Journal of Optimization Theory and Applications (2018) 179:800–819
https://doi.org/10.1007/s10957-018-1373-7

Constraint Qualifications and Stationary Conditions for Mathematical Programming with Non-differentiable Vanishing Constraints

Sajjad Kazemi[1] · **Nader Kanzi**[1]

Received: 15 October 2017 / Accepted: 4 August 2018 / Published online: 15 August 2018

Abstract
This paper aims at studying a broad class of mathematical programming with non-differentiable vanishing constraints. First, we are interested in some various qualification conditions for the problem. Then, these constraint qualifications are applied to obtain, under different conditions, several stationary conditions of type Karush/Kuhn–Tucker.

Keywords Constraint qualification · Stationary conditions · Optimality conditions · Vanishing constraints

Mathematics Subject Classification 90C46 · 90C33 · 49K10

1 Introduction

Since 2007, a difficult class of optimization problems was introduced by Kanzow and his co-authors in [1,2]; it was called mathematical programming with vanishing constraints (MPVC for short). The MPVCs received attention from different fields. Some of their applications in topological optimization have been introduced in [1]. Also, MPVCs are generalizations of another group of optimization problems, named mathematical programming with equilibrium constraints (MPEC in brief).

Since MPVCs are a generalized form of MPECs, it seems natural that we seek to generalize the results of MPECs to them. As we know, in MPECs, occurrence of situations, which could lead to necessary conditions in Karush/Kuhn–Tucker (KKT) types, is very limiting and hindering. Therefore, instead of KKT condition some other condi-

✉ Nader Kanzi
nad.kanzi@gmail.com

Sajjad Kazemi
sajjad_kazemi65@yahoo.com

1 Department of Mathematics, Payam Noor University, P. O. Box 19395-3697, Tehran, Iran

tions are substituted. They have the style and appearance of KKT condition, and also, their coefficients are more free than KKT coefficients. The recent conditions, which are called stationary conditions, can be proved under some constraint qualifications that their occurrence is more natural [3–9].

The same process has been repeated for MPVCs, and numerous constraint qualifications have been introduced, which lead to different stationary conditions. For instance, the first-order stationary conditions for MPVCs can be found in [1,10,11]. The second-order stationary conditions and exact penalty theorem are given in [2] and [12], respectively. Also, relaxation methods in [13], stability in [14], and numerical algorithms in [15] have been presented. In addition, weak and strong duality results for MPVCs could be seen in [16].

In the previous works that referenced earlier, all functions which define MPVC are continuously differentiable. Unlike the MPECs which their non-smooth form has been studied in some references as in [3,7–9,17], the earlier works for MPVCs assume the limiting criteria of continuously differentiability for functions. So, to the best of our knowledge, the current article is the first that studies the stationary conditions for MPVCs where their constraints are non-smooth, meaning that they are not necessarily differentiable. In this paper, we assume the objective function is differentiable.

Since the feasible set of an MPVC is not necessarily convex, even under the criteria of convexity of the functions which construct it (to be discussed in next sections), applying common methods of convex analysis is not applicable here. Therefore, we take the non-smooth analysis approach. To choose a suitable sub-differential, we select a sub-differential which not only is convex but also its calculation rules are known. The best selection is Clarke sub-differential. Therefore, we consider the functions which define the problem to be locally Lipschitz.

The structure of the subsequent sections of this paper is as follows: In Sect. 2, we define required definitions, theorems, and relations of non-smooth analysis. In Sect. 3, we will introduce some various constraint qualifications for a system with non-smooth vanishing constraints. Also, relations between defined constraint qualifications are presented in Sect. 3. In Sect. 4, we apply these constraint qualifications to obtain different kinds of stationary conditions for the problem.

2 Notations and Preliminaries

In this section, we provide an overview for some notations and preliminary results that will be used throughout this paper. Further details and examples of these results can be found in the books [18–20].

First, we recall that $\mathbb{R}_+$ and $\langle x, y\rangle$ denote, respectively, the nonnegative real numbers $[0, +\infty[$ and the standard inner product of vectors $x, y \in \mathbb{R}^n$. For a non-empty subset M of $\mathbb{R}^n$, let:

$$M^- := \{x \in \mathbb{R}^n : \langle x, d\rangle \leq 0, \quad \forall d \in M\},$$
$$M^s := \{x \in \mathbb{R}^n : \langle x, d\rangle < 0, \quad \forall d \in M\}.$$
$$M^\perp := M^- \cap (-M)^- = \{x \in \mathbb{R}^n : \langle x, d\rangle = 0, \quad \forall d \in M\}.$$

It can easily be shown that if $M^s \neq \emptyset$, then $cl(M^s) = M^-$, which $cl(M^s)$ denotes the topological closure of M^s. Here, it is necessary to note that for given closed convex cones $M_1,\ M_2,\ M_3$, and M_4, we have:

$$(M_1 \cup M_2)^- = M_1^- \cap M_2^-, \quad (M_1 \times M_2)^- = M_1^- \times M_2^-, \tag{1}$$

$$\left(M_1^- \times M_2^-\right) \cap \left(M_3^- \times M_4^-\right) = \left(M_1^- \cap M_3^-\right) \times \left(M_2^- \cap M_4^-\right). \tag{2}$$

The convex hull and the convex cone (containing origin) of $M \subseteq \mathbb{R}^n$ are, respectively, denoted by $\mathrm{conv}(M)$ and $\mathrm{cone}(M)$. It is easy to see that if $\{M_\kappa\}_{\kappa \in \Lambda}$ is a class of convex subsets of $\mathbb{R}^n$ with $|\Lambda| < \infty$ and with the convention $\bigcup_{\kappa \in \emptyset} M_\kappa = \emptyset$, then one has:

$$\mathrm{cone}\left(\bigcup_{\kappa \in \Lambda} M_\kappa\right) = \left\{\sum_{\kappa \in \Lambda} \alpha_\kappa d_\kappa : \alpha_\kappa \geq 0,\ d_\kappa \in M_\kappa\right\}.$$

The famous bipolar theorem says that $(M^-)^- = cl\big(\mathrm{cone}(M)\big)$ for each $\emptyset \neq M \subseteq \mathbb{R}^n$. We define the normal cone of convex set $A \subseteq \mathbb{R}^n$ at $x_0 \in cl(A)$ as

$$N(A, x_0) := \{d \in \mathbb{R}^n : \langle d, x - x_0 \rangle \leq 0, \quad \forall x \in A\}.$$

The following relation for any subset K of $\mathbb{R}^n$ is immediately from bipolar theorem:

$$N(K^-, 0) = cl\big(\mathrm{cone}(K)\big). \tag{3}$$

Let $\psi : \mathbb{R}^n \to \mathbb{R}^* := \mathbb{R} \cup \{+\infty\}$ be a locally Lipschitz function, and $x_0 \in dom\ \psi$. The Clarke sub-differential of ψ at x_0 is defined as

$$\partial_c \psi(x_0) := \{\xi \in \mathbb{R}^n : \psi^\circ(x_0; v) \geq \langle \xi, v \rangle, \quad \forall v \in \mathbb{R}^n\},$$

where $\psi^\circ(x_0; v)$ denotes the Clarke generalized directional derivative of ψ at x_0 in direction $v \in \mathbb{R}^n$,

$$\psi^\circ(x_0; v) = \limsup_{y \to x_0\ t \downarrow 0} \frac{\psi(y + tv) - \psi(y)}{t}.$$

It should be noted that if ψ is continuously differentiable at x_0, then $\partial_c \psi(x_0) = \{\nabla \psi(x_0)\}$. Consequently, the sentences expressed in terms of ∂_c are generalizations of sentence that are expressed with gradient for C^1 functions. Moreover, it can be shown that if ψ is a locally Lipschitz function, $\partial_c \psi(x_0)$ is a non-empty, convex, and compact set. Also, $\psi^\circ(x; v)$ is a convex function with respect to v.

