FIRST-ORDER APPROXIMATION OF STRONG VECTOR EQUILIBRIA WITH APPLICATION TO NONDIFFERENTIABLE CONSTRAINED OPTIMIZATION

AMOS UDERZO[1,*]

[1] Dipartimento di Matematica e Applicazioni, Università di Milano-Bicocca, Milano, ITALY

In respectful memory of Naum Zuselevich Shor (1937 - 2006)
esteemed explorer of NDOlandia

Abstract. Vector equilibrium problems are a natural generalization to the context of partially ordered spaces of the Ky Fan inequality, where scalar bifunctions are replaced with vector bifunctions. In the present paper, the local geometry of the strong solution set to these problems is investigated through its inner/outer conical approximations. More precisely, formulae for approximating the contingent cone to the set of strong vector equilibria are established, which are expressed via Bouligand derivatives of the bifunctions. These results are subsequently employed for deriving both necessary and sufficient optimality conditions for problems, whose feasible region is the strong solution set to a vector equilibrium problem, so they can be cast in mathematical programming with equilibrium constraints.
Keywords. Strong vector equilibrium; contingent cone; nondifferentiable optimization; generalized differentiation; subdifferential; mathematical programming with equilibrium constraint.
2020 Mathematics Subject Classification. 49J52, 49J53, 90C33.

1. INTRODUCTION

Given a mapping (vector-valued bifunction) $f:\mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R}^m$, with $\mathbb{R}^m$ being partially ordered by a (nontrivial) closed, convex and pointed cone $C\subset\mathbb{R}^m$, and a nonempty, closed set $K\subseteq\mathbb{R}^n$, by strong vector equilibrium problem the following problem is meant

$$(\mathrm{VEP})\qquad \text{find } x\in K \text{ such that } f(x,z)\in C,\quad \forall z\in K.$$

The set of all solutions (if any) to problem (VEP) will be denoted throughout the paper by $\mathcal{SE}$, namely

$$\mathcal{SE}=\bigcap_{z\in K} f^{-1}(\cdot,z)(C)\cap K, \tag{1.1}$$

and referred to as the set of strong vector equilibria. Clearly, strong vector equilibrium problems are a natural generalization of the well-known Ky Fan inequality to the more general context of

*Corresponding author.
E-mail address: amos.uderzo@unimib.it (A. Uderzo).
Received xx, x, xxxx; Accepted xx, x, xxxx.

©2022 Communications in Optimization Theory

partially ordered vector spaces. Similarly as their scalar counterpart, they provide a convenient format to treat in an unifying framework several different classes of problems, ranging from multicriteria optimization problems, vector Nash equilibrium problems, to vector variational inequalities and complementarity problems (see, for instance, [1, 2, 3, 5, 9, 10, 16]).

As for many problems formalized by traditional or generalized equations, for several purposes the mere knowledge of a single solution to (VEP) is not enough. Very often, once a strong vector equilibrium $\bar{x} \in \mathscr{SE}$ has been found (or shown to exist), one would need/aspire to glean insights into the behaviour of the set $\mathscr{SE}$ around $\bar{x}$. The fact that $\bar{x}$ may be an isolated element of $\mathscr{SE}$ or lie in the boundary or, instead, be an interior element of this set, might change dramatically the outcome of a further analysis, where the local geometry of $\mathscr{SE}$ around $\bar{x}$ does matter. On the other hand, finding all the solutions of (VEP) around $\bar{x}$ could be a task that one can hardly accomplish in many concrete cases. What is reasonably achievable sometimes is only a local approximation of $\mathscr{SE}$ near $\bar{x}$, yet suitable in specific circumstances. To mention one of them, with connection with the subject of the present paper, consider the successful approach to optimality conditions for constrained problems, where at a certain step an approximated representation of the feasible region already does the trick.

It is well known that in nonsmooth analysis tangent cones, working as a surrogate of derivative for sets, are the main tools for formalizing first-order (and beyond, if needed) approximations of sets. So the main aim of the present paper is to provide elements for a conical approximation of strong vector equilibria. It should be remarked that a difficulty in undertaking such a task comes from the fact that the set $\mathscr{SE}$ is not explicitly defined. Besides, if addressing this question through the reformulation of $\mathscr{SE}$ as in (1.1), classical results on the tangent cone representation of such sets as $f^{-1}(\cdot,z)(C)\cap K$, now at disposal in nonsmooth analysis as a modern development of the Lyusternik theorem (see [13, 15, 20]), seem not be readily exploitable because of the intersection over $z \in K$ appearing in (1.1).

In this context, the findings exposed in what follows are focussed on representing the contingent cone to $\mathscr{SE}$ at a given strong vector equilibrium $\bar{x}$, which is one of the most employed conical approximations in the literature devoted to variational analysis and optimization. The representation of such a cone will be performed by means of first-order approximations of the problem data, namely generalized derivatives of the bifunction f and tangent cones of the set K defining (VEP). In other words, following a principle deep-rooted in many contexts of nonlinear analysis, approximations of the solution set to a given problem are obtained by means of exact solutions to approximated problems.

The paper is structured as follows. Section 2 aims at recalling preliminary notions of nonsmooth analysis, which play a role in formulating and establishing the achievements of the paper. Section 3 contains the main results concerning the first-order approximation of the contingent cone to $\mathscr{SE}$. In Section 4, these results are applied to derive both necessary and sufficient optimality conditions for nondifferentiable optimization problems, whose constraint systems are formalized as a strong vector equilibrium problem.

Below, the basic notations employed in the paper are listed. The acronyms l.s.c., u.s.c and p.h. stand for lower semicontinuous, upper semicontinuous and positively homogeneous, respectively. $\mathbb{R}^d$ denotes the finite-dimensional Euclidean space, with dimension $d \in \mathbb{N}$. The closed ball centered at an element $x \in \mathbb{R}^d$, with radius $r \geq 0$, is denoted by $\mathrm{B}(x;r)$. In particular, $\mathbb{B} = \mathrm{B}(\mathbf{0};1)$ stands for the unit ball, whereas $\mathbb{S}$ stands for the unit sphere, $\mathbf{0}$ denoting the null vector of an Euclidean space. Given a subset $S \subseteq \mathbb{R}^d$, the distance of a point x from S is denoted by $\mathrm{dist}\,(x;S)$, with the convention that $\mathrm{dist}\,(x;\varnothing)=+\infty$. The prefix $\mathrm{int}\,S$ denotes the interior of S, $\mathrm{cl}\,S$ denotes its closure, whereas $\mathrm{cone}\,S$ its conical hull, respectively. Given two subsets A and B of the same space, the excess of A over B is indicated by $\mathrm{exc}(A;B)=\sup_{a\in A}\mathrm{dist}\,(a;B)$. By $\mathscr{PH}(\mathbb{R}^n,\mathbb{R}^m)$ the space of all continuous p.h. mappings acting between $\mathbb{R}^n$ and $\mathbb{R}^m$ is denoted, equipped with the norm $\|h\|_{\mathscr{PH}}=\sup_{u\in\mathbb{S}}\|h(u)\|$, $h\in\mathscr{PH}(\mathbb{R}^n,\mathbb{R}^m)$, while $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ denotes its subspace of all linear operators. The inner product of an Euclidean space will be denoted by $\langle\cdot,\cdot\rangle$. Whenever C is a cone in $\mathbb{R}^q$, by $C^{\ominus}=\{v\in\mathbb{R}^q:\ \langle v,c\rangle\le 0,\quad \forall c\in C\}$ the negative dual (a.k.a. polar) cone to C is denoted. Given a function $\varphi:\mathbb{R}^q\longrightarrow\mathbb{R}\cup\{\pm\infty\}$, the symbol $\partial\varphi(x)$ denotes the subdifferential of φ at x in the sense of convex analysis (a.k.a. Fenchel subdifferential). The normal cone to a set $S\subseteq\mathbb{R}^q$ at $x\in S$ in the sense of convex analysis is denoted by $\mathrm{N}(x;S)=\{v\in\mathbb{R}^q:\ \langle v,s-x\rangle,\ \forall s\in S\}$.

