On Stationarity Conditions and Constraint Qualifications for Multiobjective Optimization Problems with Cardinality Constraints

R. Garmanjani * E. H. M. Krulikovski † A. Ramos ‡

July 2022

Abstract

The purpose of this paper is the development of Pareto optimality conditions and constraint qualifications (CQs) for Multiobjective Programs with Cardinality Constraints (MOPCaC). In general, such problems are difficult to deal with, not only because they involve a cardinality constraint that is neither continuous nor convex, but also because there is conflict between the various objective functions. Thus, we reformulate the MOPCaC based on the problem with continuous variables, namely relaxed problem. Furthermore, we consider different notions of optimality (weak/strong Pareto optimal solutions). Thereby, we define new stationarity conditions that extend the classical Karush-Kuhn-Tucker (KKT) conditions of the scalar case. Moreover, we also introduce new CQs, based on a definition of multiobjective normal cone, to ensure compliance with such stationarity conditions. Important statements are illustrated by examples.

Keywords: Multiobjective optimization, Cardinality constraints, Nonlinear programming, Sparse solutions, Constraint qualifications, Stationarity.

AMS Subject Classification: 90C30, 90C33, 90C46, 90C29

1 Introduction

In this paper, we study optimization problems that can be formulated as multiobjective optimization problems (MOPs) with the requirement that the desired solutions have a small or a bounded number of nonzero components, namely sparse solutions. One way to obtain sparse solutions is imposing explicitly a cardinality constraint to the problem, as the pioneering work [5]. This approach has successfully used in many applications of optimization such as sampling signals, machine learning, subset selections and portfolio problems [3, 7, 12, 16, 22, 23, 24, 26, 27], but restricted to one single objective function, the scalar case. So, we will focus on the development of optimality conditions and constraints qualifications for MOPs with cardinality constraints (MOPCaC).

*Department of Mathematics, FCT-UNL-CMA, Campus de Caparica, 2829-516, Portugal (r.garmanjani@fct.unl.pt)

†Department of Mathematics, FCT-UNL-CMA, Campus de Caparica, 2829-516, Portugal (e.krulikovski@fct.unl.pt).

‡Alberto Ramos, Departamento de Matemática, Universidad de Tarapacá, casilla 7D, Arica, Chile (aramosf@academicos.uta.cl).

To deal with the potentially problematic cardinality constraint in MOPCaC, we reformulated it as a MOP with continuous variables, following the ideas of [10] for the scalar case. Compared with scalar optimization problem with cardinality constraints, MOPCaC offers new difficulties. The possible conflicting nature of the many objective functions, forces us to consider different notions of optimality (weak/strong Pareto optimal solutions) and as a consequence new stationarity conditions which extend the classical Karush-Kuhn-Tucker (KKT) conditions of the scalar case. Furthermore, since standard constraint qualifications (CQs) may not be sufficient to ensure the fulfillment of the new stationarity conditions at the optimal solutions, we also introduce new CQs useful for our purpose.

The paper is organized as follows: In Section 2, we present the reformulation of MOPCaC and discuss in detail the relation between Pareto optimal solutions. In addition, we recall some background material of variational analysis concerning MOPs. In Section 3, we propose new stationarity conditions for MOPCaC, which extend the classical stationarity conditions for optimization problems with cardinality constraints (CaC-M-stationarity and CaC-S-stationarity [11, Definition 4.1]) to MOPs, and comment their main properties. We devote Section 4 to the study of new CQs to obtain weak CaC-S-stationarity condition (see Definition 3.1) at local weak/strong Pareto optimal solution. Different to other approaches, our CQs rely on the r-multiobjective normal cone defined in Section 3. We end Section 4, by showing that if the constraints h and g in (1) are linear mappings, then the weak CaC-S-stationarity conditions holds at every local weak Pareto optimal solution. In Section 5, we propose a unified framework of the stationarity conditions for MOPCaC, which encodes different levels of stationarity depending on certain set of indices of I. In addition, we present in this section some applications. Finally, in Section 6, we close with some remarks.

Notation. We denote by $\mathbb{R}^n$ the n-dimensional Euclidean space. Throughout this paper, $\|\cdot\|_0$ is the so-called l_0-norm, that is, $\|x\|_0$ is the number of nonzero components of $x \in \mathbb{R}^n$. The Hadamard product $x * y$ of two vectors $x, y \in \mathbb{R}^n$ is the vector obtained by the componentwise product of x and y. We denote the set of non-negative number with $\mathbb{R}_+$. We use $(v_i)_{i=1}^r$ to denote a $n \times r$ matrix whose columns are the vectors $v_1, v_2, \ldots, v_r$ in $\mathbb{R}^n$.

We also consider the following index sets: $I_{00}(x, y) = \{i \mid x_i = 0, y_i = 0\}$, $I_{\pm 0}(x, y) = \{i \mid x_i \neq 0, y_i = 0\}$, $I_{01}(x, y) = \{i \mid x_i = 0, y_i = 1\}$, $I_{0\pm}(x, y) = \{i \mid x_i = 0, y_i \in (0, 1]\}$, $I_0(x) = \{i \mid x_i = 0\}$ and $I_{\pm}(x) = \{i \mid x_i \neq 0\}$.

For a function $g : \mathbb{R}^n \to \mathbb{R}^r$, denote $I_g(x) = \{i \mid g_i(x) = 0\}$, the set of active indices, and $\nabla g^T = (\nabla g_1, \ldots, \nabla g_r)$ the transpose of the Jacobian of g.

2 Problem statement

Consider Multiobjective Programs with Cardinality Constraints (MOPCaC) of the form:

$$
\begin{array}{ll}
\underset{x}{\text{minimize}} & F(x) \equiv (f_1(x), f_2(x), \ldots, f_r(x))^T \\
\text{subject to} & g(x) \leq 0, h(x) = 0, \\
& \|x\|_0 \leq \alpha,
\end{array}
\tag{1}
$$

where $F : \mathbb{R}^n \to \mathbb{R}^r (r \geq 1)$, $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ are continuously differentiable mappings, $0 < \alpha < n$ is a given non-negative integer (which we assume $\alpha < n$, otherwise the cardinality constraint would be redundant), and $\|x\|_0$ stands for the number of all nonzero components of x.

Due to the wide range of applications, scalar optimization problems with cardinality constraints have attracted the attention of many researchers, see [3, 8, 10, 11, 14, 18]. The cardinality constraint, $\|x\|_0 \leq \alpha$, makes (1) a very difficult optimization problem. To overcome this difficulty we follow the approach of [14] which introduces a new variable y and obtain a relaxed multiobjective optimization problem:

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & F(x) \equiv (f_1(x), f_2(x), \ldots, f_r(x))^T \\ \text{subject to} & g(x) \leq 0, h(x) = 0, \\ & n - e^T y \leq \alpha, \\ & 0 \leq y \leq e, \\ & x * y = 0, \end{array} \tag{2}$$

where $e = (1, \ldots, 1)^T \in \mathbb{R}^n$. This approach was successfully used to study optimality conditions and CQs in scalar optimization problems but, to the best of our knowledge, it has not been studied within the context of MOPs. Many issues arise when we consider MOPs. Then, in view of the possibility of conflict among the objective functions f_i's, several notions of optimality for MOPs are considered in such context.

Definition 2.1 *Consider the problem of minimizing $F(x) = (f_1(x), \ldots, f_r(x))^T$ subject to $x \in \mathcal{X}$ and let $\bar{x}$ be a feasible point in $\mathcal{X}$. Then,*

1. *We say that $\bar{x}$ is a local weak Pareto optimal (or local weak efficient) solution if there is a $\delta > 0$ such that there is no $x \in \mathcal{X} \cap B(\bar{x}, \delta)$ with $f_i(x) < f_i(\bar{x})$ for all $i = 1, \ldots, r$.*

2. *We say that $\bar{x}$ is a local strong Pareto optimal (or local efficient) solution if there is a $\delta > 0$ such that there is no $x \in \mathcal{X} \cap B(\bar{x}, \delta)$ with $f_i(x) \leq f_i(\bar{x})$, for all $i = 1, \ldots, r$ and $F(x) \neq F(\bar{x})$.*

3. *We say that $\bar{x}$ is a weak Pareto optimal (or weak efficient) solution if there is no other feasible point $x \in \mathcal{X}$ such that $f_i(x) < f_i(\bar{x})$, for all $i = 1, \ldots, r$.*

4. *We say that $\bar{x}$ is a strong Pareto optimal (or efficient) solution if there is no other feasible point $x \in \mathcal{X}$ such that $F(x) \leq F(\bar{x})$ and $F(x) \neq F(\bar{x})$.*

As in the scalar case, (1) and its relaxed formulation (2) share many characteristics. Indeed, Theorem 2.1 states that the MOPCaC problem has a weak/strong Pareto optimal solution if and only if the relaxed problem has it too. To prove this, we need two preliminaries lemmas:

Lemma 2.1 *[20, Lemma 3] Let $(\bar{x}, \bar{y})$ be a feasible point of the relaxed problem (2) such that $\|\bar{x}\|_0 = \alpha$. Then, $e^T \bar{y} = n - \alpha$, $\bar{y}_i = 0$ for $i \notin I_0(\bar{x})$, $\bar{y}_i = 1$ for $i \in I_0(\bar{x})$ and $I_{00}(\bar{x}, \bar{y}) = \emptyset$.*

Lemma 2.2 *[20, Lemma 4] Let $\bar{x} \in \mathbb{R}^n$ be a feasible point of (1), then there exists $\bar{y} \in \mathbb{R}^n$ such that $(\bar{x}, \bar{y})$ is feasible of (2). If, in addition, $\|\bar{x}\|_0 = \alpha$, then $\bar{y}$ is unique; Conversely, if $(\bar{x}, \bar{y})$ is feasible for (2), then $\bar{x}$ is feasible for (1).*

Theorem 2.1 *Consider a point $x^* \in \mathbb{R}^n$.*

1. *If x^* is a weak/strong Pareto optimal solution of the problem (1), there exists $y^* \in \mathbb{R}^n$ such that (x^*, y^*) is a weak/strong Pareto optimal solution of (2). Moreover, for the relaxed problem, every feasible pair of the form $(x^*, \bar{y})$ is a weak/strong Pareto optimal solution;*

2. If (x^*, y^*) is a weak/strong Pareto optimal solution of (2), then x^* is a weak/strong Pareto optimal solution of (1).