If ψ_1 and ψ_2 are locally Lipschitz functions, the following relationships should always hold:

$$\partial_c(\psi_1 + \psi_2)(x_0) \subseteq \partial_c\psi_1(x_0) + \partial_c\psi_2(x_0), \tag{4}$$

$$\partial_c(\lambda\psi_1)(x_0) = \lambda\partial_c\psi_1(x_0), \quad \forall\lambda \in \mathbb{R}, \tag{5}$$

$$\partial_c(\psi_1\psi_2)(x_0) \subseteq \psi_1(x_0)\partial_c\psi_2(x_0) + \psi_2(x_0)\partial_c\psi_1(x_0), \tag{6}$$

$$\psi^\circ(x; v) = \max\{\langle\xi, v\rangle : \xi \in \partial_c\psi(x_0)\}. \tag{7}$$

The following three approximate cones for $\emptyset \neq M \subseteq \mathbb{R}^n$ at $x_0 \in cl(M)$ will be requested in this article:

- The contingent cone $\Gamma(M, x_0)$,

$$\Gamma(M, x_0) := \Big\{v \in \mathbb{R}^n : \exists t_r \downarrow 0,\ \exists v_r \to v \text{ such that } x_0 + t_r v_r \in M \ \ \forall r \in \mathbb{N}\Big\}.$$

- The cone of feasible directions $\Upsilon(M, x_0)$,

$$\Upsilon(M, x_0) := \Big\{v \in \mathbb{R}^n : \exists\varepsilon > 0, \text{ such that } x_0 + tv \in M \quad \forall t \in]0, \varepsilon[\Big\}.$$

- The cone of attainable directions $\Omega(M, x_0)$,

$$\Omega(M, x_0) := \Big\{v \in \mathbb{R}^n : \forall t_r \downarrow 0,\ \exists v_r \to v \text{ such that } x_0 + t_r v_r \in M \quad \forall r \in \mathbb{N}\Big\}.$$

Remark 2.1 The cone of attainable directions of M at x_0 has an another representation as follows (see, e.g., [19]):

> $v \in \Omega(M, x_0)$, if and only if there exists a scalar $\delta > 0$ together with a mapping $\varphi : \mathbb{R} \to \mathbb{R}^n$ such that $\varphi(\beta) \in M$ for all $\beta \in]0, \delta[$, $\varphi(0) = x_0$, and $\lim_{\beta\downarrow 0} \frac{\varphi(\beta)-\varphi(0)}{\beta} = v$.

The relationship in these cones is as

$$cl\big(\Upsilon(M, x_0)\big) \subseteq \Omega(M, x_0) \subseteq \Gamma(M, x_0) \subseteq cl\big(\text{cone}\,(\Gamma(M, x_0))\big). \tag{8}$$

Note that if the continuously differentiable function $\psi(\cdot)$ attaints its minimum on $M \subseteq \mathbb{R}^n$ at $x_0 \in M$, then:

$$-\nabla\psi(x_0) \in N_F(M, x_0), \tag{9}$$

where $N_F(M, x_0)$ denotes the Fréchet normal cone of M at x_0, which is defined as $N_F(M, x_0) := (\Gamma(M, x_0))^-$.

3 Constraint Qualifications

Let $I := \{1, \ldots, m\}$, and for each $i \in I$ the functions G_i and H_i are locally Lipschitz from $\mathbb{R}^n$ to $\mathbb{R}^*$. In this section, we introduce and compare several constraint qualifications (CQ) for the following *system with vanishing constraints* (SVC in short):

$$\Pi := \begin{cases} H_i(x) \geq 0, & i \in I, \\ H_i(x)G_i(x) \leq 0, & i \in I. \end{cases}$$

The feasible set of Π is assumed to be non-empty, i.e.,

$$S := \{x \in \mathbb{R}^n : H_i(x) \geq 0 \text{ and } H_i(x)G_i(x) \leq 0 \text{ for all } i \in I\} \neq \emptyset.$$

Considering a feasible point $\hat{x} \in S$ (this point will be fixed throughout this section), we define following index sets:

$$\begin{aligned} I_{+0} &:= \{i \in I : H_i(\hat{x}) > 0,\ G_i(\hat{x}) = 0\}, \\ I_{+-} &:= \{i \in I : H_i(\hat{x}) > 0,\ G_i(\hat{x}) < 0\}, \\ I_{0+} &:= \{i \in I : H_i(\hat{x}) = 0,\ G_i(\hat{x}) > 0\}, \\ I_{00} &:= \{i \in I : H_i(\hat{x}) = 0,\ G_i(\hat{x}) = 0\}, \\ I_{0-} &:= \{i \in I : H_i(\hat{x}) = 0,\ G_i(\hat{x}) < 0\}. \end{aligned}$$

For given $I_\star \subseteq I$ and $\theta \in \{G, H\}$, let

$$\sigma^{\theta}_{I_*} := \bigcup_{i \in I_*} \partial_c \theta_i(\hat{x}).$$

Notice, if $I_{00} = \emptyset$, then Π is locally equivalent with the following system:

$$\Pi^* := \begin{cases} H_i(x) = 0, & G_i(x) > 0, & i \in I_{0+}, \\ H_i(x) \geq 0, & G_i(x) \leq 0, & i \in I_{0-} \cup I_{+-} \cup I_{+0}\ . \end{cases}$$

Since Π^* has the classic form of nonlinear systems, definition of CQs for it is so easy. But, when $I_{00} \neq \emptyset$, there is no any classic system that is equivalent to Π. Thus, we consider the *tightened system* for Π, which is a classic system that its feasible set is contained in S. For this purpose, it is easy to see that the following system is a tightened system for Π when $I_{00} = \emptyset$:

$$\Pi_* := \begin{cases} H_i(x) = 0, & G_i(x) \geq 0, & i \in I_{0+}, \\ H_i(x) \geq 0, & G_i(x) \leq 0, & i \in I_{0-} \cup I_{+-} \cup I_{+0}\ . \end{cases}$$

To add the functions H_i and G_i for $i \in I_{00}$ to Π_*, we can add them to the first line or to the second line of it. Thus, we achieve the following tightened systems:

$$\Pi_1 := \begin{cases} H_i(x) = 0, & G_i(x) \geq 0, & i \in I_{0+} \cup I_{00}, \\ H_i(x) \geq 0, & G_i(x) \leq 0, & i \in I_{0-} \cup I_{+-} \cup I_{+0}, \end{cases}$$

$$\Pi_2 := \begin{cases} H_i(x) = 0, & G_i(x) \geq 0, & i \in I_{0+}, \\ H_i(x) \geq 0, & G_i(x) \leq 0, & i \in I_{0-} \cup I_{+-} \cup I_{+0} \cup I_{00}\ . \end{cases}$$

Considering Π_1 and Π_2, the corresponding linearized cones at $\hat{x}$ are, respectively, given by $\mathcal{L}_1^-$ and $\mathcal{L}_2^-$, in which

$$\mathcal{L}_1 := \sigma^H_{I_{0+}\cup I_{00}} \cup \left(-\sigma^H_{I_{0+}\cup I_{00}}\right) \cup \left(-\sigma^H_{I_{0-}}\right) \cup \left(\sigma^G_{I_{+0}\cup I_{00}}\right),$$
$$\mathcal{L}_2 := \sigma^H_{I_{0+}} \cup \left(-\sigma^H_{I_{0+}}\right) \cup \left(-\sigma^H_{I_{0-}\cup I_{00}}\right) \cup \left(\sigma^G_{I_{+0}\cup I_{00}}\right).$$

Following [1,10,11], we also define the linearized cone $\mathcal{L}^-$ as

$$\mathcal{L} := \sigma^H_{I_{0+}} \cup \left(-\sigma^H_{I_{0+}}\right) \cup \left(-\sigma^H_{I_{0-}\cup I_{00}}\right) \cup \left(\sigma^G_{I_{+0}}\right).$$

Now, we can define some new version of Abadie, Khun–Tucker, Guignard, Zangwill, Mangasarian–Fromovitz, and linear independent CQs in extension of the CQs that are introduced in [1,2,10,11,13–15]. We also introduce some new useful CQs for Π.

Definition 3.1 We say that Π satisfies

- the MPVC-LICQ at $\hat{x}$ if Π_1 satisfies the LICQ at $\hat{x}$, i.e.,

$$0 \in \sum_{i\in I_{0+}\cup I_{0-}\cup I_{00}} \alpha_i \partial_c H_i(\hat{x}) + \sum_{i\in I_{00}\cup I_{+0}} \beta_i \partial_c G_i(\hat{x}) \Rightarrow \alpha_i = 0 \quad (\forall i \in I_{0+}\cup I_{0-}\cup I_{00}),\ \beta_i = 0 \quad (\forall i \in I_{00}\cup I_{+0}).$$

- the MPVC-MFCQ at $\hat{x}$ if Π_1 satisfies the MFCQ at $\hat{x}$, i.e.,
  (i) $0 \in \sum_{i\in I_{00}\cup I_{0+}} \alpha_i \partial_c H_i(\hat{x}) \Rightarrow \alpha_i = 0, \qquad \forall i \in I_{00}\cup I_{0+}$,
  (ii) $\left(\sigma^H_{I_{00}\cup I_{0+}}\right)^\perp \cap \left(-\sigma^H_{I_{0-}}\right)^s \cap \left(\sigma^G_{I_{+0}\cup I_{00}}\right)^s \neq \emptyset$.
- the $\widehat{GCQ}$ at $\hat{x}$ if

$$\mathcal{L}_1^- \subseteq cl\big(\text{conv}\left(\Gamma(S,\hat{x})\right)\big).$$