2. Preliminaries

2.1. **Approximation of sets.** Given a nonempty set $K\subseteq\mathbb{R}^n$ and $\bar{x}\in K$, in the sequel the following different notions of tangent cone will be mainly employed:

(i) the contingent (a.k.a. Bouligand tangent) cone to K at $\bar{x}$, which is defined by

$$\mathrm{T}(\bar{x};K)=\{v\in\mathbb{R}^n:\ \exists(v_n)_n,\ v_n\to v,\ \exists(t_n)_n,\ t_n\downarrow 0:\ \bar{x}+t_nv_n\in K,\ \forall n\in\mathbb{N}\};$$

(ii) the cone of radial (a.k.a. weak feasible) directions to K at $\bar{x}$, which is defined by

$$\mathrm{T_r}(\bar{x};K)=\{v\in\mathbb{R}^n:\ \forall\varepsilon>0\ \exists t_\varepsilon\in(0,\varepsilon):\ \bar{x}+t_\varepsilon v\in K\}.$$

Clearly, for every $K\subseteq\mathbb{R}^n$ and $\bar{x}\in K$, it is $\mathrm{T_r}(\bar{x};K)\subseteq\mathrm{T}(\bar{x};K)$. Moreover $\mathrm{T}(\bar{x};K)$ is always closed. If, in particular, K is convex, then the following representations hold

$$\mathrm{T_r}(\bar{x};K)=\mathrm{cone}\,(K-\bar{x})\quad\text{ and }\quad \mathrm{T}(\bar{x};K)=\mathrm{cl}\,(\mathrm{cone}\,(K-\bar{x}))=\mathrm{cl}\,\mathrm{T_r}(\bar{x};K) \tag{2.1}$$

(see [20, Proposition 11.1.2(d)]). Thus, in such an event, both $\mathrm{T_r}(\bar{x};K)$ and $\mathrm{T}(\bar{x};K)$ are convex. It is well known that an equivalent (variational) reformulation of the notion of contingent cone is provided by the equality

$$\mathrm{T}(\bar{x};K)=\left\{v\in\mathbb{R}^n:\ \liminf_{t\downarrow 0}\frac{\mathrm{dist}\,(\bar{x}+tv;K)}{t}=0\right\}. \tag{2.2}$$

Remark 2.1. Whenever a convex set $K\subseteq\mathbb{R}^n$ is, in particular, polyhedral, one has $\mathrm{T_r}(\bar{x};K)=\mathrm{T}(\bar{x};K)$. To see this, it suffices to exploit the formulae in (2.1) and to observe that, in the present circumstance, $\mathrm{T_r}(\bar{x};K)$ happens to be closed. The latter follows from the fact that, if S is a closed affine half-space in $\mathbb{R}^n$, then $\mathrm{T_r}(\bar{x};S)=\mathrm{cone}\,(S-\bar{x})=S-\bar{x}$ is a closed set and from the fact that, if K_1 and K_2 are convex sets with $\bar{x}\in K_1\cap K_2$, then it holds $\mathrm{T_r}(\bar{x};K_1\cap K_2)=\mathrm{T_r}(\bar{x};K_1)\cap\mathrm{T_r}(\bar{x};K_2)$ (whereas the intersection fails to be preserved by the cone of radial directions in the case of nonconvex sets, as shown in [6, Example 1.1]).

For a systematic discussion about properties of the above tangent cones and their relationships, the reader is referred for instance to [4, Chapter 4], [7, Chapter I.1], [8], [18, Chapter 2], and [20, Chapter 11].

2.2. **Approximation of scalar functions.** Given a function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$, let $\bar{x} \in \varphi^{-1}(\mathbb{R})$. The set

$$\widehat{\partial}^+\varphi(\bar{x}) = \left\{ v \in \mathbb{R}^n : \limsup_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}$$

is called (Fréchet) upper subdifferential of φ at $\bar{x}$. Any element $v \in \widehat{\partial}^+\varphi(\bar{x})$ can be characterized by the existence of a function $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\varphi(\bar{x}) = \psi(\bar{x})$, $\varphi(x) \leq \psi(x)$, for every $x \in \mathbb{R}^n$, ψ is (Fréchet) differentiable at $\bar{x}$ and $v = \nabla\psi(\bar{x})$. If $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ is concave, then $\widehat{\partial}^+\varphi(\bar{x})$ coincides with the superdifferential (a.k.a. upper subdifferential) in the sense of convex analysis, i.e. $-\partial(-\varphi)(\bar{x})$.

Whenever φ is an u.s.c. function, the upper subdifferential admits another characterization in terms of Dini-Hadamard directional derivative, in fact being equivalent to the Dini-Hadamard upper subdifferential (in finite-dimensional spaces, the Fréchet bornology is equivalent to the Hadamard bornology). More precisely, it holds

$$\widehat{\partial}^+\varphi(\bar{x}) = \{ v \in \mathbb{R}^n : \langle v, w \rangle \geq \mathrm{D}^+_H\varphi(\bar{x}; w), \quad \forall w \in \mathbb{R}^n \}, \tag{2.3}$$

where

$$\mathrm{D}^+_H\varphi(\bar{x}; w) = \limsup_{\substack{u \to w \\ t \downarrow 0}} \frac{\varphi(\bar{x} + tu) - \varphi(\bar{x})}{t}$$

denotes the Dini-Hadamard upper directional derivative of φ at $\bar{x}$, in the direction $w \in \mathbb{R}^n$ (see [15, Chapter 1.3], [19, Chapter 8.B]). Let us recall that, whenever φ is locally Lipschitz around $\bar{x}$, its Dini-Hadamard directional derivative at $\bar{x}$ takes the following simpler form

$$\mathrm{D}^+_D\varphi(\bar{x}; w) = \limsup_{t \downarrow 0} \frac{\varphi(\bar{x} + tw) - \varphi(\bar{x})}{t},$$

which is known as Dini upper directional derivative. The lower versions of these generalized derivatives are

$$\mathrm{D}^-_H\varphi(\bar{x}; w) = \liminf_{\substack{u \to w \\ t \downarrow 0}} \frac{\varphi(\bar{x} + tu) - \varphi(\bar{x})}{t},$$

called the Dini-Hadamard lower directional (a.k.a. contingent) derivative of φ at $\bar{x}$, in the direction w, and

$$\mathrm{D}^-_D\varphi(\bar{x}; w) = \liminf_{t \downarrow 0} \frac{\varphi(\bar{x} + tw) - \varphi(\bar{x})}{t},$$

called the Dini lower directional derivative of φ at $\bar{x}$, in the direction w.

The set

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

is called (Fréchet) regular subdifferential of φ at $\bar{x}$. Whenever φ is l.s.c. around $\bar{x}$, it admits the following representation in terms of Dini-Hadamard lower directional generalized derivative

$$\widehat{\partial}\varphi(\bar{x}) = \{ v \in \mathbb{R}^n : \langle v, w \rangle \leq \mathrm{D}^-_H\varphi(\bar{x}; w), \quad \forall w \in \mathbb{R}^n \}. \tag{2.4}$$

Whenever φ is Fréchet differentiable at $\bar{x}$, one has $\widehat{\partial}^+\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$, where $\nabla\varphi(\bar{x})$ denotes the gradient of φ at $\bar{x}$.

Comprehensive discussions from various viewpoints as well as detailed material about these generalized derivatives can be found in many textbooks devoted to nonsmooth analysis, among which [7, Chapter I.1], [15, Chapter 1], [18, Chapter 2], [19, Chapter 8], [20].

