ell \in J'} \gamma_\ell \nabla g_\ell(x) + \sum_{j \in J_+} \mu_j \nabla g_j(x) \colon\ \mu_j \in \mathbb{R}_+, \lambda_i, \gamma_\ell \in \mathbb{R} \right\}, \tag{2.12}$$

where $I' \subset I$, $J' \subset J_-$, and $J_+ = J(x^*) \setminus J_-$.

The restatement of the CPG condition is as follows. We say that CPG holds at x^* if there exists $I' \subset I$, $J' \subset J_-$, and a neighborhood V of x^* such that

1. the gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$ indexed by $i \in I'$ and $j \in J'$ are linearly independent;
2. the inclusion

$$K(x) \subset K_{I',J'}(x) \text{ holds } \forall x \in V. \tag{2.13}$$

Clearly, from (2.13) and the definition of J_-, we have that $K(x^*) = K_{I',J'}(x^*)$. The cone $K_{I',J'}(x)$ is outer semicontinuous at x^* as the following technical lemma shows.

LEMMA 2.1. *Let $x^* \in \Omega$, $I' \subset I, J' \subset J_-$ and $J_+ = J(x^*) \setminus J_-$ such that the gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$ indexed by $i \in I'$ and $j \in J'$ are linearly independent. Then the set-valued mapping $x \in \mathbb{R}^n \rightrightarrows K_{I',J'}(x)$ is outer semicontinuous at x^*.*

Proof. Let ω^* be an element of $\limsup_{x\to x^*} K_{I',J'}(x)$, so there are sequences x^k, ω^k such that $x^k \to x^*$, $\omega^k \to \omega^*$, and $\omega^k \in K_{I',J'}(x^k)$ with

$$\omega^k = \sum_{i\in I'} \lambda_i^k \nabla h_i(x^k) + \sum_{l\in J'} \gamma_\ell^k \nabla g_\ell(x^k) + \sum_{j\in J_+} \mu_j^k \nabla g_j(x^k) \tag{2.14}$$

for some sequence $\{\lambda_i^k \in \mathbb{R}, i \in I'; \gamma_\ell^k \in \mathbb{R}, \ell \in J'; \mu_j^k \in \mathbb{R}_+, j \in J_+\}$. Define $M_k = \max\{|\lambda_i^k|, i \in I'; |\gamma_\ell^k|, \ell \in J'; \mu_j^k, j \in J_+\}$. We have two possibilities:

- Let $\{M_k\}$ have a bounded subsequence. So we can assume, by possibly extracting an adequate subsequence, that for all $i \in I', \ell \in J'$, and $j \in J_+$ the subsequences of $\lambda_i^k, \gamma_\ell^k, \mu_j^k$ have limits $\lambda_i^*, \gamma_\ell^*, \mu_j^*$, respectively. Now, taking the limit at (2.14) we get

$$\omega^* = \sum_{i\in I'} \lambda_i^* \nabla h_i(x^*) + \sum_{l\in J'} \gamma_\ell^* \nabla g_\ell(x^*) + \sum_{j\in J_+} \mu_j^* \nabla g_j(x^*) \in K_{I',J'}(x^*).$$

- Otherwise, we have $M_k \to \infty$. Dividing (2.14) by M_k, we arrive at

$$\frac{\omega^k}{M_k} = \sum_{i\in I'} \frac{\lambda_i^k}{M_k} \nabla h_i(x^k) + \sum_{l\in J'} \frac{\gamma_\ell^k}{M_k} \nabla g_\ell(x^k) + \sum_{j\in J_+} \frac{\mu_j^k}{M_k} \nabla g_j(x^k). \tag{2.15}$$

  Since $\max\{|\lambda_i^k/M_k|, i \in I'; |\gamma_\ell^k/M_k|, \ell \in J'; \mu_j^k/M_k, j \in J_+\} = 1$ for all $k \in \mathbb{N}$, we can extract a convergent subsequence. Thus, taking limits in (2.15), we get a contradiction to the fact that $\{\nabla h_i(x^*), i \in I\}$, $\{\nabla g_\ell(x^*), \ell \in J'\}$, and $\{\nabla g_j(x^*), j \in J_+\}$ form a positive linear independent set; see Remark 1. ∎

3. Cone-continuity CQ. We will proceed with the next definition.

DEFINITION 3.1. *We say that $x^* \in \Omega$ satisfies CCP if the set-valued mapping (multifunction) $\mathbb{R}^n \ni x \rightrightarrows K(x)$, defined in* (2.11), *is outer semicontinuous at x^*, that is,*

$$\limsup_{x\to x^*} K(x) \subset K(x^*). \tag{3.1}$$

The AKKT condition is naturally associated with the CCP condition. The best way to see this is to write it in a more compact but equivalent form [3]. The AKKT condition holds at $x^* \in \Omega$ if and only if, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}_+^p$ with $\mu_j^k = 0$ for $j \notin J(x^*)$, such that $\lim_{k\to\infty} x^k = x^*$ and

$$\lim_{k\to\infty} \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j\in J(x^*)} \mu_j^k \nabla g_j(x^k) = 0. \tag{3.2}$$

The expression (3.2) says that if the AKKT condition holds the vector $-\nabla f(x^k)$ gets arbitrarily close to the cone $K(x^k)$ as k goes to ∞.