Proof.

1. Let x^* be a strong Pareto optimal solution of (1). By Lemma 2.2 there exists $y^* \in \mathbb{R}^n$ such that (x^*, y^*) is feasible for (2). Take a feasible point (x, y) for (2). By Lemma 2.2, x is feasible for (1). So, as there is no other feasible point x for (1) such that $f_i(x) \leq f_i(x^*)$, $\forall i$ with $F(x) \neq F(x^*)$, then (x^*, y^*) is a strong Pareto optimal solution of (2). Since this argument does not depend on y, we have the second statement.

2. Let (x^*, y^*) be a strong Pareto optimal solution of (2), by Lemma 2.2, x^* is feasible for (1). Now, given a feasible point x of (1), there exists $y \in \mathbb{R}^n$ such that (x, y) is feasible for (2). So, as there is no other feasible point (x, y) for (2) such that $f_i(x) \leq f_i(x^*)$, $\forall i$ with $F(x) \neq F(x^*)$, then x^* is a strong Pareto optimal solution of (1).

The proof is analogous for weak Pareto optimal solution. ■

Theorem 2.1 shows the equivalence between weak/strong Pareto optimal solutions of (1) and (2), but the situation is different for local solutions.

Theorem 2.2 *Let $x^* \in \mathbb{R}^n$ be a local weak/strong Pareto optimal solution of (1). Then there exists a vector $y^* \in \mathbb{R}^n$ such that (x^*, y^*) is a local weak/strong Pareto optimal solution of (2).*

Proof. Suppose that x^* is local strong Pareto optimal solution of (1). Define y^* by $y_i^* = 1$ if $i \in I_0(x^*)$ and $y_i^* = 0$ otherwise. Clearly, $y_i^* = 1$ if and only if $x_i^* = 0$ and hence $e^T y^* = n - \|x^*\|_0 \geq n - \alpha$. It is easy to see that (x^*, y^*) is feasible for (2). We claim that (x^*, y^*) is a local strong Pareto optimal solution of (2). Indeed, by definition, there exists a $\delta_1 > 0$ such that there is no a feasible point $x \in B(x^*, \delta_1)$ with $f_i(x) < f_i(x^*)$, $\forall i$ and $\|x\|_0 \leq \alpha$, due to the assumed local optimality of x^* for (1). Now, choose $\delta_2 = 1/2$. Then we have $y_i > 0$ for all $y \in B(y^*, \delta_2)$ and all i such that $y_i^* > 0$. So,

$$\{i \mid y_i = 0\} \subseteq \{i \mid y_i^* = 0\} \text{ for all } y \in B(y^*, \delta_2). \tag{3}$$

Now, take $\delta = \min\{\delta_1, \delta_2\}$. Let $(x, y) \in B(x^*, \delta) \times B(y^*, \delta)$ be a feasible point of (2). Then, x satisfies $g(x) \leq 0$ and $h(x) = 0$. Moreover, (3) implies $x_i \neq 0 \Rightarrow y_i = 0 \Rightarrow y_i^* = 0 \Rightarrow x_i^* \neq 0$ and therefore $\|x\|_0 \leq \|x^*\|_0$. Hence, x is feasible for (1). Since $x \in B(x^*, \delta_1)$, we obtain from the local weak Pareto optimality of x^* for (1) that there is no $x \in \Omega \cap B(x^*, \delta_1)$ with $f_i(x) < f_i(x^*)$, $\forall i$. Consequently, (x^*, y^*) is a local strong Pareto optimal solution of the relaxed problem (2). A similar proof when x^* is a local weak Pareto optimal solution of (1). ■ The next example shows that the converse of the theorem is not valid.

Example 2.1 *Consider the MOPCaC and the corresponding relaxed problem*

$$\begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & F(x) \equiv \big((x_3 - 1)^2, -x_3\big)^T \\ \text{subject to} & x_1 \leq 0, \\ & \|x\|_0 \leq 2, \end{array} \qquad \begin{array}{ll} \underset{x, y \in \mathbb{R}^3}{\text{minimize}} & F(x) \equiv \big((x_3 - 1)^2, -x_3\big)^T \\ \text{subject to} & x_1 \leq 0, \\ & y_1 + y_2 + y_3 \geq 1, \\ & x_i y_i = 0, \ i = 1, 2, 3, \\ & 0 \leq y_i \leq 1, \ i = 1, 2, 3. \end{array}$$

Set $x^* = (0,1,0)$. Fix $t \in (0,1)$ and put $y^* := (1-t, 0, t)$. We claim that (x^*, y^*) is a local weak Pareto optimal solution of the relaxed problem. Indeed, for every (x, y) sufficiently close to (x^*, y^*), $y_3 \neq 0$ which implies that $x_3 = 0$. Thus, in this case, $f_1(x) = f_1(x^*) = 1$ and $f_2(x) = f_2(x^*) = 0$, proving the claim. On the other hand, x^* is not a local weak Pareto optimal solution of the MOPCaC. Indeed, for every $x_\delta = (0,1,\delta)$ ($\delta \neq 0$) sufficiently close to $x^* = (0,1,0)$, we have $f_1(x_\delta) = (\delta-1)^2 < f_1(x^*) = 1$ and $f_2(x_\delta) = -\delta < f_2(x^*) = 0$.

Under some assumption, we have the converse of Theorem 2.2.

Theorem 2.3 *Let (x^*, y^*) be a local weak Pareto optimal solution of (2). Then $\|x^*\|_0 = \alpha$ if and only if y^* is unique, that is, there is exactly one y^* such that (x^*, y^*) is a local weak Pareto optimal solution of (2). In this case, the components of y^* are binary and x^* is a local weak efficient solution of (1).*

Proof. The proof follows [11, Proposition 3.5]. The "only if" part and the claim that y^* is a binary vector follow directly from Lemma 2.1. For the "if" part, assume that x^* is a local weak Pareto optimal solution of (1) with $\|x^*\|_0 < \alpha$. Then, $\|x^*\|_0 \leq n-2$, and we can find $k_1 \neq k_2$ such that $x^*_{k_1} = x^*_{k_2} = 0$. Now, define $\bar{y} \in \mathbb{R}^n$ as $\bar{y}_i = 1$ for $i \in I_0(x^*)$ and $\bar{y}_i = 0$ otherwise, and $\tilde{y} \in \mathbb{R}^n$ as

$$\tilde{y}_i = \begin{cases} 1/2 & \text{if } i \in \{k_1, k_2\} \\ 1 & \text{if } x^*_i = 0, i \notin \{k_1, k_2\} \\ 0 & \text{if } i \notin I_0(x^*) \end{cases}.$$

Then, we can prove as in Theorem 2.2 that $(x^*, \bar{y})$ and $(x^*, \tilde{y})$ are both local weak Pareto optimal solution of (2) contradicting the uniqueness of y^*. ■

Summarizing, (1) and its relaxed problem (2) are equivalent in terms of feasible points and weak/strong Pareto optimal solution, whereas for local weak/strong Pareto optimal solutions we require an additional assumption which says that the cardinality constraint must be active.

Now, we recall some standard notations from optimization and variational analysis. Given a cone $\mathcal{K} \subset \mathbb{R}^n$, $\mathcal{K}^\circ := \{v \in \mathbb{R}^n \mid v^T k \leq 0 \ \forall k \in \mathcal{K}\}$ is the polar set of $\mathcal{K}$. For a given set-valued mapping $\mathcal{F} : \mathbb{R}^s \rightrightarrows \mathbb{R}^n$, the *sequential Painlevé-Kuratowski outer/upper limit* of $\mathcal{F}(z)$ as $z \to z^*$ is defined as the set

$$\limsup_{z \to z^*} \mathcal{F}(z) = \{y^* \in \mathbb{R}^n \mid \exists\, (z^k, y^k) \to (z^*, y^*) \text{ with } y^k \in \mathcal{F}(z^k), \forall k \in \mathbb{N}\}.$$

We continue with some definitions and useful results of variational analysis regarding MOPs. Given a nonempty set Ω, the *tangent cone* to Ω at $\bar{x} \in \Omega$ is the set

$$T_\Omega(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \exists (x^k, t_k) \subset \Omega \times \mathbb{R}_+ \text{ with } t_k \to 0 \text{ and } \frac{x^k - \bar{x}}{t_k} \to d \right\}.$$

The (Fréchet) regular normal cone to Ω at $\bar{x} \in \Omega$ is defined as

$$\widehat{N}_\Omega(\bar{x}) = \left\{ w \in \mathbb{R}^n \mid \limsup_{x \to \bar{x},\, x \in \Omega} \frac{w^T(x - \bar{x})}{\|x - \bar{x}\|} \leq 0 \right\},$$

and the (Murdokhovich) limiting normal cone to Ω at $\bar{x} \in \Omega$ is defined by $N_\Omega(\bar{x}) = \limsup_{x \to \bar{x},\, x \in \Omega} \widehat{N}_\Omega(x)$. When $\bar{x} \notin \Omega$, we set $T_\Omega(\bar{x}) = \emptyset$, $\widehat{N}_\Omega(\bar{x}) = \emptyset$ and $N_\Omega(\bar{x}) = \emptyset$. We always have $\widehat{N}_\Omega(\bar{x}) = T_\Omega(\bar{x})^\circ$ and $\widehat{N}_\Omega(\bar{x}) \subset N_\Omega(\bar{x})$. When Ω is convex, $\widehat{N}_\Omega(\bar{x}) = N_\Omega(\bar{x}) = \{\zeta \in \mathbb{R}^n \mid \zeta^T(x - \bar{x}) \leq 0, \forall x \in \Omega\}$.

