Journal of Convex Analysis
Volume 16 (2009), No. 1, 187–210

Nondifferentiable Multiplier Rules for Optimization Problems with Equilibrium Constraints*

N. Movahedian
Department of Mathematics, University of Isfahan,
P. O. Box 81745-163, Isfahan, Iran

S. Nobakhtian†
Department of Mathematics, University of Isfahan,
P. O. Box 81745-163, Isfahan, Iran
nobakht@math.ui.ac.ir

Received: December 7, 2006
Revised manuscript received: April 16, 2008

We consider a mathematical program with equilibrium constraints (MPEC). First we obtain a Lagrange multiplier rule based on the linear subdifferential involving equality, inequality and set constraints. Then we propose new constraint qualifications for M-stationary condition to hold. Finally we establish the Fritz John and Karush-Kuhn Tucker M-stationary necessary conditions for a nonsmooth (MPEC) based on the Michel-Penot subdifferential.

Keywords: Optimization problems, necessary optimality conditions, constraint qualification, nonsmooth analysis

2000 Mathematics Subject Classification: 90C30, 90C46, 49J52

1. Introduction

A mathematical program with equilibrium constraints (MPEC) is a constrained optimization problem in which the essential constraints are defined by a parametric variational inequality or complementarity system. These problems are particularly important for various applications in operational research, engineering, mechanics and economics. The reader is referred to [12, 18] for applications and recent developments.

For MPEC problems, it is well known that the usual nonlinear programming constraint qualifications such as Mangasarian-Fromovitz constraint qualification is violated at every feasible point.

Various stationary conditions for MPECs exist in literature due to different reformulations. The M-stationary condition [17] is the most appropriate stationary condition for MPEC.

Ye in [25] showed that an M-stationary condition is a first order necessary condition and it is sufficient for global or local optimality under some MPEC generalized convexity and differentiability assumptions. Moreover, in [25] the author proposed a constraint

*This research was in part supported by a grant from IPM (No. 86900035).
†Corresponding author

ISSN 0944-6532 / $ 2.50

qualification for M-stationary conditions to hold. Later, Flegel and Kanzow [4] proved that M-stationary is a first order necessary condition under a weak Abadi-type constraint qualification for the case where all the functions involved except the objective function are smooth. Mordukhovich subdifferential was used for nonsmooth terms in [4].

In this paper, we introduce nonsmooth M-stationary condition via Michel-Penot subdifferential and show that it is a first order necessary condition for MPEC without any smooth assumptions. The reason for our selection of this subdifferential is that the Michel-Penot subdifferential is one of the smallest subdifferentials which coincides with classical derivative when the function is Gâteaux differentiable. The multiplier rules in terms of other bigger generalized gradients follow immediately. For this purpose we need a suitable Lagrange multiplier rule corresponding with equality, inequality and closed set constraints based on a subdifferential which is contained in Michel-Penot and Mordukhovich subdifferentials, and has a useful calculus. It appears that the linear subdifferential introduced by Treiman, is suitable.

In [22] Treiman introduced the linear generalized gradient. In [21] its nice calculus is described. Then in [23] Treiman proved a Lagrange multiplier rule for finite dimensional Lipschitz problems. The result of [23] was based on either both the linear generalized gradient and the generalized gradient of Mordukhovich or the linear generalized gradient and a qualification condition involving the pseudo-Lipschitz behavior of the feasible sets for equality and set constraints under perturbations.

We use a geometric constraint qualification, named in this paper by *the distance intersection property* (DIP), which is weaker than pseudo-Lipschitz condition. This notion appeared in the papers by Jourani and Thibault [9, 10] and Ioffe [8] to calculus of normal cones and subdifferentials and obtain multiplier rules. It is worth mentioning that (DIP) is the nonconvex case of local linear regularity studied in [20]. We derive the Fritz-John and Karush-Kuhn-Tucker M-stationary conditions for nonsmooth MPEC based on Michel-Penot subdifferential under (DIP) assumptions without requiring the smoothness of the objective function and constraints. This result generalizes the results of [25, 4].

The organization of this paper is as follows. In the next section we provide preliminary definitions and results to be used in the rest of the paper. In Section 3, we propose the concept of (DIP) as a constraint qualification to obtain a new multiplier rule based on the linear subdifferential. Then we compare our new constraint qualification with pseudo-Lipschitz property, calmness [5, 6] and the finite dimensional case of the normal qualification condition [16].

In Section 4 we give several examples for which (DIP) is satisfied but pseudo-Lipschitz property does not hold. Section 5 is devoted to the Fritz-John and Karush-Kuhn-Tucker M-stationary necessary conditions for nonsmooth MPECs based on Michel-Penot subdifferential.

2. Preliminaries

In finite-dimensional space $\mathbb{R}^n$, $\|.\|$ is the Euclidean norm and $\|.\|_1$ is the norm defined by

$$\|(x_1, \ldots, x_n)\|_1 = \sum_{i=1}^{n} |x_i|, \quad \text{for } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

For nonempty set $C \subset \mathbb{R}^n$, the function $d_C : \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined by

$$d_C(x) := \inf\{\|x-y\| : y \in C\}, \quad \forall x \in \mathbb{R}^n.$$

For given point $x \in \mathbb{R}^n$ and positive number $r > 0$, $\mathbb{B}(x,r)$ ($\bar{\mathbb{B}}(x,r)$) is the open (closed) ball centered at x with radius r. Also, $\mathbb{B}$ denotes the unit ball. We recall the definitions of the linear and Mordukhovich subdifferential from [3, 23]. First we define proximal normals and proximal subgradients.

Definition 2.1.

(a) Let $C \subseteq \mathbb{R}^n$ be a closed set and let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be lower semicontinuous (lsc). A $v \in \mathbb{R}^n$ is a proximal normal to C at $x \in C$ if for some $\lambda > 0$,

$$C \cap \overline{\mathbb{B}}(x + \lambda v, \lambda\|v\|) = \{x\}.$$

The set of all such vectors is called the proximal normal cone to C at x and is denoted by $N_P(C;x)$.

(b) A $w \in \mathbb{R}^n$ is a proximal subgradient of f at x if, for some $\mu \geq 0$,

$$f(y) \geq f(x) + < w, y - x > -\mu \parallel y - x \parallel^2,$$

on a neighborhood of x. The set of all such vectors is called the proximal subdifferential of f at x and is denoted by $\partial_p f(x)$.

An element of the normal cone of Mordukhovich is defined as the limit of a sequence of proximal normals [3, 11, 15, 19]. To define the linear normal cone, a restriction in the convergence of the proximal normals is used [23, 21].

Definition 2.2. A sequence of proximal normals $v^k \longrightarrow v$ to a closed set $C \in \mathbb{R}^n$ at $x^k \longrightarrow \overline{x}$ is linear if, either $x^k \neq \overline{x}$ for all k, and for some $\lambda > 0$ and all sufficiently large k,

$$C \cap \overline{\mathbb{B}}\left(x^k + \lambda \parallel x^k - \overline{x} \parallel v^k, \lambda \parallel x^k - \overline{x} \parallel\parallel v^k \parallel\right) = \left\{x^k\right\}$$

or $x^k = \overline{x}$ for all k.

By using this definition one can define the linear normal cone.

Definition 2.3. Let C be a closed subset of $\mathbb{R}^n$. The linear normal cone to C at $\overline{x}$ is

$$N_l(C;\overline{x}) = \text{cl}\{v : v \text{ is the limit of a linear sequence of proximal normals to } C \text{ at } x^k \longrightarrow \overline{x}\}.$$

The Mordukhovich normal cone to C at $\overline{x}$ is

$$N_M(C;\overline{x}) = \{v : v \text{ is the limit of a sequence of proximal normals to } C \text{ at } x^k \longrightarrow \overline{x}\}.$$

In the sequel the notation $\parallel x - y \parallel_f$ means $\parallel x - y \parallel + \mid f(x) - f(y) \mid$.

Definition 2.4. A sequence of proximal subgradients $v^k \longrightarrow v$ to a lsc function f for $\overline{x}$ is linear if either there are $x^k \longrightarrow \overline{x}$, $f(x^k) \longrightarrow f(\overline{x})$, $x^k \neq \overline{x}$ and $\mu, \delta > 0$ such that

$$f(x^k + h) \geq f(x^k) + < v^k, h > - \frac{\mu}{\| x^k - \overline{x} \|_f} \| h \|^2,$$

for $x^k + h \in \mathbb{B}(x^k, \delta \| x^k - \overline{x} \|_f)$, or v^k is a proximal subgradient to f at $\overline{x}$ for all k.

Note that if f is locally Lipschitz around $\overline{x}$ with constant $L > 0$, then for any k large enough,

$$\|x^k - \overline{x}\| \leq \|x^k - \overline{x}\|_f \leq (1 + L)\|x^k - \overline{x}\|,$$

thus $\|x^k - \overline{x}\|_f$ may be replaced by $\|x^k - \overline{x}\|$ in Definition 2.4.

Definition 2.5. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be lsc. The linear subdifferential to f at $\overline{x}$ is the set

$$\partial_l f(\overline{x}) := \text{cl}\{v : v \text{ is the limit of a linear sequence of proximal subgradients to } f \text{ for } \overline{x}\}.$$

The Mordukhovich subdifferential to f at $\overline{x}$ is the set

$$\partial_M f(\overline{x}) := \{v : v \text{ is the limit of a sequence of proximal subgradients to } f \text{ at } x^k \longrightarrow \overline{x}\}.$$

The following result concerning linear and Mordukhovich normal cone is needed in this paper. We refer the reader to [15, 22] for a proof.

Theorem 2.6. *If C be a closed subset of $\mathbb{R}^n$. Then*

$$N_l(C; \overline{x}) = \text{cl} \cup_{\lambda \geq 0} \lambda \partial_l d_C(\overline{x}),$$

and

$$N_M(C; \overline{x}) = \cup_{\lambda \geq 0} \lambda \partial_M d_C(\overline{x}).$$

We recall the following calculus rules for linear subdifferential which can be found in [15, 19]. Similar results for Mordukhovich subdifferential can also be found in [7, 14, 22].

Theorem 2.7. *Let f be a lsc function from $\mathbb{R}^n$ to $\mathbb{R}$ and let g be a locally Lipschitz function from $\mathbb{R}^n$ to $\mathbb{R}$. Then if $f(\overline{x})$ is finite,*

a) $\partial_l(f + g)(\overline{x}) \subseteq \partial_l f(\overline{x}) + \partial_l g(\overline{x})$,

b) *if* $\alpha \geq 0$, *then* $\partial_l(\alpha f)(\overline{x}) = \alpha \partial_l f(\overline{x})$,

c) *if $\overline{x}$ is a local minimizer of f then* $0 \in \partial_l f(\overline{x})$.

Theorem 2.8. *Let $g_1, g_2, \ldots, g_k$ be a finite collection of locally Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}$. Then*

$$\partial_l \left(\left(\max_i \right) g_i \right) (x) \subseteq \left\{ \sum_{i=1}^k \lambda_i \partial_l g_i(x) : \lambda_i \geq 0,\ \lambda_i = 0 \ \textit{if}\ i \notin I(x) \ \textit{and}\ \sum_{i=1}^k \lambda_i = 1 \right\},$$

where

$$I(x) = \left\{ i : \left(\max_i \right) g_i(x) = g_i(x) \right\}.$$

To state the following chain rule, we recall the notion of the linear co-derivative.