Note that the multifunction $x \in \mathbb{R}^n \rightrightarrows K(x)$ is always inner semicontinuous due to the continuity of the gradients and the definition of $K(x)$. For this reason, outer semicontinuity is sufficient for the continuity of $K(x)$ at x^*.

In the following theorem we show that CCP plays, with respect to AKKT, the same role as Guignard's CQ plays with respect to local optimality. Guignard's CQ is the weakest CQ that guarantees that local minimality implies KKT [16], in the same sense that CCP is the weakest condition on the constraints that guarantees that AKKT implies KKT.

THEOREM 3.2. *Consider the problem* (1.1) *and let* x^* *be a feasible point. The CCP condition at* x^* *is the weakest property under which AKKT implies KKT for every continuously differentiable objective function* f *that attains a minimum at* x^**. (In other words, CCP is the weakest strict CQ.)*

Proof. Let us show first that, if CCP holds, the sequential AKKT condition implies the KKT condition independently of the objective function. Let f be an objective function such that the sequential AKKT condition holds at x^*, then there are sequences $\{x^k\} \to x^*$, $\{\lambda^k\} \in \mathbb{R}^n$, $\{\mu^k\} \in \mathbb{R}^p_+$, and $\{\zeta^k\} \in \mathbb{R}^m$ such that $\mu^k_j = 0$ for $j \notin J(x^*)$ and

$$\zeta^k = \nabla f(x^k) + \sum_{i=1}^{m} \lambda^k_i \nabla h_i(x^k) + \sum_{j \in J(x^*)} \mu^k_j \nabla g_j(x^k) \to 0. \tag{3.3}$$

Define $\omega^k = \sum_{i=1}^m \lambda^k_i \nabla h_i(x^k) + \sum_{j \in J(x^*)} \mu^k_j \nabla g_j(x^k)$; we see that

$$\omega^k \in K(x^k) \quad \text{and} \quad \omega^k = \zeta^k - \nabla f(x^k). \tag{3.4}$$

Taking limits in (3.4) when k goes to infinity, using the continuity of the gradient of f, and $\zeta^k \to 0$, we get

$$-\nabla f(x^*) = \lim \omega^k \in \limsup_{k \to \infty} K(x^k) \subset \limsup_{x \to x^*} K(x) \subset K(x^*), \tag{3.5}$$

where the last inclusion follows from the CCP. Therefore, $-\nabla f(x^*) \in K(x^*)$, which is equivalent to saying that x^* satisfies the KKT condition.

Now, let us prove that, if AKKT implies the KKT condition for every objective function, then CCP holds. Take $\omega^* \in \limsup_{x \to x^*} K(x)$; by the definition of outer limit, there are sequences x^k, ω^k such that $x^k \to x^*$, $\omega^k \to \omega^*$, and $\omega^k \in K(x^k)$. Define $f(x) = -\langle w^*, x\rangle$ for all $x \in \mathbb{R}^n$. Then, AKKT holds at x^* for this function with $\{x^k\}$ as an AKKT sequence since $\nabla f(x^k) + \omega^k = -\omega^* + \omega^k \to 0$. So by hypothesis the KKT condition holds at x^*, that is, $-\nabla f(x^*) = \omega^* \in K(x^*)$. ∎

Since AKKT is a necessary optimality condition, cf. [3], we have the next corollary.

COROLLARY 3.3. *The CCP is a CQ.*

Remark 2. Certainly, some CQs are easily verifiable (for example, LICQ) and others are verifiable with different degrees of difficulty. CCP is not easily verifiable. In fact, this is not one of our objectives in the analysis of CQs. We are mainly interested in the weakness of CQs because, when a CQ is weak, the condition KKT or not-CQ is strong and, so, the corresponding sequential optimality condition is strong. Clearly, stopping an algorithm with the fulfillment of a strong optimality condition increases our chances of obtaining minimizers.

Remark 3. The book [10] introduces a version of the CCP called the U-condition, based on CCP [6], but not directly using variational analysis concepts. In several practical constrained optimization algorithms that generate AKKT sequences (for example, the sequential quadratic programming algorithm of Qi and Wei [21], the interior-point method of Chen and Goldfarb [14], and augmented Lagrangian algorithms [2, 10]) convergence to KKT points has been proved assuming different CQs. By Theorem 3.2, in all these cases the CQ employed may be replaced with the weaker CCP and cannot be improved using another CQs.

4. Relations with other CQs.

In this section, we study the relations between the CCP condition with other CQs.

4.1. CCP and CPG condition.

In this subsection we will show that the CCP is strictly weaker that CPG.