The normal and tangent cones are important objects in variational analysis, and they are useful for obtaining verifiable necessary optimality conditions. In the scalar case ($r = 1$), they are also used to define CQs, i.e., assumptions on the feasible set which guarantee the fulfillment of the KKT conditions at local minimizers. In the multiobjective case, besides the tangent and normal cone, we follow the approach [17] and we mention the multiobjective normal cone defined in [17], which is a suitable tool for studying CQs for MOPs.

Definition 2.2 *Let Ω be a closed subset of $\mathbb{R}^n$, $\bar{x} \in \Omega$ and $r \in \mathbb{N}$, the regular r-multiobjective normal cone to Ω at $\bar{x}$ is the cone defined as*

$$\widehat{N}_\Omega(\bar{x}; r) = \left\{ V = (v_i)_{i=1}^r \in \mathbb{R}^{n \times r} \mid \limsup_{x \to \bar{x},\, x \in \Omega} \min_{i=1,\ldots,r} \frac{v_i^T(x - \bar{x})}{\|x - \bar{x}\|} \leq 0 \right\}, \tag{4}$$

and the limiting r-multiobjective normal cone to Ω at $\bar{x} \in \Omega$ is defined by

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

When Ω is convex, $\widehat{N}_\Omega(\bar{x}; r) = N_\Omega(\bar{x}; r)$. Furthermore, we have

$$N_\Omega(\bar{x}; r) = \{(v_i)_{i=1}^r \in \mathbb{R}^{n \times r} \mid \min_{\ell=1,\ldots,r} v_\ell^T(x - \bar{x}) \leq 0, \forall x \in \Omega\}.$$

For more details and properties of the r-multiobjective normal cone, see [17].

3 Stationary concepts for MOPCaC

From now on, consider the feasible set of (2) denoted by

$$\Omega = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid \begin{array}{ll} g(x) \leq 0, & h(x) = 0, \\ e^T y \geq n - \alpha, & 0 \leq y \leq e, \\ x * y = 0 & \end{array} \right\}. \tag{5}$$

Multiobjective optimization problems have many stationarity concepts such as weak/strong Karush-Kuhn-Tucker (KKT) conditions which consider non-negative multipliers associated to each objective functions, see [9]. In our optimization problem, in addition to considering multipliers associated to the objective functions, we also need to consider the different type of multipliers associated to the cardinality constraint. This observation leads us to propose the next tailored stationary conditions:

Definition 3.1 *Consider $(\bar{x}, \bar{y})$ be a feasible point of (2). Suppose that there exists a vector $0 \neq (\lambda^f, \lambda^g, \lambda^h, \gamma) \in \mathbb{R}^r_+ \times \mathbb{R}^m_+ \times \mathbb{R}^p \times \mathbb{R}^n$ such that*

$$\nabla F(\bar{x})^T \lambda^f + \nabla g(\bar{x})^T \lambda^g + \nabla h(\bar{x})^T \lambda^h + \gamma = 0. \tag{6}$$

$$\lambda_i^g g_i(\bar{x}) = 0 \ \text{ for all } \ i = 1, \ldots, m. \tag{7}$$

Then, we say that $(\bar{x}, \bar{y})$ is a:

1. weak CaC-M-stationary point, *if (6) and (7) hold with $\lambda^f \neq 0$ and $\gamma_i = 0$ for all $i \in I_\pm(\bar{x})$.*

2. strong CaC-M-stationary point, *if (6) and (7) hold with* $\lambda_i^f > 0$ *for all* $i = 1, \dots, r$ *and* $\gamma_i = 0$ *for all* $i \in I_\pm(\bar{x})$*.*

Moreover, $(\bar{x}, \bar{y})$ *is a*

1. weak CaC-S-stationary point, *if* $(\bar{x}, \bar{y})$ *is a weak CaC-M-stationary point with* $\gamma_i = 0$ *for all* $i \in I_0(\bar{y})$*;*

2. strong CaC-S-stationary point, *if* $(\bar{x}, \bar{y})$ *is a strong CaC-M-stationary point with* $\gamma_i = 0$ *for all* $i \in I_0(\bar{y})$*;*

From the definition above, a weak/strong CaC-M-stationary point does not depend on the point $\bar{y}$, and every weak/strong CaC-S-stationary point is weak/strong CaC-M-stationary. Clearly, if in (6), $r = 1$ such concepts reduce to CaC-M-stationarity and CaC-S-stationarity conditions, discussed in [10]. Following the same arguments in [18], we see that given a weak/strong CaC-M-stationary point $(\bar{x}, \bar{y})$, it is possible to find another variable $\bar{z}$ such that $(\bar{x}, \bar{z})$ is a weak/strong CaC-S-stationary point. See also Proposition 5.2. Furthermore, adapting the arguments in [11, Theorem 4.7] (see also Theorem 5.1), we have:

Theorem 3.1 *A point* $(\bar{x}, \bar{z})$ *is a weak/strong CaC-S-stationary point if and only* $(\bar{x}, \bar{z})$ *is a weak/strong KKT point (see [17, Definition 2.2]) of the relaxed MOPCaC problem* (2)*.*

In the scalar case, since the standard CQs for nonlinear optimization are not sufficient to ensure that local minimizers are CaC-M-stationary and/or CaC-S-stationary points, several tailored CQs have been proposed and which exploit the special structure of the cardinality constraint. To define such CQs we need more definitions. For the feasible Ω defined in (5), we consider the *linearized cone* to Ω at $(\bar{x}, \bar{y})$ defined as

$$D_\Omega(\bar{x}, \bar{y}) = \left\{ d = (u, v) \mid \begin{array}{ll} \nabla g_i(\bar{x})^T u \leq 0 & \text{for all } i \in I_g(\bar{x}), \\ \nabla h_i(\bar{x})^T u = 0 & \text{for all } i = 1, \dots, p, \\ e^T v \geq 0 & \text{if } e^T \bar{y} = n - \alpha, \\ v_i \geq 0 & \text{for all } i \in I_{00}(\bar{x}, \bar{y}), \\ v_i \leq 0 & \text{for all } i \in I_{01}(\bar{x}, \bar{y}), \\ u_i = 0 & \text{for all } i \in I_{0\pm}(\bar{x}, \bar{y}), \\ v_i = 0 & \text{for all } i \in I_{\pm 0}(\bar{x}, \bar{y}). \end{array} \right\}. \tag{8}$$

As in [14], we consider the CaC-*linearized cone* to Ω at $(\bar{x}, \bar{y})$ defined by

$$D_\Omega^{CaC}(\bar{x}, \bar{y}) = D_\Omega(\bar{x}, \bar{y}) \cap \{d = (u, v) \mid u_i v_i = 0 \text{ for all } i \in I_{00}(\bar{x}, \bar{y})\}. \tag{9}$$

The difference between $D_\Omega^{CaC}(\bar{x}, \bar{y})$ and $D_\Omega(\bar{x}, \bar{y})$ rely on in the inclusion of the relations $u_i v_i = 0$ for all $i \in I_{00}(\bar{x}, \bar{y})$. Clearly, $D_\Omega^{CaC}(\bar{x}, \bar{y}) \subseteq D_\Omega(\bar{x}, \bar{y})$. We use the linearized cones to define very well-known CQs for cardinality constrained optimization problems.

Definition 3.2 *Let* $(\bar{x}, \bar{y})$ *be feasible of (2). Then, we say that* $(\bar{x}, \bar{y})$ *satisfies*

1. *Abadie CQ (ACQ) if* $T_\Omega(\bar{x}, \bar{y}) = D_\Omega(\bar{x}, \bar{y})$*.*

2. *Guignard CQ (GCQ) if* $T_\Omega(\bar{x}, \bar{y})^\circ = D_\Omega(\bar{x}, \bar{y})^\circ$*.*

3. *CaC-Abadie CQ (CaC-ACQ) if* $T_\Omega(\bar{x}, \bar{y}) = D_\Omega^{CaC}(\bar{x}, \bar{y})$*.*

4. CaC-Guignard CQ (CaC-GCQ) if $T_\Omega(\bar{x},\bar{y})^\circ = D_\Omega^{CaC}(\bar{x},\bar{y})^\circ$.

By [14, Proposition 3.4 and Theorem 3.7], $T_\Omega(\bar{x},\bar{y}) \subseteq D_\Omega^{CaC}(\bar{x},\bar{y}) \subseteq D_\Omega(\bar{x},\bar{y})$ and $D_\Omega(\bar{x},\bar{y})^\circ = D_\Omega^{CaC}(\bar{x},\bar{y})^\circ$. So, from the definitions presented above, ACQ implies CaC-ACQ and CaC-ACQ implies CaC-GCQ. Moreover, the implications are strict [14] and GCQ is equivalent to CaC-GCQ. We mention that in the scalar case, by [14, Theorem 4.2], we see that CaC-GCQ is sufficient to ensure that every local minimizer $(\bar{x},\bar{y})$ of (2) is a CaC-S-stationary point. Finally, we mention that in the linear case (i.e., when g and h are affine mappings), CaC-ACQ always holds, but ACQ may fail which motivates the study of the CaC-linearized cone, [14]. For more relations between others CaC-CQs stated in the literature, see Fig. 2.

4 New Constraint Qualifications for CaC-S-stationary points

In this section, we will use the r-multiobjective normal cone to define CQs for characterizing different types of stationary concepts. Here, we focus on weak/strong CaC-S-stationarity conditions, we will discuss CQs for weak/strong CaC-M-stationarity conditions in the next section. We mention that our CQs depend only on the feasible set and do not require any information of the objective functions, which differs from the regularity conditions for MOPs, see [6, 9] and the references therein.