Definition 2.9 ([22]). Let $G : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a locally Lipschitz function. The linear co-derivative of G is the multifunction $D_l^* G$ from $\mathbb{R}^m \times \mathbb{R}^n$ to $\mathbb{R}^n$ such that

$$D_l^* G(x)(y^*) = \{v^* : (v^*, -y^*) \in N_l(\operatorname{gph} G; (x, G(x)))\},$$

where $\operatorname{gph} G = \{(x, G(x)) \in \mathbb{R}^m \times \mathbb{R}^n : x \in \mathbb{R}^n\}$.

Theorem 2.10 ([22]). *Let $G : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ and $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be locally Lipschitz functions. Then*

$$\partial_l f \circ G(\overline{x}) \subseteq D_l^* G(\overline{x}) \partial_l f(G(\overline{x})).$$

The other result is a "scalarization" formula for the co-derivative.

Theorem 2.11 ([21]). *Let $F : \mathbb{R}^m :\longrightarrow \mathbb{R}^n$ be a locally Lipschitz function. Then*

$$D_l^* F(x)(y) = \partial_l < y, F > (x), \quad \forall y \in \mathbb{R}^n.$$

The same results hold if the linear subdifferential and co-derivative are replaced by those of Mordukhovich.

The following result proved by Treiman in [23], describes that the linear and also the Mordukhovich normal cone to a set of the form $\{x : h(x) = 0\}$ is contained in the positive linear combinations of the subdifferential of h, under some conditions.

Theorem 2.12. *Suppose h is a locally Lipschitz function from $\mathbb{R}^n$ to $\mathbb{R}$ with $h(\overline{x}) = 0$. Let $C = \{x : h(x) = 0\}$. If $0 \notin \partial_l h(\overline{x}) \cup \partial_l(-h)(\overline{x})$, then*

$$N_l(C; \overline{x}) \subseteq \cup_{\alpha \in \mathbb{R}} \partial_l(\alpha h)(\overline{x}).$$

If $0 \notin \partial_M h(\overline{x}) \cup \partial_M(-h)(\overline{x})$, then

$$N_M(C; \overline{x}) \subseteq \cup_{\alpha \in \mathbb{R}} \partial_M(\alpha h)(\overline{x}).$$

Definition 2.13 ([13]). Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be locally Lipschitz function around $x \in \mathbb{R}^n$. The Michel-Penot directional derivative of f at x is defined by

$$f^{\diamond}(x; h) = \sup_{u \in \mathbb{R}^n} \limsup_{t \downarrow 0} \frac{f(x + th + tu) - f(x + tu)}{t},$$

and the Michel-Penot subdifferential of f at x by

$$\partial_{\diamond} f(x) := \{\xi \in \mathbb{R}^n : f^{\diamond}(x; h) \geq \langle \xi, d \rangle \quad \forall h \in \mathbb{R}^n\}.$$

A useful property of Michel-Penot subdifferential is that

$$\partial_{\diamond}(-f)(x) = -\partial_{\diamond} f(x),$$

whenever f is locally Lipschitz around x. Also, note that for locally Lipschitz functions on finite-dimensional spaces, the closed convex hull of the linear subdifferential is the Michel-Penot subdifferential [21].

Now we recall the following results from [3]. Let C be a nonempty subset of $\mathbb{R}^n$ and $x \notin C$. The projection set of C at x, denoted by $\operatorname{proj}_C(x)$, is defined by

$$\operatorname{proj}_C(x) := \{c \in C : d_C(x) = \|x - c\|\}.$$

Theorem 2.14. *Suppose C is a nonempty subset of $\mathbb{R}^n$. Let $x \in \mathbb{R}^n$, $c \in C$ and $c \in \operatorname{proj}_C(x)$. Then for all $t \in (0,1)$,*

$$\operatorname{proj}_C(c + t(x - c)) = \{c\}.$$

Theorem 2.15. *Suppose that x is not in the closed subset C of $\mathbb{R}^n$. If the vector ζ belongs to $\partial_p d_C(x)$, Then there exists a point $c \in C$ such that the following conditions hold:*

(a) $\operatorname{proj}_C(x) = \{c\}$
(b) *The Fréchet derivative $d'_C(x)$ exists, and*

$$\{\zeta\} = \partial_p d_C(x) = \{d'_C(x)\} = \left\{\frac{x - c}{\|x - c\|}\right\}$$

(c) $\zeta \in N_p(C; c)$.

3. Constraint qualifications

To establish the main result of this section we need to prove the following Theorem.

Theorem 3.1. *Let $C_1, C_2, \ldots, C_k$ be closed nonempty subsets of $\mathbb{R}^n$ and suppose $\overline{x} \in C = \cap_{i=1}^k C_i$. Suppose that for some $\varepsilon > 0$ and $\sigma > 0$,*

$$d_C(y) \leq \frac{1}{\sigma}\left(\max_{1\leq i\leq k} d_{C_i}\right)(y), \tag{1}$$

for all $y \in \overline{x} + \varepsilon\mathbb{B}$ .Then

$$\partial_l d_C(\overline{x}) \subseteq \sum_{i=1}^k N_l(C_i; \overline{x}).$$

Proof. Consider the following sets:

$$A := \{v : v \text{ is the limit of a linear sequence} \\ \text{of proximal subgradients to } d_C \text{ at } \overline{x}\},$$

$$A' := \Big\{v : v \text{ is the limit of a linear sequence} \\ \text{of proximal subgradients to } \frac{1}{\sigma}\max_i\{d_{C_i}(.)\} \text{ at } \overline{x}\Big\}.$$

It is sufficient to prove that $A \subseteq A'$. Assume that $v \in A$. Therefore there exist sequences $x^k \longrightarrow \overline{x}$ and $v^k \longrightarrow v$, such that $v^k \in \partial_p d_C(x^k)$ and either $x^k = \overline{x}$ for all k, or $x^k \neq \overline{x}$ for all k and for some positive numbers $\mu, \delta > 0$, we have

$$d_C(x^k + h) \geq d_C(x^k) + \langle v^k, h\rangle - \frac{\mu}{\| x^k - \overline{x} \|} \| h \|^2, \tag{2}$$

for $x^k + h \in \mathbb{B}(x^k, \delta \| x^k - \overline{x} \|)$.

If $x^k = \overline{x}$ for all k, then $d_C(\overline{x}) = \frac{1}{\sigma}\max_i d_{C_i}(\overline{x}) = 0$. Furthermore for all k, $v^k \in \partial_p d_C(\overline{x})$, hence by Definition 2.1 and (1), there exist positive numbers σ_k, η_k such that for all h with $\|h\| \leq \eta_k$,

$$\begin{aligned}\frac{1}{\sigma}\max_i d_{C_i}(\overline{x}+h) &\geq d_C(\overline{x}+h)\\ &\geq d_C(\overline{x}) + < v^k, h > -\sigma_k\|h\|^2\\ &= \frac{1}{\sigma}\max_i d_{C_i}(\overline{x}) + < v^k, h > -\sigma_k\|h\|^2.\end{aligned}$$

Therefore $v^k \in \partial_p(\frac{1}{\sigma}\max_i d_{C_i}(\overline{x}))$ for all k. Thus v is the limit of a linear sequence of subgradients for $\frac{1}{\sigma}\max_i d_{C_i}$ at $\overline{x}$. Hence $v \in A'$.

For each $k \in \mathbb{N}$, according to Theorem 2.15 there exists $z^k \in C$ such that, $\operatorname{proj}_C(x^k) = \{z^k\}$. If $z^k = x^k$ for infinitely many k, then $d_C(x^k) = 0 = \frac{1}{\sigma}\max_i d_{C_i}(x^k)$ and hence (2) combined with assumption (1) gives that v^k is a linear sequence for $\frac{1}{\sigma}\max_i d_{C_i}$ at x^k, which ensures that $v \in A'$.

Suppose that $z^k \neq x^k$ for k large enough. Then according to Theorem 2.15 $v^k = (x^k - z^k)/\|x^k - z^k\|$ and hence by [1, Theorem 4.1], $v^k \in N_P(C; z^k) \cap \mathbb{B} = \partial_p d_C(z^k)$. Note that

$$\|z^k - \overline{x}\| \leq \|z^k - x^k\| + \|x^k - \overline{x}\| \leq 2\|x^k - \overline{x}\|. \tag{3}$$

If $\|h\| \leq (\delta/2)\|z^k - \overline{x}\|$, then by (2)

$$\begin{aligned}&d_C(x^k) + < v^k, h > -(\mu\|h\|^2)/\|x^k - \overline{x}\|\\ &\leq d_C(x^k+h) \leq d_C(z^k+h) + \|z^k - x^k\| = d_C(z^k+h) + d_C(x^k),\end{aligned}$$

hence combining with (3), we get

$$\begin{aligned}&\frac{1}{\sigma}\max_i d_{C_i}(z^k) + < v^k, h > -(2\mu\|h\|^2)/\|z^k - \overline{x}\|\\ =& < v^k, h > -(2\mu\|h\|^2)/\|z^k - \overline{x}\| \leq d_C(z^k+h) \leq \frac{1}{\sigma}\max_i d_{C_i}(z^k+h),\end{aligned}$$

where the last inequality being due to the assumption (1). So v^k is a linear sequence for $\frac{1}{\sigma}\max_i d_{C_i}$ at z^k. Deduce that $A \subseteq A'$ and the proof is complete. $\square$

Observe that the assumption (1) has been used by A. Jourani and L. Thibault [10] and also by A. D. Ioffe [8] to obtain the inclusion

$$\partial_M d_C(\overline{x}) \subseteq \gamma \sum_{i=1}^{k} \partial_M d_{C_i}(\overline{x}) \subseteq \sum_{i=1}^{k} N_M(C_i; \overline{x}),$$

for the Mordukhovich normal cone in Asplund space and the geometric Ioffe normal cone in general Banach space. Note that in finite dimension those cones coincide, and the following holds.

Theorem 3.2 ([8, 10]). *Let $C_1, C_2, \ldots, C_k$ be closed nonempty subsets of $\mathbb{R}^n$ and suppose $\overline{x} \in C = \cap_{i=1}^k C_i$. Suppose that for some $\varepsilon > 0$ and $\sigma > 0$,*

$$d_C(y) \leq \frac{1}{\sigma}\left(\max_{1\leq i\leq k} d_{C_i}\right)(y),$$

for all $y \in \overline{x} + \varepsilon\mathbb{B}$*. Then*

$$N_M(C;\overline{x}) \subseteq \sum_{i=1}^{k} N_M(C_i;\overline{x}).$$

Now, we introduce a simple but useful geometric property for points which belongs to intersection of a finite collection of closed sets. We show that this property is equivalent with calmness [6, 19] and weaker than pseudo-Lipschitz condition (or Aubin property) (see, e.g., [19]). We also show that this property is weaker than the finite dimensional case of normal qualification condition employed in [16].