THEOREM 4.1. *The CPG condition implies the CCP condition.*

Proof. From the definition of CPG, there is a set of indices I', J', J_+ such that the gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$, $(i,j) \in (I', J')$, are linearly independent and a neighborhood V of x^* such that

$$K(x) \subset K_{I',J'}(x) \quad \forall x \in V. \tag{4.1}$$

Now taking limits in (4.1) and using the outer semicontinuity of $K_{I',J'}(x)$ at x^*, Lemma 2.1, we get

$$\limsup_{x \to x^*} K(x) \subset \limsup_{x \to x^*} K_{I',J'}(x) \subset K_{I',J'}(x^*) = K(x^*), \tag{4.2}$$

which implies the outer semicontinuity of $K(x)$ at x^*. ∎

Now the next example shows in fact that the CCP condition is strictly weaker that CPG.

Example 1 (CCP does not imply CPG). In $\mathbb{R}^2$, consider $x^* = (0,0)$ and the inequality constraints defined by

$$g_1(x_1,x_2) = x_1,$$
$$g_2(x_1,x_2) = x_1^3 \exp(x_2^2).$$

Clearly, $x^* = (0,0)$ is a feasible point and both constraints are active at x^*. By direct calculations,

$$\nabla g_1(x_1,x_2) = (1,0) \text{ and}$$
$$\nabla g_2(x_1,x_2) = (3x_1^2 \exp(x_2^2), 2x_2x_1^3 \exp(x_2^2)) \ \forall x = (x_1,x_2) \in \mathbb{R}^2.$$

Since $\nabla g_1(x^*) = (1,0)$ and $\nabla g_2(x^*) = (0,0)$, we have that $K(x^*) = \mathbb{R}_+ \times \{0\}$ and the unique choice for a positive basis of $K(x^*)$ is $\{I' = \emptyset, J' = \emptyset, J_+ = \{1\}\}$. Thus we get $K_{I',J'}(x) = \mathbb{R}_+ \times \{0\}$ for every $x = (x_1,x_2) \in \mathbb{R}^2$. On the other hand,

$$K(x) = \{(\mu_1 + 3\mu_2 x_1^2 \exp(x_2^2), 2\mu_2 x_2 x_1^3 \exp(x_2^2)) \mid \quad \mu_1, \mu_2 \geq 0\}.$$

Thus, $K(x)$ cannot be a subset of $K_{I',J'}(x)$ in any neighborhood of x^* and CPG is not fulfilled.

Now let us prove that $K(x)$ is continuous at x^*. Let ω^* be an element of $\limsup_{x \to x^*} K(x)$. Therefore, there are sequences x^k and ω^k, such that $x^k = (x_1^k, x_2^k) \to x^*$, $\omega^k = (\omega_1^k, \omega_2^k) \to \omega^*$, and

$$\omega^k = \mu_1^k(1,0) + \mu_2^k(3(x_1^k)^2 \exp((x_2^k)^2), 2x_2^k(x_1^k)^3 \exp((x_2^k)^2)) \in K(x^k), \tag{4.3}$$

where μ_1^k, μ_2^k are nonnegative scalars. Suppose, by contradiction, that $\omega^* = (\omega_1^*, \omega_2^*)$ does not belong to $K(x^*) = \mathbb{R}_+ \times \{0\}$. So, ω_2^* must be nonzero. From (4.3) we have that there exists $\rho > 0$ such that

$$|\omega_2^k| = 2\mu_2^k |x_2^k (x_1^k)^3 \exp((x_2^k)^2)| > \rho > 0 \tag{4.4}$$

for k large enough. In particular, both x_1 and x_2 are nonzero numbers. Moreover, using (4.4) and $\mu_1^k \geq 0$, we get

$$\omega_1^k = \mu_1^k + 3\mu_2^k (x_1^k)^2 \exp((x_2^k)^2) \geq \frac{3|\omega_2^k|}{2|x_1^k x_2^k|} > \frac{3\rho}{2|x_1^k x_2^k|} > 0. \tag{4.5}$$

Taking limits in (4.5), we obtain $\omega_1^k \to \infty$. This is a contradiction with the fact that $\omega^k \to \omega^*$. Hence, ω^* must be in $K(x^*)$.

A natural question about CCP is if CCP is a stable condition. We say that a condition is stable if, whenever the condition holds at a base point, it continues to hold at nearby points. It is well known that LICQ and MFCQ are stable conditions. CCP, on the other hand, is not a stable condition. Consider the example of [5, Figure 3.2, p. 1118]. In this example, we have that CPG holds at x^* and hence CCP. But, for every point $(x_1, 0)$ with $x_1 < 0$, Guignard's CQ fails. Thus, CCP cannot be stable.

4.2. CCP and Abadie's CQ. Next we show that CCP is stronger than Abadie's CQ and independent of quasi-normality and pseudonormality conditions. The implications are based on the following two lemmas. The first one can be found in [23, Theorem 6.11].