4.1 Constraint Qualifications for weak CaC-S-stationary points

Here, we start by giving the weakest CQ for the fulfillment of weak CaC-S-stationarity condition at local weak Pareto optimal solution. In [17, Theorem 4.2] for nonlinear multiobjective optimization problems, the authors proposed the weakest CQ that guarantees that a local weak Pareto optimal solution is a weak KKT point. Due to the equivalence between weak CaC-S-stationary point and the weak KKT conditions for (2), we get the next result:

Theorem 4.1 *Let $(\bar{x},\bar{y})$ be a feasible point of (2). The weakest property, namely MOP-WCQ, under which every local weak Pareto optimal solution is a weak CaC-S-stationary point for every continuously differentiable mapping, $F(x) \equiv (f_1(x),\dots,f_r(x))^T$ is*

$$D_\Omega(\bar{x},\bar{y}) \subset \mathfrak{L}_\Omega(\bar{x},\bar{y};r) \tag{10}$$

where $\mathfrak{L}_\Omega(\bar{x},\bar{y};r) = \{d \in \mathbb{R}^{2n} \mid \min_{i=1,\dots,r} v_i^T d \le 0,$ *for all* $(v_i)_{i=1}^r \in \widehat{N}_\Omega(\bar{x},\bar{y};r)\}$.

As it stated in [17], when $r=1$, (10) is the inclusion $D_\Omega(\bar{x},\bar{y}) \subset T_\Omega(\bar{x},\bar{y})^{\circ\circ}$, which it turns is equivalent to GCQ, $D_\Omega(\bar{x},\bar{y})^\circ = T_\Omega(\bar{x},\bar{y})^\circ$. In MOP, we have a gap in scalar and multiobjective optimization, see [1, 13], which states that ACQ is sufficient for the fulfillment of the weak KKT point at every local weak Pareto optimal solution, but GCQ is not. Since ACQ implies CaC-ACQ and both implies GCQ, it is natural to ask if CaC-ACQ is sufficient from the fulfillment of the weak CaC-S-stationary point at local weak Pareto optimal solution. Another related question is if in (10) we can replace $D_\Omega(\bar{x},\bar{y})$ by $D_\Omega^{CaC}(\bar{x},\bar{y})$ and still has a CQ for the fulfillment of weak CaC-S-stationarity conditions, since in the scalar case, (10) is equivalent to GCQ and hence to CaC-GCQ. The next example provides negative answers to both questions.

Example 4.1 *(CaC-ACQ and $D_\Omega^{CaC}(\bar{x},\bar{y}) \subset \mathfrak{L}_\Omega(\bar{x},\bar{y};r)$ hold, but (10) fails).*

Consider the MOPCaC and the corresponding relaxed problem,

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2}{\text{minimize}} & F(x) \equiv (x_1, x_2)^T \\ \text{subject to} & x_1x_2 = 0, \\ & \|x\|_0 \leq 1, \end{array} \qquad \begin{array}{ll} \underset{x,y\in\mathbb{R}^2}{\text{minimize}} & F(x) \equiv (x_1, x_2)^T \\ \text{subject to} & x_1x_2 = 0, \\ & y_1 + y_2 \geq 1, \\ & x_iy_i = 0,\ i = 1, 2, \\ & 0 \leq y_i \leq 1,\ i = 1, 2. \end{array}$$

Here, $x^* = (0, 0)$ is the Pareto optimal solution of the cardinality problem. Set $y^* = (1, 0)$, in this case, (x^*, y^*) is a Pareto optimal solution of the relaxed problem which it is not a weak KKT point. This, MOP-WCQ fails at (x^*, y^*).

After some calculations, $T_\Omega(\bar{x}, \bar{y}) = D_\Omega^{CaC}(\bar{x}, \bar{y})$, that is, CaC-ACQ holds at $(\bar{x}, \bar{y})$. Furthermore, by [17, Proposition 4.1 (17)], $T_\Omega(\bar{x}, \bar{y}) \subset \mathfrak{L}_\Omega(\bar{x}, \bar{y}; r)$ and since CaC-ACQ holds, we have $D_\Omega^{CaC}(\bar{x}, \bar{y}) = T_\Omega(\bar{x}, \bar{y}) \subset \mathfrak{L}_\Omega(\bar{x}, \bar{y}; r)$.

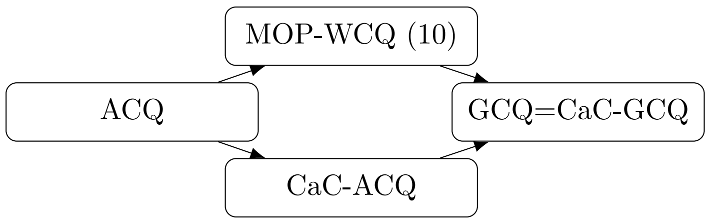

Figure 1: Relationships between the CQs for the fulfillment of MOP-WCQ.

4.2 Constraint Qualifications for weak CaC-S-stationary points: Linear case

Here, we investigate whether the CaC-S-stationarity conditions are necessary optimality conditions for (2) where h and g are affine mappings. In the scalar case, it is known that CaC-ACQ holds at every feasible point in Ω and hence GCQ, but ACQ may fails without further assumptions, see [14, Corollary 3.10]. Due to the gap between scalar and multiobjective optimization, [1, 13], GCQ does not imply (10), and by Example 4.1, CaC-ACQ may not imply (10). Thus, we cannot use the previously results to ensure (10) in the linear case. In order to be able to ensure the fulfillment of (10) when h and g are affine, we proceed as follows: First, we start with Lemma 4.1 that guarantees some kind of outer continuity of the polar sets associated with linearized cones (8). Then, using this lemma, we get the result in Theorem 4.2

We state Lemma 4.1 in a more abstract setting, indeed, we consider a general polyhedral $\mathcal{X}$ which includes Ω as a particular case. Given $a_j, b_j \in \mathbb{R}^n$, $c_j \in \mathbb{R}$, $j \in K := \{1, \ldots, m\}$, set

$$\mathcal{X} := \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid \begin{array}{ll} a_j^T x + b_j^T y \leq c_j, & \forall j \in K = \{1, \ldots, m\} \\ x_iy_i = 0, & \forall i \in J = \{1, \ldots, n\} \end{array} \right\}. \tag{11}$$

Clearly, Ω is a particular case of $\mathcal{X}$ for a proper choice of constraints. The next lemma ensures the outer continuity of the linearized cone mapping $z \rightrightarrows D_\mathcal{X}(z)$, $z \in \mathcal{X}$ at z^*, as z approaches z^* but restricted to $\mathcal{X}$.

Lemma 4.1 *Let $z^* = (x^*, y^*)$ be a feasible point within $\mathcal{X}$. Then*

$$\limsup_{z\to z^*,\ z\in\mathcal{X}} D_\mathcal{X}(z)^\circ \subset D_\mathcal{X}(z^*)^\circ.$$

Proof. Let $z^* = (x^*, y^*)$ be a point in $\mathcal{X}$ and take $(v, w) \in \limsup_{z \to z^*,\, z \in \mathcal{X}} D_{\mathcal{X}}(z)^\circ$. By definition of outer limits, there exist sequences $z^k := (x^k, y^k) \in \mathcal{X}$ and $v^k \in D_{\mathcal{X}}(z^k)^\circ$ such that $z^k \to z^*$ and $v^k \to v$. Since $(v^k, w^k) \in D_{\mathcal{X}}(z^k)^\circ$ and using the analytical form of $\mathcal{X}$, we find multipliers $\{\lambda_i^k\} \subset \mathbb{R}_+^{|K|}$ and μ_j^k such that

$$(v^k, w^k)^T = \sum_{j \in \mathcal{K}^k} \lambda_j^k (a_j, b_j)^T + \sum_{i \in J} \mu_i^k (y_i^k e_i, x_i^k e_i)^T, \tag{12}$$

with e_i the ith canonical vector in $\mathbb{R}^n$ and $\mathcal{K}^k := \{j \mid a_j^T x^k + b_j^T y^k = c_j\}$. After taking an adequate subsequence, we assume that $\mathcal{K} := \mathcal{K}^k$, $\widehat{I}_{\pm 0} = I_{\pm 0}(x^k, y^k)$, $\widehat{I}_{0\pm} = I_{0\pm}(x^k, y^k)$ and $\widehat{I}_{00} = I_{00}(x^k, y^k)$ are constant. Certainly,

$$I_{\pm 0} = I_{\pm 0}(x^*, y^*) \subset \widehat{I}_{\pm 0}, \quad I_{0\pm} = I_{0\pm}(x^*, y^*) \subset \widehat{I}_{0\pm} \text{ and } \widehat{I}_{00} \subset I_{00} = I_{00}(x^*, y^*).$$

From (12) and setting $I_{00}^* := (\widehat{I}_{\pm 0} \setminus I_{\pm 0}) \cup (\widehat{I}_{0\pm} \setminus I_{0\pm})$, we get

$$\begin{aligned} v^k &= \sum_{i \in \mathcal{K}} \lambda_i^k a_i + \sum_{j \in I_{\pm 0}} \mu_j^k y_j^k e_j + \sum_{j \in I_{0\pm}} \mu_j^k y_j^k e_j + \sum_{j \in I_{00}^*} \mu_j^k y_j^k e_j \\ w^k &= \sum_{i \in \mathcal{K}} \lambda_i^k b_i + \sum_{j \in I_{\pm 0}} \mu_j^k x_j^k e_j + \sum_{j \in I_{0\pm}} \mu_j^k x_j^k e_j + \sum_{j \in I_{00}^*} \mu_j^k x_j^k e_j. \end{aligned} \tag{13}$$

