Constraint Qualifications for Karush–Kuhn–Tucker Conditions in Multiobjective Optimization

Gabriel Haeser[1] · Alberto Ramos[2]

Abstract

The notion of a normal cone of a given set is paramount in optimization and variational analysis. In this work, we give a definition of a multiobjective normal cone, which is suitable for studying optimality conditions and constraint qualifications for multiobjective optimization problems. A detailed study of the properties of the multiobjective normal cone is conducted. With this tool, we were able to characterize weak and strong Karush–Kuhn–Tucker conditions by means of a Guignard-type constraint qualification. Furthermore, the computation of the multiobjective normal cone under the error bound property is provided. The important statements are illustrated by examples.

Keywords Multiobjective optimization · Optimality conditions · Constraint qualifications · Regularity · Weak and strong Kuhn–Tucker conditions

Mathematics Subject Classification 90C29

1 Introduction

Multiobjective optimization problems (MOPs) are a class of optimization problems involving more than one objective function to be optimized. Many real-life problems can be formulated as MOPs, which include engineering design, economics, financial

✉ Alberto Ramos
albertoramos@ufpr.br

Gabriel Haeser
ghaeser@ime.usp.br

[1] Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, São Paulo, Brazil

[2] Department of Mathematics, Federal University of Paraná, Paraná, Brasil

investment, mechanics, etc. See [1,2] and references therein. For several theoretical
and numerical issues related to MOPs, see [3–6].

Due to the possible conflicting nature of the objective functions, it is not expected
to find a solution that optimizes all of the objective functions simultaneously. Thus,
different notions of optimality are developed for MOPs, where one aims at finding
best trade-off options (strong/weak Pareto points, see Definition 2.1). As in nonlinear
programming, optimality conditions play a crucial role in the development of efficient
methods for solving MOPs. Some methods are Newton's methods [7], SQP methods
[8], trust-region methods [9,10] and scalarization methods such as weighted sum
methods [11,12], ε-constraint methods [13], Charnes–Cooper [14,15], Gerstewitz [16],
Pascoletti–Serafini [17] scalarizations, and others.

In scalar optimization problems, the Karush–Kuhn–Tucker (KKT) conditions play
a major role in theoretical and numerical optimization. To guarantee the fulfillment
of KKT conditions at a minimizer, one needs constraint qualifications (CQs), that
is, properties depending only on the functions defining the constraints, which ensure
the KKT conditions at local minimizers. For MOPs, the corresponding KKT con-
ditions take into account, simultaneously, multipliers for the constraints and for the
objective functions, with the multipliers associated with the objective functions being
nonnegative. If there exists at least one positive multiplier corresponding to the objec-
tive functions, we say that weak Karush–Kuhn–Tucker (weak KKT) conditions hold,
and when all the multipliers corresponding to the objective functions are positive, we
say that strong Karush–Kuhn–Tucker (strong KKT) conditions hold. But, differently
from the scalar case, usual CQs, in general, do not ensure weak/strong KKT condi-
tions. Thus, additional conditions, which take into account the objective functions, are
introduced: the so-called *regularity conditions* (RC). Such conditions have been used
for establishing first-order and second-order optimality conditions for smooth MOPs.
For more information, see [18–21] and references therein.

In this work, we take a new look at CQs for MOPs. We are interested in establishing
the weakest possible CQs that guarantee the validity of weak/strong KKT conditions
at locally weak Pareto solutions. To develop our new CQs, we rely on a new notion
of normal cone adapted to MOPs and some variational analysis tools. For details, see
Sect. 4.

The paper is organized as follows: In Sect. 2, we introduce the basic assumptions
and concepts useful in our analysis. In Sect. 3, we introduce the multiobjective normal
cone and we study its properties and calculus rules. In Sect. 4, we introduce new CQs
for MOP and we study their properties. Our concluding remarks are given in Sect. 5.

2 Preliminaries and Basic Assumptions

Our notation is standard in optimization and variational analysis: $\mathbb{R}^n$ is the n-
dimensional real Euclidean space, $n \in \mathbb{N}$. $\mathbb{R}_+$ is the set of nonnegative numbers
and $a^+ := \max\{0, a\}$, $a \in \mathbb{R}$. We denote by $\| \cdot \|$ the Euclidean norm in $\mathbb{R}^n$ and by
$\| \cdot \|_1$ the ℓ_1-norm, i.e., $\|z\|_1 := |z_1| + \cdots + |z_n|$, for $z \in \mathbb{R}^n$. Given a differentiable
mapping $F : \mathbb{R}^s \to \mathbb{R}^d$, we use $\nabla F(x)$ to denote the Jacobian matrix of F at x. For a
set-valued mapping $\Gamma : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the sequential Painlevé-Kuratowski outer limit of

$\Gamma(z)$ as $z \to z^*$ is denoted by

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

For a cone $\mathcal{K} \subset \mathbb{R}^s$, its polar is $\mathcal{K}^\circ := \{v \in \mathbb{R}^s : \langle v, k \rangle \leq 0 \text{ for all } k \in \mathcal{K}\}$. It is always the case that $\mathcal{K} \subset \mathcal{K}^{\circ\circ}$, and when $\mathcal{K}$ is a closed convex set, the equality $\mathcal{K}^{\circ\circ} = \mathcal{K}$ holds.

Given $X \subset \mathbb{R}^n$ and $\bar{z} \in X$, we define the *tangent cone* to X at $\bar{z}$ by

$$T(\bar{z}, X) := \{d \in \mathbb{R}^n : \exists \ t_k \downarrow 0, d^k \to d \text{ with } \bar{z} + t_k d^k \in X\},$$

and the *regular normal cone* to X at $\bar{z} \in X$ by

$$\widehat{N}(\bar{z}, X) := \left\{w \in \mathbb{R}^n : \limsup_{z \to \bar{z}, z \in X} \frac{\langle w, z - \bar{z} \rangle}{\|z - \bar{z}\|} \leq 0 \right\}.$$

The (Mordukhovich) *limiting normal cone* to X at $\bar{z} \in X$ is defined by

$$N(\bar{z}, X) := \limsup_{z \to \bar{z}, z \in X} \widehat{N}(z, X).$$

For a function $\phi : \mathbb{R}^n \to [-\infty, \infty]$ and a point $\bar{z} \in \mathbb{R}^n$ where $\phi(\bar{z})$ is finite. We write $z \to_\phi \bar{z}$ when $z \to \bar{z}$ and $\phi(z) \to \phi(\bar{z})$. The *regular subdifferential* of ϕ at $\bar{z}$ is defined as

$$\widehat{\partial}\phi(\bar{z}) := \left\{v \in \mathbb{R}^n : \liminf_{z \to \bar{z}, z \in X} \frac{\phi(z) - \phi(\bar{z}) - \langle v, z - \bar{z} \rangle}{\|z - \bar{z}\|} \geq 0 \right\}.$$

The *limiting subdifferential* of ϕ at $\bar{z}$ is defined as

$$\partial\phi(\bar{z}) := \limsup_{z \to_\phi \bar{z}} \widehat{\partial}\phi(z). \tag{1}$$

. When X is a nonempty closed set and ι_X is the indicator function, we see that $\widehat{N}(\bar{z}, X) = \widehat{\partial}\iota_X(\bar{z})$ and $N(\bar{z}, X) = \partial\iota_X(\bar{z})$. For a detailed treatment of variational analysis and optimization, we recommend [22–24].

Here, we consider the smooth constrained multiobjective optimization problem (MOP) of the form:

$$\begin{aligned} \text{minimize} \quad & f(x) := (f_1(x), f_2(x), \ldots, f_r(x)), \\ \text{subject to} \quad & h_i(x) = 0, \quad i = 1, \ldots, m; \\ & g_j(x) \leq 0, \quad j = 1, \ldots, p, \end{aligned} \tag{2}$$

where $h(\cdot) := (h_1(\cdot), \ldots, h_m(\cdot))$ and $g(\cdot) := (g_1(\cdot), \ldots, g_p(\cdot))$ are continuously differentiable on $\mathbb{R}^n$. We denote by $\Omega := \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$ the feasible set of (2), which is assumed to be a nonempty set.

Due to the possible conflicting nature of the objective functions, it is not expected to find a solution that can locally optimize all of the objective functions simultaneously. Therefore, different notions of optimality are developed for MOPs, where one aims at finding best trade-off options (local Pareto points, see Definition 2.1). We give below some important definitions.