LEMMA 4.2. *Let $\bar{x}$ be in Ω. For every $v \in T_\Omega^\circ(\bar{x})$, there exists a smooth function F such that $-\nabla F(\bar{x}) = v$ and attains its global minimum relative to Ω uniquely at $\bar{x}$.*

LEMMA 4.3. *For all $\bar{x} \in \Omega$ and $v \in T_\Omega^\circ(\bar{x})$, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}_+^p$ with $x^k \to \bar{x}$ such that*

(i) $\omega^k := \sum_{i=1}^m \lambda_j^k \nabla h_j(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k)$ *converges to v;*

(ii) $\lambda_j^k = k h_j(x^k)$ *for all $i = 1, \ldots, m$ and $\mu_j^k = k g_j(x^k)^+$ for all $j = 1, \ldots, p$.*

Proof. Let v be an element of $T_\Omega^\circ(\bar{x})$. By Lemma 4.2, there exists a smooth function F such that $-\nabla F(\bar{x}) = v$ and F attains its global minimum relative to Ω uniquely at $\bar{x}$. Consider, for each $k \in \mathbb{N}$, the following optimization problem:

$$\begin{array}{lll} & \text{Minimize} & F_k(x) = F(x) + \frac{k}{2}\left(\sum_{j=1}^m h_j(x)^2 + \sum_{j=1}^p (g_j(x)^+)^2\right) \\ (4.6) & \text{subject to} & x \in \mathbb{B}(\bar{x}, \eta). \end{array}$$

Since $\mathbb{B}(\bar{x}, \eta)$ is a compact set and $F_k(x)$ is continuous, by Weierstrass' theorem, there is at least one solution for (4.6), namely, x^k. Since x^k is a solution of (4.6) we have (4.7)

$$F(x^k) \leq F(x^k) + \frac{k}{2}\left(\sum_{j=1}^m h_j(x^k)^2 + \sum_{j=1}^p (g_j(x^k)^+)^2\right) = F_k(x^k) \leq F_k(\bar{x}) = F(\bar{x}).$$

By (4.7) and using the fact that F is bounded in the compact set $\mathbb{B}(\bar{x}, \eta)$, we get

$$\lim \left|g_j(x^k)^+\right| = 0 \;\; \forall \;\; j = 1, \ldots, p \quad \text{and} \quad \lim \left|h_j(x^k)\right| = 0 \;\; \forall j = 1, \ldots, m.$$

By the continuity of the constraints we have that every limit point of $\{x^k\}$ is feasible. Moreover, the sequence x^k converges to $\bar{x}$. In fact, if x^∞ is a limit point of the sequence, by (4.7) we have that $F(x^k) \leq F(\bar{x})$ and, taking limits in this inequality, $F(x^\infty) \leq F(\bar{x})$. Therefore, x^∞ is also a global minimizer of F. Since $\bar{x}$ is the unique global minimizer, $x^\infty = \bar{x}$. Thus, for k large enough, $x^k \in \text{int}\mathbb{B}(\bar{x}, \eta)$. Since x^k is a solution of (4.6) and $x^k \in \text{int}\,\mathbb{B}(\bar{x}, \eta)$, we get

$$\nabla F_k(x^k) = \nabla F(x^k) + \sum_{j=1}^m k h_j(x^k) \nabla h_i(x^k) + \sum_{j=1}^p k g_j(x^k)^+ \nabla g_j(x^k) = 0. \tag{4.8}$$

Define $\lambda_j^k := kh_i(x^k)$, $\mu_j^k := kg_j(x^k)^+$, and $\omega^k := \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k)$. From (4.8) and the continuity of $\nabla F(x)$ we have $\omega^k = -\nabla F(x^k) \to -\nabla F(x^*) = v$. Therefore, the items (i) and (ii) hold. ∎

THEOREM 4.4. *CCP implies Abadie's CQ.*

Proof. Our aim is to prove $T_\Omega(x^*) = L_\Omega(x^*)$. The inclusion $T_\Omega(x^*) \subset L_\Omega(x^*)$ is well known. To show that $L_\Omega(x^*) \subset T_\Omega(x^*)$, we will first prove that $N_\Omega(x^*) \subset L_\Omega(x^*)^\circ$, which is equivalent to $N_\Omega(x^*) \subset K(x^*)$ (by Farkas' lemma $L_\Omega(x^*)^\circ = K(x^*)$).

Let $v \in N_\Omega(x^*)$, so by definition of the normal cone (2.7) there are sequences $\{x^k\} \in \Omega$, $\{v^k\}$ such that

$$x^k \to x^*, \quad v^k \to v, \quad \text{ and } \; v^k \in T_\Omega^\circ(x^k).$$

By Lemma 4.3, for each $v^k \in T_\Omega^\circ(x^k)$ there exist sequences $\{x^{k,\ell}\}$ and $\{\omega^{k,\ell}\}$ satisfying the items (i) and (ii) of Lemma 4.3. This means that, for all $k \in \mathbb{N}$, we have

$$\lim_{\ell\to\infty} \omega^{k,\ell} := \lim_{\ell\to\infty} \sum_{j=1}^m \lambda_j^{k,\ell} \nabla h_j(x^{k,\ell}) + \sum_{j=1}^p \mu_j^{k,\ell} \nabla g_j(x^{k,\ell}) = v^k,$$

where

$$\mu_j^{k,\ell} = \ell g_j(x^{k,\ell})^+ \;\forall j = 1,\dots,p \text{ and } \lambda_j^{k,\ell} = \ell h_j(x^{k,\ell}) \;\forall j = 1,\dots,m.$$