- the VC-MFCQ at $\hat{x}$ if
  (i) $0 \in \sum_{i\in I_{00}\cup I_{0+}} \alpha_i \partial_c H_i(\hat{x}) \Rightarrow \alpha_i = 0, \qquad \forall i \in I_{00}\cup I_{0+}$,
  (ii) $\left(\sigma^H_{I_{00}\cup I_{0+}}\right)^\perp \cap \left(-\sigma^H_{I_{0-}}\right)^s \cap \left(\sigma^G_{I_{+0}}\right)^s \neq \emptyset$.
- the VC-GCQ at $\hat{x}$ if

$$\mathcal{L}_2^- \subseteq cl\big(\text{conv}\left(\Gamma(S,\hat{x})\right)\big).$$

- the weak GCQ (resp., ACQ, KTCQ, ZCQ), denoted shortly by WGCQ (resp., WACQ, WKTCQ, WZCQ), at $\hat{x}$ if

$$\mathcal{L}^- \subseteq cl\big(\text{conv}\left(\Gamma(S,\hat{x})\right)\big)\ \Big(\text{resp., } \subseteq \Gamma(S,\hat{x}),\ \subseteq \Omega(S,\hat{x}),\ \subseteq cl\big(\Upsilon(S,\hat{x})\big)\Big).$$

Remark 3.1

(i) Since the feasible set of Π_2 (resp. Π_1) is included in S and is not necessarily equal with it, therefore VC-GCQ (resp. $\widehat{GCQ}$) for Π is not equivalent to GCQ for Π_2 (resp. Π_1). Note that GCQ for Π_2 (resp. Π_1) implies the VC-GCQ (resp. $\widehat{GCQ}$) for Π.

(ii) There are no implication relations between the VC-MFCQ for Π and the MFCQ for Π_2 at $\hat{x}$. Note that the MFCQ for Π_2 is irrelevant for us and we did not consider.

(iii) Unlike the Ref. [1], the inclusion $\mathcal{L}^- \subseteq cl\big(\text{conv}\,\big(\Gamma(S,\hat{x})\big)\big)$ is not named GCQ in above definition. The reason of this naming is that the classic GCQ for Π is formalized as

$$\Big(-\sigma^H_{I_{0+}\cup I_{0-}\cup I_{00}}\Big)^- \cap \Bigg(\bigcup_{i\in I_{+0}\cup I_{0+}\cup I_{0-}\cup I_{00}} \partial_c\big(H_iG_i\big)(\hat{x})\Bigg)^- \subseteq cl\big(\text{conv}\,\big(\Gamma(S,\hat{x})\big)\big). \tag{10}$$

On the other hand, owing to

$$\partial_c(G_iH_i)(\hat{x}) \subseteq G_i(\hat{x})\partial_c H_i(\hat{x}) + H_i(\hat{x})\partial_c G_i(\hat{x}),$$

we conclude that:

$$\begin{aligned}
&\Big(-\sigma^H_{I_{0+}\cup I_{0-}\cup I_{00}}\Big)^- \cap \Bigg(\bigcup_{i\in I_{+0}\cup I_{0+}\cup I_{0-}\cup I_{00}} \partial_c\big(H_iG_i\big)(\hat{x})\Bigg)^- \\
&\supseteq \bigcap_{i\in I_{0+}} \Big[\Big(-\partial_c H_i(\hat{x})\Big)^- \cap \Big(G_i(\hat{x})\partial_c H_i(\hat{x})\Big)^-\Big] \cap \bigcap_{i\in I_{00}} \Big(-\partial_c H_i(\hat{x})\Big)^- \cap \\
&\quad \bigcap_{i\in I_{0-}} \Big[\Big(-\partial_c H_i(\hat{x})\Big)^- \cap \Big(G_i(\hat{x})\partial_c H_i(\hat{x})\Big)^-\Big] \cap \bigcap_{i\in I_{+0}} \Big(H_i(\hat{x})\partial_c G_i(\hat{x})\Big)^- \\
&= \bigcap_{i\in I_{0+}} \Big(-\partial_c H_i(\hat{x})\Big)^- \cap \bigcap_{i\in I_{0+}} \Big(\partial_c H_i(\hat{x})\Big)^- \cap \bigcap_{i\in I_{00}} \Big(-\partial_c H_i(\hat{x})\Big)^- \cap \\
&\quad \bigcap_{i\in I_{0-}} \Big(-\partial_c H_i(\hat{x})\Big)^- \cap \bigcap_{i\in I_{+0}} \Big(\partial_c G_i(\hat{x})\Big)^- = \mathcal{L}^-.
\end{aligned}$$

Therefore, $\mathcal{L}^- \subseteq cl\big(\text{conv}\,\big(\Gamma(S,\hat{x})\big)\big)$ is weaker than (10). Note that if the functions H_i and G_i are continuously differentiable, [1, Lemma 4] shows that equality holds in last inclusion, and so $\mathcal{L}^- \subseteq cl\big(\text{conv}\,\big(\Gamma(S,\hat{x})\big)\big)$ is equivalent to (10).

We recall that by adding the index set I_{00} as a whole to system Π_*, we attained to systems Π_1 and Π_2. Now, in another way, we add the set I_{00} to systems Π_* in which I_{00} has been divided to two parts. To this aim, we call the ordered pair (β_1, β_2), a partition of I_{00} and we write $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, if β_1 and β_2 are non-empty subsets of I_{00} such that $\beta_1 \cap \beta_2 = \emptyset$ and $\beta_1 \cup \beta_2 = I_{00}$. Motivated by [10], we add β_1 to the first line and β_2 to the second line of system Π_*, and we get the below system:

$$\Pi^{(\beta_1,\beta_2)} := \begin{cases} H_i(x) = 0, & G_i(x) \geq 0, \quad i \in I_{0+} \cup \beta_1, \\ H_i(x) \geq 0, & G_i(x) \leq 0, \quad i \in I_{0-} \cup I_{+-} \cup I_{+0} \cup \beta_2 . \end{cases}$$

The feasible set of $\Pi^{(\beta_1,\beta_2)}$ is denoted by $S^{(\beta_1,\beta_2)}$. Also, we consider the linearized cone $\left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^-$ at $\hat{x}$, in which

$$\left(\mathcal{L}^{(\beta_1,\beta_2)}\right) := \left(\sigma^H_{I_{0+}\cup\beta_1}\right) \cup \left(-\sigma^H_{I_{0+}\cup\beta_1}\right) \cup \left(-\sigma^H_{I_{0-}\cup\beta_2}\right) \cup \left(\sigma^G_{I_{+0}\cup\beta_2}\right).$$

Following [10,11], we introduce another linearized cone for Π as $\mathcal{L}^{\triangle} := \mathcal{L}^- \cap \mathcal{T}$, where

$$\mathcal{T} := \{d \in \mathbb{R}^n : H_i^\circ(\hat{x}; d)\, G_i^\circ(\hat{x}; d) \leq 0, \ \ i \in I_{00}\}.$$

It is worth mentioning that the product of convex functions $d \to H_i^\circ(\hat{x}; d)$ and $d \to G_i^\circ(\hat{x}; d)$ is, in general, non-convex. So, unlike to $\mathcal{L}^-$, $\mathcal{L}_1^-$, $\mathcal{L}_2^-$, and $\left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^-$, the linearized cone $\mathcal{L}^{\triangle}$ is not convex.

Since the proof of following lemma is just similar to the proof of [11, Lemma 2.4], we do not repeat it.

Lemma 3.1 *The following equality is always true:*

$$\Gamma(S, \hat{x}) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \Gamma\left(S^{(\beta_1,\beta_2)}, \hat{x}\right).$$

Lemma 3.2 *The following inclusion holds true:*

$$\bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^- \subseteq \mathcal{L}^{\triangle}.$$

Proof Assume that $d \in \left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^-$. Then,

$$d \in \left(\sigma^H_{I_{0+}}\right)^\perp \cap \left(\sigma^H_{\beta_1}\right)^\perp \cap \left(-\sigma^H_{I_{0-}}\right)^- \cap \left(-\sigma^H_{\beta_2}\right)^- \cap \left(\sigma^G_{I_{+0}}\right)^- \cap \left(\sigma^G_{\beta_2}\right)^-.$$

It is enough to prove that:

$$d \in \left(-\sigma^H_{I_{00}}\right)^- \cap \mathcal{T}.$$

From $d \in \left(\sigma^H_{\beta_1}\right)^\perp \cap \left(-\sigma^H_{\beta_2}\right)^-$, it can be concluded that $d \in \left(-\sigma^H_{I_{00}}\right)^-$.

From $d \in \left(-\sigma^H_{\beta_2}\right)^-$, we get $\langle d, s\rangle \geq 0$ for each $i \in \beta_2$ and $s \in \sigma^H_{\beta_2}$. Therefore, $H_i^\circ(\hat{x}; d) \geq 0$ for each $i \in \beta_2$. Similarly, from $d \in (\sigma^G_{\beta_2})^-$ we can conclude that $G_i^\circ(\hat{x}; d) \leq 0$ for each $i \in \beta_2$. Hence,

$$H_i^\circ(\hat{x}; d)\, G_i^\circ(\hat{x}; d) \leq 0, \quad \forall i \in \beta_2. \tag{11}$$

Also, from $d \in (\sigma^H_{\beta_1})^\perp$, we deduce that $H_i^\circ(\hat{x}; d) = 0$ for each $i \in \beta_1$. This equality and (11) imply that:

$$H_i^\circ(\hat{x}; d)\, G_i^\circ(\hat{x}; d) \leq 0, \quad i \in \beta_1 \cup \beta_2 = I_{00} \implies d \in \mathcal{T},$$

and the lemma is proved. □

In [11, lemma 2.4], the remarkable fact was shown that, if functions H_i and G_i are smooth, then equality holds in Lemma 3.2. The following example shows that inclusion of Lemma 3.2 can be strict in non-smooth case.