Definition 2.1 Let $\bar{x}$ be a feasible point of (2). Then,

1. We say that $\bar{x}$ is a local *(strong[1]) Pareto* optimal point if there is a $\delta > 0$ such that there is no $x \in \Omega \cap B(\bar{x}, \delta)$ with $f_\ell(x) \leq f_\ell(\bar{x})$, $\forall \ell \in \{1, \ldots, r\}$ and $f(x) \neq f(\bar{x})$;
2. We say that $\bar{x}$ is a local *weak Pareto* optimal point if there is a $\delta > 0$ such that there is no $x \in \Omega \cap B(\bar{x}, \delta)$ with $f_\ell(x) < f_\ell(\bar{x})$, $\forall \ell \in \{1, \ldots, r\}$.

There is a natural way to relate local weak/strong Pareto optimal points with a scalar optimization problem. Indeed, given a set of not all zero scalars (*weights*) $\theta_1 \geq 0, \ldots, \theta_r \geq 0$, all local minimizers of the scalarized problem of minimizing $\theta_1 f_1(x) + \cdots + \theta_r f_r(x)$ subject to $x \in \Omega$ are local weak Pareto points, while the reverse implication holds under convexity. See [3]. By considering the scalarized problem with positive weights $\theta_1 > 0, \ldots, \theta_r > 0$, all its local minimizers are local Pareto solutions, and, under convexity and compacity assumptions, a dense subset of the local Pareto solutions (so-called proper local Pareto solutions) coincides with the local minimizers of this scalarized problem. See [26]. This suggests the following definitions of Karush–Kuhn–Tucker (KKT) notions associated with local weak and strong Pareto solutions. See [21,27].

Definition 2.2 Let x be a feasible point of (2) and assume that f is continuously differentiable. Suppose that there exists a vector $0 \neq (\theta, \lambda, \mu) \in \mathbb{R}_+^r \times \mathbb{R}^m \times \mathbb{R}_+^p$ such that $\mu_j g_j(x) = 0$, for $j = 1, \ldots, p$ and

$$\sum_{\ell=1}^r \theta_\ell \nabla f_\ell(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{j=1}^p \mu_j \nabla g_j(x) = 0. \tag{3}$$

Then, we say that:

a) x is a *weak KKT point*, if (3) holds with $\theta \neq 0$.
b) x is a *strong KKT point*, if (3) holds with $\theta_\ell > 0$, $\forall \ell = 1, \ldots, r$.

When $r = 1$, both concepts reduce to the classical KKT conditions, which are known to hold at a local minimizer only under a CQ. Among the most known ones, we mention the *linear independence CQ* (LICQ), which states the linear independence of all the gradients of active constraints, and the *Mangasarian-Fromovitz CQ* (MFCQ), which states the positive linear independence of such gradients, see [28]. We mention that necessary optimality conditions are not only useful for identifying possible candidate solutions to a problem, but also crucial in designing numerical methods

---

[1] We sometimes use the term strong Pareto to refer to a Pareto point only to emphasize the contrast with the notion of a weak Pareto point. This is not related to the notion of strong minimality defined in [25].

for solving such problems. Also, many CQs (but not all) can be used in the convergence analysis of numerical methods for scalar optimization problems. See [29–31] and references therein.

Finally, we recall some basic results on MOPs. In this paper, we use $V = (v_\ell)_{\ell=1}^r$ to denote a $n \times r$ matrix whose columns are the vectors $v_1, v_2, \ldots, v_r$ in $\mathbb{R}^n$. Thus, $V\theta = \theta_1 v_1 + \cdots + \theta_r v_r$ for $\theta := (\theta_1, \ldots, \theta_r) \in \mathbb{R}^r$. From [27,32], we have the following

Lemma 2.1 *Let $\bar{x}$ be a local weak Pareto point of* (2). *Consider the scalar optimization problem*

$$minimize\ \Phi(x) := \max_{\ell=1,\ldots,r} (f_\ell(x) - f_\ell(\bar{x})),\ subject\ to\ x \in \Omega. \tag{4}$$

Then, $\bar{x}$ is a local solution of (4). *Furthermore, if f is a C^1 function, $\max_{\ell=1,\ldots,r} \langle \nabla f_\ell(\bar{x}), d \rangle \geq 0, \forall d \in T(\bar{x}, \Omega)$.*

Remark 2.1 Under the assumptions of Lemma 2.1, it is not difficult to see that there exists $\delta > 0$ such that $\bar{x}$ is the unique minimizer of

$$minimize\ \max_{\ell=1,\ldots,r} (f_\ell(x) - f_\ell(\bar{x})) + \frac{1}{2} \|x - \bar{x}\|^2,\ s.t. x \in \Omega, \|x - \bar{x}\| \leq \delta. \tag{5}$$

From [33], we have the following theorem of alternative

Lemma 2.2 *Let A, B, C be given matrices with n columns and $A \neq 0$. Then, exactly one of the following statements is true:*

(s1) $Az > 0$, $Bz \geq 0$, $Cz = 0$ has a solution z;
(s2) $A^T y_1 + B^T y_2 + C^T y_3 = 0$, $y_1 \geq 0$, $y_2 \geq 0$, $y_1 \neq 0$ has solution y_1, y_2, y_3.

3 A Multiobjective Normal Cone

In this section, we will introduce a notion of normal cone for multiobjective optimization problems and derive its properties. To the best of our knowledge, this has not been previously considered. We start by taking a closed subset Ω of $\mathbb{R}^n$ and a point $\bar{x} \in \Omega$.

Definition 3.1 Given $\bar{x} \in \Omega$ and $r \in \mathbb{N}$, the *regular r-multiobjective normal cone* to Ω at $\bar{x}$ is the cone defined as

$$\widehat{N}(\bar{x}, \Omega; r) := \left\{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \limsup_{x \to \bar{x}, x \in \Omega} \min_{\ell=1,\ldots,r} \frac{\langle v_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}. \tag{6}$$

For a lower semicontinuous mapping f having lower semicontinuous extended real-valued components $f_\ell : \mathbb{R}^n \to] -\infty, \infty]$, $\ell = 1, \ldots, r$, we define the *regular r-multiobjective subdifferential* of f at $\bar{x}$, denoted by $\widehat{\partial} f(\bar{x}; r)$, as the set

$$\left\{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \liminf_{x \to \bar{x}} \left(\max_{\ell=1,\ldots,r} \frac{f_\ell(x) - f_\ell(\bar{x}) - \langle v_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} \right) \geq 0 \right\}.$$

Furthermore, we define the *limiting r-multiobjective normal cone* to Ω at $\bar{x}$ by

$$N(\bar{x}, \Omega; r) := \limsup_{x \to \bar{x}, x \in \Omega} \widehat{N}(x, \Omega; r)$$

and the *limiting r-multiobjective subdifferential* of f at $\bar{x}$

$$\partial f(\bar{x}; r) := \limsup_{x \to_f \bar{x}} \widehat{\partial} f(x; r).$$

where $x \to_f \bar{x}$ means that $x \to \bar{x}$ with $f(x) \to f(\bar{x})$. Clearly, when $r = 1$, the (limiting/regular) r-multiobjective normal cones and (limiting/regular) r-multiobjective subdifferential coincide with their scalar counterparts which are very well studied in [22,23]. In the rest of this section, we will analyze the main properties of the multiobjective normal cone. Thus, we start by computing them for some simple cases.

Proposition 3.1 *Let Ω be a closed convex set, $\bar{x} \in \Omega$, and $r \in \mathbb{N}$. Then, we have that $N(\bar{x}, \Omega; r) = \widehat{N}(\bar{x}, \Omega; r)$. Furthermore,*

$$\widehat{N}(\bar{x}, \Omega; r) = \{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \min_{\ell=1,\dots,r} \langle v_\ell, x - \bar{x} \rangle \leq 0, \ \forall x \in \Omega \}. \quad (7)$$

Proof First, we will show the equality in (7). Take $V \in \widehat{N}(\bar{x}, \Omega; r)$ and $x \in \Omega$. For every $t \in [0, 1]$, set $x(t) := \bar{x} + t(x - \bar{x}) \in \Omega$. Thus, since $x(t) - \bar{x} = t(x - \bar{x})$, we obtain that

$$\min_{\ell=1,\dots,r} \frac{\langle v_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} = \lim_{t \to 0} \min_{\ell=1,\dots,r} \frac{\langle v_\ell, x(t) - \bar{x} \rangle}{\|x(t) - \bar{x}\|} \leq \limsup_{z \to \bar{x}, z \in \Omega} \min_{\ell=1,\dots,r} \frac{\langle v_\ell, z - \bar{x} \rangle}{\|z - \bar{x}\|} \leq 0,$$

which implies that $\min_{\ell=1,\dots,r} \langle v_\ell, x - \bar{x} \rangle \leq 0$.