2.3. **Approximation of mappings and bifunctions.** A mapping $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be B-differentiable at $\bar{x} \in \mathbb{R}^n$ if there exists a mapping $\mathrm{D}_B g(\bar{x}) \in \mathscr{PH}(\mathbb{R}^n, \mathbb{R}^m)$ such that

$$\lim_{x \to \bar{x}} \frac{\|g(x) - g(\bar{x}) - \mathrm{D}_B g(\bar{x})(x - \bar{x})\|}{\|x - \bar{x}\|} = 0.$$

As a consequence of the continuity of $\mathrm{D}_B g(\bar{x})$, it is readily seen that if g is B-differentiable at $\bar{x}$, it is also continuous at the same point. Notice that, when, in particular, $\mathrm{D}_B g(\bar{x}) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$, g turns out to be (Fréchet) differentiable at $\bar{x}$. In such an event, its derivative, represented by its Jacobian matrix, will be indicated by $\nabla g(\bar{x})$. Given a nonempty set $K \subseteq \mathbb{R}^n$, a bifunction $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be B-differentiable at $\bar{x} \in K$, uniformly on K, if there exists a family $\{\mathrm{D}_B f(\bar{x}, z) \in \mathscr{PH}(\mathbb{R}^n, \mathbb{R}^m) :\ z \in K\}$ such that for every $\varepsilon > 0$ $\exists \delta_\varepsilon > 0$ such that

$$\sup_{z \in K} \frac{\|f(x, z) - f(\bar{x}, z) - \mathrm{D}_B f(\bar{x}, z)(x - \bar{x})\|}{\|x - \bar{x}\|} < \varepsilon, \quad \forall x \in \mathrm{B}(\bar{x}; \delta_\varepsilon).$$

It should be clear that the above notion of generalized differentiation for bifunctions is a kind of partial differentiation, in considering variations of a mapping with respect to changes of one variable only.

Example 2.2. (i) Separable mappings: let us consider mappings $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$, which can be expressed in the form

$$f(x, z) = f_1(x) + f_2(z),$$

for proper $f_1, f_2 : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Whenever f_1 is B-differentiable at $\bar{x}$, with B-derivative $\mathrm{D}_B f_1(\bar{x})$, the bifunction f is B-differentiable at $\bar{x}$ uniformly on K, with $\{\mathrm{D}_B f(\bar{x}, z) :\ z \in K\} = \{\mathrm{D}_B f_1(\bar{x})\}$.

(ii) Factorable mappings: whenever a mapping $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ can be factorized as

$$f(x, z) = \alpha(z) g(x),$$

where $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is B-differentiable at $\bar{x}$, with B-derivative $\mathrm{D}_B g(\bar{x})$, and $\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}$ is bounded on K, the bifunction f is B-differentiable at $\bar{x}$ uniformly on K, with $\{\mathrm{D}_B f(\bar{x}, z) :\ z \in K\} = \{\alpha(z) \mathrm{D}_B g(\bar{x}) : z \in \mathbb{R}^n\}$.

(iii) Composition with differentiable mappings: if $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^p$ is B-differentiable at $\bar{x}$ uniformly on K and $g : \mathbb{R}^p \longrightarrow \mathbb{R}^m$ is Fréchet differentiable at each point $f(\bar{x}, z)$, with $z \in K$, then their composition $g \circ f$ turns out to be B-differentiable at $\bar{x}$ uniformly on K, with $\{\mathrm{D}_B (g \circ f)(\bar{x}, z) :\ z \in K\} = \{\nabla g(f(\bar{x}, z)) \mathrm{D}_B f(\bar{x}, z) :\ z \in K\}$.

2.4. **Distance from strong vector equilibria.** The function $\nu : \mathbb{R}^n \longrightarrow [0, +\infty]$, defined by

$$\nu(x) = \sup_{z \in K} \mathrm{dist}(f(x, z); C), \tag{2.5}$$

can be exploited as a natural measure of the distance of a given point $x \in \mathbb{R}^n$ from being a solution to (VEP). Clearly it is $\mathscr{SE} = \nu^{-1}(0) \cap K$, while positive values of ν quantify the violation of the strong equilibrium condition in (VEP).

A local error bound (in terms of ν) is said to be valid near $\bar{x} \in \mathscr{SE}$ for problem (VEP) if there exist positive κ and δ such that

$$\mathrm{dist}(x; \mathscr{SE}) \le \kappa \nu(x), \quad \forall x \in \mathrm{B}(\bar{x}; \delta) \cap K. \tag{2.6}$$

Notice that, whereas for computing $\mathrm{dist}(x;\mathscr{SE})$ one needs to know all the solutions to (VEP) near $\bar{x}$, the value of $\nu(x)$ can be computed directly by means of problem data. A study of sufficient conditions for the error in bound in (2.6) to hold has been recently undertaken in [21]. In particular, the following global error bound condition under an uniform B-differentiability assumption on f is known to hold.

Proposition 2.3 ([21])**.** *With reference to a problem* (VEP)*, suppose that:*

(i) *each function $x \mapsto f(x,z)$ is C-u.s.c. on K, for every $z \in K$;*
(ii) *the set-valued mapping $x \rightsquigarrow f(x,K)$ takes C-bounded values on K;*
(iii) *K is convex;*
(iv) *f is B-differentiable uniformly on K at each point of $K\backslash\mathscr{SE}$;*
(v) *there exists $\sigma > 0$ with the property that for every $x_0 \in K\backslash\mathscr{SE}$ there is $u_0 \in \mathbb{S} \cap \mathrm{cone}(K - x_0)$ such that*

$$\mathrm{D}_B f(x_0,z)(u_0) + \sigma\mathbb{B} \subseteq C, \quad \forall z \in K.$$

Then, $\mathscr{SE}$ is nonempty, closed and the following estimate holds true

$$\mathrm{dist}(x;\mathscr{SE}) \le \frac{\nu(x)}{\sigma}, \quad \forall x \in K.$$

The error bound in (2.6) will be exploited as a crucial qualification condition for (VEP) in providing the conical inner approximation of $\mathscr{SE}$.

3. Tangential approximation of $\mathscr{SE}$

As a first result, a one-side conical approximation from inside of the contingent cone to $\mathscr{SE}$ is presented.

Theorem 3.1 (Inner approximation)**.** *With reference to a problem* (VEP)*, let $\bar{x} \in \mathscr{SE}$. Suppose that:*

(i) *f is B-differentiable at $\bar{x}$, uniformly on K, with $\{\mathrm{D}_B f(\bar{x},z) :\ z \in K\}$;*
(ii) *a local error bound such as* (2.6) *is valid near $\bar{x}$.*

Then, it holds

$$\bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K) \subseteq \mathrm{T}(\bar{x};\mathscr{SE}). \tag{3.1}$$

Proof. Let us start with observing that, since it is $\mathrm{D}_B f(\bar{x},z) \in \mathscr{PH}(\mathbb{R}^n,\mathbb{R}^m)$ for every $z \in K$, and C is a cone, each set $\mathrm{D}_B f(\bar{x},z)^{-1}(C)$ turns out to be a cone containing $\mathbf{0}$, as well as $\mathrm{T_r}(\bar{x};K)$ does by definition. Thus, if taking $v = \mathbf{0} \in \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K)$, the inclusion $v \in \mathrm{T}(\bar{x};\mathscr{SE})$ obviously holds as the latter cone is closed. So, take an arbitrary $v \in \left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K)\right) \backslash \{\mathbf{0}\}$. Since both the sets in the inclusion in (3.1) are cones, one can assume without any loss of generality that $\|v\| = 1$. In the light of the characterization via (2.2), v is proven to belong to $\mathrm{T}(\bar{x};\mathscr{SE})$ if one shows that

$$\liminf_{t\downarrow 0} \frac{\mathrm{dist}(\bar{x}+tv;\mathscr{SE})}{t} = 0. \tag{3.2}$$

Showing the equality in (3.2) amounts to show that for every $\tau > 0$ and $\varepsilon > 0$ there exists $t_0 \in (0,\tau)$ such that

$$\frac{\operatorname{dist}(\bar{x}+t_0v;\mathscr{SE})}{t_0} \le \varepsilon. \tag{3.3}$$