Thus, for all $k \in \mathbb{N}$, there exists $\ell(k)$ such that

- $\|x^k - x^{k,\ell(k)}\| < 1/2^k$;
- $\omega^{k,\ell(k)} = \sum_{j=1}^m \lambda_j^{k,\ell(k)} \nabla h_j(x^{k,\ell(k)}) + \sum_{j=1}^p \mu_j^{k,\ell(k)} \nabla g_j(x^{k,\ell(k)})$;
- $\|v^k - w^{k,\ell(k)}\| < 1/2^k$;
- $\mu_j^{k,\ell(k))} = \ell(k) g_j(x^{k,\ell(k)})^+$ for all $j = 1,\dots,p$, and $\lambda_j^{k,\ell(k)} = \ell(k) h_j(x^{k,\ell(k)})$ for all $j = 1,\dots,m$.

Clearly,

$$\lim_{k\to\infty} x^{k,\ell(k)} = x^* \;\text{ and }\; \lim_{k\to\infty} \omega^{k,\ell(k)} = v.$$

Furthermore, for k large enough, $\mu_j^{k,\ell(k)} = \ell(k) g_j(x^{k,\ell(k)})^+ = 0$ for $j \notin J(x^*)$. Therefore, $\omega^{k,\ell(k)}$ belongs to $K(x^{k,\ell(k)})$ if k is large enough. So, we have found sequences such that $x^{k,\ell(k)} \to x^*$ and $\omega^{k,\ell(k)} \to v$, with $\omega^{k,\ell(k)} \in K(x^{k,\ell(k)})$. By the CCP and the definition of outer limit we have that $v \in \limsup_{x\to x^*} K(x) \subset K(x^*)$. Therefore, $N_\Omega(x^*) \subset K(x^*) = L_\Omega(x^*)^\circ$. Taking polar conjugation we deduce

$$L_\Omega(x^*) = K(x^*)^\circ \subset N_\Omega(x^*)^\circ. \tag{4.9}$$

Finally, by [23, Theorems 6.28(b) and 6.26], we have that $N_\Omega(x^*)^\circ \subset T_\Omega(x^*)$. Therefore, by (4.9), $L_\Omega(x^*) \subset T_\Omega(x^*)$, as we wanted to prove. ∎

Now we wish to show that Abadie's CQ is strictly weaker than CCP. Note that there is no contradiction with the fact that CCP is the weakest SCQ, because Abadie's CQ is not an SCQ. We are going to show an example in which Abadie's CQ holds but CCP does not. In fact, Example 3 does the job. In this example, the pseudonormality CQ is fulfilled but CCP fails. Note that the fulfillment of pseudonormality is enough to prove that Abadie's CQ also holds [11].

4.3. CCP and quasi-normality. Let us recall the definition of quasi-normality [17, 11]. We say that the quasi-normality CQ holds at $x^* \in \Omega$ if whenever $\sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x^*) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_j \in \mathbb{R}_+$ for every $j \in J(x^*)$, there is no sequence $x^k \to x^*$ such that for every $k \in \mathbb{N}$, $\lambda_i h_i(x^k) > 0$ when λ_i is nonzero and $g_j(x^k) > 0$ when $\mu_j > 0$. The next example shows that CCP does not imply quasi-normality.

Example 2 (CCP does not imply quasi-normality). Consider $x^* = (0,0)$ and the inequality constraints defined by

$$g_1(x_1,x_2) = x_1^3,$$
$$g_2(x_1,x_2) = x_1 \exp x_2.$$

Clearly, x^* is feasible and both constraints are active at x^*. The gradients are $\nabla g_1(x_1,x_2) = (3x_1^2, 0)$ and $\nabla g_2(x_1,x_2) = (\exp x_2, x_1 \exp x_2)$ for all (x_1,x_2) in $\mathbb{R}^2$.

Let us show that CCP holds at $x^* = (0,0)$. First,

$$\begin{aligned} K(x^*) &= \{\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) : \mu_1, \mu_2 \geq 0\} \\ &= \{\mu_1(0,0) + \mu_2(1,0) : \mu_1,\mu_2 \geq 0\} = \mathbb{R}_+ \times \{0\}. \end{aligned}$$