Clearly, if $V = (v_\ell)_{\ell=1}^r$ satisfies $\min_{\ell=1,\dots,r} \langle v_\ell, x - \bar{x} \rangle \leq 0 \ \forall x \in \Omega$, we have that $V \in \widehat{N}(\bar{x}, \Omega; r)$. From (7), we obtain that $N(\bar{x}, \Omega; r) = \widehat{N}(\bar{x}, \Omega; r)$. $\square$

When Ω is a polyhedral cone, we can obtain a nice characterization of the r-multiobjective normal cone.

Proposition 3.2 *Suppose that $\Omega = \{x \in \mathbb{R}^n : Ax \leq 0\}$ for some matrix $A \in \mathbb{R}^{m \times n}$. Then, for every $\bar{x} \in \Omega$, $\widehat{N}(\bar{x}, \Omega; r) = N(\bar{x}, \Omega; r)$ holds and*

$$N(\bar{x}, \Omega; r) = \left\{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \begin{array}{l} V\theta = A^T \lambda, \ \text{with } \lambda \in N(A\bar{x}, \mathbb{R}_-^m) \\ \text{and } 0 \neq \theta \in \mathbb{R}_+^r \end{array} \right\}. \quad (8)$$

Proof This will follow by the more general result proved in Corollary 4.1, for every constraint set satisfying the error bound property. $\square$

It is well known that the regular normal cone can be characterized by viscosity subgradients [22, Theorem 6.11]. Here, we have a variational characterization of the regular r-multiobjective normal cone, which extends the viscosity characterization of the classical normal cone when $r = 1$.

Theorem 3.1 *Let $\bar{x}$ be a feasible point of* (2). *Then, the following statements are equivalent*

(a) *$V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r}$ belongs to $\widehat{N}(\bar{x}, \Omega; r)$;*
(b) *There is a C^1 mapping $f(x) = (f_\ell(x))_{\ell=1}^r$ such that $\bar{x}$ is a local weak Pareto optimal point relative to Ω, and $v_\ell = -\nabla f_\ell(\bar{x})$, for every $\ell = 1, \ldots, r$;*
(c) *For every $d \in T(\bar{x}, \Omega)$, we have that $\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \leq 0$.*

As a consequence

$$\widehat{N}(\bar{x}, \Omega; r) = \{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \leq 0, \ \forall d \in T(\bar{x}, \Omega) \}. \quad (9)$$

Proof First, we will show that (a) implies (b). Take $V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r)$. We will use an argument similar to [22, Theorem 6.11]. We define

$$\eta_0(r) := \text{Sup} \{ \min_{\ell=1,\ldots,r} \langle v_\ell, x - \bar{x} \rangle : x \in \Omega, \|x - \bar{x}\| \leq r \}. \quad (10)$$

Clearly, $\eta_0(r) \leq r \min_{\ell=1,\ldots,r} \|v_\ell\|$ and $0 = \eta_0(0) \leq \eta_0(r) \leq o(r)$. By using [22, Theorem 6.11], there is $\eta : \mathbb{R}_+ \to \mathbb{R}$ continuously differentiable on $\mathbb{R}_+$ with $\eta_0(r) \leq \eta(r)$, $\eta'(r) \to 0$ and $\eta(r)/r \to 0$ as $r \to 0$. Thus, for every $\ell \in \{1, \ldots, r\}$ define $f_\ell(x) := -\langle v_\ell, x - \bar{x} \rangle + \eta(\|x - \bar{x}\|)$ and $f(x) = (f_\ell(x))_{\ell=1}^r$. Clearly, f is a C^1 mapping with $f(\bar{x}) = 0$. To show that $\bar{x}$ is a local weak Pareto point for f relative to Ω, suppose that this is not the case. Then, there is a sequence $x^k \to \bar{x}$ with $x^k \in \Omega$, $\forall k \in \mathbb{N}$ such that $f_\ell(x^k) < f_\ell(\bar{x}) = 0, \forall \ell = 1, \ldots, r$. Since $\eta_0(r) \leq \eta(r)$ and, by definition of $f_\ell(x)$, we see that $-\langle v_\ell, x^k - \bar{x} \rangle + \eta_0(\|x^k - \bar{x}\|) < 0, \forall \ell$ and hence $\eta_0(\|x^k - \bar{x}\|) < \min_{\ell=1,\ldots,r} \langle v_\ell, x^k - \bar{x} \rangle$ for k large enough, which is a contradiction with (10).

To show that (b) implies (c), suppose that there exists a C^1 mapping $f(x) = (f_\ell(x))_{\ell=1}^r$ such that $\bar{x}$ is a local weak Pareto optimal point relative to Ω. From Lemma 2.1, $\max_{\ell=1,\ldots,r} \langle \nabla f_\ell(\bar{x}), d \rangle \geq 0, \forall d \in T(\bar{x}, \Omega)$. Thus, if
    $V := (v_\ell)_{\ell=1}^r$ with $v_\ell := -\nabla f_\ell(\bar{x}), \forall \ell$, we get that $\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \leq 0$.

Finally, to see that (c) implies (a), assume that V does not belong to $\widehat{N}(\bar{x}, \Omega; r)$. Thus, there exist a scalar $\alpha > 0$ and a sequence $\{x^k\} \subset \Omega$ converging to $\bar{x}$ such that

$$\lim_{k \to \infty} \min_{\ell=1,\ldots,r} \frac{\langle v_\ell, x^k - \bar{x} \rangle}{\|x^k - \bar{x}\|} > \alpha > 0. \quad (11)$$

Taking a further subsequence of $\{x^k\}$, we may assume that $(x^k - \bar{x})/\|x^k - \bar{x}\|$ converges to some vector $d \in \mathbb{R}^n$. Clearly, $d \in T(\bar{x}, \Omega)$. From (11), we get that $\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \geq \alpha > 0$, a contradiction. Thus, $V \in \widehat{N}(\bar{x}, \Omega; r)$. □

Remark 3.1 When $r = 1$, the right-hand side of (9) coincides with $T(\bar{x}, \Omega)^\circ$. Thus, $\widehat{N}(\bar{x}, \Omega) = \widehat{N}(\bar{x}, \Omega; 1) = T(\bar{x}, \Omega)^\circ$.

We have also the following

Lemma 3.1 *Consider f and g two lower semicontinuous mappings having lower semicontinuous extended-real-valued components, and Ω a closed subset of $\mathbb{R}^n$. Then, we always have that*

1. *$\widehat{\partial} \bar{\iota}_{\Omega}(x; r) = \widehat{N}(x, \Omega; r)$ where $\bar{\iota}_{\Omega}$ is the indicator mapping, that is, a mapping defined by $\bar{\iota}_{\Omega}(x) = 0$ if $x \in \Omega$ and $\bar{\iota}_{\Omega}(x) = \infty$ otherwise.*
2. *$V = (v_\ell)_{\ell=1}^r \in \widehat{\partial} f(x; r)$ iff there is a C^1 mapping h such that x is a local weak Pareto solution of minimizing $f - h$, and $v_\ell = \nabla h_\ell(x), \forall \ell = 1, \dots, r$.*
3. *If x is a local weak Pareto solution of minimizing f, then $0 \in \widehat{\partial} f(x; r)$.*
4. *If f is a C^1 mapping, then*

$$\widehat{\partial} f(x; r) = \{U \in \mathbb{R}^{n \times r} : U\theta = \nabla f(x)^T \theta, \ \|\theta\|_1 = 1, \theta \geq 0\}.$$

5. *If f is a C^1 mapping, then*

$$\widehat{\partial}(f + g)(x; r) = \nabla f(x)^T + \widehat{\partial} g(x; r).$$

Proof First, it is straightforward to see that item 1 holds. Also, observe that item 2 is just a modification of Theorem 3.1.

Item 3: Since $\bar{x}$ is a local weak Pareto solution of minimizing f, we have that $\max_{\ell=1,\dots,r}(f_\ell(x) - f_\ell(\bar{x})) \geq 0$ for every x near $\bar{x}$, (see Lemma 2.1), and thus, $0 \in \widehat{\partial} f(\bar{x}; r)$.

Item 4: Take $V = (v_\ell)_{\ell=1}^r \in \widehat{\partial} f(\bar{x}; r)$. By item 2, there exists a C^1 mapping h such that $\bar{x}$ is a local weak Pareto solution of minimizing $f - h$ and $v_\ell = \nabla h_\ell(\bar{x})$, $\forall \ell = 1, \dots, r$. Then, $\max_{\ell=1,\dots,r}(f_\ell(x) - h_\ell(x) - (f_\ell(\bar{x}) - h_\ell(\bar{x}))) \geq 0$, for every x near $\bar{x}$. Thus, by Fermat's rule and computing Clarke's subdifferential, there exists $\theta \geq 0$ with $\|\theta\|_1 = 1$ such that $\sum_{\ell=1}^r \theta_\ell(\nabla f_\ell(\bar{x}) - \nabla h_\ell(\bar{x})) = 0$, i.e., $V\theta = \nabla f(\bar{x})^T \theta$.