Example 3.1 Take in Π,

$$\begin{aligned} H_1(x_1, x_2) &= -x_2^2, \qquad & G_1(x_1, x_2) &= |x_1| - x_2,\\ H_2(x_1, x_2) &= x_1^2, \qquad & G_2(x_1, x_2) &= x_1 - |x_2|. \end{aligned}$$

It is easy to see that the feasible set of this system is $S = \mathbb{R}_- \times \{0\}$. Considering $\hat{x} = (0, 0)$, we deduce that $I_{00} = \{1, 2\}$ and $\mathcal{P}(I_{00}) = \big\{(\beta_1, \beta_2),\ (\beta_2, \beta_1)\big\}$, in which $\beta_1 = \{1\}$ and $\beta_2 = \{2\}$. By a short calculation, we get,

$$\begin{aligned} &\partial_c H_1(\hat{x}) = \{(0, 0)\}, \quad H_1^0(\hat{x}; d) = 0,\\ &\partial_c H_2(\hat{x}) = \{(0, 0)\}, \quad H_2^0(\hat{x}; d) = 0,\\ &\partial_c G_1(\hat{x}) = [-1, 1] \times \{-1\}, \quad G_1^0(\hat{x}; d) = |d_1| - d_2,\\ &\partial_c G_2(\hat{x}) = \{1\} \times [-1, 1], \quad G_2^0(\hat{x}; d) = d_1 - |d_2|. \end{aligned}$$

Thus, we conclude that

$$\begin{aligned} &\left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^- = \left(\sigma^H_{\beta_1}\right)^\perp \cap \left(-\sigma^H_{\beta_2}\right)^- \cap \left(\sigma^G_{\beta_2}\right)^- = \mathbb{R}_- \times \{0\},\\ &\left(\mathcal{L}^{(\beta_2,\beta_1)}\right)^- = \left(\sigma^H_{\beta_2}\right)^\perp \cap \left(-\sigma^H_{\beta_1}\right)^- \cap \left(\sigma^G_{\beta_1}\right)^- = \{0\} \times \mathbb{R}_+,\\ &\mathcal{L}^- = \left(-\sigma^H_{I_{00}}\right)^- = \mathbb{R} \times \mathbb{R},\\ &\mathcal{T} = \{d \in \mathbb{R}^2 :\ 0 \times G_i^0(\hat{x}; d) \leq 0,\ \ i = 1, 2\} = \mathbb{R} \times \mathbb{R}. \end{aligned}$$

Therefore,

$$\begin{aligned} \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^- &= (\mathbb{R}_- \times \{0\}) \cup (\{0\} \times \mathbb{R}_+)\\ &= \mathbb{R}_- \times \mathbb{R}_+ \subsetneqq \mathbb{R} \times \mathbb{R} = \mathcal{L}^- \cap \mathcal{T} = \mathcal{L}^\triangle. \end{aligned}$$

Corollary 3.1 *The following assertion is valid:*

$$\Gamma(S, \hat{x}) \subseteq \mathcal{L}^{\triangle}.$$

Proof According to Lemmas 3.1 and 3.2, we can conclude that:

$$\Gamma(S, \hat{x}) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \Gamma\left(S^{(\beta_1,\beta_2)}, \hat{x}\right) \overset{(*)}{\subseteq} \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \Gamma\left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^{-} \subseteq \mathcal{L}^{\triangle},$$

in which the inclusion $(*)$ always holds (see, e.g., [19]). □

The last corollary leads us to the following definition.

Definition 3.2 We say that Π satisfies the MPVC-GCQ (resp., MPVC-ACQ, MPVC-KTCQ, MPVC-ZCQ) at $\hat{x}$ if

$$\mathcal{L}^{\triangle} \subseteq cl\big(\text{conv}\left(\Gamma(S, \hat{x})\right)\big) \Big(\text{resp.,} \subseteq \Gamma(S, \hat{x}), \subseteq \Omega(S, \hat{x}), \subseteq cl\big(\Upsilon(S, \hat{x})\big)\Big).$$

Now, we can state the main theorem of this section.

Theorem 3.1 *The following diagram summarizes the relationships between the defined CQs:*

$$\begin{array}{ccccccc}
MPVC\text{-}LICQ & & & & & & \\
{\scriptstyle (a)}\downarrow & & & & & & \\
MPVC\text{-}MFCQ & \overset{(b)}{\longrightarrow} & VC\text{-}MFCQ & & & & \\
 & & {\scriptstyle (c)}\downarrow & {\scriptstyle (d)}\searrow & & & \\
 & & WKTCQ & \overset{(e)}{\longleftarrow} & WZCQ & & \\
 & & {\scriptstyle (f)}\downarrow & {\scriptstyle (g)}\searrow & & {\scriptstyle (h)}\searrow & \\
 & & WACQ & & MPVC\text{-}KTCQ & \overset{(i)}{\longleftarrow} & MPVC\text{-}ZCQ \\
 & & {\scriptstyle (j)}\downarrow & {\scriptstyle (k)}\searrow & {\scriptstyle (l)}\downarrow & & \\
 & & WGCQ & & MPVC\text{-}ACQ & & \\
 & & & {\scriptstyle (m)}\searrow & {\scriptstyle (n)}\downarrow & & \\
\widehat{GCQ} & \overset{(q)}{\longleftarrow} & VC\text{-}GCQ & \overset{(p)}{\longleftarrow} & MPVC\text{-}GCQ, & &
\end{array} \tag{12}$$

where (d) holds when $I_{00} = \emptyset$ and H_i functions are linear for $i \in I_{0+}$, and implication (c) holds if $I_{00} = \emptyset$ and H_i functions for $i \in I_{0+}$ are continuously differentiable at $\hat{x}$.

Proof

(a) and (b): Straightforward.
(e), (f), (j), (i), (l) and (n): The conclusions are immediate from (8).
(m), (h), (g) and (k): The implications are straightforward consequences of $\mathcal{L}^{\triangle} \subseteq \mathcal{L}^{-}$.

(p): It is enough to prove that $\mathcal{L}_2^- \subseteq \mathcal{L}^\triangle$. To this end, let $d \in \mathcal{L}_2^- = \left(\sigma^H_{I_{0+}}\right)^\perp \cap \left(-\sigma^H_{I_{0-}\cup I_{00}}\right)^- \cap \left(\sigma^G_{I_{+0}\cup I_{00}}\right)^-$. From $d \in (-\sigma^H_{I_{00}})^-$ and $d \in (\sigma^G_{I_{00}})^-$, we can conclude that:

$$\begin{aligned} \langle d, \zeta \rangle &\le 0, \quad \forall \zeta \in \sigma^G_{I_{00}} \implies G_i^\circ(\hat{x}; d) \le 0, \quad i \in I_{00}, \\ \langle d, \xi \rangle &\ge 0, \quad \forall \xi \in \sigma^H_{I_{00}} \implies H_i^\circ(\hat{x}; d) \ge 0, \quad i \in I_{00}. \end{aligned}$$

Therefore, for each $i \in I_{00}$ we have:

$$H_i^\circ(\hat{x}; d)\, G_i^\circ(\hat{x}; d) \le 0 \implies d \in \mathcal{T}.$$

This means $d \in \mathcal{L}^\triangle$ which gives the assertion.
(q): It is direct consequence of inclusion $\mathcal{L}_2 \subseteq \mathcal{L}_1$.
(d): Let

$$d \in \left(\sigma^H_{I_{0+}}\right)^\perp \cap \left(-\sigma^H_{I_{0-}}\right)^s \cap \left(\sigma^G_{I_{+0}}\right)^s. \tag{13}$$

Assume that $H_i(x) = a_i x$ for each $i \in I_{0+}$. From $d \in \left(\sigma^H_{I_{0+}}\right)^\perp = \{a_i : i \in I_{0+}\}^\perp$ and $a_i\hat{x} = 0$ (since $H_i(\hat{x}) = 0$ for all $i \in I_{0+}$), we get:

$$H_i(\hat{x} + td) = a_i\hat{x} + ta_i d = 0, \quad \forall i \in I_{0+}. \tag{14}$$