So, let us fix ad libitum τ and ε. Hypothesis (ii) ensures the existence of $\delta,\ \kappa > 0$ as in (2.6). By virtue of hypothesis (i), corresponding to ε/κ, there exists $\delta_\varepsilon > 0$ such that

$$f(x,z) \in f(\bar{x},z) + \mathrm{D}_B f(\bar{x},z)(x-\bar{x}) + \kappa^{-1}\varepsilon\|x-\bar{x}\|\mathbb{B}, \quad \forall x \in \mathrm{B}(\bar{x};\delta_\varepsilon),\ \forall z \in K,$$

and hence, in particular,

$$f(\bar{x}+tv,z) \in f(\bar{x},z) + t\mathrm{D}_B f(\bar{x},z)(v) + \kappa^{-1}\varepsilon t\mathbb{B}, \quad \forall t \in (0,\delta_\varepsilon),\ \forall z \in K.$$

By taking into account that $\bar{x} \in \mathscr{SE}$ and $v \in \mathrm{D}_B f(\bar{x},z)^{-1}(C)$ for every $z \in K$, the above inclusion implies

$$f(\bar{x}+tv,z) \in C + tC + \kappa^{-1}\varepsilon t\mathbb{B} \subseteq C + \kappa^{-1}\varepsilon t\mathbb{B}, \quad \forall t \in (0,\delta_\varepsilon),\ \forall z \in K.$$

In terms of the residual function ν introduced in (2.5), this means

$$\begin{aligned}\nu(\bar{x}+tv) = \sup_{z\in K}\operatorname{dist}(f(\bar{x}+tv,z);C) &\le \operatorname{exc}(C+\kappa^{-1}\varepsilon t\mathbb{B};C) = \operatorname{exc}(\kappa^{-1}\varepsilon t\mathbb{B};C)\\ &\le \kappa^{-1}\varepsilon t, \quad \forall t \in (0,\delta_\varepsilon),\end{aligned} \tag{3.4}$$

where the second equality holds because C is a convex cone. On the other hand, according to hypothesis (ii) there exists $\delta_0 \in (0,\min\{\tau,\delta,\delta_\varepsilon\})$ such that

$$\operatorname{dist}(x;\mathscr{SE}) \le \kappa\nu(x), \quad \forall x \in \mathrm{B}(\bar{x};\delta_0)\cap K. \tag{3.5}$$

Since it is $v \in \mathrm{T_r}(\bar{x};K)$, for some $t_* \in (0,\delta_0)$ it happens

$$\bar{x}+t_*v \in K \cap \mathrm{B}(\bar{x};\delta_0),$$

and therefore, by inequality (3.5), one obtains

$$\operatorname{dist}(\bar{x}+t_*v;\mathscr{SE}) \le \kappa\nu(\bar{x}+t_*v). \tag{3.6}$$

By combining inequalities (3.4) and (3.6), as it is $t_* < \delta_0 < \delta_\varepsilon$, one obtains

$$\operatorname{dist}(\bar{x}+t_*v;\mathscr{SE}) \le \kappa\cdot\kappa^{-1}\varepsilon t_* = \varepsilon t_*.$$

The last inequality shows that (3.3) is true for $t_0 = t_* \in (0,\tau)$, thereby completing the proof. □

The inclusion in (3.1) states that, under proper assumptions, any solution of the (approximated) problem

$$\text{find } v \in \mathrm{T_r}(\bar{x};K) \text{ such that } \mathrm{D}_B f(\bar{x};z)(v) \in C, \quad \forall z \in K, \tag{3.7}$$

provides a vector, which is tangent to $\mathscr{SE}$ at $\bar{x}$ in the sense of Bouligand. Notice that problem (3.7) is almost in the form (VEP) (it would be exactly in the form (VEP) if $\mathrm{T_r}(\bar{x};K) = K$). Roughly speaking, all of this means that if the problem data of (VEP) are properly approximated (K by its radial direction cone, f by its generalized derivatives in the sense of Bouligand, respectively) near a reference solution $\bar{x}$, then the solutions of the resulting approximated problem (3.7) work as a first-order approximation of the solution set to the original problem (VEP). Problem (3.7) is typically expected to be easier than (VEP) by virtue of the structural properties of its data. Basically, (3.7) can be regarded as a cone constrained p.h. vector inequality system, so its solution set is a cone. Furthermore, if K is convex and $\mathrm{D}_B f(\bar{x},z) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is C-concave for every $z \in K$, the latter meaning that

$$\mathrm{D}_B f(\bar{x},z)(v_1) + \mathrm{D}_B f(\bar{x},z)(v_2) \leq_C \mathrm{D}_B f(\bar{x},z)(v_1+v_2), \quad \forall v_1, v_2 \in \mathbb{R}^n,$$

where $\leq_C$ denotes the partial ordering on $\mathbb{R}^m$ induced in the standard way by the cone C, then the solution set to problem (3.7) is a convex cone.

As a further comment to Theorem 3.1, it must be remarked that the inclusion in (3.1) provides only a one-side approximation of $\mathrm{T}(\bar{x};\mathscr{SE})$, which may happen to be rather rough. This fact is illustrated by the next example.

Example 3.2 (Inclusion (3.1) may be strict)**.** Consider the problem (VEP) defined by the following data: $K = C = \mathbb{R}^2_+ = \{x = (x_1,x_2) \in \mathbb{R}^2 :\ x_1 \geq 0,\ x_2 \geq 0\}$ and a vector-valued bifunction $f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$f(x_1,x_2,z_1,z_2) = \begin{pmatrix} \frac{1}{2}(-m_z^- x_1 + x_2 + 1)^2 \\ \\ \frac{1}{2}(m_z^+ x_1 - x_2 + 1)^2 \end{pmatrix},$$

where

$$m_z^- = 1 - \frac{1}{\|z\|^2+1} \qquad \text{and} \qquad m_z^+ = 1 + \frac{1}{\|z\|^2+1}, \quad z \in \mathbb{R}^2.$$

Since $f(x,z) \in \mathbb{R}^2_+$ for every $(x,z) \in \mathbb{R}^2 \times \mathbb{R}^2$, it is clear that $\mathscr{SE} = K = \mathbb{R}^2_+$. Fix $\bar{x} = \mathbf{0} \in \mathscr{SE}$, so one has

$$\mathrm{T_r}(\mathbf{0};K) = \mathrm{T}(\mathbf{0};\mathscr{SE}) = \mathbb{R}^2_+.$$

In view of the next calculations, it is convenient to observe that

$$f(x,z) = (g \circ h)(x,z),$$

where the mappings $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and $h : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ are given respectively by

$$g(y) = \begin{pmatrix} y_1^2/2 \\ y_2^2/2 \end{pmatrix} \qquad \text{and} \qquad h(x,z) = \begin{pmatrix} -m_z^- x_1 + x_2 + 1 \\ m_z^+ x_1 - x_2 + 1 \end{pmatrix}.$$

To check that the bifunction h is B-differentiable at $\mathbf{0}$ uniformly on $\mathbb{R}^2_+$, with

$$\left\{ \mathrm{D}_B h(\mathbf{0},z) = \nabla h(\mathbf{0},z) = \begin{pmatrix} -m_z^- & 1 \\ m_z^+ & -1 \end{pmatrix},\ z \in \mathbb{R}^2_+ \right\}$$

it suffices to observe that

$$\|h(x,z) - h(\mathbf{0},z) - \mathrm{D}_B h(\mathbf{0},z)(x)\| =$$
$$= \left\| \begin{pmatrix} -m_z^- x_1 + x_2 + 1 \\ m_z^+ x_1 - x_2 + 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} -m_z^- & 1 \\ m_z^+ & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| = 0, \quad \forall z \in \mathbb{R}^2_+.$$

Thus, since g is Fréchet differentiable at each point of $\mathbb{R}^2$ and

$$\nabla g(y) = \begin{pmatrix} y_1 & 0 \\ 0 & y_2 \end{pmatrix},$$

according to what remarked in Example 2.2(iii), the mapping $f = g \circ h$ turns out to be B-differentiable at $\mathbf{0}$ uniformly on $\mathbb{R}^2_+$, with

$$\mathrm{D}_B f(\mathbf{0},z) = \nabla g(h(\mathbf{0},z)) \circ \mathrm{D}_B h(\mathbf{0},z) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -m_z^- & 1 \\ m_z^+ & -1 \end{pmatrix} = \begin{pmatrix} -m_z^- & 1 \\ m_z^+ & -1 \end{pmatrix},\ z \in \mathbb{R}^2_+.$$

Notice that a local error bound as in (2.6) is evidently valid near $\mathbf{0}$ because it is $\mathscr{SE} = K$. Thus, all the hypotheses of Theorem 3.1 are satisfied.