Definition 3.3. Let $C_1, C_2, \ldots, C_k$ be closed subsets of $\mathbb{R}^n$ and suppose $\overline{x} \in C = \cap_{i=1}^k C_i$. We say that $C_1, C_2, \ldots, C_k$ has the distance intersection property (DIP) at $\overline{x}$, if there exist positive numbers σ, ε such that,

$$d_C(y) \leq \frac{1}{\sigma} \max_i d_{C_i}(y),$$

for all $y \in \overline{x} + \varepsilon\mathbb{B}$.

Remark 3.4. We may assume $\sigma \leq 1$, since $\max_i d_{C_i}(x) \leq d_C(x)$.

Now, we recall the definitions of pseudo-Lipschitz property and calmness [5, 6, 19, 23].

Definition 3.5. Let $C_1, C_2, \ldots, C_k$ be closed subsets of $\mathbb{R}^n$. Suppose that $\Phi(v)$ is the set-valued function from $\mathbb{R}^{nk}$ to $\mathbb{R}^n$ defined by

$$\Phi(v) := \Phi(v_1, v_2, \ldots, v_k) = \cap_{i=1}^k (C_i + v_i), \tag{4}$$

for all $(v_1, v_2, \ldots, v_k) \in \mathbb{R}^{nk}$.

(a) We say that Φ is calm at $(\overline{v}, \overline{x}) \in \operatorname{gph}\Phi$, if there exist $c > 0$ and neighborhoods V of $\overline{v}$ and Z of $\overline{x}$ such that for any $v \in V$ and $x \in \Phi(v) \cap Z$,

$$d_{\Phi(\overline{v})}(x) \leq c\|v - \overline{v}\|.$$

(b) We say that Φ is pseudo-Lipschitz (or has Aubin property) at $(\overline{v}, \overline{x}) \in \operatorname{gph}\Phi$, if there exist $c > 0$ and neighborhoods V of $\overline{v}$ and Z of $\overline{x}$ such that for any $v_1, v_2 \in V$ and $x \in \Phi(v_2) \cap Z$,

$$d_{\Phi(v_1)}(x) \leq c\|v_1 - v_2\|.$$

In the next theorem and its corollary, we prove that (DIP) condition is equivalent with calmness and weaker than pseudo-Lipschitz property.

Theorem 3.6. *Let $C_1, C_2, \ldots, C_k$ be nonempty closed subsets of $\mathbb{R}^n$ and $\overline{x} \in C = \cap_{i=1}^k C_i$. Suppose that Φ is defined as* (4)*. Let $\overline{0} = (0, \ldots, 0) \in \mathbb{R}^{nk}$. The multifunction Φ is calm at $(\overline{0}, \overline{x}) \in \operatorname{gph}\Phi$, if and only if $\overline{x}$ is a* (DIP) *point for $C_1, C_2, \ldots, C_k$.*

Proof. In the following argument we consider the $\|.\|_\infty$ in $\mathbb{R}^n$ and $\mathbb{R}^{nk}$. The extension to an arbitrary norm is trivial since all the norms are equivalent in finite dimensions.

First, we assume that Φ is calm at $(\overline{0}, \overline{x}) \in \operatorname{gph}\Phi$. Therefore, there exist positive numbers c and ε such that for every $v \in \varepsilon\mathbb{B}$ and $y \in \Phi(v) \cap (\overline{x} + \varepsilon\mathbb{B})$,

$$d_{\Phi(\overline{0})}(y) \leq c\|v\|_\infty. \tag{5}$$

It is clear that $\Phi(\overline{0}) = \cap_{i=1}^{k} C_i = C$. Without loss of generality, we may assume $c \geq 1$. Let $y \in \overline{x} + \varepsilon\mathbb{B}$. Assume that for each i, $c_i \in \operatorname{proj}_{C_i}(y)$ and $v_i := y - c_i$. Then, $y \in \cap_{i=1}^{k}(C_i + v_i) = \Phi(v_1, v_2, \ldots, v_k)$. On the other hand,

$$\|v_i\|_\infty = d_{C_i}(y) \leq \|y - \overline{x}\|_\infty \leq \varepsilon.$$

Consequently $\|(v_1, \ldots, v_k)\|_\infty \leq \varepsilon$. By (5), we have,

$$d_C(y) \leq c\|(v_1, \ldots, v_k)\|_\infty = c \max_i\{\|v_i\|_\infty\} = c \max_i\{d_{C_i}(y)\}.$$

Setting $\frac{1}{\sigma} = c$, we deduce that $\overline{x}$ is a (DIP) point for $C_1, C_2, \ldots, C_k$.

Conversely, suppose that for some $\varepsilon > 0$ and any $y \in \overline{x} + \varepsilon\mathbb{B}$, one has $d_C(y) \leq \frac{1}{\sigma} \max_i d_{C_i}(y)$. Then for $y \in \Phi(v) \cap (\overline{x} + \varepsilon\mathbb{B})$, $d_C(y) \leq \frac{1}{\sigma} \max_i \|v_i\|$. This concludes the proof. □

It is clear that the Aubin property of Φ at $(\overline{0}, \overline{x})$, implies the calmness of Φ at the same point. Therefore the following result is a direct consequence of Theorem 3.6.

Corollary 3.7. *Let $C_1, C_2, \ldots, C_k$, $\overline{x}$ and Φ be as in Theorem* 3.6. *If Φ is pseudo-Lipschitz at $(\overline{0}, \overline{x}) \in \operatorname{gph}\Phi$, then $\overline{x}$ is a* (DIP) *point for $C_1, C_2, \ldots, C_k$.*

Now we prove a new Fritz-John type necessary condition via linear generalized gradient, for Lipschitz problems with inequality, equality and nonconvex set constraints. Consider the following optimization problem (P)

$$(P) \qquad \min f(x) \quad \text{subject to} \quad \begin{array}{ll} g_i(x) \leq 0, & i = 1, 2, \ldots, m \\ h_j(x) = 0, & j = 1, 2, \ldots, k \end{array} \qquad x \in U,$$

where f, g_i $(i = 1, 2, \ldots, m)$ and h_j $(j = 1, \ldots k)$ are locally Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}$ and U is a closed subset of $\mathbb{R}^n$.

Let for $j = 1, 2, \ldots, k$, $C_j = \{x \in \mathbb{R}^n : h_j(x) = 0\}$, and $C_{k+1} = U$. Take $\overline{x}$ to be a feasible point of (P). Treiman in [23] stated a Fritz-John type Lagrange multipliers result via linear generalized gradient for (P) under pseudo-Lipschitzness of the function Φ defined in (4). We derive the result of Treiman based on (DIP) condition.

Theorem 3.8. *Let $\overline{x}$ be a local minimizer for (P), which is a* (DIP) point for $C_1, C_2, \ldots,$ C_{k+1}. Then there exist $\lambda_0 \geq 0, \lambda_i \geq 0, \alpha_j$, not all zero, such that

$$0 \in \lambda_0 \partial_l f(\overline{x}) + \sum_{i=1}^{m} \lambda_i \partial_l g_i(\overline{x}) + \sum_{j=1}^{k} \partial_l(\alpha_j h_j)(\overline{x}) + N_l(U; \overline{x}).$$

$$\lambda_i g_i(\overline{x}) = 0, \quad i = 1, 2, \ldots, m.$$

Proof. Let

$$F(x) = \max\{f(x) - f(\overline{x}), g_1(x), g_2(x), \ldots, g_m(x)\}.$$

Then $\overline{x}$ is a local minimizer for the following problem (P'),

$$(P') \qquad \min F(x) \quad \text{subject to} \quad x \in C,$$

where $C = \cap_{j=1}^{k+1} C_j$. Since $f, g_1, g_2, \dots, g_m$ are locally Lipschitz, it follows that F is locally Lipschitz. By exact penalty Theorem (see e.g. [3], Theorem 1.6.3), there exists $L > 0$ such that $\overline{x}$ is a local minimizer for the function $F + Ld_C$ on $\mathbb{R}^n$. Therefore by Theorem 2.7, $0 \in \partial_l (F + Ld_C)(\overline{x})$ and,

$$0 \in \partial_l F(\overline{x}) + \partial_l (Ld_C)(\overline{x}) = \partial_l F(\overline{x}) + L\partial_l d_C(\overline{x}). \tag{6}$$

Since, $\overline{x}$ is a (DIP) point for $C_1, C_2, \dots, C_{k+1}$, by Theorem 3.1 we have,

$$L\partial_l d_C(\overline{x}) \subseteq \sum_{i=1}^{k+1} N_l(C_i; \overline{x}). \tag{7}$$

On the other hand Theorem 2.8 implies

$$\partial_l F(\overline{x}) \subseteq \left\{ \lambda_0 \partial_l f(\overline{x}) + \sum_{i=1}^{m} \lambda_i \partial_l g_i(\overline{x}),\ \lambda_i \geq 0,\ \sum_{i=0}^{m} \lambda_i = 1 \right\}, \tag{8}$$

and $\lambda_i = 0$ if $g_i(\overline{x}) < 0$. From (6)–(8), we obtain

$$0 \in \lambda_0 \partial_l f(\overline{x}) + \sum_{i=1}^{m} \lambda_i \partial_l g_i(\overline{x}) + \sum_{j=1}^{k} N_l(C_j; \overline{x}) + N_l(U; \overline{x}).$$

If $0 \in \partial_l h_j(\overline{x}) \cup \partial_l(-h_j)(\overline{x})$ for some j, then the result holds true. Suppose $0 \notin \cup_{j=1}^{k} (\partial_l h_j(\overline{x}) \cup \partial_l(-h_j)(\overline{x}))$. Theorem 2.12 implies that there exist $\lambda_0 \geq 0, \lambda_i \geq 0, \alpha_j$, not all zero, such that $\lambda_i g_i(\overline{x}) = 0$ and

$$0 \in \lambda_0 \partial_l f(\overline{x}) + \sum_{i=1}^{m} \lambda_i \partial_l g_i(\overline{x}) + \sum_{j=1}^{k} \partial_l(\alpha_j h_j)(\overline{x}) + N_l(U; \overline{x}).$$

Hence the proof is complete. □

The following proposition appears as an exercise in [3]. We provide a short proof for the reader's convenience. In fact, the assumption of this proposition is known to entail a stronger result. (See Remark 3.11 at the end of this section.)