Now, take $U = (u_\ell)_{\ell=1}^r$ such that $U\theta = \nabla f(x)^T \theta$ for $\theta \geq 0$, $\|\theta\|_1 = 1$. Observe that

$$
\begin{aligned}
\max_{\ell=1,\dots,r} \frac{f_\ell(x) - f_\ell(\bar{x}) - \langle u_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} &\geq \sum_{\ell=1}^r \frac{\theta_\ell f_\ell(x) - \theta_\ell f_\ell(\bar{x}) - \theta_\ell \langle u_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} \\
&\geq \sum_{\ell=1}^r \frac{\theta_\ell f_\ell(x) - \theta_\ell f_\ell(\bar{x}) - \theta_\ell \langle \nabla f_\ell(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \\
&\geq \sum_{\ell=1}^r \theta_\ell \left(\frac{f_\ell(x) - f_\ell(\bar{x}) - \langle \nabla f_\ell(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \right).
\end{aligned}
$$

Taking limit inferior as $x \to \bar{x}$ in the last expression, we get $U \in \widehat{\partial} f(x; r)$.

Item 5: By item 2 of this lemma, we see that $V = (v_\ell)_{\ell=1}^r \in \widehat{\partial}(f + g)(x; r)$. iff there is a C^1 mapping h such that x is a local weak Pareto solution of $(f + g) - h$ and $v_\ell = \nabla h_\ell(x), \forall \ell$ iff x is also a local weak Pareto solution of $g - (h - f)$ where $h - f$ is a C^1 mapping iff (using again the item 2) $V - \nabla f(x)^T = (\nabla h_\ell(x) - \nabla f_\ell(x))_{\ell=1}^r$ is in $\widehat{\partial} g(x; r)$. □

We end this section with the following observation: If $\bar{x}$ is a local Pareto solution of minimizing $f(x) = (f_1(x), \ldots, f_r(x))$ subject to $x \in \Omega$, then $\bar{x}$ is also a local Pareto solution of the unconstrained MOP of minimizing $f(x) + \bar{\iota}_\Omega(x)$, where $\bar{\iota}_\Omega(x)$ is the indicator mapping defined in Lemma 3.1, Item 1. Thus, by Lemma 3.1, Item 3, we obtain that $0 \in \widehat{\partial}(f + \bar{\iota}_\Omega)(\bar{x}; r)$. If we assume that $f(x)$ is a C^1 mapping, by Lemma 3.1, Item 5, we obtain that $0 \in (\nabla f_1(x), \ldots, \nabla f_r(x)) + \widehat{N}(\bar{x}, \Omega; r)$.

Corollary 3.1 *If $\bar{x}$ is a local weak Pareto optimal point of* (2)*, then*

$$0 \in (\nabla f_1(x), \ldots, \nabla f_r(x)) + \widehat{N}(x, \Omega; r). \tag{12}$$

Thus, (12) can be interpreted as Fermat's rule using the r-multiobjective normal cone.

4 On Constraint Qualifications for Multiobjective Optimization

In this section, we will focus on the use of the r-multiobjective normal cone to define CQs to characterize the weak/strong KKT conditions. Later, we will analyze the relations of the r-multiobjective cones with the tangent cone and normal cone of the feasible set.

Let us consider the tangent cone to Ω at $\bar{x}$, $T(\bar{x}, \Omega)$, and its first-order approximation, $L(\bar{x}, \Omega)$, which is called the *linearized tangent cone to Ω at $\bar{x}$*, and is defined by

$$L(\bar{x}, \Omega) := \left\{ d \in \mathbb{R}^n : \begin{array}{ll} \langle \nabla h_i(\bar{x}), d \rangle = 0, & i = 1, \ldots, m \\ \langle \nabla g_j(\bar{x}), d \rangle \leq 0, & j \in A(\bar{x}) \end{array} \right\},$$

where $A(\bar{x}) := \{ j = 1, \ldots, p : g_j(\bar{x}) = 0 \}$. Given a feasible point $\bar{x} \in \Omega$, *Guignard's CQ* is stated as $T(\bar{x}, \Omega)^\circ = L(\bar{x}, \Omega)^\circ$, while *Abadie's CQ* states the stronger equality $T(\bar{x}, \Omega) = L(\bar{x}, \Omega)$.

In this section, we define new CQs for weak/strong KKT conditions based on the multiobjective normal cone, and we analyze their properties.

4.1 Constraint Qualifications for Weak KKT Points

Here, we start by giving the weakest CQ that guarantees that a local weak Pareto minimizer is, in fact, a weak KKT point. This is characterized by the fact that the fulfillment of the CQ is equivalent to the fact that, independently of the objective function, local weak Pareto implies weak KKT.

Theorem 4.1 *Let $\bar{x}$ be a feasible point of* (2)*. Then, the weakest property under which every local weak Pareto point satisfies the weak KKT conditions, for every continuously differentiable mapping $f(x) = (f_1(x), \ldots, f_r(x))$, is*

$$L(\bar{x}, \Omega) \subset \{ d \in \mathbb{R}^n : \min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \leq 0, \text{ for all } V = (v_i)_{i=1}^r \in \widehat{N}(\bar{x}, \Omega; r) \}. \tag{13}$$

Proof Assume that every local weak Pareto point is actually a weak KKT point. We will show that inclusion (13) holds. Indeed, take $d \in L(\bar{x}, \Omega)$ and $V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r)$. By Theorem 3.1, there is a C^1 mapping f such that $\bar{x}$ is a local weak Pareto minimizer of f over Ω and $v_\ell = -\nabla f_\ell(\bar{x})$, for every $\ell = 1, \ldots, r$. By hypotheses, we get that $\bar{x}$ is a weak KKT point and thus, there exist $\theta \in \mathbb{R}_+^m$, $\mu \in \mathbb{R}_+^p$ and $\lambda \in \mathbb{R}^m$ such that

$$\sum_{\ell=1}^r \theta_\ell v_\ell = \sum_{i=1}^m \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^p \mu_j \nabla g_j(\bar{x}), \quad \text{and} \quad \mu_j = 0, \quad \forall j \notin A(\bar{x}).$$

Furthermore, we can assume that $\|\theta\|_1 = 1$. Since $d \in L(\bar{x}, \Omega)$, we see that

$$\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \le \sum_{\ell=1}^r \theta_\ell \langle v_\ell, d \rangle = \sum_{i=1}^m \lambda_i \langle \nabla h_i(\bar{x}), d \rangle + \sum_{j=1}^p \mu_j \langle \nabla g_j(\bar{x}), d \rangle \le 0.$$

Now, assume that (13) holds. Thus, take $f(x) = (f_1(x), \ldots, f_r(x))$ a C^1 mapping such that $\bar{x}$ is a local weak Pareto optimal point over Ω. Define $v_\ell := -\nabla f_\ell(\bar{x})$, $\forall \ell = 1, \ldots, r$. By Lemma 2.2, we see that $\bar{x}$ is a weak KKT point, if the following system

$$\nabla h_i(\bar{x})^T d = 0, \quad i = 1, \ldots, m; \quad \nabla g_j(\bar{x})^T d \le 0, \quad j \in A(\bar{x}),$$
$$\nabla f_\ell(\bar{x})^T d < 0, \quad \ell = 1, \ldots, r \tag{14}$$

has no solution $d \in \mathbb{R}^n$. Now, if we suppose that system (14) admits a solution d, we get that $d \in L(\bar{x}, \Omega)$, and by (13), $\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \le 0$, which is equivalent to $\max_{\ell=1,\ldots,r} \langle \nabla f_\ell(\bar{x}), d \rangle \ge 0$, a contradiction with the second line of (14). Thus, (14) has no solution. $\qquad\square$

Recall that expression (13) is a true CQ, since it only depends on the feasible set, and no information of the objective functions is required. To continue our discussion, define

$$\mathfrak{L}(\bar{x}, \Omega; r) = \left\{ d \in \mathbb{R}^n : \min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle \le 0, \forall V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r) \right\}. \tag{15}$$

Using (15), Theorem 4.1 states that the weakest CQ guaranteeing that a local weak Pareto minimizer is a weak KKT point is the inclusion $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$. Note that in the scalar case ($r = 1$), inclusion (13) reduces to Guignard's CQ. Indeed, when $r = 1$, we see that $\mathfrak{L}(\bar{x}, \Omega; 1) = \widehat{N}(\bar{x}, \Omega)^\circ = T(\bar{x}, \Omega)^{\circ\circ}$. Thus, (13) is equivalent to $L(\bar{x}, \Omega) \subset T(\bar{x}, \Omega)^{\circ\circ}$, which is equivalent to the classical Guignard's CQ.