Now, one readily sees that

$$\mathrm{D}_B f(\mathbf{0},z)(v) = \begin{pmatrix} -m_z^- v_1 + v_2 \\ m_z^+ v_1 - v_2 \end{pmatrix} \in \mathbb{R}^2_+ \qquad \text{iff} \qquad \begin{cases} -m_z^- v_1 + v_2 \geq 0 \\ \\ m_z^+ v_1 - v_2 \geq 0. \end{cases}$$

This leads to find

$$\mathrm{D}_B f(\mathbf{0},z)^{-1}(\mathbb{R}^2_+) = \{v \in \mathbb{R}^2 :\ m_z^- v_1 \leq v_2 \leq m_z^+ v_1\}, \quad \forall z \in \mathbb{R}^2_+.$$

Since one has

$$\lim_{\|z\|\to\infty} m_z^- = 1^- = 1 = 1^+ = \lim_{\|z\|\to\infty} m_z^+,$$

it results in

$$\bigcap_{z\in\mathbb{R}^2_+} \mathrm{D}_B f(\mathbf{0},z)^{-1}(\mathbb{R}^2_+) \cap \mathrm{T_r}(\mathbf{0};\mathbb{R}^2_+) = \{v \in \mathbb{R}^2_+ :\ v_2 = v_1\} \subsetneqq \mathbb{R}^2_+ = \mathrm{T}(\mathbf{0};\mathscr{SE}).$$

The above example motivates the interest in outer approximations of $\mathscr{SE}$. Below, a result in this direction is presented. In what follows, recall that a family of mappings $\{h_z \in \mathscr{PH}(\mathbb{R}^n,\mathbb{R}^m) :\ z \in K\}$ is said to be equicontinuous at $x_0 \in \mathbb{R}^n$ if for every $\varepsilon > 0$ there exists $\delta > 0$ (not depending on $z \in K$) such that

$$\|h_z(x) - h_z(x_0)\| < \varepsilon, \quad \forall z \in K,\ \forall x \in \mathrm{B}(x_0;\delta).$$

Theorem 3.3 (Outer approximation)**.** *With reference to a problem* (VEP)*, let $\bar{x} \in \mathscr{SE}$. Suppose that:*

(i) *f is B-differentiable at $\bar{x}$, uniformly on K, with $\{\mathrm{D}_B f(\bar{x},z) :\ z \in K\}$;*
(ii) *the family of mappings $\{\mathrm{D}_B f(\bar{x},z) :\ z \in K\}$ is equicontinuous at each point of $\mathbb{R}^n$.*

Then, it holds

$$\mathrm{T}(\bar{x};\mathscr{SE}) \subseteq \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(\mathrm{T}(f(\bar{x},z);C)) \cap \mathrm{T}(\bar{x};K). \tag{3.8}$$

Proof. Since it is $\mathrm{D}_B f(\bar{x},z) \in \mathscr{PH}(\mathbb{R}^n,\mathbb{R}^m)$ for every $z \in K$, one has

$$\mathrm{D}_B f(\bar{x},z)(\mathbf{0}) = \mathbf{0} \in \mathrm{T}(f(\bar{x},z);C), \quad \forall z \in K.$$

Therefore, it clearly holds

$$\mathbf{0} \in \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(\mathrm{T}(f(\bar{x},z);C)) \cap \mathrm{T}(\bar{x};K).$$

So take an arbitrary $v \in \mathrm{T}(\bar{x};\mathscr{SE})\backslash\{\mathbf{0}\}$. As all the sets involved in inclusion (3.8) are cones, without loss of generality it is possible to assume that $\|v\| = 1$. According to the definition of contingent cone, there exist $(v_n)_n$, with $v_n \longrightarrow v$ and $(t_n)_n$, with $t_n \downarrow 0$, such that $\bar{x} + t_n v_n \in \mathscr{SE} \subseteq K$. Notice that this inclusion in particular implies that $v \in \mathrm{T}(\bar{x};K)$. What remains to be shown is that

$$v \in \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(\mathrm{T}(f(\bar{x},z);C)). \tag{3.9}$$

Fix an arbitrary $\varepsilon > 0$. By virtue of hypothesis (i), there exists $\delta_\varepsilon > 0$ such that

$$f(x,z) - f(\bar{x},z) - \mathrm{D}_B f(\bar{x},z)(x - \bar{x}) \in \varepsilon\|x - \bar{x}\|\mathbb{B}, \quad \forall z \in K,\ \forall x \in \mathrm{B}(\bar{x};\delta_\varepsilon)$$

and hence

$$\mathrm{D}_B f(\bar{x},z)(x-\bar{x}) \in f(x,z) - f(\bar{x},z) + \varepsilon\|x-\bar{x}\|\mathbb{B}, \quad \forall z \in K,\ \forall x \in \mathrm{B}\left(\bar{x};\delta_\varepsilon\right). \tag{3.10}$$

Since it is $\bar{x}+t_n v_n \longrightarrow \bar{x}$ as $n \to \infty$ (as a converging sequence $(v_n)_n$ must be bounded), for some $n_\varepsilon \in \mathbb{N}$ it is true that $\bar{x}+t_n v_n \in \mathrm{B}\left(\bar{x};\delta_\varepsilon\right)$ for every $n \geq n_\varepsilon$. Thus, by taking $x = \bar{x}+t_n v_n$ in (3.10), one finds

$$t_n \mathrm{D}_B f(\bar{x},z)(v_n) \in f(\bar{x}+t_n v_n,z) - f(\bar{x},z) + \varepsilon t_n \|v_n\|\mathbb{B}, \quad \forall z \in K,\ \forall n \geq n_\varepsilon,$$

whence it follows

$$\mathrm{D}_B f(\bar{x},z)(v_n) \in \frac{f(\bar{x}+t_n v_n,z) - f(\bar{x},z)}{t_n} + \varepsilon\|v_n\|\mathbb{B}, \quad \forall z \in K,\ \forall n \geq n_\varepsilon.$$

By taking into account that $v_n \longrightarrow v$ as $n \to \infty$ and $\|v\| = 1$, one has that $\|v_n\| \leq 2$ for all $n \geq n_\varepsilon$, up to a proper increase in the value of n_ε, if needed. Thus, from the last inclusion one obtains

$$\mathrm{D}_B f(\bar{x},z)(v_n) \in \frac{f(\bar{x}+t_n v_n,z) - f(\bar{x},z)}{t_n} + 2\varepsilon\mathbb{B}, \quad \forall z \in K,\ \forall n \geq n_\varepsilon. \tag{3.11}$$

By hypothesis (ii) the family $\{\mathrm{D}_B f(\bar{x},z) :\ z \in K\}$ is equicontinuous at v. This means that there exists $n_* \in \mathbb{N}$ (independent of z), with $n_* \geq n_\varepsilon$, such that

$$\|\mathrm{D}_B f(\bar{x},z)(v_n) - \mathrm{D}_B f(\bar{x},z)(v)\| \leq \varepsilon, \quad \forall z \in K,\ \forall n \geq n_*,$$

or, equivalently,

$$\mathrm{D}_B f(\bar{x},z)(v) \in \mathrm{D}_B f(\bar{x},z)(v_n) + \varepsilon\mathbb{B}, \quad \forall z \in K,\ \forall n \geq n_*.$$

By recalling (3.11), from the last inclusion one gets

$$\mathrm{D}_B f(\bar{x},z)(v) \in \frac{f(\bar{x}+t_n v_n,z) - f(\bar{x},z)}{t_n} + 3\varepsilon\mathbb{B}, \quad \forall z \in K,\ \forall n \geq n_*.$$