Proposition 3.9. *Let C_1, C_2 be closed subsets of $\mathbb{R}^n$, $\overline{x} \in C_1 \cap C_2$ and*

$$N_M(C_1; \overline{x}) \cap (-N_M(C_2; \overline{x})) = \{0\}.$$

Then for some positive numbers ε, σ, and for each $y \in \overline{x} + \varepsilon B$,

$$d_C(y) \leq \frac{1}{\sigma} \max\{d_{C_1}(y), d_{C_2}(y)\}.$$

Proof. We proceed by contradiction. Suppose that for each $\varepsilon > 0$ and $\sigma > 0$ there exists a point $y_\varepsilon \in \overline{x} + \varepsilon\mathbb{B}$ such that

$$\sigma d_C(y_\varepsilon) > \max\{d_{C_1}(y_\varepsilon), d_{C_2}(y_\varepsilon)\}. \tag{9}$$

Claim. There exist sequences $c_1^k \in C_1$ and $c_2^k \in C_2$ and v^k such that $c_1^k, c_2^k \longrightarrow \overline{x}$ and

$$\|v^k\| = 1, \quad v^k \in N_P(C_1 : c_1^k) \cap (-N_P(C_2; c_2^k)).$$

Proof of the Claim. Fix $k \in \mathbb{N}$, by (9) for all $n \in \mathbb{N}$ there exists a point $y_k^{(n)} \in B(\overline{x}; \frac{1}{k})$ such that

$$d_C\left(y_k^{(n)}\right) > n \max\left\{d_{C_1}\left(y_k^{(n)}\right), d_{C_2}\left(y_k^{(n)}\right)\right\}. \tag{10}$$

Now take $c_{k(1)}^{(n)} \in \operatorname{proj}_{C_1}(y_k^{(n)}), c_{k(2)}^{(n)} \in \operatorname{proj}_{C_2}(y_k^{(n)})$. From (10),

$$\left\|y_k^{(n)} - \frac{1}{2}\left(c_{k(1)}^{(n)} + c_{k(2)}^{(n)}\right)\right\| \leq \frac{1}{n} d_C\left(y_k^{(n)}\right) \leq \frac{1}{nk}. \tag{11}$$

Since sequences $y_k^{(n)}, c_{k(1)}^{(n)}$ and $c_{k(2)}^{(n)}$ are bounded we can extract convergent subsequences (without relabelling) $y_k^{(n)} \longrightarrow y^k, c_{k(1)}^{(n)} \longrightarrow c_1^k$ and $c_{k(2)}^{(n)} \longrightarrow c_2^k$ as $n \to \infty$. For $j = 1, 2$, we have $c_j^k \in C_j$ and $d_{C_j}(y^k) = \|y^k - c_j^k\|$, which implies that $C_j \cap \mathbb{B}(y^k; \|y^k - c_j^k\|) = \{c_j^k\}$ or $y^k - c_j^k \in N_P(C_j; c_j^k)$. Using (11) we obtain,

$$\left\|y_k^{(n)} - c_{k(j)}^{(n)} - \frac{1}{2}\left(c_{k(i)}^{(n)} - c_{k(j)}^{(n)}\right)\right\| \leq \frac{1}{nk}, \ \forall n \in \mathbb{N}, \ \forall i, j \in \{1, 2\}.$$

By taking limit as $n \longrightarrow \infty$, we obtain $y^k - c_j^k = \frac{1}{2}(c_i^k - c_j^k)$ for $i, j = 1, 2$. Thus, $\frac{1}{2}(c_i^k - c_j^k) \in N_P(C_j; c_j^k)$. Take $v^k = \frac{c_2^k - c_1^k}{\|c_2^k - c_1^k\|}$. It follows that $\|v^k\| = 1$ and $v^k \in N_P(C_1; c_1^k) \cap (-N_P(C_2; c_2^k))$. Suppose that $v^k \longrightarrow v$ as $k \longrightarrow \infty$. Thus $\|v\| = 1$ and $v \in N_M(C_1; \overline{x}) \cap (-N_M(C_2; \overline{x}))$, which is a contradiction. Hence the proof is complete. □

Theorem 3.10. *Suppose that h_j is a locally Lipschitz function from $\mathbb{R}^n$ to $\mathbb{R}$, for $j = 1, 2, .., k$ and*

$$C_j = \{x : h_j(x) = 0\}, \quad \overline{x} \in \cap_{j=1}^k C_j = C.$$

Furthermore assume that the following nonsmooth linear independence condition holds true,

"for any $\gamma_1, \gamma_2, \ldots, \gamma_k$ with $\gamma_j \in \partial_M h_j(\overline{x}) \cup \partial_M(-h_j)(\overline{x})$, if there exists $\alpha_j \geq 0$ such that $\sum_{j=1}^k \alpha_j \gamma_j = 0$ then $\alpha_j = 0$ for $j = 1, 2, \ldots, k$."

Then $\overline{x}$ is a (DIP) *point for $C_1, C_2, \ldots, C_k$.*

Proof. We prove by induction. For $k = 2$, by using Theorem 3.9, it is sufficient to prove

$$N_M(C_1; \overline{x}) \cap (-N_M(C_2; \overline{x})) = \{0\}.$$

Let $\xi \in N_M(C_1; \overline{x}) \cap (-N_M(C_2; \overline{x}))$ and $\xi \neq 0$. The nonsmooth linear independence condition implies that, $0 \notin \cup_{j=1}^2(\partial_M h_j(\overline{x}) \cup \partial_M(-h_j)(\overline{x}))$. Theorem 2.12 implies that there exist some $\alpha_j \in \mathbb{R}/\{0\}$, $j = 1, 2$, such that

$$\begin{aligned} \xi &\in \partial_M(\alpha_1 h_1)(\overline{x}) \cap (-\partial_M(\alpha_2 h_2)(\overline{x}) \\ &= |\alpha_1| \partial_M\left(\frac{\alpha_1}{|\alpha_1|} h_1\right) \cap \left(-|\alpha_2| \partial_M\left(\frac{\alpha_2}{|\alpha_2|} h_2\right)(\overline{x})\right). \end{aligned}$$

Thus there exist some $\gamma_j \in \partial_M h_j(\overline{x}) \cup \partial_M(-h_j)(\overline{x})$ such that $\xi = |\alpha_1|\gamma_1 = -|\alpha_2|\gamma_2$,. Thus $\sum_{j=1}^{2} |\alpha_j|\gamma_j = 0$, and by the nonsmooth linear independence assumption, $\alpha_j = 0, j = 1, 2$, which implies that $\xi = 0$. By contradiction, $N_M(C_1;\overline{x}) \cap (-N_M(C_2;\overline{x})) = \{0\}$.

Now suppose for $j = 1, 2, \dots k-1$ the result holds. Let $C' = \cap_{j=1}^{k-1} C_j$. Therefore there exist positive ε', σ' with $\sigma' \leq 1$ (see Remark 3.4) such that for any $y \in \overline{x} + \varepsilon'\mathbb{B}$,

$$d_{C'}(y) \leq \frac{1}{\sigma'} \left(\max_{1 \leq j \leq k-1} \right) d_{C_j}(y). \tag{12}$$

By virtue of Theorem 3.2

$$N_M(C';\overline{x}) \subseteq \sum_{j=1}^{k-1} N_M(C_j;\overline{x}). \tag{13}$$

By linear independence condition, $0 \notin \cup_{j=1}^{k-1} (\partial_M h_j(\overline{x}) \cup \partial_M(-h_j)(\overline{x}))$. It immediately follows from Theorem 2.12 that, for each j,

$$N_M(C_j;\overline{x}) \subseteq \cup_{\alpha \in \mathbb{R}} \partial_M(\alpha h_j)(\overline{x}). \tag{14}$$

If $\xi \in N_M(C';\overline{x}) \cap (-N_M(C_k;\overline{x}))$ then by similar argument with the first part of proof, $\xi = \sum_{j=1}^{k-1} \alpha_j \gamma_j = -\alpha_k \gamma_k$, for some $\gamma_j \in \partial_M(h_j)(\overline{x}) \cup \partial_M(-h_j)(\overline{x})$ and $\alpha_j \geq 0$. It follows that $\sum_{j=1}^{k} \alpha_j \gamma_j = 0$, and by the nonsmooth linear independence assumption, $\alpha_j = 0$ for all j. Hence $\xi = 0$. Proposition 3.9 yields $d_C(y) \leq \frac{1}{\sigma} \max\{d_{C'}(y), d_{C_k}(y)\}$, for all $y \in \overline{x} + \varepsilon\mathbb{B}$ and for some $\varepsilon, \sigma > 0$. Set $\sigma_0 = \sigma\sigma'$ and $\varepsilon_0 = \min\{\varepsilon, \varepsilon'\}$. Then by inequality (12) we have $d_C(y) \leq \frac{1}{\sigma_0} \max_{1 \leq i \leq k}\{d_{C_i}(y)\}$, for all $y \in \overline{x} + \varepsilon_0\mathbb{B}$. Hence $\overline{x}$ is a (DIP) point for $C_1, C_2, \dots, C_k$. □

Remark 3.11. (a) The condition

$$N_M(C_1;\overline{x}) \cap (-N_M(C_2;\overline{x})) = \{0\},$$

for more than two sets takes the form:

$$\left[\sum_i x_i^* = 0, \quad \text{with } x_i^* \in N_M(C_i;\overline{x}) \right] \Rightarrow x_i^* = 0, \quad \forall i. \tag{15}$$

In [8, 9, 10] it has been proven that condition (15) ensures that there is some constant $\gamma \geq 0$ such that for y near $\overline{x}$ and u_i near 0 one has

$$d_{\cap_i (C_i + u_i)}(y) \leq \gamma \left(\max_i \right) d_{C_i + u_i}(y).$$

(b) It is obvious that when the sets C_i are the null sets of locally Lipschitz functions, the condition (15) coincides with nonsmooth linear independence employed in Theorem 3.10.

4. Examples

The following examples satisfy (DIP) but do not pseudo-Lipschitz condition. For both of these examples, the normal qualification condition does not hold.

Example 4.1. Define

$$C_1 = \left\{(x,y) \in \mathbb{R}^2 : x^2 + (y-1)^2 = 1\right\},$$

and

$$C_2 = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1,\ y \geq 0\right\} \cup \{(0,0)\}.$$

Furthermore define $\Phi : \mathbb{R}^4 \rightrightarrows \mathbb{R}^2$ as

$$\Phi(v_1, v_2) = (C_1 + v_1) \cap (C_2 + v_2).$$

Take $\bar{0} = (0,0)$ and $(v_1^{(n)}, v_2^{(n)}) = (0,0,0,\frac{1}{n^2})$, for each $n \in \mathbb{N}$. Then $(0,0) \in \Phi(\bar{0},\bar{0})$ and $\Phi(v_1^{(n)}, v_2^{(n)})$ contains only two points with the same distance of the origin, i.e. the intersection points of two circles with centers $(0,1)$ and $(0,\frac{1}{n^2})$, and of radius one, and it is easy to see that

$$d_{\Phi\left(v_1^{(n)},v_2^{(n)}\right)}((0,0)) = \left(1 + \frac{1}{n^2}\right)^{\frac{1}{2}} > \frac{1}{n} = n\left\|\left(v_1^{(n)}, v_2^{(n)}\right)\right\|, \quad \forall n \in \mathbb{N}.$$

Since $(v_1^{(n)}, v_2^{(n)}) \longrightarrow (\bar{0},\bar{0})$, Φ is not pseudo-Lipschitz at $(\bar{0},\bar{0},\bar{0}) \in \operatorname{gph}\Phi$. Furthermore, C_1, C_2 do not satisfy the normal qualification condition at $(0,0)$, because $N_M(C_1;(0,0)) = \{(x,y) \in \mathbb{R}^2 : x = 0, y \leq 0\}$, while $N_M(C_2;(0,0)) = \mathbb{R}^2$.