We proceed by analyzing the relation of (13) with Abadie's and Guignard's CQs. It is known that Abadie's CQ is enough to ensure that every local weak Pareto point satisfies the weak KKT conditions, see, for instance, [34], but Guignard's CQ does not ensure the validity of the weak KKT conditions at local weak Pareto points, [35]. This is the so-called gap in scalar and multiobjective optimization, which has attracted the

attention of many researchers. For more details, see also [36–38]. Here, we will try to elucidate this "gap" by using the cone $\mathfrak{L}(\bar{x}, \Omega; r)$. First, consider Proposition 4.1, where some properties about $\mathfrak{L}(\bar{x}, \Omega; r)$ are stated and another proof of the fact that Abadie's CQ ensures the validity of weak KKT at local weak Pareto points is provided [see (18)].

Proposition 4.1 *Given $\bar{x} \in \Omega$ and $r \in \mathbb{N}$, we always have*

$$\mathfrak{L}(\bar{x}, \Omega; r) \subset \widehat{N}(\bar{x}, \Omega)^{\circ}, \tag{16}$$

$$T(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r), \tag{17}$$

$$\textit{Abadie's CQ at } \bar{x} \textit{ implies } L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r), \tag{18}$$

$$L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r) \textit{implies Guignard's CQ.} \tag{19}$$

Proof To see that (16) holds, take $d \in \mathfrak{L}(\bar{x}, \Omega; r)$ and $w \in \widehat{N}(\bar{x}, \Omega)$. By Theorem 3.1 in the case that $r = 1$, we see that there exists a C^1 function $h(x)$ such that $\bar{x}$ is a local minimizer of $h(x)$ over Ω and $w = -\nabla h(\bar{x})$. Now, define the C^1 mapping $f(x) = (f_1(x), \ldots, f_r(x))$, where each component $f_\ell(x) = h(x), \forall x, \forall \ell = 1, \ldots, r$. It is not difficult to see that $\bar{x}$ is a local weak Pareto minimizer for $f(x)$ over Ω. Since $\nabla f_\ell(\bar{x}) = \nabla h(\bar{x})$ for every ℓ, we see that $\langle w, d \rangle = \langle -\nabla h(\bar{x}), d \rangle = \min_{\ell=1,\ldots,r} \langle -\nabla f_\ell(\bar{x}), d \rangle \leq 0$. Thus, we obtain that $d \in \widehat{N}(\bar{x}, \Omega)^{\circ}$ and hence (16) holds. Inclusion (17) is a simple consequence of (9). Implication (18) follows directly from inclusion (17), when Abadie's CQ holds at $\bar{x}$ (i.e., $T(\bar{x}, \Omega) = L(\bar{x}, \Omega)$). To show (19), from (16) and the equality $\widehat{N}(\bar{x}, \Omega) = T(\bar{x}, \Omega)^{\circ}$, we see that

$$T(\bar{x}, \Omega) \subset L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r) \subset \widehat{N}(\bar{x}, \Omega)^{\circ} = T(\bar{x}, \Omega)^{\circ\circ}. \tag{20}$$

Taking polar in the last sequence of inclusions, we get $L(\bar{x}, \Omega)^{\circ} = T(\bar{x}, \Omega)^{\circ}$, which is Guignard's CQ at $\bar{x}$. □

Finally, we show that Abadie's CQ, Guignard's CQ and inclusion (13) are independent CQs.

Example 4.1 (*Guignard's CQ is strictly weaker than $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$*) In $\mathbb{R}^2$, take $\bar{x} = (0, 0)$ and the multiobjective problem

$$\text{minimize } (f_1(x_1, x_2) := x_1, \quad f_2(x_1, x_2) := x_2) \text{ s.t. } h(x_1, x_2) := x_1 x_2 = 0.$$

It is not difficult to see that $\bar{x}$ is a local weak Pareto point, which is not a weak KKT point. Thus, Theorem 4.1 says that (13) fails. Now, we also observe that $L(\bar{x}, \Omega) = \mathbb{R}^2$ and $T(\bar{x}, \Omega) = \{(d_1, d_2) \in \mathbb{R}^2 : d_1 d_2 = 0\}$, and thus $L(\bar{x}, \Omega)^{\circ} = T(\bar{x}, \Omega)^{\circ} = \{0\}$ holds, that is, Guignard's CQ is valid at $\bar{x}$.

Clearly, Abadie's CQ is strictly stronger than inclusion (13), since in the scalar case, Guignard's CQ is equivalent to (13).

Finally, we have the following

Proposition 4.2 *Let $\bar{x}$ be a feasible point of (2) such that $\mathfrak{L}(\bar{x}, \Omega; r)$ is a closed convex cone. Then, Guignard's CQ holds at $\bar{x}$ iff $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$ holds.*
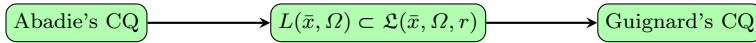
Fig. 1 Gap between Abadie's and Guignard's CQ for weak KKT points

Proof By Proposition 4.1, inclusion $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$ implies Guignard's CQ holds. For the other implication, from (16), (17) and Guignard's CQ, we get that $T(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r) \subset T(\bar{x}, \Omega)^{\circ\circ} = L(\bar{x}, \Omega)$. Taking the polar and using again Guignard's CQ, we see that $L(\bar{x}, \Omega)^{\circ} \subset \mathfrak{L}(\bar{x}, \Omega; r)^{\circ} \subset T(\bar{x}, \Omega)^{\circ} = L(\bar{x}, \Omega)^{\circ}$ implies that $L(\bar{x}, \Omega) = \mathfrak{L}(\bar{x}, \Omega; r)^{\circ\circ} = \mathfrak{L}(\bar{x}, \Omega; r)$. $\square$

4.2 Constraint Qualifications for Strong KKT Points

Now, we focus on CQs and the fulfillment of the strong KKT conditions. It is well known that Abadie's CQ is not sufficient to ensure the validity of strong KKT at local weak Pareto minimizers. Using the multiobjective normal cone, similarly to Theorem 4.1, we obtain the weakest CQ needed to guarantee that a local weak Pareto minimizer is a strong KKT point.

Theorem 4.2 *Let $\bar{x}$ be a feasible point of* (2)*. The weakest property under which every local weak Pareto point is a strong KKT point for every continuously differentiable mapping $f(x) = (f_1(x), \ldots, f_r(x))$ is*

$$L(\bar{x}, \Omega) \subset \left\{ d \in \mathbb{R}^n : \begin{array}{l} \text{for every } V = (v_i)_{i=1}^r \in \widehat{N}(\bar{x}, \Omega; r), \text{ we have that} \\ \min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle < 0 \text{ or } \langle v_\ell, d \rangle = 0, \forall \ell = 1, \ldots, r \end{array} \right\}. \tag{21}$$

Proof To prove inclusion (21), take $V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r)$ and a direction $d \in L(\bar{x}, \Omega)$. By Theorem 3.1, there is a C^1 mapping $f(x)$ having $\bar{x}$ as a local weak Pareto point over Ω, such that $-\nabla f_\ell(\bar{x}) = v_\ell$ for every $\ell = 1, \ldots, r$. By hypotheses, $\bar{x}$ must be a strong KKT point. So, there exist $\theta \in \mathbb{R}_+^r$, $\mu \in \mathbb{R}_+^p$ and $\lambda \in \mathbb{R}^m$, with $\theta_\ell > 0$, $\forall \ell = 1, \ldots, r$, such that $\|\theta\|_1 = 1$ and

$$\sum_{\ell=1}^r \theta_\ell v_\ell = \sum_{i=1}^m \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^p \mu_j \nabla g_j(\bar{x}), \quad \mu_j = 0, \quad \forall j \notin A(\bar{x}). \tag{22}$$

Now, from (22) and since $d \in L(\bar{x}, \Omega)$, we see that

$$\sum_{\ell=1}^r \theta_\ell \langle v_\ell, d \rangle = \sum_{i=1}^m \lambda_i \langle \nabla h_i(\bar{x}), d \rangle + \sum_{j=1}^p \mu_j \langle \nabla g_j(\bar{x}), d \rangle \leq 0. \tag{23}$$

From (23), we get that $\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle < 0$ or $\min_{\ell=1,\ldots,r} \langle v_\ell, d \rangle = 0$. In the last case, since $\theta_\ell > 0$, $\forall \ell$, we get $\langle v_\ell, d \rangle = 0$, for every $\ell = 1, \ldots, r$.