The definition of Clarke directional derivative and $d \in \left(\sigma^G_{I_{+0}}\right)^s$ imply that for some $\delta_1 > 0$ we have:

$$G_i^0(\hat{x}; v) < 0 \Rightarrow G_i(\hat{x} + td) - \underbrace{G_i(\hat{x})}_{=0} < 0, \;\; \forall t \in]0, \delta_1[, \;\; \forall i \in I_{+0}. \tag{15}$$

Put $\widehat{H}_i(x) := -H_i(x)$ for $i \in I_{0-}$. According to $d \in \left(-\sigma^H_{I_{0-}}\right)^s$, we conclude that

$$\left\langle d, \xi \right\rangle < 0, \quad \forall \xi \in \partial_c \widehat{H}_i(\hat{x}), \; \forall i \in I_{0-}.$$

This means $\widehat{H}_i^0(\hat{x}; d) < 0$, and similar to (15), we find a scalar $\delta_2 > 0$ such that for all $t \in]0, \delta_2[$ and $i \in I_{0-}$ one has $\widehat{H}_i(\hat{x} + td) < 0$. Thus,

$$H_i(\hat{x} + td) > 0, \quad \forall i \in I_{0-}, \; \forall t \in]0, \delta_2[. \tag{16}$$

Taking $\delta := \min\{\delta_1, \delta_2\} > 0$, from (14)-(16) we deduce that $\hat{x} + td \in S$ for each $t \in]0, \delta[$. This means that $d \in \Upsilon(S, \hat{x})$, and so, with regard to (13), we get:

$$\left(\sigma^H_{I_{0+}}\right)^\perp \cap \left(-\sigma^H_{I_{0-}}\right)^s \cap \left(\sigma^G_{I_{+0}}\right)^s \subseteq \Upsilon(S, \hat{x}).$$

Therefore, one has:

$$\begin{aligned}\mathcal{L}^- &= \left(\sigma^H_{I_{0+}}\right)^{\perp} \cap \left(-\sigma^H_{I_{0-}}\right)^{-} \cap \left(\sigma^G_{I_{+0}}\right)^{-} \\ &= cl\Big(\left(\sigma^H_{I_{0+}}\right)^{\perp} \cap \left(-\sigma^H_{I_{0-}}\right)^{s} \cap \left(\sigma^G_{I_{+0}}\right)^{s}\Big) \subseteq cl\big(\Upsilon(S,\hat{x})\big) \subseteq \Gamma(S,\hat{x}).\end{aligned}$$

(c): Let

$$d \in \{\nabla H_i(\hat{x}) \colon i \in I_{0+}\}^{\perp} \cap \left(-\sigma^H_{I_{0-}}\right)^{s} \cap \left(\sigma^G_{I_{+0}}\right)^{s}.$$

According to [21, Theorem F], and recalling the linearly independence of $\{\nabla H_i(\hat{x}) \colon i \in I_{0+}\}$, we can find a neighborhood U of $\hat{x}$ and a function $\psi : U \longrightarrow \mathbb{R}^n$ such that ψ is continuous on U and differentiable at $\hat{x}$ with

$$\begin{aligned}&\psi(\hat{x}) = 0, \qquad \nabla\psi(\hat{x}) = 0, \\ &H_i\big(x + \psi(x)\big) = \big\langle \nabla H_i(\hat{x}), x - \hat{x}\big\rangle, \qquad \forall i \in I_{0+} \quad \forall x \in U.\end{aligned} \tag{17}$$

For each $\alpha \in \mathbb{R}$ with $\hat{x} + \alpha d \in U$, let $\varphi(\alpha) := \hat{x} + \alpha d + \psi(\hat{x} + \alpha d)$. Owing to $d \in \{\nabla H_i(\hat{x}) \colon i \in I_{0+}\}^{\perp}$, there exists $\varepsilon > 0$ such that for each $\eta \in]0, \varepsilon[$ one has:

$$H_i\big(\varphi(\eta)\big) = H_i\big(\hat{x} + \eta d + \psi(\hat{x} + \eta d)\big) = \eta\big\langle \nabla H_i(\hat{x}), d\big\rangle = 0, \quad i \in I_{0+}. \tag{18}$$

Taking $\widehat{H}_i(x) := -H_i(x)$ for $i \in I_{0-}$, and according to $d \in \left(-\sigma^H_{I_{0-}}\right)^{s} = \left(\bigcup_{i\in I_{0-}} \partial_c \widehat{H}_i(\hat{x})\right)^{s}$, we get $\widehat{H}_i^0(\hat{x}; d) < 0$. Consequently, there exist $\varepsilon_1 > 0$ and $\beta > 0$ such that:

$$\widehat{H}_i(\hat{x} + td) - \overbrace{\widehat{H}_i(\hat{x})}^{=0} \leq -\beta t, \quad \forall t \in]0, \varepsilon_1[. \tag{19}$$

On the other hand, if we denote the Lipschitz constance of $\widehat{H}_i(\cdot)$ around $\hat{x}$ by ρ_i, then there exists $\varepsilon_2 > 0$ such that:

$$\widehat{H}_i(\hat{x} + t\widetilde{d}) - \widehat{H}_i(\hat{x} + td) \leq \rho t \|\widetilde{d} - d\|, \quad \forall t \in]0, \varepsilon_2[, \ \forall \widetilde{d} \in \mathbb{R}^n, \tag{20}$$

where $\rho := \max\{\rho_i \colon i \in I_{0-}\}$. Since

$$0 = \nabla\psi(\hat{x}) = \lim_{d\to 0} \frac{\psi(\hat{x} + td) - \overbrace{\psi(\hat{x})}^{=0}}{t} = \lim_{d\to 0} \frac{\psi(\hat{x} + td)}{t}, \tag{21}$$

we find a positive scalar $\varepsilon_3 > 0$ such that:

$$\left\| \frac{\psi(\hat{x} + td)}{t} \right\| < \frac{\beta}{2\rho}, \quad \forall t \in]0, \varepsilon_3[. \tag{22}$$

Taking $\eta \in]0, \bar{\varepsilon}[$ in which $\bar{\varepsilon} := \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$, we conclude that for all $i \in I_{0-}$ one has

$$\begin{aligned}\widehat{H}_i\big(\varphi(\eta)\big) &= \widehat{H}_i\left(\hat{x} + \eta\left(d + \frac{\psi(\hat{x}+\eta d)}{\eta}\right)\right) - \widehat{H}_i(\hat{x}+\eta d) + \widehat{H}_i(\hat{x}+\eta d)\\ &\overset{(20)}{\leq} \rho\eta \left\|\frac{\psi(\hat{x}+\eta d)}{\eta}\right\| + \widehat{H}_i(\hat{x}+\eta d)\\ &\overset{(22)}{\leq} \rho\eta\,\frac{\beta}{2\rho} + \widehat{H}_i(\hat{x}+\eta d) \overset{(19)}{\leq} \frac{\eta\beta}{2} - \eta\beta = -\frac{\eta\beta}{2} < 0,\end{aligned}$$

which ensure that

$$H_i\big(\varphi(\eta)\big) < 0, \qquad \forall \eta \in]0, \bar{\varepsilon}[,\ \forall i \in I_{0-}. \tag{23}$$

By $d \in \left(\sigma^G_{I_{+0}}\right)^s$ and the latter proof, we can find a scalar $\tilde{\varepsilon} > 0$ such that:

$$G_i\big(\varphi(\eta)\big) < 0, \qquad \forall \eta \in]0, \tilde{\varepsilon}[,\ \forall i \in I_{+0}.$$

The last inequality together (18) and (23) implies that for all $\eta \in \big]0, \min\{\varepsilon, \bar{\varepsilon}, \tilde{\varepsilon}\}\big[$ we have $\varphi(\eta) \in S$. Also,

$$\lim_{\eta\to 0} \frac{\varphi(\eta)-\varphi(0)}{\eta} = \lim_{\eta\to 0} \frac{\hat{x}+\eta d+\psi(\hat{x}+\eta d)-\hat{x}-\psi(\hat{x})}{\eta} \overset{(21)}{=} 0,$$

and so $d \in \Omega(S, \hat{x})$.