We show that $(0,0)$ is a (DIP) point for C_1, C_2. Let $(x,y) \in B((0,0);\frac{1}{3})$. If $y \geq 0$ then the closest point of C_2 to (x,y) is the origin. Therefore $d_C(x,y) = d_{C_2}(x,y)$. If $y < 0$ and $0 \leq x \leq \frac{1}{3}$, the closest point of $C_2 - \{(0,0)\}$ to (x,y) is $(1,0)$, and we have

$$\|(x,y)\| < \|(1-x,-y)\| = \text{the distance of } (x,y) \text{ to } (1,0).$$

If $y < 0$ and $0 \geq x \geq \frac{-1}{3}$, then the closest point of $C_2 - \{(0,0)\}$ to (x,y) is $(-1,0)$, and we obtain,

$$\|(x,y)\| < \|(1+x,y)\| = \text{the distance of } (x,y) \text{ to } (-1,0).$$

Deduce that, $d_{C_2}(x,y) = \|(x,y)\| = d_C(x,y)$. Thus

$$d_C(x,y) = \max\{d_{C_1}(x,y), d_{C_2}(x,y)\}, \forall (x,y) \in B\left((0,0);\frac{1}{3}\right),$$

which implies that $(0,0)$ is a (DIP) point for C_1, C_2.

Example 4.2. Define:

$$C_1 := \left\{(x,y) \in \mathbb{R}^2 : y = 0\right\}$$

and,

$$C_2 = \left\{(x,y) \in \mathbb{R}^2 : y = -\frac{1}{\sqrt{3}}x, x \leq 0\right\} \cup \left\{(x,y) \in \mathbb{R}^2 : y = \frac{1}{\sqrt{3}}x, x \geq 0\right\}$$

and,

$$C_3 = \left\{(x,y) \in \mathbb{R}^2 : y = -\sqrt{3}x, x \le 0\right\} \cup \left\{(x,y) \in \mathbb{R}^2 : y = \sqrt{3}x, x \ge 0\right\}.$$

Furthermore define $\Phi : \mathbb{R}^6 \rightrightarrows \mathbb{R}^2$ as

$$\Phi(v_1, v_2, v_3) = \cap_{i=1}^3 (C_i + v_i).$$

Take $\overline{v}^n = ((0, \frac{1}{n}), (0,0), (0,0))$. We have $\Phi(\overline{v}^n) = \emptyset$, and for each $y \in \mathbb{R}^2$, $d_{\Phi(\overline{v}^n)}(y) = \infty$. Therefore Φ is not pseudo-Lipschitz at $(0,0)$.
Moreover, $(0,1) \in N_M(C_1;(0,0))$ and $(0, \frac{-1}{2}) \in N_M(C_2;(0,0)) \cap N_M(C_3;(0,0))$, and $(0,1) + (0, \frac{-1}{2}) + (0, \frac{-1}{2}) = (0,0)$. Thus the normal qualification condition is not satisfied for these sets at $(0,0)$.

We show that $(0,0)$ is a (DIP) point for C_1, C_2, C_3 with $\sigma = \frac{1}{2}$. Take $(x,y) \in \mathbb{R}^2$. Let $x \ge 0$. If $y \le -\sqrt{3}x$ then

$$d_C(x,y) = d_{C_2}(x,y) = d_{C_3}(x,y) = \max_i d_{C_i}(x,y).$$

Now, let $y \ge -\sqrt{3}x$ and for $j \in \{1,2,3\}$, denote the angle between the line joining (x,y) to the origin and the right half-line of C_j by θ_j. Then for some j, $\theta_j > \frac{\pi}{6}$. We have $d_{C_j}(x,y) = \sin\theta_j d_C(x,y)$, which implies that $d_C(x,y) \le 2\max_i d_{C_i}(x,y)$. In the case $x \le 0$ by symmetry with respect to y-axis, the proof is similar.

Treiman in [21] by a given example showed that the general Lagrange multiplier rule does not hold with a set constraint and equality constraint even under the smoothness of functions. We recall this example and show that (DIP) condition is not satisfied for it.

Example 4.3 ([21]). Consider the problem (P)

$$(P) \qquad \min\ f(x,y) \quad \text{subject to} \quad \begin{array}{l} g(x,y) = x^2 + (y+2)^2 - 4 = 0 \\ (x,y) \in \overline{\mathbb{B}}((0,1);1) \cup \overline{\mathbb{B}}((0,-1);1), \end{array}$$

with singleton feasible set $\{(0,0)\}$. Take $C = \overline{\mathbb{B}}((0,1);1) \cup \overline{\mathbb{B}}((0,-1);1)$ and $C_1 = \{(x,y) : g(x,y) = 0\}$. The point $(x,y) = (0,0)$, is the optimal solution of (P). It is easy to verify that g is C^∞, $\nabla g(0,0) = (0,4)$ and $N_l(C,(0,0)) = \{(0,0)\}$. Suppose that there exist $\lambda_0 \ge 0$ and λ_1, not both zero, such that

$$(0,0) \in \lambda_0 \partial_l f(0,0) + \partial_l(\lambda_1 g)(0,0) + N_l(C,(0,0)) = \lambda_0 \partial_l f(0,0) + \lambda_1(0,4).$$

Let $f(x,y) = x$. Then $(0,0) = \lambda_0(1,0) + \lambda_1(0,4)$ and $\lambda_0 = \lambda_1 = 0$.

Now, we show that $(0,0)$ is not a (DIP) point for the sets C and C_1. Take $(x_n, y_n) = (\frac{1}{n}, 0)$. Then $\max\{d_C(x_n,y_n), d_{C_1}(x_n,y_n)\} < \frac{1}{2n^2}$. Therefore for any $\sigma, \varepsilon > 0$, if $n > \max\{\frac{1}{2\sigma}, \frac{1}{\varepsilon}\}$ we have,

$$\frac{1}{\sigma}\max\{d_C(x_n,y_n), d_{C_1}(x_n,y_n)\} < \|(x_n,y_n)\| = d_{C\cap C_1}(x_n,y_n),$$

for some $(x_n,y_n) \in B((0,0),\varepsilon)$.

5. M-stationary necessary conditions for nonsmooth MPEC

In this section we consider the following program, known across the literature as a mathematical program with complementary-or often also equilibrium-constraints, MPEC for short:

$$(1) \qquad \min f(z) \quad \text{s.t.} \quad \begin{array}{ll} g(z) \le 0, & h(z) = 0, \\ G(z) \ge 0, & H(z) \ge 0, \\ G(z)^T H(z) = 0, & \end{array}$$

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $G : \mathbb{R}^n \longrightarrow \mathbb{R}^l$, and $H : \mathbb{R}^n \longrightarrow \mathbb{R}^l$ are locally Lipschitz functions.

Definition 5.1. Let z^* be a feasible point of the MPEC (1). We divide the indices of G and H into three sets:

$$\begin{aligned} \alpha := \alpha(z^*) &:= \{i : G_i(z^*) = 0, H_i(z^*) > 0\}, \\ \beta := \beta(z^*) &:= \{i : G_i(z^*) = 0, H_i(z^*) = 0\}, \\ \gamma := \gamma(z^*) &:= \{i : G_i(z^*) > 0, H_i(z^*) = 0\}. \end{aligned}$$

We call z^* nonsmooth M-stationary if there exist

$$\lambda^f \ge 0, \lambda^g \in \mathbb{R}^m, \qquad \lambda^h \in \mathbb{R}^p, \qquad \lambda^G \in \mathbb{R}^l, \qquad \lambda^H \in \mathbb{R}^l,$$

not all zero, such that

$$0 \in \lambda^f \partial_\diamond f(z^*) + \sum_{i=1}^{m} \lambda_i^g \partial_\diamond g_i(z^*) + \sum_{i=1}^{p} \lambda_i^h \partial_\diamond h_i(z^*), - \sum_{i=1}^{l} \left[\lambda_i^G \partial_\diamond G_i(z^*) + \lambda_i^H \partial_\diamond H_i(z^*)\right]$$

$$\begin{array}{lcr} \lambda_\alpha^G \text{ free}, & & \lambda_\gamma^G = 0, \\ & \left(\lambda_i^G > 0 \wedge \lambda_i^H > o\right) \vee \lambda_i^G \lambda_i^H = 0, \quad \forall i \in \beta, & \\ \lambda_\gamma^H \text{ free}, & & \lambda_\alpha^H = 0, \\ & g(z^*) \le 0, \qquad \lambda^g \ge 0, \qquad g(z^*)^T \lambda^g = 0. & \end{array}$$

In this section we will apply results of Section 3 to prove Fritz-John and KKT type M-stationary results for MPEC involving locally Lipschitz functions, based on Michel-Penot subdifferential. To get a Fritz-John type result, in addition to Flegel's condition we impose an extra assumption. However, the KKT type result is proved just under the same condition employed by Flegel.

To succeed in dealing with the complementary term in the constraints of the MPEC, we require a result which investigates the Mordukhovich normal cone to a complementary set. This result was originally stated in a slightly different format by Outrata in [17], see also [24].

Theorem 5.2. *Let the set*

$$\mathcal{C} := \left\{(\xi, \eta) \in \mathbb{R}^l \times \mathbb{R}^l : \xi \ge 0, \eta \ge 0, \xi^T \eta = 0\right\}$$

be given. Then, for an arbitrary but fixed $(\xi,\eta)\in\mathcal{C}$*, define*

$$\mathcal{I}_\xi = \{i : \xi_i = 0, \eta_i > 0\},$$
$$\mathcal{I}_\eta = \{i : \xi_i > 0, \eta_i = 0\},$$
$$\mathcal{I}_{\xi\eta} = \{i : \xi_i = 0, \eta_i = 0\}.$$

Then the Mordukhovich normal cone to $\mathcal{C}$ *at* (ξ,η) *is given by*

$$N_M(\mathcal{C};(\xi,\eta)) = \big\{(x,y)\in\mathbb{R}^l\times\mathbb{R}^l : x_{\mathcal{I}_\eta}=0, y_{\mathcal{I}_\xi}=0,\\ (x_i<0\wedge y_i<0)\vee x_iy_i=0,\ \forall i\in\mathcal{I}_{\xi\eta}\big\}.$$

In order to apply Theorem 3.8, we introduce slack variables ξ and η, in the following equivalent reformulation of the MPEC,

$$(P_0)\qquad \min \widetilde{f}(z,\xi,\eta)=f(z)\quad\text{subject to}\quad \begin{array}{l}\widetilde{g}(z,\xi,\eta)=g(z)\le 0,\\ \widetilde{h}(z,\xi,\eta)=h(z)=0,\\ \Gamma(z,\xi,\eta):=\begin{pmatrix}G(z)-\xi\\ H(z)-\eta\end{pmatrix}=0,\\ (z,\xi,\eta)\in\mathbb{R}^n\times\mathcal{C},\end{array}$$

where

$$\mathcal{C}=\left\{(\xi,\eta)\in\mathbb{R}^l\times\mathbb{R}^l : \xi\ge 0,\eta\ge 0,\xi^T\eta=0\right\}.$$