Notice that $I_{00}^* \subset \widehat{I}_{\pm 0} \cup \widehat{I}_{0\pm}$ and that $(y_j^k e_j, x_j^k e_j)^T$, $j \in \widehat{I}_{\pm 0} \cup \widehat{I}_{0\pm}$ are linearly independent vectors. By applying Carathéodory theorem [2, Lemma 1] to (13), we find subset $\widehat{\mathcal{K}}^k \subset \mathcal{K}$ and multipliers $\widehat{\lambda}_i^k$ with $\widehat{\lambda}_i^k > 0$, $\forall j \in \widehat{\mathcal{K}}^k$ such that

$$\begin{aligned} v^k &= \sum_{i \in \widehat{\mathcal{K}}^k} \widehat{\lambda}_i^k a_i + \sum_{j \in I_{\pm 0}} \mu_j^k y_j^k e_j + \sum_{j \in I_{0\pm}} \mu_j^k y_j^k e_j + \sum_{j \in I_{00}^*} \mu_j^k y_j^k e_j, \\ w^k &= \sum_{i \in \widehat{\mathcal{K}}^k} \widehat{\lambda}_i^k b_i + \sum_{j \in I_{\pm 0}} \mu_j^k x_j^k e_j + \sum_{j \in I_{0\pm}} \mu_j^k x_j^k e_j + \sum_{j \in I_{00}^*} \mu_j^k x_j^k e_j \end{aligned} \tag{14}$$

and $\{(a_i, b_i)^T, (y_j^k e_j, x_j^k e_j)^T \mid j \in \widehat{I}_{\pm 0} \cup \widehat{I}_{0\pm}, i \in \widehat{\mathcal{K}}^k\}$ are linearly independent vectors. Moreover, we will take a subsequence that $\widehat{\mathcal{K}}^k$ is constant, and we will denote this set by $\widehat{\mathcal{K}}$.

Set $M_k = \max\{\widehat{\lambda}_i^k, |\mu_j^k \mid i \in \widehat{\mathcal{K}}, j \in \widehat{I}_{\pm 0} \cup \widehat{I}_{0\pm}\}$. Now, let us suppose that M_k is unbounded, $M_k \to \infty$. Then, dividing (14) by M_k and assuming $M_k^{-1}(\widehat{\lambda}^k, \mu^k) \to (\lambda, \mu)$ with $\|(\lambda, \mu)\| \neq 0$, we get

$$0 = \sum_{i \in \widehat{\mathcal{K}}} \widehat{\lambda}_i (a_i, b_i)^T + \sum_{j \in I_{\pm 0}} \mu_j (y_j^* e_j, x_j^* e_j)^T + \sum_{j \in I_{0\pm}} \mu_j (y_j^* e_j, x_j^* e_j)^T, \tag{15}$$

where we use that $(y_j^k e_j, x_j^k e_j)^T \to (0, 0)$ as $k \to \infty$ for every $j \in I_{00}^* \subset I_{00}$. Note that (15) and $\|(\lambda, \mu)\| \neq 0$ imply that $\{(a_i, b_i)^T, (y_j^* e_j, x_j^* e_j)^T \mid j \in I_{\pm 0} \cup I_{0\pm}, i \in \widehat{\mathcal{K}}\}$ are linearly dependent vectors, which is a contradiction with the fact that $\{(a_i, b_i)^T, (y_j^k e_j, x_j^k e_j)^T \mid j \in \widehat{I}_{\pm 0} \cup \widehat{I}_{0\pm}, i \in \widehat{\mathcal{K}}\}$ are linearly independent, since for every $j \in I_{\pm 0} \cup I_{0\pm}$, the vector $(y_j^* e_j, x_j^* e_j)^T$ is parallel to

$(y_j^k e_j, x_j^k e_j)^T$. Indeed, this follow from the next observation: Take $j \in I_{0\pm}$, then $x_j^* = 0$ and $y_j^* \neq 0$. From, $I_{0\pm} \subset \widehat{I}_{0\pm}$, we also get $x_j^k = 0$ and $y_j^k \neq 0$. Thus, we get

$$(y_j^* e_j, x_j^* e_j)^T = (y_j^* e_j, 0)^T = \frac{y_j^*}{y_j^k}(y_j^k e_j, 0)^T = \frac{y_j^*}{y_j^k}(y_j^k e_j, x_j^k e_j)^T.$$

Similarly, for every $j \in I_{\pm 0}$, $(y_j^* e_j, x_j^* e_j)^T$ is parallel to $(y_j^k e_j, x_j^k e_j)^T$.

Thus, M_k must be bounded. Taking an adequate subsequence, we assume that $(\widehat{\lambda}^k, \mu^k) \to (\widehat{\lambda}, \mu)$. Taking limit in (14),

$$(v, w)^T = \sum_{i \in \widehat{\mathcal{K}}} \widehat{\lambda}_i \begin{pmatrix} a_i \\ b_i \end{pmatrix} + \sum_{j \in I_{\pm 0}} \mu_j \begin{pmatrix} y_j^* e_j \\ x_j^* e_j \end{pmatrix} + \sum_{j \in I_{0\pm}} \mu_j \begin{pmatrix} y_j^* e_j \\ x_j^* e_j \end{pmatrix}, \tag{16}$$

with $\widehat{\lambda}_i \geq 0, \forall i \in \widehat{\mathcal{I}}$. Clearly, $\widehat{\mathcal{K}} \subset \{j \mid a_j^T x^* + b_j^T y^* = c_j\}$. Thus, (16) implies that $(v, w) \in D_{\mathcal{X}}(z^*)^\circ$. ■

We continue with the main result of this section, which ensures the validity of $D_\Omega(\bar{x}, \bar{y}) \subset \mathfrak{L}_\Omega(\bar{x}, \bar{y}; r)$ when h and g are affine mappings.

Theorem 4.2 *Consider the feasible set Ω when h and g are affine mappings. Then, $D_\Omega(\bar{x}, \bar{y}) \subset \mathfrak{L}_\Omega(\bar{x}, \bar{y}; r)$ holds for every feasible point $(\bar{x}, \bar{y}) \in \Omega$.*

In particular, if $(\bar{x}, \bar{y})$ is a weak local Pareto optimal solution for some smooth mapping F. Then, $(\bar{x}, \bar{y})$ is a weak CaC-S-stationary point.

Proof. Let $\bar{z} = (\bar{x}, \bar{y})$ be a feasible point of Ω. By using the equivalence in Theorem 4.1, we assume that $\bar{z} = (\bar{x}, \bar{y})$ is a weak local Pareto optimal solution of (2). Since, $\bar{z}$ is a weak local Pareto optimal solution, there exists $\delta > 0$ such that $\bar{z}$ is the unique solution of

$$\min_{w} \max_{i=1,\ldots,r} \{f_i(w) - f_i(\bar{z})\} + \frac{1}{2}\|w - \bar{z}\|^2 \text{ subject to } \|w - \bar{z}\| \leq \delta, w \in \Omega. \tag{17}$$

Following the proof of [17], we consider a smoothing function approximating $\max\{f_i(z) - f_i(\bar{z})\}$, namely, for $\eta > 0$, consider the smoothing function approximation $g_\eta(z)$ defined as

$$g_\eta(z) := \eta \ln \left\{ \sum_{i=1}^{r} \exp\left(\frac{f_i(z) - f_i(\bar{z})}{\eta} \right) \right\} - \eta \ln r, \text{ for every } z \in \mathbb{R}^{2n}. \tag{18}$$

Note for each $z \in \mathbb{R}^{2n}$, we get $g_\eta(z) \to \max_{i=1,\ldots,r} \{f_i(z) - f_i(\bar{z})\}$ as $\eta \to 0$ and

$$\nabla g_\eta(z) = \sum_{i=1}^{r} \theta_i(z) \nabla f_i(z) \text{ with } \theta_i(z) = \frac{\exp((f_i(z) - f_i(\bar{z}))/\eta)}{\sum_{i=1}^{r} \exp((f_i(z) - f_i(\bar{z}))/\eta)}. \tag{19}$$

From (19), $\|\theta(w) = (\theta_1(z), \ldots, \theta_r(z))\|_1 = 1$. Now, consider the smooth optimization problem

$$\text{Minimize } g_\eta(w) + \frac{1}{2}\|w - \bar{z}\|^2 \text{ subject to } \|w - \bar{z}\| \leq \delta, \quad w \in \Omega. \tag{20}$$

Let $\{\eta^k\}$ be a sequence of positive parameters converging to 0 and denote by $z^k \in \Omega$ the global minimizer of (20) with $\eta = \eta^k$. From the proof of [17], $z^k \to \bar{z}$ and thus, for k large enough $\|z^k - \bar{z}\| < \delta$. So, by optimality of z^k and applying the Fermat's rule, we get

$$v^k := -\nabla g_{\eta^k}(z^k) - (z^k - \bar{z}) = -\sum_{i=1}^{r} \theta_i(z^k)\nabla f_i(z^k) - (z^k - \bar{z}) \in \widehat{N}_\Omega(z^k) = T_\Omega(z^k)^\circ \tag{21}$$

Without loss of generality (after the possibility of taking an additional subsequence) we assume that θ^k converges to θ with $\|\theta\|_1 = 1$ and $\theta \in \mathbb{R}^r_+$. Taking limit in (21), since $z^k \to \bar{z}$, we get $v^k \to v := -\sum_{i=1}^{r} \theta_i \nabla f_i(\bar{z})$.