Thus, we proved that:

$$\{\nabla H_i(\hat{x})\colon i \in I_{0+}\}^{\perp} \cap \left(-\sigma^H_{I_{0-}}\right)^s \cap \left(\sigma^G_{I_{+0}}\right)^s \subseteq \Omega(S, \hat{x}),$$

which ensure that

$$\mathcal{L}^- = cl\Big(\{\nabla H_i(\hat{x})\colon i \in I_{0+}\}^{\perp} \cap \left(-\sigma^H_{I_{0-}}\right)^s \cap \left(\sigma^G_{I_{+0}}\right)^s\Big) \subseteq cl\big(\Omega(S, \hat{x})\big) \subseteq \Omega(S, \hat{x}).$$

□

Example 3.2 Consider the following system,

$$\widehat{\Pi} := \begin{cases} x_2 \geq 0,\\ x_2(x_2 - |x_1|) \leq 0.\end{cases}$$

with following data:

$$H_1(x_1, x_2) = x_2, \quad G_1(x_1, x_2) = x_2 - |x_1|, \quad \hat{x} = (0,0), \quad I = \{1\}.$$

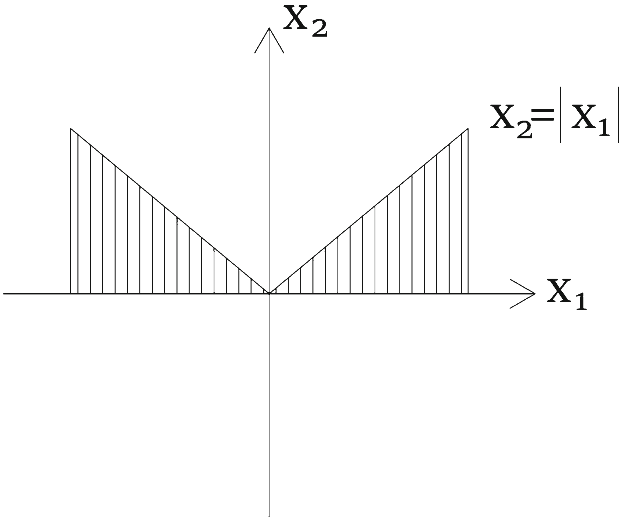

Fig. 1 Feasible set of $\widehat{\Pi}$

We have drawn the feasible set of $\widehat{\Pi}$ in Fig. 1. With a simple calculation, it could be seen that:

$$I_{00} = \{1\}, \quad \partial_c H_1(\hat{x}) = \{(0, 1)\}, \quad \partial_c G_1(\hat{x}) = [-1, 1] \times \{1\}.$$

The MPVC-LICQ is not satisfied at $\hat{x}$, since we can choose $\alpha_1 = 1$ and $\beta_1 = -1$ in the following inclusion,

$$0 \in \alpha_1\{(0, 1)\} + \beta_1\big([-1, 1] \times \{1\}\big).$$

According to

$$\begin{aligned} &0 \in \alpha_1\{(0, 1)\} \implies \alpha_1 = 0, \\ &\{(0, 1)\}^{\perp} \cap \Big([-1, 1] \times \{1\}\Big)^s = \big(\mathbb{R} \times \{0\}\big) \cap \big(\{0\}\times]-\infty, 0[\,\big) = \emptyset, \end{aligned} \tag{24}$$

we conclude that the MPVC-MFCQ is not satisfied at $\hat{x}$, while the implication (24) and equality $\{(0, 1)\}^{\perp} = \mathbb{R} \times \{(0, 0)\} \neq \emptyset$ show that the VC-MFCQ holds at $\hat{x}$. Thus, VC-MFCQ is strictly weaker than MPVC-MFCQ for non-smooth SVCs. Also,

$$\begin{aligned} &\mathcal{L}^- = \{-(0, 1)\}^- = \mathbb{R} \times \mathbb{R}_+, \\ &\mathcal{L}_1^- = \{(0, 1)\}^{\perp} \cap \Big([-1, 1] \times \{1\}\Big)^- = \big(\mathbb{R} \times \{0\}\big) \cap \big(\{0\} \times \mathbb{R}_-\big) = \{(0, 0)\}, \\ &\mathcal{L}_2^- = \{-(0, 1)\}^- \cap \Big([-1, 1] \times \{1\}\Big)^- = \{(0, 0)\}, \\ &\mathcal{L}^{\triangle} = \mathcal{L}^- \cap \{d \in \mathbb{R}^2 \colon d_2 \geq 0,\ d_2(d_2 - |d_1|) \leq 0\} = S. \end{aligned}$$

Above equalities show that WZCQ, WKTCQ, and WACQ fail, whereas WGCQ, MPVC-ZCQ, MPVC-KTCQ, MPVC-ACQ, MPVC-GCQ, VC-GCQ, and $\widehat{GCQ}$ hold at $\hat{x}$. Thus, the inverse implications of $[WZCQ \Rightarrow MPVC\text{-}ZCQ]$, $[WKTCQ \Rightarrow MPVC\text{-}KTCQ]$, $[WACQ \Rightarrow MPVC\text{-}ACQ]$, and $[WACQ \Rightarrow WGCQ]$ do not hold. Also, the assumption $I_{00} = \emptyset$ could not be eliminated in the implications $[VC\text{-}MFCQ \Rightarrow WKTCQ]$ and $[VC\text{-}MFCQ \Rightarrow WZCQ]$.

4 Stationarity Conditions

In this section, motivated by [12,14,15], we will be mainly concerned with the mathematical problem with vanishing constraints (MPVC) given as

$$\min \quad f(x) \qquad \text{s.t.} \quad x \in S, \tag{P}$$

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a continuously differentiable function and S is defined as previous section. It should be noted that the general form of a MPVC which has been considered in [1,2,10,11,13,16] includes inequality constraints $g_j(x) \leq 0,\ j \in J$ and equality constraints $h_t(x) = 0,\ t \in T$ for some finite index sets J and T. Since adding these constraints to problem (P) does not increase the technical problems of the issue and just prolongs the formulas, we ignore them and just deal with problem (P).

Now, we can state the KKT type necessary optimality condition for problem (P). The following theorem is non-smooth version of [1, Theorem 1].

Theorem 4.1 *Let $\hat{x}$ be a local solution of (P) such that WGCQ holds at $\hat{x}$. If* $\text{cone}(\mathcal{L})$ *is a closed cone, and the objective function $f(\cdot)$ is continuously differentiable at $\hat{x}$, then there exist coefficients λ_i^H and λ_i^G, for $i \in I$, such that:*

$$0 \in \nabla f(\hat{x}) + \sum_{i \in I} \Big(-\lambda_i^H \partial_c H_i(\hat{x}) + \lambda_i^G \partial_c G_i(\hat{x}) \Big), \tag{25}$$

and

$$\lambda_i^H = 0 \quad (i \in I_{+0} \cup I_{+-}), \quad \lambda_i^H \geq 0 \quad (i \in I_{0-} \cup I_{00}), \quad \lambda_i^H \text{ free} \quad (i \in I_{0+}), \tag{26}$$

$$\lambda_i^G = 0 \quad (i \in I_{0+} \cup I_{0-} \cup I_{00} \cup I_{+-}), \quad \lambda_i^G \geq 0 \quad (i \in I_{+0}). \tag{27}$$

Proof Owing to (9), WGCQ, bipolar theorem, and closedness of $\text{cone}(\mathcal{L})$, we conclude that

$$-\nabla f(\hat{x}) \in N_F(S, \hat{x}) = \big(\Gamma(S, \hat{x})\big)^- = \big(cl\big(\text{conv}\,(\Gamma(S, \hat{x}))\big)\big)^- \subseteq \mathcal{L}^{--} = cl\big(\text{cone}(\mathcal{L})\big) = \text{cone}(\mathcal{L}).$$

Considering the structure of convex cones, we can find nonnegative scalars

$$a_i \ (i \in I_{0+}), \quad b_i \ (i \in I_{0+}), \quad c_i \ (i \in I_{00} \cup I_{0-}), \quad d_i \ (i \in I_{+0}),$$

such that

$$-\nabla f(\hat{x}) \in \sum_{i\in I_{0+}} a_i \partial_c H_i(\hat{x}) + \sum_{i\in I_{0+}} -b_i \partial_c H_i(\hat{x}) + \sum_{i\in I_{00}\cup I_{0-}} -c_i \partial_c H_i(\hat{x}) + \sum_{i\in I_{+0}} d_i \partial_c G_i(\hat{x}).$$

Tacking

$$\lambda_i^H := \begin{cases} -(a_i - b_i), & \text{if } i \in I_{0+}, \\ c_i & \text{if } i \in I_{00} \cup I_{0-}, \\ 0 & \text{if } i \in I_{+-} \cup I_{+0}, \end{cases}$$
$$\lambda_i^G := \begin{cases} d_i & \text{if } i \in I_{+0}, \\ 0 & \text{if } i \in I_{+-} \cup I_{0+} \cup I_{0-} \cup I_{00}, \end{cases}$$

the result is proved. □

Conditions (25)–(27) were named the *KKT conditions* and *strongly stationary conditions* for (P) in [1,10], respectively. We will call them the strongly stationary conditions (SSC in short) too.

The following example shows that we cannot replace the smoothness condition of $f(\cdot)$ by its Lipschitzian condition in above theorem, and we cannot replace the $\nabla f(\hat{x})$ by $\partial_c f(\hat{x})$ in (25) too.