Also define

$$C_i := \left\{(z,\xi,\eta)\in\mathbb{R}^n\times\mathbb{R}^l\times\mathbb{R}^l : \widetilde{h}_i(z,\xi,\eta)=0\right\},\ i=1,2,\dots,p,$$
$$C_i := \left\{(z,\xi,\eta)\in\mathbb{R}^n\times\mathbb{R}^l\times\mathbb{R}^l : G_{i-p}(z)=\xi_{i-p}\right\},\ i=p+1,\dots,p+l,$$
$$C_i := \left\{(z,\xi,\eta)\in\mathbb{R}^n\times\mathbb{R}^l\times\mathbb{R}^l : H_{i-p-l}(z)=\eta_{i-p-l}\right\},\ i=p+l+1,\dots,p+2l,$$

and

$$C_{p+2l+1} := \mathbb{R}^n\times\mathcal{C}.$$

Theorem 5.3. *Let* z^* *be a local minimizer of MPEC* (1)*, where all functions are locally Lipschitz around* z^**. Suppose that* $(z^*,G(z^*),H(z^*))$ *is a* (DIP) *point for* $C_1,C_2,\dots,C_{p+2l+1}$*. Then* z^* *is M-stationary.*

Proof. Set $\xi^*:=G(z^*)$ and $\eta^*:=H(z^*)$. Then applying Theorem 3.8 to (P_0), there exist $\lambda^f\ge 0,\lambda^g\ge 0,\lambda^h$ and $\lambda^\Gamma=(-\lambda^G,-\lambda^H)$ not all zero, such that $g(z^*)^T\lambda^g=0$ and

$$\begin{aligned}0 &\in \lambda^f\partial_l\widetilde{f}(z^*,\xi^*,\eta^*)+\sum_{i=1}^m\lambda_i^g\partial_l\widetilde{g}_i(z^*,\xi^*,\eta^*)+\sum_{i=1}^p\partial_l\left(\lambda_i^h\widetilde{h}_i\right)(z^*,\xi^*,\eta^*)\\ &\quad+\sum_{i=1}^{2l}\partial_l\left(\lambda_i^\Gamma\Gamma_i\right)(z^*,\xi^*,\eta^*)+N_l((z^*,\xi^*,\eta^*),\mathbb{R}^n\times\mathcal{C})\\ &\subseteq\begin{pmatrix}\lambda^f\partial_\diamond f(z^*)\\0\\0\end{pmatrix}+\begin{pmatrix}\sum_{i=1}^m\lambda_i^g\partial_\diamond g_i(z^*)\\0\\0\end{pmatrix}+\begin{pmatrix}\sum_{i=1}^p\lambda_i^h\partial_\diamond h_i(z^*)\\0\\0\end{pmatrix}\\ &\quad-\begin{pmatrix}\sum_{i=1}^1\left[\lambda_i^G\partial_\diamond G_i(z^*)+\lambda_i^H\partial_\diamond H_i(z^*)\right]\\-\lambda^G\\-\lambda^H\end{pmatrix}+\begin{pmatrix}0\\N_M((\xi^*,\eta^*),\mathcal{C})\end{pmatrix}.\end{aligned}\qquad (*)$$

We now take a closer look at those components pertaining to ξ and η in (*),

$$\left(-\lambda^G, -\lambda^H\right) \in N_M((\xi^*, \eta^*), \mathcal{C}).$$

From Theorem 5.2, we obtain the following rules for the components of λ^G and λ^H:

$$\left(\lambda_i^G, \lambda_i^H\right) \in \begin{cases} \{(a,b) : \text{a free}, b = 0\}, & \text{if } \xi_i^* = 0,\ \eta_i^* > 0, \\ \{(a,b) : (a > 0 \wedge b > 0) \vee ab = 0\} & \text{if } \xi_i^* = 0,\ \eta_i^* = 0, \\ \{(a,b) : a = 0, b \text{ free}\}, & \text{if } \xi_i^* > 0,\ \eta_i^* = 0. \end{cases}$$

Taking into account that $\xi_i^* = G_i(z^*)$, that $\eta_i^* = H_i(z^*)$ and the definitions of α, β and γ, yield the statement of the theorem. □

Now by considering some properties of the sets $C_1, C_2, \ldots, C_{p+2l+1}$, we try to shorten the hypotheses of the above theorem, to facilitate its application in the special cases. First, we take a closer look at these sets. Set $B = \cap_{i=1}^p C_i$ and

$$C_i' = \{z \in \mathbb{R}^n : h_i(z) = 0\}, \quad \text{for } i = 1, 2, \ldots, p, \qquad B' = \cap_{i=1}^p C_i'.$$

Then $C_i = C_i' \times \mathbb{R}^l \times \mathbb{R}^l$ for $i = 1, 2, \ldots, p$, and $B = B' \times \mathbb{R}^l \times \mathbb{R}^l$. If $(u, v, w) \in N_M(B; (z^*, \xi^*, \eta^*))$, then $v = w = 0$. Since $C_{p+2l+1} = \mathbb{R}^n \times \mathcal{C}$, if $(u, v, w) \in N_M(C_{p+2l+1}; (z^*, \xi^*, \eta^*))$ then $u = 0$ and we obtain,

$$N_M(B; (z^*, \xi^*, \eta^*)) \cap (-N_M(C_{p+2l+1}; (z^*, \xi^*, \eta^*))) = \{(0, 0, 0)\}. \tag{16}$$

Now, define

$$\Gamma_i(z, \xi, \eta) = G_i(z) - \xi_i, \quad \text{for } i = 1, 2, \ldots, l$$
$$\Gamma_i(z, \xi, \eta) = H_{i-l}(z) - \eta_{i-l}, \quad \text{for } i = l+1, \ldots, 2l.$$

Thus we have,

$$\partial_M \Gamma_i(z^*, \xi^*, \eta^*) = \begin{pmatrix} \partial_M G_i(z^*) \\ -e_i \\ 0 \end{pmatrix}, \ i = 1, \ldots, l,$$

and

$$\partial_M \Gamma_i(z^*, \xi^*, \eta^*) = \begin{pmatrix} \partial_M H_{i-l}(z^*) \\ 0 \\ -e_{i-l} \end{pmatrix}, \ i = l+1, \ldots, 2l,$$

where $e_i \in \mathbb{R}^l$ is the ith unite vector. Suppose $\sum_{i=1}^{2l} \alpha_i \gamma_i = 0$ with $\alpha_i \geq 0$ and $\gamma_i \in \partial_M \Gamma_i(z^*, \xi^*, \eta^*)$, for $i \in \{1, \ldots, 2l\}$. Then

$$\sum_{i=1}^{l} -\alpha_i e_i = \sum_{i=l+1}^{2l} -\alpha_i e_{i-l} = 0.$$

It follows that $\alpha_i = 0$ for any i. Theorem 3.10 implies that (z^*, ξ^*, η^*) is a (DIP) point for $C_{p+1}, \ldots, C_{p+2l}$.

Theorem 5.4. *Let z^* be a local minimizer of MPEC* (1). *Suppose that,* (*a*) z^* *is a* (DIP) *point for* $C'_1, \dots, C'_p$,

(*b*) (z^*, ξ^*, η^*) *is a* (DIP) *point for* A_1, A_2, *where*

$$A_1 = B \cap C_{p+2l+1}, \qquad A_2 = \cap_{i=p+1}^{2l} C_i.$$

Then z^* *is M-stationary.*

Proof. By Theorem 5.3, it is sufficient to show that (z^*, ξ^*, η^*) is a (DIP) point for $C_1, \dots, C_{P+2l+1}$. It is clear that for $z' \in \mathbb{R}^n$ and $(\xi', \eta') \in \mathbb{R}^l \times \mathbb{R}^l$,

$$d_{C'_i}(z') = d_{C_i}(z', \xi', \eta'), \qquad d_{B'}(z') = d_B(z', \xi', \eta'),\ i = 1, 2, \dots, p.$$

Since z^* is a (DIP) point for $C'_1, \dots, C'_p$, therefore there exist $\sigma_1, \varepsilon_1 > 0$ with $\sigma_1 \le 1$ such that for each $(z', \xi', \eta') \in \mathbb{B}((z^*, \xi^*, \eta^*); \varepsilon_1)$:

$$d_B(z', \xi', \eta') \le \frac{1}{\sigma_1} \left(\max_{1 \le i \le p} \right) \{ d_{C_i}(z', \xi', \eta') \}. \tag{17}$$

From (16) and Theorem 3.9, it follows that for some positive numbers σ_2, ε_2 with $\sigma_2 \le 1$ and for each $(z', \xi', \eta') \in \mathbb{B}((z^*, \xi^*, \eta^*); \varepsilon_2)$,

$$d_{B \cap C_{p+2l+1}}(z', \xi', \eta') \le \frac{1}{\sigma_2} \max\{ d_B(z', \xi', \eta'), d_{C_{p+2l+1}}(z', \xi', \eta') \}. \tag{18}$$

By the argument before the theorem, (z^*, ξ^*, η^*) is a (DIP) point for $C_{p+1}, \dots, C_{P+2l}$, hence for some positive numbers σ_3, ε_3 with $\sigma_3 \le 1$ and for each $(z', \xi', \eta') \in \mathbb{B}((z^*, \xi^*, \eta^*); \varepsilon_3)$,

$$d_{A_2}(z', \xi', \eta') \le \frac{1}{\sigma_3} \left(\max_{p+1 \le i \le p+2l} \right) \{ d_{C_i}(z', \xi', \eta') \}. \tag{19}$$

Finally, by the assumption (*b*) there exist some positive numbers σ_4, ε_4 with $\sigma_4 \le 1$ such that for all $(z', \xi', \eta') \in \mathbb{B}((z^*, \xi^*, \eta^*); \varepsilon_4)$,

$$d_{A_1 \cap A_2}(z', \xi', \eta') \le \frac{1}{\sigma_4} \max\{ d_{A_1}(z', \xi', \eta'), d_{A_2}(z', \xi', \eta') \}. \tag{20}$$

Let $\varepsilon = (\min_{1 \le j \le 4})\{\varepsilon_j\}$ and $\sigma = \sigma_4 \sigma_3 \sigma_2 \sigma_1$. If $(z', \xi', \eta') \in B((z^*, \xi^*, \eta^*); \varepsilon)$,

$$\begin{aligned} d_{\cap_{i=1}^{p+2l+1} C_i}(z', \xi', \eta') &= d_{A_1 \cap A_2}(z', \xi', \eta') \\ &\le \frac{1}{\sigma} \left(\max_{1 \le i \le p+2l+1} \right) \{ d_{C_i}(z', \xi', \eta') \}. \end{aligned}$$

According to Theorem 5.3, this assertion completes the proof. □

We introduce the multifunction $\Psi : \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l \rightrightarrows \mathbb{R}^n$ as follows,

$$\begin{aligned} \Psi(t, q, r, s) =& \{ z \in \mathbb{R}^n : g(z) + t \le 0, h(z) + q = 0, G(z) + r \ge 0, \\ & H(z) + s \ge 0, (G(z) + r)^T (H(z) + s) = 0 \}. \end{aligned}$$

Note that the feasible set of the MPEC is equal to $\Psi = \Psi(0, 0, 0, 0)$. Flegel [4] used this multifunction to define an MPEC variant of calmness, for the case where all functions except the objective function are smooth. Let us recall this definition.