Since it is $\bar{x}+t_n v_n \in \mathscr{SE}$ for every $n \in \mathbb{N}$, this implies

$$\mathrm{D}_B f(\bar{x},z)(v) \in \frac{C - f(\bar{x},z)}{t_n} + 3\varepsilon\mathbb{B} \in \mathrm{cone}\left(C - f(\bar{x},z)\right) + 3\varepsilon\mathbb{B}, \quad \forall z \in K,\ \forall n \geq n_*.$$

Since C is convex so $\mathrm{T}(f(\bar{x},z);C) = \mathrm{cl}\,\mathrm{cone}\left(C - f(\bar{x},z)\right)$, it results in

$$\mathrm{D}_B f(\bar{x},z)(v) \in \mathrm{T}(f(\bar{x},z);C) + 3\varepsilon\mathbb{B}, \quad \forall z \in K.$$

The arbitrariness of ε and the fact $\mathrm{T}(f(\bar{x},z);C)$ is closed allow one to assert that

$$\mathrm{D}_B f(\bar{x},z)(v) \in \mathrm{T}(f(\bar{x},z);C), \quad \forall z \in K,$$

which proves the validity of (3.9). Thus the proof is complete. □

Remark 3.4. (i) In the case in which $\mathrm{int}\,C \neq \varnothing$, it is useful to remark that the formula in (3.8) can be equivalently rewritten as

$$\mathrm{T}(\bar{x};\mathscr{SE}) \subseteq \{\mathbf{0}\} \cup \left(\bigcap_{z \in K \cap f^{-1}(\bar{x},\cdot)(\mathrm{bd}\,C)} \mathrm{D}_B f(\bar{x},z)^{-1}(\mathrm{T}(f(\bar{x},z);C)) \cap \mathrm{T}(\bar{x};K) \right),$$

with the convention that an intersection over an empty index set is the empty set. Indeed, whenever it happens $f(\bar{x},z) \in \mathrm{int}\,C$, one has $\mathrm{T}(f(\bar{x},z);C) = \mathbb{R}^m$, with the consequence that $\mathrm{D}_B f(\bar{x},z)^{-1}(\mathrm{T}(f(\bar{x},z);C)) = \mathbb{R}^n$.

(ii) It is worth noticing that for all those $z_0 \in K$ such that $f(\bar{x}, z_0) = \mathbf{0}$ (if any), the formula in (3.8) entails

$$\mathrm{T}(\bar{x}; \mathscr{SE}) \subseteq \mathrm{D}_B f(\bar{x}, z_0)^{-1}(C) \cap \mathrm{T}(\bar{x}; K),$$

as it is $\mathrm{T}(f(\bar{x}, z_0); C) = \mathrm{T}(\mathbf{0}; C) = C$.

(iii) The reader should observe that, in contrast to Theorem 3.1, Theorem 3.3 has been established without assuming any error bound type qualification condition.

The next example shows that also the outer one-side approximation of $\mathrm{T}(\bar{x}; \mathscr{SE})$ provided by Theorem 3.3 may happen to be rather rough.

Example 3.5 (Inclusion (3.8) may be strict)**.** Consider the (actually scalar) problem (VEP) defined by the following data: $K = \mathbb{R}$, $C = [0, +\infty)$, $f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f(x,z) = \frac{x^2 z}{z^2 + 1}.$$

It is clear that with the above data one has $\mathscr{SE} = \{0\}$. So, fix $\bar{x} = 0$. In order for checking that f is B-differentiable at 0 uniformly on $\mathbb{R}$, with $\{\mathrm{D}_B f(0,z) \equiv 0 : \ z \in \mathbb{R}\}$, according to the definition it suffices to observe that, fixed an arbitrary $\varepsilon > 0$, one has

$$\sup_{z \in \mathbb{R}} \frac{|f(x,z) - f(\bar{x},z)|}{|x - \bar{x}|} = \sup_{z \in \mathbb{R}} \frac{\left| \dfrac{x^2 z}{z^2+1} \right|}{|x|} = \sup_{z \in \mathbb{R}} \frac{|z|}{z^2+1} \cdot |x| \leq \frac{|x|}{2} \leq \varepsilon, \quad \forall x \in \mathrm{B}(0; \varepsilon).$$

As the family $\{\mathrm{D}_B f(0,z) \equiv 0 : \ z \in \mathbb{R}\}$ is actually independent of $z \in \mathbb{R}$, also the hypothesis (ii) of Theorem 3.3 is satisfied.

Since $f(0,z) = 0$ for every $z \in \mathbb{R}$, so it is $\mathrm{T}(f(0,z); [0, +\infty)) = [0, +\infty)$, one finds

$$\mathrm{D}_B f(0,z)^{-1} \left(\mathrm{T}(f(0,z); [0, +\infty)) \right) = \mathbb{R}, \quad \forall z \in \mathbb{R}.$$

Consequently, in the current case, one obtains

$$\mathrm{T}(0; \mathscr{SE}) = \{0\} \subsetneqq \mathbb{R} \cap \mathbb{R} = \bigcap_{z \in \mathbb{R}} \mathrm{D}_B f(0,z)^{-1} (\mathrm{T}(f(0,z); [0, +\infty))) \cap \mathrm{T}(0; \mathbb{R}).$$

Relying on both the preceding approximations, the next result singles out a sufficient condition, upon which one can establish an exact representation of $\mathrm{T}(\bar{x}; \mathscr{SE})$.

Corollary 3.6. *With reference to a problem* (VEP)*, let* $\bar{x} \in \mathscr{SE}$*. Suppose that:*

(i) *K is polyhedral;*
(ii) $f(\bar{x}, z) = \mathbf{0}, \quad \forall z \in K$*;*
(iii) *f is B-differentiable at $\bar{x}$, uniformly on K, with* $\{\mathrm{D}_B f(\bar{x}, z) : \ z \in K\}$*;*
(iv) *the family of mappings* $\{\mathrm{D}_B f(\bar{x}, z) : \ z \in K\}$ *is equicontinuous at each point of* $\mathbb{R}^n$*;*
(v) *a local error bound such as in* (2.6) *is valid near $\bar{x}$.*

Then, it holds

$$\mathrm{T}(\bar{x}; \mathscr{SE}) = \bigcap_{z \in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(C) \cap \mathrm{T}(\bar{x}; K).$$

Proof. The above assumptions enable one to apply both Theorem 3.1 and Theorem 3.3. From the former one, in the light of Remark 2.1 and hypothesis (i), one obtains

$$\bigcap_{z \in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(C) \cap \mathrm{T}(\bar{x}; K) \subseteq \mathrm{T}(\bar{x}; \mathscr{SE}). \tag{3.12}$$

From the latter, in the light of hypothesis (ii) and Remark 3.4(ii), one obtains

$$\mathrm{T}(\bar{x};\mathscr{SE}) \subseteq \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T}(\bar{x};K). \tag{3.13}$$

By combining inclusions (3.12) and (3.13) one gets the equality in the thesis. □

4. Applications to constrained optimization

This section deals with first-order optimality conditions for optimization problems, whose feasible region is formalized as a set of strong vector equilibria. As such, these problems can be cast in mathematical programming with equilibrium constraints, a well-recognized topic and active area of research (see, among others, [11, 12, 14, 17, 22]). Thus, the optimization problems here considered take the following form

$$\text{(MPVEC)} \qquad \min \vartheta(x) \quad \text{subject to} \quad x \in \mathscr{SE},$$

where $\vartheta : \mathbb{R}^n \longrightarrow \mathbb{R}$ is the objective function formalizing the criterion used for comparing variables, while $\mathscr{SE}$ is the feasible region of the problem, denoting as in the previous sections the solution sets to an inner (lower level) problem (VEP). Throughout this section ϑ will be assumed to be continuous around $\bar{x}$, but possibly nondifferentiable, as well as the bifunction f defining (VEP).

In constrained nondifferentiable optimization, first-order optimality conditions are typically obtained by locally approximating the objective function and the feasible region of a given problem. In this vein, the fact stated in the next lemma is widely known to hold, which has been used as a starting point for various, more elaborated, optimality conditions. For a direct proof see, for instance, [20, Chapter 7.1]. To a deeper view, it can be restored as a special case of an axiomatic scheme of analysis, which was developed in [6, 8] (see [6, Theorem 2.1]).