Example 4.1 Consider the optimization problem as

$$\min \ -x_2 + |x_1 - x_2| \ \text{ s.t. } \ x \in S = (\{0\} \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times \{0\}).$$

We can formalize this problem as a MPVC with following data,

$$H_1(x_1, x_2) = x_2, \quad H_2(x_1, x_2) = x_1, \quad G_1(x_1, x_2) = -x_1, \quad G_2(x_1, x_2) = -x_2,$$
$$f(x_1, x_2) = -x_2 + |x_1 - x_2|.$$

We observe that $\hat{x} = (0, 0)$ is an optimal solution of problem, and $I_{00} = \{1, 2\}$. Since

$$\mathcal{L}^- = \left(-\sigma^H_{I_{00}}\right)^- = (\{(-1, 0), (0, -1)\})^- = \mathbb{R}_+ \times \mathbb{R}_+ = \text{conv}(\Gamma(S, \hat{x})),$$

the WGCQ is satisfied at $\hat{x}$. It is easy to show that the cone$(\mathcal{L})$ is closed, and also, the below SSC type relations do not hold for any nonnegative scalars λ_1^H, λ_2^H, λ_1^G, and λ_2^G:

$$\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}\gamma\\-1-\gamma\end{pmatrix} - \lambda_1^H\begin{pmatrix}0\\1\end{pmatrix} - \lambda_2^H\begin{pmatrix}1\\0\end{pmatrix} + \lambda_1^G\begin{pmatrix}-1\\0\end{pmatrix} + \lambda_2^G\begin{pmatrix}0\\1\end{pmatrix},$$
$$\gamma \in [-1, 1], \qquad \lambda_1^G = \lambda_2^G = 0.$$

The following example shows the assumption of closedness of cone$(\mathcal{L})$ in Theorem 4.1 cannot be waived.

Example 4.2 Let $H_1(x_1, x_2) = 1$, and $G_1(x_1, x_2)$ is the support function of

$$Q := \{(-q, -q^2) \colon 0 \le q \le 1\}.$$

Thus, $S = \{(x_1, x_2) \in \mathbb{R}^2 \colon x_1 \ge 0, \quad x_1 + x_2 \ge 0\}$. Clearly, $\hat{x} = (0, 0)$ is the optimal solution of the problem $\min_{x \in S} f(x)$, in which $f(x_1, x_2) = x_1$. By a short calculation, we get $I_{+0} = \{1\}$, and

$$\begin{aligned}
&\mathcal{L} = \partial_c G_1(\hat{x}) = Q = \{(x_1, x_2) \in \mathbb{R}^2 \colon -1 < x_1 < 0,\ x_1 \le x_2 < -x_1^2\} \cup \{(0,0), (-1,-1)\},\\
&\mathcal{L}^- = S = cl\Big(\mathrm{conv}\big(\Gamma(S, \hat{x})\big)\Big),\\
&\mathrm{cone}\big(\mathcal{L}\big) = \{(x_1, x_2) \in \mathbb{R}^2 \colon x_1 \le x_2,\ x_1 < 0,\ x_2 < 0\} \cup \{(0,0)\}.
\end{aligned}$$

Therefore, the WGCQ is satisfied at $\hat{x}$, and $\mathrm{cone}\big(\mathcal{L}\big)$ is not closed. Obviously, following SSC type relations do not hold for each of the nonnegative scalars λ_1^H and λ_1^G:

$$(0, 0) \in \{(1, 0)\} - \lambda_1^H \{(0, 0)\} + \lambda_1^G Q, \quad \lambda_1^H = 0.$$

As diagram (12) shows, VC-GCQ is weaker than WGCQ. In [1, page 93], with an example shows that WGCQ is strictly stronger than VC-GCQ. As a result, it seems that one cannot get from VC-GCQ (resp. $\widehat{GCQ}$) to SSC, but it can reach to some weaker conditions which in future, we call them VC-stationary conditions (reap. weakly stationary conditions).

Since the proof of following theorem is similar to the proof of Theorem 4.1, we ignore it.

Theorem 4.2 *Suppose that $\hat{x}$ is a local solution of (P) and VC-GCQ (resp. $\widehat{GCQ}$) holds at $\hat{x}$. If* $\mathrm{cone}(\mathcal{L}_2)$ *(resp.* $\mathrm{cone}(\mathcal{L}_1)$*) be a closed cone, then there exist real coefficients λ_i^H and λ_i^G, for $i \in I$, which satisfy in (25) and*

$$\begin{array}{l}
\lambda_i^H = 0 \quad (i \in I_{+-} \cup I_{+0}), \quad \lambda_i^H \text{ free } (i \in I_{0+}), \quad \lambda_i^H \ge 0 \quad (i \in I_{0-} \cup I_{00}),\\
\lambda_i^G = 0 \quad (i \in I_{+-} \cup I_{0+} \cup I_{0-}), \quad \lambda_i^G \ge 0 \quad (i \in I_{+0} \cup I_{00}).
\end{array} \tag{28}$$

$$\left(\text{resp.} \begin{array}{l}
\lambda_i^H = 0 \quad (i \in I_{+-} \cup I_{+0}), \quad \lambda_i^H \text{ free } (i \in I_{0+} \cup I_{00}), \quad , \lambda_i^H \ge 0 \quad (i \in I_{0-}),\\
\lambda_i^G = 0 \quad (i \in I_{+-} \cup I_{0+} \cup I_{0-}), \quad \lambda_i^G \ge 0 \quad (i \in I_{+0} \cup I_{00}).
\end{array}\right) \tag{29}$$

As mentioned before, conditions (25) and (28) are refereed by *VC-stationary condition* (VCSC in brief), and conditions (25) and (29) are named *weakly stationary conditions* (shortly, WSC).

Theorem 4.3 *Assume that $\hat{x}$ is an optimal solution of (P), and MPVC-GCQ is satisfied at $\hat{x}$. If* $\mathrm{cone}\left(\mathcal{L}^{(\beta_1^*, \beta_2^*)}\right)$ *and* $\mathrm{cone}\left(\mathcal{L}^{(\beta_2^*, \beta_1^*)}\right)$ *are closed for some $(\beta_1^*, \beta_2^*) \in \mathcal{P}(I_{00})$, then the strongly stationary conditions hold at $\hat{x}$.*

Proof Considering the classic formula (9), MPVC-GCQ, Lemma 3.2, and bipolar theorem, one can see that

$$
\begin{aligned}
&-\nabla f(\hat{x}) \in \left(\Gamma(S,\hat{x})\right)^- \subseteq \left(\mathcal{L}^{\triangle}\right)^- \subseteq \left(\bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^-\right)^- \\
&= \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \left(\mathcal{L}^{(\beta_1,\beta_2)}\right)^{--} = \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} cl\Big(\text{cone}\left(\mathcal{L}^{(\beta_1,\beta_2)}\right)\Big) \\
&\subseteq cl\left(\text{cone}\left(\mathcal{L}^{(\beta_1^*,\beta_2^*)}\right)\right) \cap cl\left(\text{cone}\left(\mathcal{L}^{(\beta_2^*,\beta_1^*)}\right)\right) = \text{cone}\left(\mathcal{L}^{(\beta_1^*,\beta_2^*)}\right) \cap \text{cone}\left(\mathcal{L}^{(\beta_2^*,\beta_1^*)}\right).
\end{aligned}
\tag{30}
$$

On the other hand, according to the definition of $\mathcal{L}^{(\beta_1^*,\beta_2^*)}$ and $\mathcal{L}^{(\beta_2^*,\beta_1^*)}$ we have

$$
\begin{aligned}
\text{cone}\left(\mathcal{L}^{(\beta_1^*,\beta_2^*)}\right) = \bigcup \Bigg\{ &\sum_{i\in I_{0+}\cup\beta_1^*} -\lambda_i^H \partial_c H_i(\hat{x}) + \sum_{i\in I_{0-}\cup\beta_2^*} -\lambda_i^H \partial_c H_i(\hat{x}) \\
&+ \sum_{i\in I_{+0}\cup\beta_2^*} \lambda_i^G \partial_c G_i(\hat{x}) : \\
&\lambda_i^H \text{ free } \; i \in I_{0+}\cup\beta_1^*;\; \lambda_i^H \geq 0 \;\; i \in I_{0-}\cup\beta_2^* \\
&\lambda_i^G \geq 0 \;\; i \in I_{+0}\cup\beta_2^* \Bigg\},
\end{aligned}
$$