Definition 5.5 ([4]). The MPEC (1) is said to satisfy MPEC-calmness in z^*, if for some $\varepsilon, \mu \geq 0$ and for all $(t, q, r, s) \in \mathbb{B}(0; \varepsilon) \subseteq \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ and all $z \in \Psi(t, q, r, s) \cap \mathbb{B}(z^*; \varepsilon)$ it holds that

$$f(z^*) \leq f(z) + \mu \|(t, q, r, s)\|. \tag{21}$$

Proposition 5.6. *Assume that z^* is a local minimizer of MPEC (1) and that MPEC (1) is MPEC-calm at z^*. Then there exists $\mu > 0$ such that the vector $(z^*, 0, 0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^l$ is a local minimizer of*

$$\begin{array}{lll} & & f(z) + \mu(\|(g_1^+(z), \ldots, g_m^+(z))\| + \|h(z)\| + \|(r, s)\|), \\ \min(z, r, s) & \textit{subject to} & G(z) + r \geq 0, \qquad H(z) + s \geq 0, \\ & & (G(z) + r)^T (H(z) + s) = 0, \end{array} \tag{22}$$

where $\alpha^+ := \max\{0, \alpha\}$ *for* $\alpha \in \mathbb{R}$.

Proof. By continuity of h and g there exists $0 < \delta \leq \frac{\varepsilon}{3}$ such that

$$\|h(z)\| \leq \frac{\varepsilon}{3}, \qquad \|(g_1^+(z), \ldots, g_m^+(z))\| \leq \frac{\varepsilon}{3}, \quad \forall z \in \mathbb{B}(z^*; \delta).$$

Let $(z, r, s) \in \mathbb{B}((z^*, 0, 0); \delta)$ be a feasible point for (22). Then set $t_i := -g_i^+(z)$ for $i = 1, \ldots, m$ and $q := -h(z)$. With this choice of t and q, we have $(t, q, r, s) \in \mathbb{B}(0; \varepsilon)$ and $z \in \Psi(t, q, r, s) \cap \mathbb{B}(z^*; \varepsilon)$. Since z^* satisfies the MPEC-calmness, the condition (21) holds and we obtain,

$$\begin{aligned} f(z^*) &\leq f(z) + \mu \|(t, q, r, s)\| \\ &\leq f(z) + \mu(\|(g_1^+(z), \ldots, g_m^+(z))\| + \|h(z)\| + \|(r, s)\|), \end{aligned}$$

for some $\mu > 0$. Thus $(z^*, 0, 0)$ is a local minimizer of (22). $\square$

Theorem 5.7. *Suppose that all the functions of MPEC (1) are locally Lipschitz and that z^* is a local minimizer of MPEC (1) at which the MPEC is MPEC-calm. Furthermore assume that*

$$N_M(C_{2l+1}; (z^*, 0, 0, \xi^*, \eta^*)) \cap (-N_M(\cap_{i=1}^{2l} C_i; (z^*, 0, 0, \xi^*, \eta^*))) = \{(0, 0, 0, 0, 0)\},$$

where

$$C_i := \{(z, r, s, \xi, \eta) \in \mathbb{R}^{n+2l+2l} : G_i(z) + r_i = \xi_i\}, \ i = 1, \ldots, l,$$

and

$$C_i := \{(z, r, s, \xi, \eta) \in \mathbb{R}^{n+2l+2l} : H_{i-l}(z) + s_{i-l} = \eta_{i-l}\}, \ i = l + 1, \ldots, 2l,$$

and

$$C_{2l+1} := \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^l \times \mathcal{C}.$$

Then z^* is M-stationary.

Proof. We will prove this theorem in two steps.

Step 1: Let us show that $(z^*, 0, 0, \xi^*, \eta^*)$ is a (DIP) point for $C_1, \ldots, C_{2l+1}$.

Define,

$$\widehat{h}_i(z,r,s,\xi,\eta) := G_i(z) + r_i - \xi_i, \quad i = 1,\dots,l,$$
$$\widehat{h}_i(z,r,s,\xi,\eta) := H_{i-l}(z) + s_{i-l} - \eta_{i-l}, \quad i = l+1,\dots,2l.$$

Then

$$\partial_M \widehat{h}_i(z^*,0,0,\xi^*,\eta^*) = \begin{pmatrix} \partial_M G_i(z^*) \\ e_i \\ 0 \\ -e_i \\ 0 \end{pmatrix}$$

for $i = 1,..,l$, and

$$\partial_M \widehat{h}_i(z^*,0,0,\xi^*,\eta^*) = \begin{pmatrix} \partial_M H_{i-l}(z^*) \\ 0 \\ e_{i-l} \\ 0 \\ -e_{i-l} \end{pmatrix}$$

for $i = l+1,..,2l$.

From Theorem 3.10, $(z^*,0,0,\xi^*,\eta^*)$ is a (DIP) point for $C_1,\dots,C_{2l}$. Thus, for some $\varepsilon',\sigma' > 0$ with $\sigma' \le 1$ and for each $y \in (z^*,0,0,\xi^*,\eta^*) + \varepsilon'\mathbb{B}_{\mathbb{R}^{2l+1}}$,

$$d_{C'}(y) \le \frac{1}{\sigma'}\left(\max_{1\le j\le 2l}\right) d_{C_j}(y),$$

where $C' = \cap_{i=1}^{2l} C_i$. Also, by Proposition 3.9 and by our second assumption, for some $\varepsilon,\sigma > 0$ with $\sigma \le 1$,

$$d_C(y) \le \frac{1}{\sigma}\max\{d_{C'}(y), d_{C_{2l+1}}(y)\},$$

for all $y \in (z^*,0,0,\xi^*,\eta^*) + \varepsilon\mathbb{B}_{\mathbb{R}^l\times\mathbb{R}^l\times\mathbb{R}}$, where $C = \cap_{i=1}^{2l+1} C_i$. If $\sigma_0 = \sigma\sigma'$ and $\varepsilon_0 = \min\{\varepsilon,\varepsilon'\}$, then

$$d_C(y) \le \frac{1}{\sigma_0}\left(\max_{1\le i\le 2l+1}\right)\{d_{C_i}(y)\},$$

for all $y \in (z^*,0,0,\xi^*,\eta^*) + \varepsilon_0\mathbb{B}_{\mathbb{R}^l\times\mathbb{R}^l\times\mathbb{R}}$. Thus $(z^*,0,0,\xi^*,\eta^*)$ is a (DIP) point for $C_1,\dots,C_{2l+1}$.

Step 2: Let us show that z^* is M-stationary.

Since the MPEC satisfies MPEC-calmness at z^*, therefore by Theorem 5.6, $(z^*,0,0)$ is a local minimizer of (22) for some $\rho > 0$. Consider (22) as a MPEC with equilibrium constraints $\widehat{G}(z,r,s) = G(z) + r$ and $\widehat{H}(z,r,s) = H(z) + s$. Then we deduce, by an argument similar to that of Theorem 5.3, that there exists $(\lambda^{\hat{f}},\lambda^G,\lambda^H) \ne (0,0,0)$ such that $\lambda^{\hat{f}} \ge 0, \lambda^G$ and λ^H satisfies the sign conditions

$$\lambda^G_\alpha \text{ free}, \qquad \lambda^G_\gamma = 0,$$
$$\left(\lambda^G_i > 0 \wedge \lambda^H_i > o\right) \vee \lambda^G_i\lambda^H_i = 0, \quad \forall i \in \beta,$$
$$\lambda^H_\gamma \text{ free}, \qquad \lambda^H_\alpha = 0,$$

and

$$0 \in \lambda^{\hat{f}} \partial_l \widehat{f}(z^*, 0, 0) - \sum_{i=1}^{l} \left[\lambda_i^G \begin{pmatrix} \partial_\diamond G_i(z^*) \\ e_i \\ 0 \end{pmatrix} + \lambda_i^H \begin{pmatrix} \partial_\diamond H_i(z^*) \\ 0 \\ e_i \end{pmatrix} \right],$$

where $\widehat{f}(z, r, s) := f(z) + \rho(\|(g_1^+(z), \ldots, g_m^+(z))\|_1 + \|h(z)\|_1 + \|(r, s)\|_1)$, is a locally Lipschitz function. If we assume $\lambda^{\hat{f}} = 0$, it follows immediately from the last relation that $\lambda^G = \lambda^H = 0$, which is a contradiction. Therefore, without loss of generality, we may assume $\lambda^{\hat{f}} = 1$. By Theorem 2.7, we get

$$\partial_l \widehat{f}(., 0, 0)(z^*) \subseteq \partial_l f(z^*) + \rho \partial_l(\|(g_1^+(.), \ldots, g_m^+(.))\|_1)(z^*) + \rho \partial_l(\|h(.)\|_1)(z^*).$$

From Theorems 2.10 and 2.11, we obtain

$$\partial_l \widehat{f}(., 0, 0)(z^*) \subseteq \partial_l f(z^*) + \sum_{i=1}^{m} \rho \eta_i^g \partial_l g_i(z^*) + \sum_{i=1}^{p} \rho \eta_i^h \partial_l h_i(z^*),$$

for some $\eta_i^g \in \partial_l(g_i^+(z^*))$ and $\eta_i^h \in \partial_l(|.|)(h_i(z^*))$. Observe that $g(z^*) \leq 0$ and $h(z^*) = 0$, so by computing the last subdifferentials and by taking $\lambda^g = \rho\eta^g$ and $\lambda^h = \rho\eta^h$, we have, $g(z^*)^T \lambda^g = 0$ and,

$$\begin{aligned} 0 &\in \partial_l f(z^*) + \sum_{i=1}^{m} \lambda_i^g \partial_l g_i(z^*) + \sum_{i=1}^{p} \lambda_i^h \partial_l h_i(z^*) - \sum_{i=1}^{l} \lambda_i^G \partial_\diamond G_i(z^*) + \lambda_i^H \partial_\diamond H_i(z^*) \\ &\subseteq \partial_\diamond f(z^*) + \sum_{i=1}^{m} \lambda_i^g \partial_\diamond g_i(z^*) + \sum_{i=1}^{p} \lambda_i^h \partial_\diamond h_i(z^*) - \sum_{i=1}^{l} \lambda_i^G \partial_\diamond G_i(z^*) + \lambda_i^H \partial_\diamond H_i(z^*). \end{aligned}$$

This concludes the proof . □

Theorem 5.8. *Suppose that all the functions of MPEC* (1) *are locally Lipschitz and that z^* is a local minimizer of MPEC* (1) *at which the MPEC is MPEC-calm. Then z^* is M-stationary.*

Proof. By Theorem 5.7, it is sufficient to show that the condition

$$N_M(C_{2l+1}; (z^*, 0, 0, \xi^*, \eta^*)) \cap (-N_M(C'; (z^*, 0, 0, \xi^*, \eta^*))) = \{(0, 0, 0, 0, 0)\}, \tag{23}$$

holds true, when $C' = \cap_{i=1}^{2l} C_i$.

If $(u, v, w, x, y) \in N_M(C_{2l+1}; (z^*, 0, 0, \xi^*, \eta^*))$, then by [19, Proposition 6.41], $u = v = w = 0$ and we have $(x, y) \in N_M(\mathcal{C}; (\xi^*, \eta^*))$.