By [14, Corollary 3.10], GCQ holds at z^k for Ω, so $T_\Omega(z^k)^\circ = D_\Omega(z^k)^\circ$. From (21), $\{v^k\}$ is a sequence in $D_\Omega(z^k)^\circ$ converging to $v \in \limsup_{z\to\bar{z},\, z\in\Omega} D_\Omega(z)^\circ$. By Lemma 4.1, $\limsup_{z\to\bar{z},\, z\in\Omega} D_\Omega(z)^\circ \subset D_\Omega(\bar{z})^\circ$, so $v = -\sum_{i=1}^{r} \theta_i \nabla f_i(\bar{z}) \in D_\Omega(\bar{z})^\circ$, which in turn implies that $\bar{z}$ is a weak CaC-S-stationary point. ■

4.3 Constraint Qualifications for strong CaC-S-stationary points

Here, we focus on CQs for the fulfillment of strong CaC-S-stationary point. By [17, Theorem 4.2] and the equivalence of strong CaC-S-stationary point with strong KKT points, we establish the following theorem.

Theorem 4.3 *Let $(\bar{x}, \bar{y})$ be a feasible point of the problem (2). The weakest property, namely MOP-SCQ, under which every local weak Pareto optimal solution is a strong CaC-S-stationary point for every continuously differentiable mapping, $F(x) \equiv (f_1(x), \dots, f_r(x))^T$ is*

$$D_\Omega(\bar{x}, \bar{y}) \subset \mathfrak{L}^S_\Omega(\bar{x}, \bar{y}; r) \tag{22}$$

where

$$\mathfrak{L}^S_\Omega(\bar{x}, \bar{y}; r) = \{d \in \mathbb{R}^{2n} \mid \forall (v_i)_{i=1}^r \in \widehat{N}_\Omega(\bar{x}, \bar{y}; r), v_i^T d = 0, \forall i \text{ or } \min_{i=1,\dots,r} v_i^T d < 0\}.$$

Clearly, when $r = 1$, (22) reduces to $D_\Omega(\bar{x}, \bar{y}) \subset \widehat{N}_\Omega(\bar{x}, \bar{y})^\circ = T_\Omega(\bar{x}, \bar{y})^{\circ\circ}$ which it is equivalent to GCQ. Regrettably, inclusion (22) is too restrictive, and it may fail even in very well-behaved constrained systems where *Linear Independence* CQ (LICQ) holds, [17, Examples 4.2, 4.3]. In our case, inclusion (22) fails even if the constraints h and g are linear mappings with linearly independent gradients, as the next example shows.

Example 4.2 *Consider the MOPCaC and the corresponding relaxed problem*

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2}{\text{minimize}} & F(x) \equiv \left(x_1^2, x_1^2 + x_2\right)^T \\ \text{subject to} & x_1 + x_2 \le 0, \\ & \|x\|_0 \le 1, \end{array} \qquad \begin{array}{ll} \underset{x,y\in\mathbb{R}^2}{\text{minimize}} & F(x) \equiv \left(x_1^2, x_1^2 + x_2\right)^T \\ \text{subject to} & x_1 + x_2 \le 0, \\ & y_1 + y_2 \ge 1, \\ & x_i y_i = 0,\ i = 1, 2, \\ & 0 \le y_i \le 1,\ i = 1, 2. \end{array}$$

Set $x^* = (0, 0)$ and $y^* = (1, 0)$. Here, (x^*, y^*) is a Pareto optimal solution of the relaxed problem which it is not a strong CaC-S-stationary point. However, the inequality constraint $g(x_1, x_2) := x_1 + x_2 \leq 0$ is linear with non-null gradient $\nabla g(x_1, x_2) = (1, 1) \neq (0, 0)$ for every (x_1, x_2).

We close this section with the Fig. 2 establishing the relations among some CQs discussed here. See [14] for the corresponding definitions and relations.

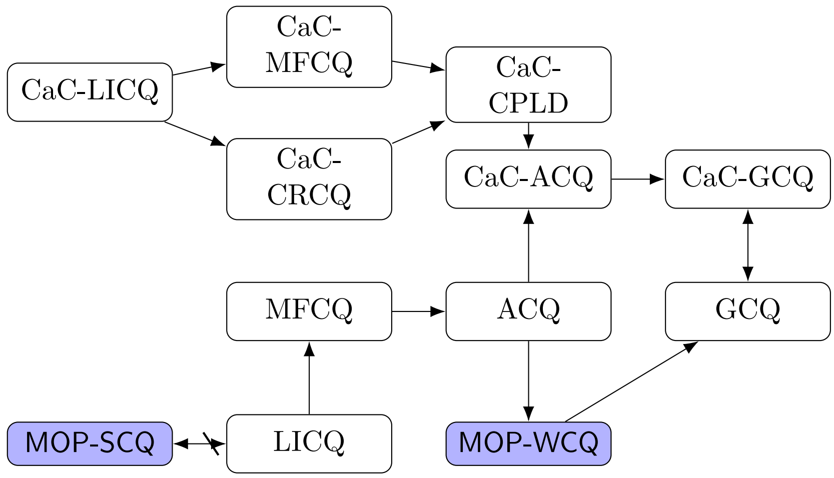

Figure 2: Relationships between the CQs discussed in this work apply to the constrained system Ω. An arrow indicates a strict implication between two conditions.

5 Unifying the stationarity conditions for multiobjective problems with cardinality constraints

In this section, for MOP, we discuss the possibility of describing the concepts of CaC-M-stationarity in a unified way. Here, we follow the approach of [20] which defines different levels of stationarity depending on how the continuous parameter y determines the cardinality of the variable x.

Definition 5.1 *Let $(\bar{x}, \bar{y})$ be a feasible point of the problem (2) and I be an index set such that*

$$I_{\pm}(\bar{x}) \subset I \subset I_0(\bar{y}). \tag{23}$$

Suppose that there exists a vector $0 \neq (\lambda^f, \lambda^g, \lambda^h, \gamma) \in \mathbb{R}^r_+ \times \mathbb{R}^m_+ \times \mathbb{R}^p \times \mathbb{R}^n$ such that

$$\nabla F(\bar{x})^T \lambda^f + \nabla g(\bar{x})^T \lambda^g + \nabla h(\bar{x})^T \lambda^h + \gamma = 0. \tag{24}$$

$$\lambda^g_i g_i(\bar{x}) = 0 \text{ for all } i = 1, \ldots, m. \tag{25}$$

Then, we say that $(\bar{x}, \bar{y})$ is a:

1. weak CaC-M$_I$-stationary point, *if (24) and (25) hold with $\lambda^f \neq 0$ and $\gamma_i = 0$ for all $i \in I$.*

2. strong CaC-M$_I$-stationary point, *if (24) and (25) hold with* $\lambda_i^f > 0$ *for all* $i = 1, \ldots, r$ *and* $\gamma_i = 0$ *for all* $i \in I$.

Clearly, if $I = I_\pm(\bar{x})$, then weak/strong CaC-M_I-stationarity conditions coincide with the weak/strong CaC-M-stationarity conditions, and if $I = I_0(\bar{y})$ they coincide with weak/strong CaC-S-stationarity conditions. As we see, there are different levels of stationarity, corresponding to the set I in (23). As a direct consequence of the definitions, we get Proposition 5.1, see Fig. 3.

Proposition 5.1 *Let* $(\bar{x}, \bar{y})$ *be feasible of (2). Let* I_1 *and* I_2 *be two sets of indexes such that* $I_\pm(\bar{x}) \subset I_1 \subset I_2 \subset I_0(\bar{y})$. *Then, weak/strong CaC-*M_{I_2}*-stationarity implies weak/strong CaC-*M_{I_1}*-stationarity.*

In particular, a weak/strong CaC-S-stationary point is a CaC-M_I*-stationary point for every* I *such that* $I_\pm(\bar{x}) \subset I \subset I_0(\bar{y})$.

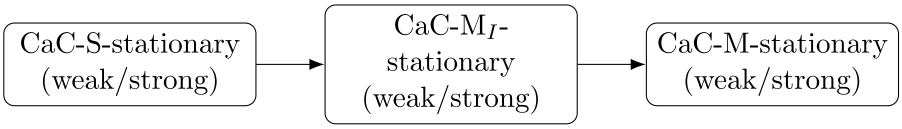

Figure 3: Relationships between different stationary concepts in MOP with cardinality constraints, for a fixed feasible point $(\bar{x}, \bar{y})$ and an index set I with $I_\pm(\bar{x}) \subset I \subset I_0(\bar{y})$.

When $I_{00}(\bar{x}, \bar{y}) = \emptyset$, we say that $(\bar{x}, \bar{y})$ satisfies the *strict complementarity* and in this case, all weak/strong CaC-M_I-stationarity conditions collapse to the weak/strong CaC-S-stationarity condition. Similarly to the scalar case [20], we can characterize the different forms of weak/strong M_I-stationarity with the weak/strong KKT stationarity of a certain MOPCaC. For this purpose, for every feasible point $(\bar{x}, \bar{y})$ of (2) and for every I satisfying (23), we define the *Tight Nonlinear Problem* at $(\bar{x}, \bar{y})$ with respect to the index I by

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & F(x) \equiv (f_1(x), f_2(x), \ldots, f_r(x))^T \\ \text{subject to} & g(x) \leq 0, h(x) = 0, \\ & n - e^T y \leq \alpha, \\ & 0 \leq y \leq e, \\ & y_i = 0, i \in I_0(\bar{y}) \setminus I \\ & x_i y_i = 0, i \in I \cup I_\pm(\bar{y}). \end{array} \tag{26}$$

We denote this problem by TNLP$_I(\bar{x}, \bar{y})$ and its feasible set is denoted by $\Omega_I(\bar{x}, \bar{y})$. Clearly, $(\bar{x}, \bar{y}) \in \Omega_I(\bar{x}, \bar{y})$. Furthermore, as a direct consequence of the definition of TNLP$_I(\bar{x}, \bar{y})$, we have the next lemma:

Lemma 5.1 *Consider the problem TNLP*$_I(\bar{x}, \bar{y})$ *given by (26). Then,*

1. *Every feasible point of (26) is feasible for (2);*

2. *Every local weak/strong Pareto optimal solution of the problem (2) is also a local weak/strong Pareto optimal solution of (26).*

The constrained system given by the constraints, $y_i = 0$, $i \in I_0(\bar{y}) \setminus I$; $x_i y_i = 0$, $i \in I \cup I_\pm(\bar{y})$, near to $(\bar{x}, \bar{y})$ is locally the same as $y_i = 0$, $i \in I_0(\bar{y})$; $x_i = 0$, $i \in I_\pm(\bar{y})$. This is the reason why we call TNLP$_I$ as the tightened problem.