Take $\omega^* \in \limsup K(x)$. From the definition of outer limit there are sequences x^k and ω^k such that $x^k = (x_1^k, x_2^k) \to x^* = (0,0)$, $\omega^k = (\omega_1^k, \omega_2^k) \to \omega^*$, and

$$\omega^k = \mu_1^k(3(x_1^k)^2, 0) + \mu_2^k(\exp(x_2^k), x_1^k \exp x_2^k) \in K(x^k), \tag{4.10}$$

where μ_1^k, μ_2^k are nonnegative scalars. Suppose, by contradiction, that $\omega^* = (\omega_1^*, \omega_2^*)$ does not belong to $K(x^*) = \mathbb{R}_+ \times \{0\}$, so ω_2^* must be nonzero. By (4.10), we have that for k large enough

$$|\omega_2^k| = \mu_2^k |x_1^k \exp x_2^k| > \rho > 0, \tag{4.11}$$

where $\rho = |\omega_2^*|/2 > 0$. In particular $x_1^k \neq 0$. Since $\mu_1^k \geq 0$, using (4.11) we get

$$\omega_1^k = 3\mu_1^k(x_1^k)^2 + \mu_2^k \exp x_2^k \geq \mu_2^k \exp x_2^k \geq \frac{|\omega_2^k|}{|x_1^k|} > \frac{\rho}{|x_1^k|} > 0. \tag{4.12}$$

Taking limits in (4.12) we obtain $\omega_1^k \to \infty$, a contradiction with its convergence. Then, ω^* must be in $K(x^*)$.

Now, let us show that quasi-normality does not hold at $x^* = (0,0)$. Define $x_1^k = x_2^k = 1/k$, $\mu_1 = 1$, and $\mu_2 = 0$. With this choice $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = 1.(0,0) + 0.(1,0) = (0,0)$ and $\mu_1 g_1(x_1^k, x_2^k) = (x_1^k)^3 > 0$ for all $k \in \mathbb{N}$. So quasi-normality does not hold.

Example 3 will show that pseudonormality does not imply CCP, and as a consequence neither does quasi-normality. Thus, CCP and quasi-normality are independent CQs.

4.4. CCP and pseudonormality. In this subsection we will prove that CCP and pseudonormality are independent of each other. We say that the pseudonormality CQ holds at $x^* \in \Omega$, [11, 12], if whenever $\sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x^*) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_j \in \mathbb{R}_+$ for every $j \in J(x^*)$, there is no sequence $x^k \to x^*$ such that $\sum_{i=1}^m \lambda_j h_j(x^k) + \sum_{j\in J(x^*)} \mu_j g_j(x^k) > 0$ for all $k \in \mathbb{N}$. Trivially, pseudonormality implies quasi-normality. Since CCP does not imply quasi-normality (Example 2), it turns out that CCP does not imply pseudonormality either. In order to show that pseudonormality does not imply CCP, consider the following example.

Example 3 (pseudonormality does not imply CCP). Consider $x^* = (0,0)$ and the inequality constraints defined by

$$g_1(x_1, x_2) = -x_1,$$
$$g_2(x_1, x_2) = x_1 - x_1^2 x_2^2.$$

The point x^* is clearly feasible and both constraints are active at x^*. The gradients are given by $\nabla g_1(x_1, x_2) = (-1, 0)$ and $\nabla g_2(x_1, x_2) = (1 - 2x_1 x_2^2, -2x_1^2 x_2)$ for all (x_1, x_2) in $\mathbb{R}^2$.

Let us show that pseudonormality holds at $x^* = (0,0)$. Suppose that $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = (0,0)$ for some positive scalars μ_1 and μ_2, then $\mu_1 = \mu_2 = \mu$, but $\mu_1 g_1(x_1, x_2) + \mu_2 g_2(x_1, x_2) = -\mu x_1^2 x_2^2 \le 0$ for every $(x_1, x_2) \in \mathbb{R}^2$. So there no sequence that contradicts the pseudonormality condition.

Now, we will show that CCP does not hold at x^*. In x^* we have $K(x^*) = \mathbb{R} \times \{0\}$. Take $x_1^k = x_2^k = 1/k$ and define $\mu_2^k := (2(x_1^k)^2 x_2^k)^{-1}$, $\mu_1^k := \mu_2^k(1 - 2x_1^k(x_2^k)^2)$, and $\omega^k := \mu_1^k \nabla g_1(x_1^k, x_2^k) + \mu_2^k \nabla g_2(x_1^k, x_2^k)$. Clearly, $\omega^k \in K(x_1^k, x_2^k)$ for all $k \in \mathbb{N}$ and $\omega^k = \mu_1^k(-1, 0) + \mu_2^k(1 - 2x_1^k(x_2^k)^2, -2(x_1^k)^2 x_2^k) = (0, -1)$. So $\lim_{k\to\infty} \omega^k = (0, -1) \notin K(x^*)$. Thus, CCP does not hold at $x^* = (0,0)$.

As consequence of Examples 2 and 3 we conclude that CCP does not imply and neither does pseudonormality.

We end this section with the next example.