Now suppose that $(0, 0, 0, x, y) \in N_M(C'; (z^*, 0, 0, \xi^*, \eta^*)$. Namely, there exist sequences $\{(z_n, r_n, s_n, \xi_n, \eta_n)\}$ and $\{(u_n, v_n, w_n, x_n, y_n)\}$ such that,

$$\begin{gathered} (z_n, r_n, s_n, \xi_n, \eta_n) \in C', \\ (z_n, r_n, s_n, \xi_n, \eta_n) \longrightarrow (z^*, 0, 0, \xi^*, \eta^*), \\ (u_n, v_n, w_n, x_n, y_n) \longrightarrow (0, 0, 0, x, y) \end{gathered}$$

and

$$(u_n, v_n, w_n, x_n, y_n) \in N_P(C'; (z_n, r_n, s_n, \xi_n, \eta_n)).$$

For each n, there exists $\sigma_n > 0$ such that for all $(z, r, s, \xi, \eta) \in C'$

$$\begin{aligned} &< (u_n, v_n, w_n, x_n, y_n),\ (z - z_n, r - r_n, s - s_n, \xi - \xi_n, \eta - \eta_n) > \\ &\leq \sigma_n \|(z, r, s, \xi, \eta) - (z_n, r_n, s_n, \xi_n, \eta_n)\|^2. \end{aligned} \tag{24}$$

Let $\varepsilon > 0$ be arbitrary and for each integer $i \in \{1, 2, \ldots, l\}$ set

$$0_i^{\varepsilon} = \underbrace{(0, \ldots, \overbrace{\varepsilon}^{ith}, \ldots 0)}_{l \ times}$$

$$0_i^{-\varepsilon} = \underbrace{(0, \ldots, \overbrace{-\varepsilon}^{ith}, \ldots 0)}_{l \ times}.$$

Then we have

$$(z_n, r_n + 0_i^{\varepsilon}, s_n, \xi_n + 0_i^{\varepsilon}, \eta_n),\ (z_n, r_n + 0_i^{-\varepsilon}, s_n, \xi_n + 0_i^{-\varepsilon}, \eta_n) \in C'$$

Now by (24), it follows that $v_i^{(n)} + x_i^{(n)} \leq 2\sigma_n\varepsilon$ and $-(v_i^{(n)} + x_i^{(n)}) \leq 2\sigma_n\varepsilon$, where $v_i^{(n)}$ and $x_i^{(n)}$ are the ith component of v and x. Since ε is arbitrary, then $v_i^{(n)} + x_i^{(n)} = 0$, $\forall n$. On the other hand $\lim_{n \longrightarrow \infty} v_i^{(n)} = 0$, thus $x_i = \lim_{n \longrightarrow \infty} x_i^{(n)} = 0$. By similar argument we can prove that $y = 0$ and this completes the proof . □

The next example illustrates a nonsmooth MPEC which satisfies assumptions of Theorem 5.4 at a minimizer point and is not MPEC-calm at that point.

Example 5.9. Consider the following problem:

$$\min f(x,y) = x^{\frac{1}{3}} \quad \text{s.t.} \quad \begin{array}{l} h(x,y) = \sin(\|(x,y)\|) = 0, \\ G(x,y) = x \geq 0, \quad H(x,y) = y \geq 0, \quad G(x,y)H(x,y) = 0, \\ (x,y) \in \mathbb{R}^2. \end{array}$$

The point $(0,0)$ is a local minimum for the above problem. For simplicity, let us apply $\|.\|_1$ in our argument. Since there is only one equality constraint in this system, the condition 5.4(a) is trivial. We have,

$$A_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : \sin(\|(x_1, x_2)\|) = 0, x_3 \geq 0, x_4 \geq 0, x_3 x_4 = 0\},$$

and

$$\begin{aligned} A_2 &= \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_3 = G(x_1, x_2),\ x_4 = H(x_1, x_2)\} \\ &= \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_3 = x_1, x_4 = x_2\}. \end{aligned}$$

We claim that $(0,0,0,0)$ is a (DIP) point for A_1 and A_2. If $\varepsilon < \pi$, then $A_1 \cap A_2 \cap \varepsilon\mathbb{B}_{\mathbb{R}^4} = \{(0,0,0,0)\}$. Suppose that $(y_1, y_2, y_3, y_4) \in \varepsilon\mathbb{B}_{\mathbb{R}^4}$. We have, $d_{A_1 \cap A_2}((y_1, y_2, y_3, y_4)) = |y_1| + |y_2| + |y_3| + |y_4|$. It is easy to verify that,

$$d_{A_1}(y_1, y_2, y_3, y_4) = \begin{cases} |y_1| + |y_2| + |y_3| + |y_4| & \text{if } y_3 < 0, y_4 < 0 \\ |y_1| + |y_2| + |y_3| & \text{if } y_3 < 0 \leq y_4 \text{ or } 0 \leq y_3 \leq y_4 \\ |y_1| + |y_2| + |y_4| & \text{if } y_4 < 0 \leq y_3 \text{ or } 0 \leq y_4 \leq y_3 \end{cases}$$

and $d_{A_2}(y_1, y_2, y_3, y_4) = |y_1 - y_3| + |y_2 - y_4|$. We claim that

$$d_{A_1 \cap A_2}(y_1, y_2, y_3, y_4) \leq 3 \max\{d_{A_1}(y_1, y_2, y_3, y_4), d_{A_2}(y_1, y_2, y_3, y_4)\}. \tag{25}$$

If $y_3 < 0, y_4 < 0$, then (25) holds. If

$$d_{A_1}(y_1, y_2, y_3, y_4) = |y_1| + |y_2| + |y_3| \leq d_{A_2}(y_1, y_2, y_3, y_4),$$

then

$$\begin{aligned} d_{A_1 \cap A_2}(y_1, y_2, y_3, y_4) &\leq 3(|y_1 - y_3| + |y_2 - y_4|) \\ &= 3 \max\{d_{A_1}(y_1, y_2, y_3, y_4), d_{A_2}(y_1, y_2, y_3, y_4)\}. \end{aligned}$$

If $d_{A_2}(y_1, y_2, y_3, y_4) \leq d_{A_1}(y_1, y_2, y_3, y_4) = |y_1| + |y_2| + |y_3|$, then

$$\begin{aligned} d_{A_1 \cap A_2}(y_1, y_2, y_3, y_4) &\leq |y_1| + |y_2| + |y_3| + |y_2| + |y_2 - y_4| \\ &\leq 3(|y_1| + |y_2| + |y_3|) \\ &= 3 \max\{d_{A_1}(y_1, y_2, y_3, y_4), d_{A_2}(y_1, y_2, y_3, y_4)\}. \end{aligned}$$

The argument is similar in other cases. Now, we prove that the above system is not MPEC-calm at $(0,0)$. For $n > 2$ take $(x_n, y_n) = (\frac{-1}{n^6}, 0)$ and $(q_n, r_n, s_n) = (-\sin(\frac{1}{n^6}), \frac{1}{n^6}, 0)$. Then $(x_n, y_n) \in \Psi(q_n, r_n, s_n)$ and $(x_n, y_n) \longrightarrow (0,0)$ and $(q_n, r_n, s_n) \longrightarrow (0,0,0)$. On the other hand,

$$\begin{aligned} f(x_n, y_n) + n\|(q_n, r_n, s_n)\| &= \frac{-1}{n^2} + n\left\{\sin(\frac{1}{n^6}) + \frac{1}{n^6}\right\} \\ &< 0 = f(0,0). \end{aligned}$$

Thus this system is not MPEC-calm at $(0,0)$.

Acknowledgements. The authors would like to thanks Professors B. S. Mordukhovich and L. Thibault and anonymous referee for their helpful comments on improving the first version of this paper.

References

[1] M. Bounkhel, L. Thibault: On various notions of regularity of sets in nonsmooth analysis, Nonlinear Anal., Theory Methods Appl. 48 (2002) 223–246.

[2] F. H. Clarke: Optimization and Nonsmooth Analysis, Wiley, New York (1983).

[3] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski: Nonsmooth Analysis and Control Theory, Graduate Texts in Mathematics 178, Springer, New York (1998).

[4] M. L. Flegel, C. Kanzow: On M-stationary points for mathematical programs with equilibrium constraints, J. Math. Anal. Appl. 310 (2005) 286–302.

[5] R. Henrion, J. Outrata: A subdifferential condition for calmness of multifunctions, J. Math. Anal. Appl. 258 (2001) 110–130.

[6] R. Henrion, J. Outrata: Calmness of constraint systems with applications, Math. Program., Ser. 104(B) (2005) 437–464.

[7] A. D. Ioffe: Approximate subdifferentials of nonconvex functions, Cahier 8120, CEREMADE, Université de Paris 'Dauphine', Paris (1981).

[8] A. D. Ioffe: Metric regularity and subdifferential calculus, Russ. Math. Surv. 55 (2000) 501–558.

[9] A. Jourani, L. Thibault: Metric regularity for strongly compactly Lipschitzian mappings, Nonlinear Anal., Theory Methods Appl. 24 (1995) 229–240.

[10] A. Jourani, L. Thibault: Extensions of subdifferential calculus rules in Banach spaces, Canad. J. Math. 48 (1996) 834–848.

[11] A. Ja. Kruger, B. S. Mordukhovich: Minimization of nonsmooth functionals in optimal control problems, Izv. Akad. Nauk SSSR, Tekh. Kibern. 1978 (1978) 176–183.

[12] Z. Q. Luo, J. S. Pang, D. Ralph: Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge (1997).

[13] P. Michel, J. P. Penot: A generalized derivative for calm and stable functions, Differ. Integral Equ. 5 (1992) 189–196.

[14] B. S. Mordukhovich: Maximum principle in problem of time optimal response with nonsmooth constraints, J. Appl. Math. Mech. 40 (1976) 960–969.

[15] B. S. Mordukhovich: Approximation Methods in Problems of Optimization and Control, Nauka, Moscow (1988) (in Russian).

[16] B. S. Mordukhovich: Variational Analysis and Generalized Differentiation. I. Basic Theory. Grundlehren der Mathematischen Wissenschaften 330, Springer, Berlin (2006).

[17] J. V. Outrata: Optimality conditions for a class of mathematical programs with equilibrium constraints, Math. Oper. Res. 24 (1999) 627–644.

[18] J. V. Outrata, M. Kocvara, J. Zowe: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Theory, Applications and Numerical Results, Kluwer, Dordrecht (1998).

[19] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Grundlehren der Mathematischen Wissenschaften 317, Springer, Berlin (1998).

[20] W. Song: Calmness and error bounds for convex constraint systems, SIAM J. Optim. 17 (2006) 353–371.

[21] J. S. Treiman: The linear nonconvex generalized gradient and Lagrange multipliers, SIAM J. Optim. 5 (1995) 670–680.

[22] J. S. Treiman: The linear nonconvex generalized gradient, World Congress of Nonlinear Analysis (Tampa, 1992), V. Lakshmikantham (ed.), de Gruyter, Berlin (1996) 2325–2336.

[23] J. S. Treiman: Lagrange multipliers for nonconvex generalized gradients with equality, inequality and set costraints, SIAM J. Control Optimization 37 (1999) 1313–1329.

[24] J. J. Ye: Constraint qualifications and necessary optimality conditions with variational inequality constraints, SIAM J. Optim. 10 (2000) 943–962.

[25] J. J. Ye: Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints, J. Math. Anal. Appl. 307 (2005) 350–369.