and

$$
\begin{aligned}
\text{cone}\left(\mathcal{L}^{(\beta_2^*,\beta_1^*)}\right) = \bigcup \Bigg\{ &\sum_{i\in I_{0+}\cup\beta_2^*} -\lambda_i^H \partial_c H_i(\hat{x}) + \sum_{i\in I_{0-}\cup\beta_1^*} -\lambda_i^H \partial_c H_i(\hat{x}) \\
&+ \sum_{i\in I_{+0}\cup\beta_1^*} \lambda_i^G \partial_c G_i(\hat{x})) : \\
&\lambda_i^H \text{ free } \; i \in I_{0+}\cup\beta_2^*;\; \lambda_i^H \geq \; 0\, i \in I_{0-}\cup\beta_1^*; \\
&\lambda_i^G \geq 0 \;\; i \in I_{+0}\cup\beta_1^* \Bigg\}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\text{cone}\left(\mathcal{L}^{(\beta_1^*,\beta_2^*)}\right) \cap \text{cone}\left(\mathcal{L}^{(\beta_2^*,\beta_1^*)}\right) = \bigcup \Bigg\{ \sum_{i\in I_{0+}\cup I_{0-}\cup I_{00}} -\lambda_i^H \partial_c H_i(\hat{x}) + \sum_{i\in I_{+0}} \lambda_i^G \partial_c G_i(\hat{x}) : \\
&\quad \lambda_i^H \text{ free } i \in I_{0+};\; \lambda_i^H \geq 0 \;\; i \in \beta_1^*\cup\beta_2^* = I_{00};\; \lambda_i^H \geq 0 \;\; i \in I_{0-};\;\; \lambda_i^G \geq 0 \;\; i \in I_{0+} \Bigg\} \\
&= \bigcup \Bigg\{ \sum_{i\in I} (-\lambda_i^H \partial_c H_i(\widehat{x}) + \lambda_i^G \partial_c G_i(\widehat{x})) : \lambda_i^H \text{ free } \;\; i \in I_{0+};\; \lambda_i^H = 0 \;\; i \in I_{+0}\cup I_{+-}; \\
&\quad \lambda_i^H \geq 0 \;\; i \in I_{00}\cup I_{0-};\; \lambda_i^G \geq 0 \;\; i \in I_{+0};\; \lambda_i^G = 0 \;\; i \in I_{0+}\cup I_{0-}\cup I_{00}\cup I_{+-} \Bigg\}.
\end{aligned}
\tag{31}
$$

Combining the virtues of (30) and (31) , the result is proved. □

As we saw in proof of Theorem 4.3,

$$\text{cone}(\mathcal{L}^{(\beta_1,\beta_2)}) \cap \text{cone}(\mathcal{L}^{(\beta_2,\beta_1)}) = \text{cone}(\mathcal{L}),$$

for each $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, and so, the closedness of $\text{cone}(\mathcal{L}^{(\beta_1,\beta_2)})$ and $\text{cone}(\mathcal{L}^{(\beta_2,\beta_1)})$ leads us to closedness of $\text{cone}(\mathcal{L})$, while closedness of $\text{cone}(\mathcal{L})$ does not lead to them. There are two different assumptions in getting the SSC:

($\mathfrak{a}$) WGCQ and closedness of $\text{cone}(\mathcal{L})$ (Theorem 4.1).
($\mathfrak{b}$) MPVC-GCQ and closedness of $\text{cone}(\mathcal{L}^{(\beta_1,\beta_2)})$ and $\text{cone}(\mathcal{L}^{(\beta_2,\beta_1)})$ (Theorem 4.3).

Since the MPVC-GCQ is weaker than WGCQ (see diagram (12)) and closedness condition in ($\mathfrak{b}$) is stronger than in ($\mathfrak{a}$), none of Theorems 4.1 and 4.3 are better than the other. It is worth mentioning that if H_i and G_i are continuously differentiable, their sub-differentials contain single element, and so, the closedness conditions in ($\mathfrak{a}$) and ($\mathfrak{b}$) automatically hold. Hence, as mentioned in [11], Theorem 4.3 is strictly better than Theorem 4.1 for MPVCs with smooth constraints.

5 Conclusions

Calculating the Clarke sub-differential of a function $\hbar : \mathbb{R}^n \to \mathbb{R}$ might be difficult in general. However, the functions considered in our work are locally Lipschitz and it empowers the numerical viability of our presented necessary conditions. This is due to an important corollary of the Rademacher's Theorem [18] which says: Given $\bar{x} \in \mathbb{R}^n$,

$$\partial_c \hbar(\bar{x}) = \text{conv}\left(\{\lim_{r\to\infty} \nabla f(x_r) : \ x_r \to \bar{x}, x_r \in Q\}\right), \tag{32}$$

in which Q is any full-measure subset of $\mathbb{R}^n$ that includes all the points where $\hbar$ is differentiable. According to (32), the gradient sampling approach [22] can be utilized to generate a suitable approximation of $\partial_c \hbar(\bar{x})$ and applying in our results. Here, we address this approach briefly. Considering an appropriate ball centered at $\bar{x}$, say $U_{\bar{x}}$, a large number of points in $U_{\bar{x}}$ are randomly generated. At the second steep, among randomly generated points, the points at which $\hbar$ is differentiable are selected. Assume that these points are $x_1, x_2, \ldots, x_k$. According to (32),

$$C_k := \text{conv}\left(\{\nabla \hbar(x_r) : r = 1, 2, \ldots, k\}\right)$$

is considered as an approximation of $\partial_c \hbar(\bar{x})$ and is applied in our results. The accuracy of this approximation can be improved by increasing k. See [22] for more details about gradient sampling techniques. To compute $\partial_c \hbar(\bar{x})$, another sampling method has been presented in [23]. It works based on ε-sub-differential introduced by Goldstein.

Acknowledgements The authors are very grateful to the referees for their constructive comments. Also, we express our thanks to Professor M. Soleimani-damaneh for suggesting the conclusions of the paper.

References

1. Achtziger, W., Kanzow, C.: Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications. Math. Program. **114**, 69–99 (2007)
2. Hoheisel, T., Kanzow, C.: First- and second-order optimality conditions for mathematical programs with vanishing constraints. Appl. Math. **52**, 495–514 (2007)
3. Ansari Ardali, A., Movahedian, N., Nobakhtian, S.: Optimality conditions for nonsmooth mathematical programs with equilibrium constraints, using convexificators. Optimization **65**, 67–85 (2016)
4. Kanzi, N., Nobakhtian, S.: Nonsmooth multiobjective semi-infinite problems with mixed constraints. Pac. J. Optim. **13**, 43–53 (2017)
5. Pandey, Y., Mishra, S.K.: Duality for nonsmooth optimization problems with equilibrium constraints, using convexificators. J. Optim. Theory Appl. **171**, 694–707 (2016)
6. Bigi, G., Pappalardo, M., Passacantando, M.: Optimization tools for solving equilibrium problems with nonsmooth data. J. Optim. Theory Appl. **171**, 887–905 (2016)
7. Chieu, N.H., Lee, G.M.: Constraint qualifications for mathematical programs with equilibrium constraints and their local preservation property. J. Optim. Theory Appl. **163**, 755–776 (2014)
8. Movahedian, N.: Bounded Lagrange multiplier rules for general nonsmooth problems and application to mathematical programs with equilibrium constraints. J. Global Optim. **67**, 829–850 (2017)
9. Luu, D.V.: Optimality condition for local efficient solutions of vector equilibrium problems via convexificators and applications. J. Optim. Theory Appl. **171**, 643–665 (2016)
10. Hoheisel, T., Kanzow, C.: Stationarity conditions for mathematical programs with vanishing constraints using weak constraint qualifications. J. Math. Anal. Appl. **337**, 292–310 (2008)
11. Hoheisel, T., Kanzow, C.: On the Abadie and Guignard constraint qualifications for mathematical programs with vanishing constraints. Optimization **58**, 431–448 (2009)
12. Hoheisel, T., Kanzow, C., Outrata, J.: Exact penalty results for mathematical programs with vanishing constraints. Nonlinear Anal. **72**, 2514–2526 (2010)
13. Achtziger, W., Hoheisel, T., Kanzow, C.: A smoothing-regularization approach to mathematical programs with vanishing constraints. Comput. Optim. Appl. **55**, 733–767 (2013)
14. Izmailov, A.F., Solodov, M.V.: Mathematical programs with vanishing constraints: optimality conditions, sensitivity, and a relaxation method. J. Optim. Theory Appl. **142**, 501–532 (2009)
15. Izmailov, A.F., Pogosyan, A.L.: Optimality conditions and Newton-type methods for mathematical programs with vanishing constraints. Comput. Math. Math. Phys. **49**, 1128–1140 (2009)
16. Mishra, S.K., Singh, V., Laha, V.: On duality for mathematical programs with vanishing constraints. Annal. Oper. Res. **243**, 249–272 (2016)
17. Mordukhovich, B.: Optimization and equilibrium problems with equilibrium constraints in infinite-dimensional spaces. Optimization **57**, 715–741 (2008)
18. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, Hoboken (1983)
19. Giorgi, G., Gwirraggio, A., Thierselder, J.: Mathematics of Optimization; Smooth and Nonsmooth cases. Elsevier, Amsterdam (2004)
20. Hiriart- Urruty, J.B., Lemarechal, C.: Convex Analysis and Minimization Algorithms, I & II. Springer, Berlin (1991)
21. Halkin, H.: Implicit functions and optimization problems without continuous differentiability of the data. SIAM J. Control **12**, 229–236 (1974)
22. Burke, J.V., Lewis, A.S., Overton, M.L.: Approximating subdifferentials by random sampling of gradients. Math. Oper. Res. **27**, 567–584 (2002)
23. Burke, J.V., Lewis, A.S., Overton, M.L.: A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. SIAM J. Optim. **15**, 751–779 (2005)