Example 4 (AKKT methods may not converge to KKT points under quasi-normality or Abadie's CQ). In $\mathbb{R}^2$, consider the following optimization problem

(4.13)
$$\text{minimize } f(x_1, x_2) = x_2 \text{ subject to } h(x_1, x_2) = x_1 x_2 = 0, g(x_1, x_2) = -x_1 \le 0.$$

The constraints satisfy quasi-normality at $x^* = (0,1)$ and as a consequence Abadie's CQ but not the CCP condition. Now we will prove this fact. The point x^* is feasible, both constraints are active, $\nabla g(x_1, x_2) = (-1, 0)$, and $\nabla h(x_1, x_2) = (x_2, x_1)$. First, let us show that quasi-normality holds at $x^* = (0,1)$. We have that $\nabla g(x^*) = -\nabla h(x^*) = (-1, 0)$, so if $\mu \nabla g(x^*) + \lambda \nabla h(x^*) = (0,0)$ with nonnull coefficients, we have $\mu = \lambda > 0$. Assume by contradiction that there is a sequence $(x_1^k, x_2^k) \to (0,1)$, such that $\lambda h(x_1^k, x_2^k) > 0$ and $\mu g(x_1^k, x_2^k) > 0$ for all $k \in \mathbb{N}$. From $\mu g(x_1^k, x_2^k) = -\mu x_1^k > 0$ we get $x_1^k < 0$ and from $\lambda h(x_1^k, x_2^k) = \lambda x_1^k x_2^k > 0$, we get $x_2^k < 0$ for all $k \in \mathbb{N}$. This is impossible since $x_2^k \to 1$. Therefore, quasi-normality holds at x^*.

Now, let us prove that CCP does not hold at x^*. We observe that $K(x^*) = \{\lambda \nabla h(x^*) + \mu \nabla g(x^*) : \lambda \in \mathbb{R}, \mu \in \mathbb{R}_+\} = \mathbb{R}_- \times \{0\}$. To show that CCP does not hold at x^*, we must find $\omega^* \in \limsup_{x\to x^*} K(x)$ such that $\omega^* \notin K(x^*)$. Define $x^k = (x_1^k, x_2^k)$ as $x_1^k = 1/k$, $x_2^k = 1$, and $\lambda^k = \mu^k = k$ for all $k \in \mathbb{N}$. With this choice define $\omega^k := \lambda^k \nabla h(x^k) + \mu^k \nabla g(x^k) \in K(x^k)$. Clearly, $\omega^k = \lambda^k(1, 1/k) + \mu^k(-1, 0) = (0,1)$ for all $k \in \mathbb{N}$, thus we get $w^* = \lim_{k\to\infty} \omega^k = (0,1) \in \limsup_{x\to x^*} K(x)$. However, $(0,1)$ does not belong to the cone $K(x^*) = \mathbb{R}_- \times \{0\}$.

Furthermore, $x^* = (0,1)$ is an AKKT point for (4.13): To see this, take $x_1^k = 1/k$, $x_2^k = 1$, $\lambda^k = k$, and $\mu^k = k$. Calculating we obtain that $\nabla f(x^k) + \mu^k \nabla g(x^k) + \lambda^k \nabla h(x^k) = (0,1) + k(-1,0) + k(1, 1/k) = (0,0)$. In spite of $x^* = (0,1)$ being an AKKT point, it means nothing for the optimization problem (4.13). The point x^* is neither an optimal solution point nor a stationary point. But it can be attained by an algorithm that generates AKKT points (as an augmented Lagrangian method, for instance). This means that the point $(0, 1)$ fulfills any sensible practical KKT test and the algorithm will accept a point which has no relation with the optimization problem (4.13); certainly this cannot happen if instead of the quasi-normality condition the point satisfies any CQ which implies the CCP condition such as LICQ, MFCQ, CRSC, CPG, or even CCP itself.

5. Concluding remarks. Many constrained optimization algorithms (but not all, see [7]) generate points that satisfy the AKKT condition. This optimality condition is known to be strong because it implies "KKT or not-CQ" for many popular CQs such as LICQ, MFCQ, CPLD, CRCQ, RCRCQ, RCPLD, CRSC, and the pleasantly weak CPG CQ [3, 4, 5, 8]. In this paper we asked for the weakest CQ for which AKKT implies "KKT or not-CQ." We found that the CCP condition is this weakest CQ, being even weaker than CPG. Therefore, the fact that AKKT implies "KKT or not-CCP" gives us the most accurate measure of the strength of AKKT as a stopping criterion for practical algorithms. In addition, we estabished the relations of CCP with other CQs that do not enjoy the "strictness property" (pseudonormality, quasi-normality, and Abadie's CQ).

The updated landscape of CQs and SCQs is given in Figure 2, where arrows indicate implications. Note that in Figure 1 we may identify the small rectangle with the points that satisfy CCP and the big rectangle with the points that satisfy Guignard's CQ.

Future work includes the analysis of strict CQs associated with different sequential optimality conditions and, so, with different stopping criteria for constrained optimization algorithms. Among these, we can mention the approximate gradient projection (AGP) condition [19], the L-AGP condition [3], and the complementary AGP condition [9].