Theorem 5.1 *Let $(\bar{x}, \bar{y})$ be feasible of (2) and I be an index set satisfying $I_\pm(\bar{x}) \subset I \subset I_0(\bar{y})$. Then, $(\bar{x}, \bar{y})$ is a weak/strong M_I-stationary point, if and only if, it is a weak/strong KKT point for the tightened problem (26).*

Proof. Suppose that $(\bar{x}, \bar{y})$ is a weak KKT point of $\Omega_I(\bar{x}, \bar{y})$. From the definition of weak KKT condition, there are multipliers λ^f, λ^g, λ^h, θ, ν^+, ν^-, γ^ξ with $\lambda^f \in \mathbb{R}^r_+$, $\lambda^f \neq 0$ such that

$$\begin{aligned} \nabla F(\bar{x})^T \lambda^f + \nabla g(\bar{x})^T \lambda^g + \nabla h(\bar{x})^T \lambda^h + \sum_{i \in I \cup I_\pm(\bar{y})} \gamma_i^\xi \bar{y}_i e_i = 0, \\ -\theta e + \sum_{i \in I_0(\bar{y}) \setminus I} \eta_i e_i + (\nu^+ - \nu^-) + \sum_{i \in I \cup I_\pm(\bar{y})} \gamma_i^\xi \bar{x}_i e_i = 0, \\ \lambda_j^g \geq 0, \quad \lambda_j^g g_j(\bar{x}) = 0, \forall j, \\ \theta \geq 0, \quad \theta(\bar{y}^T e - n + \alpha) = 0, \\ \nu^+ \geq 0, \quad \nu_i^+(\bar{y}_i - 1) = 0, \forall i, \\ \nu^- \geq 0, \quad \nu_i^- \bar{y}_i = 0, \forall i. \end{aligned}$$

Now, set $\gamma \in \mathbb{R}^n$ as $\gamma_i := \gamma_i^\xi \bar{y}_i$, if $i \in I \cup I_\pm(\bar{y})$, and $\gamma_i := 0$, otherwise. It is clear from the definition that $\gamma_i = 0$ for every $i \in I_0(\bar{y})$, since $I \subset I_0(\bar{y})$ and $I_0(\bar{y}) \cap I_\pm(\bar{y}) = \emptyset$. Thus, $(\bar{x}, \bar{y})$ is a weak CaC-M$_I$-stationary point.

Now, for the other inclusion, assume that $(\bar{x}, \bar{y})$ is a weak CaC-M$_I$-stationary point for some $\lambda^f \in \mathbb{R}^r_+$, $\lambda^f \neq 0$, λ^g, λ^h and γ. Clearly, $\gamma_i = 0$, $i \in I$. Define γ^ξ as $\gamma_i^\xi = \gamma_i / \bar{y}_i$, $i \in I_\pm(\bar{y})$ and $\gamma_i^\xi = 0$, otherwise. It is easy to see that $\gamma_i^\xi \bar{y}_i = \gamma_i$, $i \in I \cup I_\pm(\bar{y})$. Furthermore, $\gamma_i^\xi \bar{x}_i = 0$ for $i \in I \cup I_\pm(\bar{y})$ since $(\bar{x}, \bar{y})$ is feasible. Thus, putting $\theta := 0$, $\nu^+ := 0$, $\nu^- := 0$ and $\eta := 0$, we see that $(\bar{x}, \bar{y})$ is a weak KKT point. The same conclusions follow if $(\bar{x}, \bar{y})$ is a strong KKT point. ■

Proposition 5.2 *Let $(\bar{x}, \bar{y})$ be feasible for (2) and I be an index set satisfying* (23)*. If $(\bar{x}, \bar{y})$ is a weak/strong CaC-M_I-stationary point, then there exists $\bar{z}$ such that $(\bar{x}, \bar{z})$ is weak/strong CaC-S-stationary point.*

Proof. We suppose that $(\bar{x}, \bar{y})$ is a weak CaC-M_I-stationary point. The proof when $(\bar{x}, \bar{y})$ is a strong CaC-M_I-stationary point is similar. Now, from the definition, there exists a vector $0 \neq (\lambda^f, \lambda^g, \lambda^h, \gamma) \in \mathbb{R}^r_+ \times \mathbb{R}^m_+ \times \mathbb{R}^p \times \mathbb{R}^n$ with $\lambda^f \neq 0$, $\lambda_i^g g_i(\bar{x}) = 0$, $\forall i$ such that

$$\nabla F(\bar{x})^T \lambda^f + \nabla g(\bar{x})^T \lambda^g + \nabla h(\bar{x})^T \lambda^h + \gamma = 0 \quad \text{and} \quad \gamma_i = 0 \;\; \forall i \in I.$$

Now, define $\bar{z}$ as $\bar{z}_i = 0$, if $i \in I$ and $\bar{z}_i = 1$, if $i \notin I$. Note that $I = I_0(\bar{z})$. Since $I_\pm(\bar{x}) \subset I$, we get $x_i = 0$, $i \notin I$ and so $x_i z_i = 0, \forall i$. With this choice of $\bar{z}$, it is not difficult to see that $(\bar{x}, \bar{z})$ is a weak CaC-S-stationary point. ■

Remark 5.1 *Proposition 5.2 says that if $\bar{x}$ is a point such that $(\bar{x}, \bar{y})$ is a CaC-M_I-stationary point for some $\bar{y}$ and I satisfying* (23)*, then, there exists $\bar{z}$ such that $(\bar{x}, \bar{z})$ is a CaC-M_I-stationary point for every I with $I_\pm(\bar{x}) \subset I \subset I_0(\bar{z})$. Thus, from the point of view of the variable*

x, the only relevant stationarity concept is the weak/strong CaC-M-stationarity since the others can be achieved with a proper choice of the companion variable y. But this does not make the parameter y irrelevant, indeed, for the scalar case, several numerical methods for solving (1) *use the parameter y in its formulation. Furthermore, it is possible that some CQ holds at some (x, y) but fail if we change y by another value, as the Example 5.1 shows.*

Example 5.1 *Consider the problem of minimizing $F(x) = (x_1 + x_2, x_1 + x_2)^T$ subject to $-x_1 + x_2^2 \leq 0$, and $\|x\|_0 \leq 1$. Here, $x^* = (0, 0)$ is its Pareto optimal solution. Set $y^* = (1, 0)$ and $I = I_0(x^*)$. Now, consider the associated relaxed problem and the corresponding $TNLP_I(\bar{x}, \bar{y})$.*

$$\begin{array}{ll} \underset{x,y\in\mathbb{R}^2}{\text{minimize}} & F(x) \equiv (x_1 + x_2, x_1 + x_2)^T \\ \text{subject to} & -x_1 + x_2^2 \leq 0, \\ & y_1 + y_2 \geq 1, \\ & x_i y_i = 0,\ i = 1, 2, \\ & 0 \leq y_i \leq 1,\ i = 1, 2. \end{array} \qquad \begin{array}{ll} \underset{x,y\in\mathbb{R}^2}{\text{minimize}} & F(x) \equiv (x_1 + x_2, x_1 + x_2)^T \\ \text{subject to} & -x_1 + x_2^2 \leq 0, \\ & y_1 + y_2 \geq 1, \\ & x_2 = 0, \quad y_1 = 0, \\ & 0 \leq y_i \leq 1,\ i = 1, 2. \end{array}$$

Here, (x^*, y^*) is a local Pareto optimal solution of the relaxed problem, but it is not a weak CaC-S-stationary point. On the other hand, (x^*, y^*) is a CaC-M_I-stationary point and $\Omega_I(\bar{x}, \bar{y})$ satisfies ACQ at this point.

Example 5.1 leads us to consider CQs for the fulfillment of the CaC-M$_I$-stationarity condition at local minimizer for a given I. Using the equivalence given in Theorem 5.1 and the characterizations of the weak CQ for weak/strong KKT in [17, Theorem 4.1, Theorem 4.2], we have

Theorem 5.2 *Let $(\bar{x}, \bar{y})$ be a feasible point of* (2) *and I be an index set such that $I_\pm(\bar{x}) \subset I \subset I_0(\bar{y})$. Set $\Omega_I := \Omega_I(\bar{x}, \bar{y})$. Then,*

(a) *The weakest property such that every local weak Pareto optimal solution is a weak CaC-S-stationary point for every continuously differentiable mapping, $F(x) \equiv (f_1(x), \ldots, f_r(x))^T$ is*

$$D_{\Omega_I}(\bar{x}, \bar{y}) \subset \mathfrak{L}_{\Omega_I}(\bar{x}, \bar{y}; r), \tag{27}$$

where $\mathfrak{L}_{\Omega_I}(\bar{x}, \bar{y}; r) = \{d \in \mathbb{R}^{2n} \mid \min_{i=1,\ldots,r} v_i^T d \leq 0,$ for all $(v_i)_{i=1}^r \in \widehat{N}_{\Omega_I}(\bar{x}, \bar{y}; r)\}$.