Now, let $f(x) = (f_1(x), \ldots, f_r(x))$ be a C^1 mapping such that $\bar{x}$ is a local weak Pareto optimal point over Ω. Set $v_\ell := -\nabla f_\ell(\bar{x})$, $\forall \ell = 1, \ldots, r$. Using Lemma 2.2,

we see that $\bar{x}$ is a strong KKT point, if the following system

$$\nabla h_i(\bar{x})^T d = 0, \quad i = 1, \ldots, m; \quad \nabla g_j(\bar{x})^T d \leq 0, \quad j \in A(\bar{x});$$
$$\nabla f_\ell(\bar{x})^T d \leq 0, \quad \ell = 1, \ldots, r; \quad \nabla f_\ell(\bar{x})^T d < 0, \text{ for some } \ell \in \{1, \ldots, r\} \quad (24)$$

has no solution $d \in \mathbb{R}^n$. Indeed, if (24) admits a solution d, then $d \neq 0$ belongs to $L(\bar{x}, \Omega)$. From (21), we get a contradiction. Thus, $\bar{x}$ is a strong KKT point. $\qquad \square$

When $r = 1$, the right-hand side of (21) collapses to $\widehat{N}(\bar{x}, \Omega)^\circ$. Thus, (21) coincides with $L(\bar{x}, \Omega) \subset T(\bar{x}, \Omega)^{\circ\circ}$ which is Guignard's CQ.

Observe that inclusion (21) is a true CQ, since it does not require any information of the objective functions for the formulation; moreover, it characterizes the fulfillment of the strong KKT conditions at local weak Pareto points, which coincides with Guignard's CQ when $r = 1$. Unfortunately, for multiobjective problems ($r > 1$), inclusion (21) is too strong, in the sense that even in well-behaved constrained system, where LICQ holds, it may fail.

Example 4.2 (*Inclusion* (21) *does not imply LICQ*) Consider [39, Example 4.18]: In $\mathbb{R}$, take $\bar{x} = 0$ and the constraints $g_1(x) := x$ and $g_2(x) := -\exp(x) + 1$. Note that $\nabla g_1(x) = 1$ and $\nabla g_2(x) = -\exp(x)$, and thus LICQ fails. Now, from $L(\bar{x}, \Omega) = \{0\}$, inclusion (21) holds.

Example 4.3 (*LICQ does not imply the inclusion* (21)) Indeed, in $\mathbb{R}^2$, take $\bar{x} = (0, 0)$, $a \in \mathbb{R}$ and the multiobjective problem

$$\min (f_1(x_1, x_2) := x_1^2, \quad f_2(x_1, x_2) := (x_1 - a)^2) \text{ s.t. } g(x_1, x_2) := x_1^2 - x_2 \leq 0.$$

It is not difficult to see that $\bar{x}$ is a local weak Pareto point. Straightforward calculations show that $\bar{x}$ is not a strong KKT point. Thus, Theorem 4.2 says that (21) fails. Now, since $\nabla g(x_1, x_2) := (2x_1, -1)^T$, LICQ holds at $\bar{x}$.

From the examples above, we see that usual CQs are not sufficient to obtain strong KKT points at local weak Pareto solutions. Besides this observation, it is still possible to use CQs for obtaining strong KKT, but for a stronger notion of optimality. We focus on the concept of proper Pareto efficiency in the sense of Geoffrion [40].

Definition 4.1 We say that a feasible point $\bar{x}$ is a local *Geoffrion-properly efficient* solution, if it is a local strong Pareto solution, within a neighborhood $\mathcal{N}$, and there exists a scalar $M > 0$ such that, for each ℓ,

$$\frac{f_\ell(x) - f_\ell(\bar{x})}{f_j(\bar{x}) - f_j(x)} \leq M,$$

for some j such that $f_j(\bar{x}) < f_j(x)$, whenever $x \in \Omega \cap \mathcal{N}$ and $f_\ell(\bar{x}) > f_\ell(x)$.

Proposition 4.3 *Let $\bar{x}$ be a feasible point of* (2). *If $\bar{x}$ is a Geoffrion-properly efficient point, with Abadie's CQ holding at $\bar{x}$, then $\bar{x}$ is a strong KKT point.*

Proof By Lemma 2.2, $\bar{x}$ is a strong KKT point of (2), if the following system

$$
\begin{aligned}
&\nabla f_\ell(\bar{x})^T d \le 0, \quad \ell = 1, \ldots, r; \quad \nabla f_\ell(\bar{x})^T d < 0, \text{ for some } \ell \in \{1, \ldots, r\}; \\
&\nabla h_i(\bar{x})^T d = 0, \quad i = 1, \ldots, m; \quad \nabla g_j(\bar{x})^T d \le 0, \quad j \in A(\bar{x})
\end{aligned} \tag{25}
$$

has no solution $d \in \mathbb{R}^n$. Thus, suppose that (25) admits a nontrivial solution d. From Abadie's CQ, there is a sequence $x^k := \bar{x} + t_k d^k \in \Omega$, with $t_k \to 0$, $d^k \to d$ and $x^k \to \bar{x}$. Thus, set $\mathcal{I}_- := \{\ell : \nabla f_\ell(\bar{x})^T d < 0\}$ and $\mathcal{I}_0 := \{\ell : \nabla f_\ell(\bar{x})^T d = 0\}$. Now, consider the set $\mathcal{K} := \{k \in \mathbb{N} : f_\ell(x^k) > f_\ell(\bar{x}) \text{ for some } \ell \in \mathcal{I}_0\}$. The set $\mathcal{K}$ is infinite; otherwise, we get a contradiction with the fact that $\bar{x}$ is a local strong Pareto solution. Moreover, taking a further subsequence if necessary, there exists $\ell_0 \in \mathcal{I}_0$ such that $f_{\ell_0}(x^k) > f_{\ell_0}(\bar{x}), \forall k \in \mathcal{K}$. Take $\ell_1 \in \mathcal{I}_-$. By using the Taylor expansion, we get

$$
\begin{aligned}
f_{\ell_1}(x^k) &= f_{\ell_1}(\bar{x}) + t_k \nabla f_{\ell_1}(\bar{x})^T d^k + o(\|t_k d^k\|), \\
f_{\ell_0}(x^k) &= f_{\ell_0}(\bar{x}) + t_k \nabla f_{\ell_0}(\bar{x})^T d^k + o(\|t_k d^k\|).
\end{aligned}
$$

From $f_{\ell_0}(x^k) > f_{\ell_0}(\bar{x}), \forall k \in \mathcal{K}$, we see that $\nabla f_{\ell_0}(\bar{x})^T d^k + o(\|t_k d^k\|) t_k^{-1} > 0$ and converges to $\nabla f_{\ell_0}(\bar{x})^T d = 0$ from the right. Thus,

$$
\frac{f_{\ell_1}(x^k) - f_{\ell_1}(\bar{x})}{f_{\ell_0}(\bar{x}) - f_{\ell_0}(x^k)} = \frac{-\nabla f_{\ell_1}(\bar{x})^T d^k - \frac{o(\|t_k d^k\|)}{t_k}}{\nabla f_{\ell_0}(\bar{x})^T d^k + \frac{o(\|t_k d^k\|)}{t_k}} \to \infty \text{ as } k \to \infty,
$$

which contradicts the definition of a Geoffrion-properly efficient point. □

4.3 Computation of the Multiobjective Normal Cone Under the Error Bound Property

Here, we compute the multiobjective normal cone under the error bound property, see Definition 4.2. The error bound property has a large range of applications, which include convergence analysis of iterative algorithms, stability of inequality systems, to mention a few of them. For more information, see [41,42] and references therein.