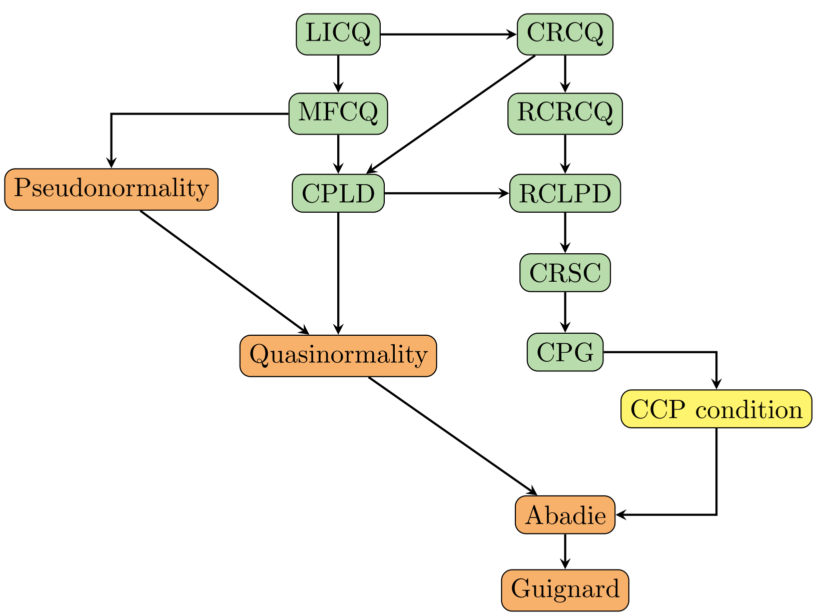

Fig. 2. *Updated landscape of CQs and SCQs.*

Acknowledgment. The authors would like to thank the anonymous referees whose comments and suggestions have greatly improved the quality of this work.

REFERENCES

[1] J. Abadie, *On the Kuhn-Tucker theorem*, in Nonlinear Programming, J. Abadie, ed. North-Holland, Amsterdam, 1967.

[2] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt, *On augmented Lagrangian methods with general lower-level constraints*, SIAM J. Optim., 18 (2008), pp. 1286–1309.

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

[4] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva, *A relaxed constant positive linear dependence constraint qualification and applications*, Math. Program., 135 (2012), pp. 255–273.

[5] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva, *Two new weak constraint qualifications and applications*, SIAM J. Optim., 22 (2012), pp. 1109–1135.

[6] R. Andreani, J. M. Martínez, A. Ramos, and P. J. S. Silva, *Constraint qualifications and approximate KKT-sequences*, contributed presentation at X Brazilian Workshop on Continuous Optimization, Florianópolis, Santa Catarina, Brazil, 2014.

[7] R. Andreani, J. M. Martínez, L. T. Santos, and B. F. Svaiter, *On the behaviour of constrained optimization methods when Lagrange multipliers do not exist*, Optim. Methods Softw., 29 (2014), pp. 646–657.

[8] R. Andreani, J. M. Martínez, and M. L. Schuverdt, *On the relation between constant positive linear dependence condition and quasinormality constraint qualification*, J. Optim. Theory Appl., 125 (2005), pp. 473–485.

[9] R. Andreani, J. M. Martínez, and B. F. Svaiter, *A new sequential optimality condition for constrained optimization and algorithmic consequences*, SIAM J. Optim., 20 (2010), pp. 3533–3554.

[10] E. G. Birgin and J. M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*, Fundam. Algorithms 10, SIAM, Philadelphia, 2014.

[11] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, Belmont, MA, 1999.

[12] D. P. Bertsekas and A. Ozdaglar, *Pseudonormality and a Lagrange multiplier theory for constrained optimization*, J. Optim. Theory Appl., 114 (2002), pp. 287–343.

[13] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, CMS Books in Mathematics, Springer, 2006.

[14] L. Chen and D. Goldfarb, *Interior-point ℓ_2-penalty methods for nonlinear programming with strong global convergence properties*, Math. Program., 108 (2006), pp. 1–36.

[15] M. Guignard, *Generalized Kuhn–Tucker conditions for mathematical programming problems in Banach space*, SIAM J. Control, 7 (1969), pp. 232–241.

[16] F. J. Gould and J. W. Tolle, *A necessary and sufficient qualification for constrained optimization*, SIAM J. Appl. Math., 20 (1971), pp. 164–172.

[17] M. R. Hestenes, *Optimization Theory: The Finite Dimensional Case*, Wiley, New York, 1975.

[18] O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in presence of equality and inequality constraints*, J. Math. Anal. Appl., 17 (1967), pp. 37–47.

[19] J. M. Martínez and B. F. Svaiter, *A practical optimality condition without constraint qualifications for nonlinear programming*, J. Optim. Theory Appl., 118 (2003), pp. 117–133.

[20] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* I. *Basis Theory*, Springer, Berlin, 1998.

[21] L. Qi and Z. Wei, *On the constant positive linear dependence condition and its application to SQP methods*, *SIAM J. Optim.*, 10 (2000), pp. 963–981.

[22] R. T. Rockafellar, *Lagrange multipliers and optimality*, SIAM Rev., 35 (1993), pp. 183–238.

[23] R. T. Rockafellar and R. Wets, *Variational Analysis*, Springer, Berlin, 1998.