(b) *The weakest property under which every local weak Pareto optimal solution is a strong CaC-M_I-stationary point for every continuously differentiable mapping, $F(x) \equiv (f_1(x), \ldots, f_r(x))^T$ is*

$$D_{\Omega_I}(\bar{x}, \bar{y}) \subset \mathfrak{L}^S_{\Omega_I}(\bar{x}, \bar{y}; r), \tag{28}$$

where

$$\mathfrak{L}^S_{\Omega_I}(\bar{x}, \bar{y}; r) = \{d \in \mathbb{R}^{2n} \mid \forall (v_i)_{i=1}^r \in \widehat{N}_{\Omega_I}(\bar{x}, \bar{y}; r), v_i^T d = 0, \forall i \text{ or } \min_{i=1,\ldots,r} v_i^T d < 0\}.$$

We know that ACQ is enough to ensure that every local weak Pareto optimal solution satisfies the weak KKT conditions then let $(\bar{x}, \bar{y})$ be a feasible point of (2) and I be an index set with $I_\pm(\bar{x}) \subset I \subset I_0(\bar{y})$, a sufficient CQ to ensure the fulfillment of (27) is that ACQ holds for the constraint set $\Omega_I(\bar{x}, \bar{y})$ at $(\bar{x}, \bar{y})$. In this case, we say that ACQ(I) holds at $(\bar{x}, \bar{y})$. Using an argument similar to [20, Theorem 5], it is not difficult to see ACQ(I) holds at $(\bar{x}, \bar{y})$ if $I_{00}(\bar{x}, \bar{y}) \cap (I \cup I_\pm(\bar{y})) = I_{00}(\bar{x}, \bar{y}) \cap I = \emptyset$.

Before wrapping up this section, to show the relevance of MOPCaC, we present some applications in which it appears. The first application is portfolio selection, which originally proposed by H. Markowitz [22] (see [19] for a recent survey on the subject). Motivated by a multiobjective formulation of the portfolio selection [15, 21], we present a MOPCaC formulation of it which imposes a bound on the number of assets in the portfolio:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & F(x) \equiv (f_1(x), \dots, f_r(x))^T \\ \text{subject to} & e^T x = 1, \quad 0 \le x \le e, \quad \|x\|_0 \le \alpha. \end{array} \tag{29}$$

The variable x is the investment portfolio, where x_i represents a fraction of the investment allocated to an asset i and α represents the bound on the number of assets in the portfolio. A mean-variance version of the problem, which is resulted from minimizing the risk and the negative return, is derived by setting $r = 2$ in (29), and choosing $f_1(x) = \sum_{i,j=1}^n \sigma_{ij} x_i x_j$ and $f_2(x) = -\sum_{i=1}^n \mu_i x_i$, where μ_i is the expected return of the asset i, σ_{ij} represents the covariance between the return on asset i and j. A mean-variance-skewness version of the problem is derived by setting $r = 3$ in (29) and choosing $f_3(x) = \sum_{i,j,k=1}^n s_{ijk} x_i x_j x_k$, where s_{ijk} represents the skewness [15, 21].

Other important problem is the elastic net [25] regularized version of logistic regression [4]. Due to the conflict between obtaining a good fit to the data set and the need for avoiding overfitting the data, and also the necessity for selecting a set of the most important features, it can be formulated as the following MOPCaC:

$$\begin{array}{ll} \underset{\beta\in\mathbb{R}^n}{\text{minimize}} & F(\beta) \equiv \left(\dfrac{1}{M} \displaystyle\sum_{i=1}^{M} \log\left(1 + \exp(-q^i \beta^T p^i)\right), \|\beta\|^2 \right)^T \\ \text{subject to} & \|\beta\|_0 \le \alpha, \end{array} \tag{30}$$

where $\{p^1, \dots, p^M\}$ is the set of sample data with n features, and $\{q^1, \dots, q^M\}$ with $q^i \in \{-1, 1\}$, represents their corresponding labels.

6 Conclusions and Final Remarks

In this paper, we proposed and analyzed Pareto optimality conditions and constraint qualifications for Multiobjective Programs with Cardinality Constraints (MOPCaC). For this purpose, we take advantage of the approach for the scalar case in [20], which uses a recently developed continuous formulation. In view of the possibility of conflict among the objective functions, we consider different notions of optimality (weak/strong Pareto optimal solutions).

Several theoretical results were proposed. First, we proved the equivalence between weak/strong Pareto optimal solution of the MOPCaC problem and its relaxed problem. We emphasize that for local solutions, we guarantee the equivalence if cardinality constraint is active.

We defined tailored stationarity conditions, namely weak/strong CaC-M-stationarity and weak/strong CaC-S-stationarity, considering multipliers associated with objective functions and cardinality constraint. Furthermore, we obtained the equivalence between weak/strong CaC-S-stationary and weak/strong KKT point of the relaxed MOPCaC problem.

In addition, we used the r-multiobjective normal cone defined in [17] to study CQs for MOPCaC to characterize the stationarity concepts presented here. In addition, we compared these CQs with other well-known CQs for cardinality constrained optimization problems. Thus, we established the weakest CQ, namely MOP-WCQ, to guarantee the fulfillment of the weak CaC-S-stationarity condition at local weak Pareto optimal solution. In particular, we proved that in the linear case, this CQ holds for all feasible points. Moreover, we proposed the weakest CQ, namely MOP-SCQ, to guarantee the fulfillment of strong CaC-S-stationarity condition at local weak Pareto optimal solution. We also proposed weak/strong CaC-M-stationarity in a unified way.

Acknowledgments

Funding support for Rohollah Garmanjani was provided by national funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications) and PTDC/MAT-APL/28400/2017. Support for Evelin H. M. Krulikovski was funded by national funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications) and the scope of project UI/297/2020-5/2021. Support for Alberto Ramos was provided by CNPq (grant 309437/2016-4).

References

[1] B Aghezzaf and M Hachimi. On a gap between multiobjective optimization and scalar optimization. 2001.

[2] R Andreani, G Haeser, M L Schuverdt, and P JS Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, 135(1):255–273, 2012.

[3] A Beck and Y C Eldar. Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM J. Optim.*, 23(3):1480–1509, 2013.

[4] D Bertsimas and A King. Logistic regression: From art to science. *Statistical Science*, pages 367–384, 2017.

[5] D Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Math. Program.*, 74(2):121–140, 1996.

[6] G Bigi and M Pappalardo. Regularity conditions in vector optimization. *Journal of Optimization Theory and Applications*, 102(1):83–96, 1999.

[7] M Branda, M Bucher, M Červinka, and A Schwartz. Convergence of a Scholtes-type regularization method for cardinality-constrained optimization problems with an application in sparse robust portfolio optimization. *Comput. Optim. Appl.*, 70(2):503–530, 2018.

[8] M Bucher and A Schwartz. Second-order optimality conditions and improved convergence results for regularization methods for cardinality-constrained optimization problems. *J. Optim. Theory Appl.*, 178:383–410, 2018.

[9] R. S. Burachick and M. M. Rizvi. On weak and strong kuhn-tucker conditions for smooth multiobjective optimization. *Journal of Optimization Theory and Applications*, 155:477–491, 2012.

[10] O Burdakov, C Kanzow, and A Schwartz. On a reformulation of mathematical programs with cardinality constraints. In D Gao, N Ruan, and W Xing, editors, *Advances in Global Optimization.* Springer Proceedings in Mathematics and Statistics, 2015.

[11] O Burdakov, C Kanzow, and A Schwartz. Mathematical programs with cardinality constraints: reformulation by complementarity-type conditions and a regularization method. *SIAM J. Optim.*, 26(1):397–425, 2016.

[12] E J Candès and M B Wakin. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, 25(2):21–30, 2008.

[13] M Castellani and M Pappalardo. About a gap between multiobjective optimization and scalar optimization. *Journal of optimization theory and applications*, 109(2):437–439, 2001.

[14] M Červinka, C Kanzow, and A Schwartz. Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. *Math. Program.*, 160:353–377, 2016.

[15] C Chen and Y Zhou. Robust multiobjective portfolio with higher moments. *Expert Systems with Applications*, 100:165–181, 2018.

[16] A d'Aspremont, F Bach, and L E Ghaoui. Optimal solutions for sparse principal component analysis. *J. Mach. Learn. Res.*, 9:1269–1294, 2008.

[17] G Haeser and A Ramos. Constraint qualifications for Karush–Kuhn–Tucker conditions in multiobjective optimization. *Journal of Optimization Theory and Applications*, 187(2):469–487, 2020.

[18] C Kanzow, A B Raharja, and A Schwartz. An augmented Lagrangian method for cardinality-constrained optimization problems. *Journal of Optimization Theory and Applications*, 189(3):793–813, 2021.

[19] P N Kolm, R Tütüncü, and FJ Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356–371, 2014.

[20] E H M Krulikovski, A A Ribeiro, and M Sachine. On the weak stationarity conditions for mathematical programs with cardinality constraints: A unified approach. *Appl. Math. Optim.*, (84):3451–3473, 2021.

[21] T. Y Lai. Portfolio selection with skewness: a multiple-objective approach. *Review of Quantitative Finance and Accounting*, 1(3):293–305, 1991.

[22] H Markowitz. Portfolio selection. *The Journal of Finance*, 7:77–91, 1952.

[23] A Miller. *Subset Selection in Regression.* Chapman and Hall/CRC Press, Florida, 2002.

[24] R Ruiz-Torrubiano, S García-Moratilla, and A Suárez. Optimization problems with cardinality constraints. In Y Tenne and C K Goh, editors, *Computational Intelligence in Optimization*, pages 105–130. Academic Press, Berlin, 2010.

[25] Y Shen, B Han, and E Braverman. Stability of the elastic net estimator. *Journal of Complexity*, 32(1):20–39, 2016.

[26] X Sun, X Zheng, and D Li. Recent advances in mathematical programming with semi-continuous variables and cardinality constraint. *J. Oper. Res. Soc. China*, 1:55–77, 2013.

[27] R Tibshirani. Regression shrinkage and selection via the Lasso. *J. R. Statist. Soc.*, 58:267–288, 1996.