Lemma 4.1. *Let $\bar{x} \in \mathscr{SE}$ be a local optimal solution to problem* (MPVEC). *Then, it holds*

$$\mathrm{D}_D^+\vartheta(\bar{x};w) \geq 0, \quad \forall w \in \mathrm{T_r}(\bar{x};\mathscr{SE}) \tag{4.1}$$

and

$$\mathrm{D}_H^+\vartheta(\bar{x};w) \geq 0, \quad \forall w \in \mathrm{T}(\bar{x};\mathscr{SE}). \tag{4.2}$$

Remark 4.2. Since from their very definition one sees that

$$\mathrm{D}_D^+\vartheta(\bar{x};w) \leq \mathrm{D}_H^+\vartheta(\bar{x};w), \quad \forall w \in \mathbb{R}^n,$$

whereas it is $\mathrm{T_r}(\bar{x};\mathscr{SE}) \subseteq \mathrm{T}(\bar{x};\mathscr{SE})$, none of the conditions (4.1) and (4.2) can imply in general the other, unless either ϑ is locally Lipschitz near $\bar{x}$ or it is $\mathrm{T_r}(\bar{x};\mathscr{SE}) = \mathrm{T}(\bar{x};\mathscr{SE})$. Thus, the author does not agree with what asserted in [20, pag. 132]. For the purposes of the present analysis, only the condition in (4.2) will be actually exploited.

Theorem 4.3 (Necessary optimality condition)**.** *Let $\bar{x} \in \mathscr{SE}$ be a local optimal solution to problem* (MPVEC). *Suppose that:*

(i) *f is B-differentiable at $\bar{x}$, uniformly on K, with $\{\mathrm{D}_B f(\bar{x},z) :\ z \in K\}$;*
(ii) *a local error bound such as in* (2.6) *is valid near $\bar{x}$.*

Then, it holds

$$-\widehat{\partial}^{+}\vartheta(\bar{x}) \subseteq \left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K)\right)^{\ominus}. \tag{4.3}$$

Proof. Under the above assumptions, by Theorem 3.1 the inclusion in (3.1) holds true. Consequently, since $\bar{x} \in \mathscr{SE}$ is a local optimal solution to (MPVEC), according to condition (4.2) it must be

$$\mathrm{D}^{+}_{H}\vartheta(\bar{x};w) \geq 0, \quad \forall w \in \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K).$$

If $\widehat{\partial}^{+}\vartheta(\bar{x}) = \varnothing$ the thesis becomes trivial. Otherwise, by taking into account the representation in (2.3), which is valid because the function ϑ is in particular u.s.c. around $\bar{x}$, for an arbitrary $v \in \widehat{\partial}^{+}\vartheta(\bar{x})$ one finds

$$\langle v, w\rangle \geq 0, \quad \forall w \in \bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K),$$

which amounts to say that

$$-v \in \left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \mathrm{T_r}(\bar{x};K)\right)^{\ominus}.$$

The arbitrariness of $v \in \widehat{\partial}^{+}\vartheta(\bar{x})$ completes the proof. □

Remark 4.4. To assess the role of the optimality condition formulated in Theorem 4.3, notice that it does not carry useful information whenever $\widehat{\partial}\vartheta(\bar{x}) = \varnothing$. This happens, for example, if ϑ is a convex continuous function, which is nondifferentiable at $\bar{x}$. Nevertheless, the upper subdifferential is nonempty for large classes of functions, including the class of semiconcave ones (see [14]). In all such cases, condition (4.3) provides a necessary optimality condition, which may be more efficient than those expressed in terms of more traditional lower subdifferentials. This because it requires that all elements in $-\widehat{\partial}^{+}\vartheta(\bar{x})$ belong to the set in the right-side of (4.3), in contrast to a mere nonempty intersection requirement, which is typical for the lower subdifferential case.

Corollary 4.5. *Under the same assumptions of Theorem 4.3, if the following additional hypotheses are satisfied:*

- (i) *K is polyhedral;*
- (ii) *$\mathrm{D}_B f(\bar{x},z) \in \mathscr{PH}(\mathbb{R}^n,\mathbb{R}^m)$ is C-concave for every $z \in K$;*
- (iii) *the qualification condition holds*

$$\bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C) \cap \operatorname{int}\mathrm{T}(\bar{x};K) \neq \varnothing, \tag{4.4}$$

then the inclusion in (4.3) takes the simpler form

$$-\widehat{\partial}^{+}\vartheta(\bar{x}) \subseteq \left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x},z)^{-1}(C)\right)^{\ominus} + \mathrm{N}(\bar{x};K).$$

Proof. It is well know that if S_1 and S_2 are closed convex cones, then $(S_1 \cap S_2)^{\ominus} = \mathrm{cl}\,(S_1^{\ominus} + S_2^{\ominus})$ (see [20, Lemma 2.4.1]). On the other hand, if $S_1 - S_2 = \mathbb{R}^n$, then $S_1^{\ominus} + S_2^{\ominus}$ is closed (see [20, Proposition 2.4.3] If the qualification condition $S_1 \cap \mathrm{int}\, S_2 \neq \varnothing$ happens to be satisfied, then $S_1 - S_2 = \mathbb{R}^n$ (see [20, Lemma 2.4.4]). It follows from hypothesis (ii) that each $\mathrm{D}_B f(\bar{x}, z)^{-1}(C)$, and then hence $\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(C)$, is convex. Thus, since $\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(C)$ and $\mathrm{T}(\bar{x}; K)$ are closed convex cone, by virtue of (4.4) and the assumption (i), one obtains

$$\left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(C) \cap \mathrm{T}_{\mathrm{r}}(\bar{x}; K)\right)^{\ominus} = \left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(C)\right)^{\ominus} + \mathrm{T}(\bar{x}; K)^{\ominus}.$$

Then, in order to achieve the inclusion in the thesis it suffices to recall that $\mathrm{T}(\bar{x}; K)^{\ominus} = \mathrm{N}(\bar{x}; K)$ (see [20, Lemma 11.2.2]). □

Whenever that data of (MPVEC) happen to be smooth, i.e. ϑ is Fréchet differentiable at $\bar{x}$ and $\mathrm{D}_B f(\bar{x}, z) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ (and therefore is C-concave) for every $z \in K$, under the assumptions of Corollary 4.5 the condition in (4.3) becomes a more standard generalized equation of the form

$$\mathbf{0} \in \nabla\vartheta(\bar{x}) + \left(\bigcap_{z\in K} \nabla f(\bar{x}, z)^{-1}(C)\right)^{\ominus} + \mathrm{N}(\bar{x}; K).$$

Now, let us consider sufficient optimality conditions, a topic usually investigated in a subsequent step of analysis.

The next lemma provides a sufficient (strict) optimality condition for (MPVEC) in the case the objective function is locally Lipschitz. For its proof see [7, Lemma 1.3, Chapter V]. Notice that for the statement of Lemma 4.6, the hypothesis on the feasible region of the problem to allow a first-order uniform conical approximation in the sense of Demyanov-Rubinov is not needed (see [7, Remark 1.6, Chapter V]).

Lemma 4.6. *With reference to* (MPVEC)*, suppose that ϑ is locally Lipschitz around $\bar{x} \in \mathscr{SE}$. If it holds*

$$\mathrm{D}_D^{-}\vartheta(\bar{x}; w) > 0, \quad \forall w \in \mathrm{T}(\bar{x}; \mathscr{SE})\backslash\{\mathbf{0}\}, \tag{4.5}$$

then $\bar{x}$ is a strict local solution to (MPVEC).

On the base of the above lemma, one is in a position to establish the next result.