Definition 4.2 We say that the *Error Bound Property* (EBP) holds for Ω at the feasible point $\bar{x}$, if there is a neighborhood $\mathcal{U}$ of $\bar{x}$ such that

$$
\text{dist}(x, \Omega) \le L \max\{\|h(x)\|, \max\{0, g(x)\}\}, \quad \text{for every } x \in \mathcal{U}.
$$

Under EBP, it is possible to get a nice characterization of the multiobjective normal cone. Indeed, we have that

Theorem 4.3 *Suppose that EBP holds at $\bar{x}$. Then, there is a neighborhood $\mathcal{U}$ of $\bar{x}$ such that for every feasible point $x \in \mathcal{U}$, we have that:*

If $V = (v_\ell)_{\ell=1}^r \in N(x, \Omega; r)$, then there exist $\theta \in \mathbb{R}_+^r$ with $\|\theta\|_1 = 1$ and multipliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}_+^p$ with $\mu_j g_j(x) = 0$, $\forall j$ such that

$$V\theta = \nabla h(x)^T \lambda + \nabla g(x)^T \mu, \text{ and } \|(\lambda, \mu)\| \leq L \|V\theta\|.$$

Proof First, suppose that $V \in \widehat{N}(x, \Omega; r)$. By Theorem 3.1, there exists a C^1 mapping f having x as a local weak Pareto solution. Thus, $v_\ell = -\nabla f_\ell(x)$, $\forall \ell$. By Remark 2.1, there exists $\delta > 0$, such that x is the unique solution of

$$\text{Minimize } \max_{\ell=1,\dots,r} (f_\ell(z) - f_\ell(x)) + \frac{1}{2}\|z - x\|^2 \text{ s.t. } \|z - x\| \leq \delta, \ z \in \Omega. \quad (26)$$

If, in (26), we apply Fermat's rule at x by using Clarke's subdifferentials, we obtain an optimality condition based on Clarke's normal cone, which is usually larger than the limiting normal cone. To avoid the use of Clarke's subdifferentials, we proceed as follows:

For each $\eta > 0$, consider the smoothing approximation of the max function, $g_{\eta,x}(z)$, defined as

$$g_{\eta,x}(z) := \eta \ln \left\{ \sum_{\ell=1}^r \exp\left(\frac{f_\ell(z) - f_\ell(x)}{\eta}\right) \right\} - \eta \ln r, \text{ for every } z \in \mathbb{R}^n.$$

Note that for each z, $g_{\mu,x}(z) \to \max_{\ell=1,\dots,r}(f_\ell(z) - f_\ell(x))$ as $\mu \to 0$, and

$$\nabla g_{\eta,x}(z) = \sum_{\ell=1}^r \theta_\ell(z) \nabla f_\ell(z), \text{ with } \theta_\ell(z) := \frac{\exp((f_\ell(z) - f_\ell(x))/\eta)}{\sum_{\ell=1}^r \exp((f_\ell(z) - f_\ell(x))/\eta)}.$$

Note that $\|\theta_\ell(z)\|_1 = 1$. Furthermore, we see that

$$\max_{\ell=1,\dots,r}(f_\ell(z) - f_\ell(x)) \leq g_{\eta,x}(z) + \eta \ln r \leq \max_{\ell=1,\dots,r}(f_\ell(z) - f_\ell(x)) + \eta \ln r,$$
$$\forall z \in \mathbb{R}^n. \quad (27)$$

For applications of the smoothing approximation technique in optimization, we mention [43–45] and reference therein.

Now, let $\{\eta^k\}$ be a sequence of positive parameters converging to 0 and denote by $x^k \in \Omega$ the global minimizer of following smooth minimization problem

$$\text{Minimize } g_{\eta^k,x}(z) + \frac{1}{2}\|z - x\|^2 \text{ subject to } \|z - x\| \leq \delta, \ z \in \Omega. \quad (28)$$

We continue by showing that $x^k \to x$.

In fact, using the optimality of x^k and (27), we get that

$$g_{\eta^k,x}(x^k) + \frac{1}{2}\|x^k - x\|^2 \leq g_{\eta^k,x}(x) \leq \max_{\ell=1,\dots,r}(f_\ell(x) - f_\ell(x)) = 0. \quad (29)$$

But, from (27), (29) and since x is a weak Pareto point for minimizing f, we see that

$$\frac{1}{2}\|x^k - x\|^2 - \eta^k \ln r \leq \max_{\ell=1,\dots,r} (f_\ell(x^k) - f_\ell(x)) + \frac{1}{2}\|x^k - x\|^2 - \eta^k \ln r$$

$$\leq g_{\eta^k,x}(x^k) + \frac{1}{2}\|x^k - x\|^2 \leq 0.$$

Taking limit in the last expression and since $\eta^k \to 0$, we get that $x^k \to x$.

Consider k large enough such that $\|x^k - x\| < \delta$. By optimality of x^k and applying Fermat's rule for limiting normal cones, we obtain that

$$-\nabla g_{\eta^k,x}(x^k) - (x^k - x) = -\sum_{\ell=1}^{r} \theta_\ell(x^k)\nabla f_\ell(x^k) - (x^k - x) \in N(x^k, \Omega).$$

$$(30)$$

Thus, by [46, Theorem 3] applied to $N(x^k, \Omega)$, there are multipliers $\lambda^k \in \mathbb{R}^m$, $\mu^k \in \mathbb{R}_+^p$ with $\mu_j^k g_j(x^k) = 0$, $\forall j$ such that

$$V^k\theta^k - (x^k - x) = \nabla h(x^k)^T\lambda^k + \nabla g(x^k)^T\mu^k, \tag{31}$$

$$\|(\lambda^k, \mu^k)\| \leq L\|V^k\theta^k - (x^k - x)\| \text{ and } \|\theta^k\|_1 = 1, \tag{32}$$

where $V^k := (-\nabla f_\ell(x^k))_{\ell=1}^r$. Clearly, $V^k \to V$. From (31), there is no loss of generality (after possibly taking a further subsequence), if we assume that λ^k and θ^k converge to λ and θ respectively. Note that $\|\theta\|_1 = 1$ and $\mu_j g_j(x) = 0$, $\forall j$. Thus, taking limits in (31), (32), we get the result when V is in $\widehat{N}(x, \Omega; r)$.

For the general case, take $V \in N(x, \Omega; r)$. Thus, there are sequences $\{V^k\} \subset \widehat{N}(x^k, \Omega; r)$, $\{x^k\} \subset \Omega$ such that $V^k \to V$, and $x^k \to x$. Using the above results for each V^k and after taking a further subsequence, we obtain the desired result for V. □

Corollary 4.1 *If EBP holds at $\bar{x}$, then $N(\bar{x}, \Omega; r) = \widehat{N}(\bar{x}, \Omega; r)$ and*

$$N(\bar{x}, \Omega; r) = \left\{ V \in \mathbb{R}^{n \times r} : \begin{array}{l} V\theta = \nabla h(x)^T\lambda + \nabla g(x)^T\mu, \text{ with} \\ \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}_+^p \text{ and } 0 \neq \theta \in \mathbb{R}_+^r \end{array} \right\}. \tag{33}$$

Proof By Theorem 4.3, $N(\bar{x}, \Omega; r)$ is included in the right-hand side set of (33). Now, take $V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r}$ such that $V\theta = \nabla h(x)^T\lambda + \nabla g(x)^T\mu$, with $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}_+^p$ and $0 \neq \theta \in \mathbb{R}_+^r$. Without loss of generality, we assume that $\|\theta\|_1 = 1$. Consider $d \in T(\bar{x}, \Omega)$, which implies that $\langle \nabla h_i(\bar{x}), d \rangle = 0$, $i = 1, \dots, m$, $\langle \nabla g_j(x), d \rangle \leq 0$, $j \in A(\bar{x})$. Thus,

$$\min_{\ell=1,\dots,r} \langle v_\ell, d \rangle \leq \sum_{\ell=1}^{r} \theta_\ell \langle v_\ell, d \rangle = \langle \sum_{\ell=1}^{r} \theta_\ell v_\ell, d \rangle = \langle V\theta, d \rangle = \langle \nabla h(x)^T\lambda + \nabla g(x)^T\mu, d \rangle$$

$$= \sum_{i=1}^{m} \lambda_i \langle \nabla h_i(\bar{x}), d \rangle + \sum_{j=1}^{p} \mu_j \langle \nabla g_j(\bar{x}), d \rangle \leq 0.$$

Thus, the last expression shows that item (c) of Theorem 3.1 is satisfied and hence $V \in \widehat{N}(\bar{x}, \Omega; r) \subset N(\bar{x}, \Omega; r)$, which complete the proof. $\square$

5 Conclusions

Constraint qualifications have an important role in nonlinear programming (NLP). These conditions are not only useful for ensuring the validity of the KKT conditions, as they are employed also in stability analysis and convergence of methods for solving NLPs, but their role for MOPs is not so clear, since different notions of solutions and KKT conditions are presented. In this work, we addressed the role of CQs for MOPs. Our main contribution is the use of the concept of multiobjective normal cone to obtain new CQs for MOPs. With this tool, we were able to characterize the weakest CQ possible to guarantee the fulfillment of weak/strong KKT conditions at local weak Pareto minimizers. To analyze the strength of such CQs, we compared them with other CQs stated in the literature. Furthermore, we provided a characterization of the multiobjective normal cone under the error bound property.