Theorem 4.7 (Sufficient optimality condition)**.** *With reference to* (MPVEC)*, assume that ϑ is locally Lipschitz around $\bar{x} \in \mathscr{SE}$. Suppose that:*

(i) *f is B-differentiable at $\bar{x}$, uniformly on K, with $\{\mathrm{D}_B f(\bar{x}, z) : \ z \in K\}$;*
(ii) *the family of mappings $\{\mathrm{D}_B f(\bar{x}, z) : \ z \in K\}$ is equicontinuous at each point of $\mathbb{R}^n$.*

If the condition

$$\mathbf{0} \in \widehat{\partial}\vartheta(\bar{x}) + \mathrm{int}\left[\left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(\mathrm{T}(f(\bar{x}, z); C)) \cap \mathrm{T}(\bar{x}; K)\right)^{\ominus}\right], \tag{4.6}$$

is satisfied, then $\bar{x}$ is a strict local solution to (MPVEC).

Proof. Observe first that if for a given cone $S \subseteq \mathbb{R}^n$ it is $v \in \operatorname{int}(S^{\ominus})$, then it must be

$$\langle v, s\rangle < 0, \quad \forall s \in S \backslash \{\mathbf{0}\}.$$

Indeed, there exists $\delta > 0$ such that $v + \delta\mathbb{B} \subseteq S^{\ominus}$, and therefore it holds

$$\langle v + \delta u, s\rangle \le 0, \quad \forall u \in \mathbb{B},\ \forall s \in S.$$

Thus, for any $s \in S\backslash\{\mathbf{0}\}$, the last inequality implies

$$\sup_{u\in\mathbb{B}}\langle v + \delta u, s\rangle = \langle v, s\rangle + \delta \sup_{u\in\mathbb{B}}\langle u, s\rangle = \langle v, s\rangle + \delta\|s\| \le 0,$$

whence one gets

$$\langle v, s\rangle \le -\delta\|s\| < 0.$$

Consequently, the condition (4.6) implies that there exists $v \in \widehat{\partial}\vartheta(\bar{x})$ such that it is

$$\langle v, w\rangle > 0, \quad \forall w \in \left[\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(\mathrm{T}(f(\bar{x}, z); C)) \cap \mathrm{T}(\bar{x}; K)\right] \backslash \{\mathbf{0}\}.$$

By recalling the representation of $\widehat{\partial}\vartheta(\bar{x})$ in (2.4), from the last inequality one obtains

$$\mathrm{D}^-_D\vartheta(\bar{x}; w) = \mathrm{D}^-_H\vartheta(\bar{x}; w) > 0, \quad \forall w \in \left[\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(\mathrm{T}(f(\bar{x}, z); C)) \cap \mathrm{T}(\bar{x}; K)\right] \backslash \{\mathbf{0}\}.$$

Since under the above assumptions Theorem 3.3 can be applied, then by virtue of the inclusion in (3.8) one can state that condition (4.5) turns out to be satisfied. Thus, the thesis of the theorem follows from Lemma 4.6. □

Remark 4.8. (i) As it is possible to see by elementary examples (see [15, Chapter 1]), $\widehat{\partial}\vartheta(\bar{x})$ may happen to be empty even though ϑ is locally Lipschitz around $\bar{x}$. In these circumstances, the condition in (4.6) can never be satisfied. On the other hand, whenever the p.h. function $\mathrm{D}^-_H\vartheta(\bar{x}; \cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}$ is sublinear (and hence continuous), then $\widehat{\partial}\vartheta(\bar{x}) = \partial\mathrm{D}^-_H\vartheta(\bar{x}; \cdot)(\mathbf{0}) \neq \varnothing$. This happens e.g. (but not only) when $\vartheta : \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex, in which case one has $\widehat{\partial}\vartheta(\bar{x}) = \partial\vartheta(\bar{x})$.

(ii) The local Lipschitz continuity of ϑ near $\bar{x}$ might lead to believe that the Clarke subdifferential may come into play in the current context. Recall that the latter is defined by

$$\partial_C\vartheta(\bar{x}) = \left\{ v \in \mathbb{R}^n : \ \langle v, w\rangle \le \limsup_{\substack{x\to\bar{x}\\ t\downarrow 0}} \frac{\vartheta(x + tw) - \vartheta(x)}{t}, \quad \forall w \in \mathbb{R}^n \right\}.$$

Since, if ϑ is locally Lipschitz around $\bar{x}$, then it is $\widehat{\partial}\vartheta(\bar{x}) \subseteq \partial_C\vartheta(\bar{x})$ (see, for instance, [15, Chapter 1]), it follows that the condition

$$\mathbf{0} \in \partial_C\vartheta(\bar{x}) + \operatorname{int}\left[\left(\bigcap_{z\in K} \mathrm{D}_B f(\bar{x}, z)^{-1}(\mathrm{T}(f(\bar{x}, z); C)) \cap \mathrm{T}(\bar{x}; K)\right)^{\ominus}\right] \tag{4.7}$$

does not imply in general the condition in (4.6).

Acknowledgements
The author would like to express his gratitude to an anonymous referee for relevant remarks, which helped him to improve the quality of the paper.

REFERENCES

[1] Q. H. Ansari, I.V. Konnov, J.C. Yao, Characterizations of solutions for vector equilibrium problems, J. Optim. Theory Appl. 113 (2002) 435–447.
[2] Q. H. Ansari, E. Köbis, J.-C. Yao, Vector variational inequalities and vector optimization. Theory and applications. Vector Optimization, Springer, Cham, 2018.
[3] Q. H. Ansari, W. Oettli, D. Schläger, A generalization of vectorial equilibria, Math. Methods Oper. Res. 46 (1997) 147–152.
[4] J.-P. Aubin, H. Frankowska, Set-valued analysis, Birkhäuser Boston, Boston, MA, 2009.
[5] M. Bianchi, N. Hadjisavvas, and S. Schaible, Vector equilibrium problems with generalized monotone bifunctions, J. Optim. Theory Appl. 92 (1997) 527–542.
[6] M. Castellani, M. Pappalardo, First-order cone approximations and necessary optimality conditions, Optimization 35 (1995) 113–126.
[7] V.F. Demyanov, A.M. Rubinov, Constructive nonsmooth analysis, Peter Lang, Frankfurt am Main, 1995.
[8] K.-H. Elster, J. Thierfelder, Abstract cone approximations and generalized differentiability in nonsmooth optimization, Optimization 19 (1988), 315–341.
[9] X.H. Gong, Strong vector equilibrium problems, J. Global Optim. 36 (2006) 339–349.
[10] X.H. Gong, K. Kimura, J.-C. Yao, Sensitivity analysis of strong vector equilibrium problems, J. Nonlinear Convex Anal. (2008) 83–94.
[11] Z.-Q. Luo, J.-S. Pang, D. Ralph, Mathematical programs with equilibrium constraints, Cambridge University Press, Cambridge, 1996.
[12] Z.-Q. Luo, J.-S. Pang, D. Ralph, S.-Q. Wu, Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints, Math. Programming 75 (1996) Ser. A 19–76.
[13] B.S. Mordukhovich, Variational analysis and generalized differentiation. I. Basic theory, Springer-Verlag, Berlin, 2006.
[14] B.S. Mordukhovich, Variational analysis and generalized differentiation. II. Applications, Springer-Verlag, Berlin, 2006.
[15] B.S. Mordukhovich, Variational analysis and applications, Springer, Cham, 2018.
[16] W. Oettli, A remark on vector-valued equilibria and generalized monotonicity, Acta Math. Vietnam. 22 (1997) 213–221.
[17] J.V. Outrata, M. Kočvara, J. Zowe, Nonsmooth approach to optimization problems with equilibrium constraints. Theory, applications and numerical results, Nonconvex Optimization and its Applications, 28. Kluwer Academic Publishers, Dordrecht, 1998.
[18] J.P. Penot, Calculus without derivatives, Springer, New York, 2013.
[19] R.T. Rockafellar and R.J.-B. Wets, Variational Analysis, Springer-Verlag, Berlin, 1998.
[20] W. Schirotzek, Nonsmooth analysis, Springer, Berlin, 2007.
[21] A. Uderzo, Some enhanced existence results for strong vector equilibrium problems, to appear on Pure and Applied Functional Analysis.
[22] J.J. Ye, Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints, J. Math. Anal. Appl. 307 (2005), 350–369.