As further subject of investigation, we point out the relation of the multiobjective normal cone with the sequential optimality conditions stated in [47], since sequential optimality conditions for NLPs have been widely used for improving and unifying the global convergence analysis of several algorithms. Another interesting sequel to this study is to extend these ideas for multiobjective optimization problems with complementarity constraints, considering similar developments for single-objective problems [48–51].

References

1. Deb, K., Datta, R.: Hybrid evolutionary multiobjective optimization and analysis of machining operations. Eng. Optim. **44**(6), 685–706 (2012)
2. Deb, K.: Multi-objective Optimization Using Evolutionary Algorithms. Wiley, New York (2001)
3. Jahn, J.: Vector Optimization: Theory, Applications, and Extensions. Springer, Berlin (2011)
4. Miettinen, K.: Nonlinear Multiobjective Optimization. Springer, Berlin (1999)
5. Ehrgott, M.: Multicriteria Optimization. Springer, Berlin (2005)
6. Gopfert, A., Tammer, C., Riahi, H., Zalinescu, C.: Variational Methods in Partially Ordered Spaces. CMS Books in Mathematics. Springer, Berlin (2003)
7. Fliege, J., Drummond, L.M.G., Svaiter, B.F.: Newton's method for multiobjective optimization. SIAM J. Optim. **20**, 602–626 (2009)
8. Fliege, J., Vaz, A.I.F.: A method for constrained multiobjective optimization based on SQP techniques. SIAM J. Optim. **26**(4), 2091–2119 (2016)
9. Carrizo, G.A., Lotito, P.A., Maciel, M.C.: Trust region globalization strategy for the nonconvex unconstrained multiobjective optimization problem. Math. Progr. **159**, 339–369 (2016)
10. Qu, S., Goh, M., Liang, B.: Trust region methods for solving multiobjective optimization. Optim. Methods Softw. **28**(4), 796–811 (2013)
11. Gass, S., Saaty, T.: The computational algorithm for the parametric objective function. Nav. Res. Logist. Q. **2**, 39–45 (1955)
12. Fliege, J., Vaz, A.I.F., Vicente, L.N.: A new scalarization and numerical method for constructing weak Pareto front of multi-objective optimization problems. Optimization **60**(8), 1091–1104 (2011)
13. Chankong, V., Haimes, Y.Y.: Multiobjective Decision Making: Theory and Methodology. North-Holland, Amsterdam (1983)

14. Kesarwani, P., Dutta, J.: Charnes-Cooper scalarization and convex vector optimization. Optim. Lett. (2019). https://doi.org/10.1007/s11590-019-01502-0

15. Burachik, R.S., Kaya, C.Y., Rizvi, M.M.: A new scalarization technique and new algorithms to generate Pareto fronts. SIAM J. Optim. **27**(2), 1010–1034 (2017)

16. Gerstewitz, C.: Nichtkonvexe Dualität in der Vektoroptimierung. Wiss. Zeitschr. Tech. Hochsch. Leuna-Merseburg **25**, 357–364 (1983)

17. Pascoletti, A., Serafini, P.: Scalarizing vector optimization problems. J. Optim. Theory Appl. **42**, 499–524 (1984)

18. Maciel, M.C., Santos, S.A., Sottosantos, G.N.: Regularity conditions in differentiable vector optimization revisited. J. Optim. Theory Appl. **142**, 385–398 (2009)

19. Maeda, T.: Constraint qualifications in multiobjective optimization problems: differentiable case. J. Optim. Theory Appl. **80**, 483–500 (1994)

20. Chandra, S., Dutta, J., Lalitha, C.S.: Regularity conditions and optimality in vector optimization. Numer. Funct. Anal. Optim. **25**, 479–501 (2004)

21. Bigi, G., Pappalardo, M.: Regularity conditions in vector optimization. J. Optim. Theory Appl. **102**, 83–96 (1999)

22. Rockafellar, R.T., Wets, R.: Variational Analysis. Series: Grundlehren der mathematischen Wissenschaften, vol. 317. Springer, Berlin (2009)

23. Murdokhovich, B.S.: Variational Analysis and Generalized Differentiation I: Basic Theory. Series: Grundlehren der mathematischen Wissenschaften, vol. 330. Springer, Berlin (2005)

24. Borwein, J.M., Zhu, Q.J.: Techniques of Variational Analysis. CMS Books in Mathematics. Springer, New York (2005)

25. Bot, R.I., Grad, S.M., Wanka, G.: Duality in Vector Optimization. Springer, Berlin (2009)

26. Petschke, M.: On a theorem of Arrow, Barankin, and Blackwell. SIAM J. Control Optim. **28**(2), 395–401 (1990)

27. Bigi, G.: Optimality and Lagrangian Regularity in Vector Optimzation. Ph.D. Thesis, University of Pisa, Pisa (1999)

28. Mangasarian, O.L.: Nonlinear Programming. SIAM, Philadelphia (1994)

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

30. Andreani, R., Martínez, J., Ramos, A., Silva, P.: A cone-continuity constraint qualification and algorithmic consequences. SIAM J. Optim. **26**(1), 96–110 (2016)

31. Andreani, R., Martínez, J., Ramos, A., Silva, P.: Strict constraint qualifications and sequential optimality conditions for constrained optimization. Math. Oper. Res. **C2**, 693–1050 (2018)

32. Corley, H.W.: On optimality conditions for maximizations with respect to cones. J. Optim. Theory Appl. **46**(1), 67–78 (1985)

33. Ben-Israel, A.: Motzkin's transposition theorem, and the related theorems of Farkas, Gordan and Stiemke. In: Encyclopedia of Mathematics, Supplement III. Kluwer Academic Publishers, Dordrecht (2001)

34. Singh, C.: Optimality conditions in multiobjective differentiable programming. J. Optim. Theory Appl. **53**(1), 115–123 (1987)

35. Aghezzaf, B., Hachimi, M.: On a gap between multiobjective optimization and scalar optimization. J. Optim. Theory Appl. **109**, 431–435 (2001)

36. Castellani, M., Pappalardo, M.: About a gap between multiobjective optimization and scalar optimization. J. Optim. Theory Appl. **109**, 437–439 (2001)

37. Wang, S.Y., Yang, F.M.: A gap between multiobjective optimization and scalar optimization. J. Optim. Theory Appl. **68**, 389–391 (1991)

38. Giorgi, G.: A note of the Guignard constraint qualification and the Guignard regularity condition in vector optimization. Appl. Math. **4**(4), 734–740 (2012)

39. Andreani, R., Haeser, G., Ramos, A., Silva, P.: A second-order sequential optimality condition associated to the convergence of optimization algorithms. IMA J. Numer. Anal. **47**, 53–63 (2017)

40. Geoffrion, A.M.: Proper efficiency and the theory of vector optimization. J. Math. Anal. **22**, 618–630 (1968)

41. Pang, J.S.: Error bounds in mathematical programming. Math. Program. **79**, 299–332 (1997)

42. Bolte, J., Nguyen, T.P., Peypouquet, J., Suter, B.W.: From error bounds to the complexity of first-order descent methods for convex functions. Math. Program. **165**(2), 471–507 (2017)

43. Chen, X.: Smoothing methods for nonsmooth, nonconvex minimization. Math. Program. **134**, 71–99 (2012)
44. Nesterov, Y.: Smoothing technique and its applications in semidefinite optimization. Math. Program. **110**, 245–259 (2007)
45. Beck, A., Teboulle, M.: Smoothing and first order methods: a unified framework. SIAM J. Optim. **22**(2), 557–580 (2012)
46. Gfrerer, H., Ye, J.J.: New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis. SIAM J. Optim. **27**, 842–865 (2017)
47. Giorgi, G., Jimenez, B., Novo, V.: Approximate Karush Kuhn Tucker condition in multiobjective optimization. J. Optim. Theory Appl. **171**, 70–89 (2016)
48. Zhang, P., Zhang, J., Lin, G.H., Ying, X.: Constraint qualifications and proper pareto optimality conditions for multiobjective problems with equilibrium constraints. J. Optim. Theory Appl. **176**, 763–782 (2018)
49. Andreani, R., Haeser, G., Secchin, L.D., Silva, P.J.S.: New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences. SIAM J. Optim. **29**(4), 3201–3230 (2019)
50. Ramos, A.: Mathematical programs with equilibrium constraints: a sequential optimality condition, new constraint qualifications and algorithmic consequences. Optim. Method. Softw. (2019). https://doi.org/10.1080/10556788.2019.1702661
51. Ramos, A.: Two new weak constraint qualifications for mathematical programs with equilibrium constraints and applications. J. Optim. Theory Appl. **183**(2), 566–591 (2019)