

## IMPROVED ENHANCED FRITZ JOHN CONDITION AND CONSTRAINTS QUALIFICATIONS USING CONVEXIFICATORS

ABEKA KHARE AND TRILOKI NATH\*

**Abstract.** In this paper, we use the concept of convexifiers to derive enhanced Fritz John optimality condition for nonsmooth optimization problems having equality, inequality and abstract set constraint, where involved functions admit convexifiers. This necessary optimality condition provides some more information about the extremal point in terms of converging sequences towards it. Then we employ this optimality condition to study enhanced Karush–Kuhn–Tucker (KKT) condition and to define associated  $\partial^*$ -pseudonormality and  $\partial^*$ -quasinormality concepts in terms of convexifiers. Later, sufficiency for  $\partial^*$ -pseudonormality and some more results based on these concepts are investigated.

**Mathematics Subject Classification.** 90C30, 90C46, 90C99.

Received February 11, 2019. Accepted September 4, 2019.

### 1. INTRODUCTION

In this paper, we consider an optimization problem of the form:

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

where the feasible region  $\Omega$  is

$$\Omega = \{x \in \mathcal{C} \mid h_1(x) = 0, \dots, h_p(x) = 0, g_1(x) \leq 0, \dots, g_q(x) \leq 0\}.$$

Unless otherwise stated, throughout the paper we will proceed with the assumption that  $f$ ,  $h_i$  ( $i = 1, \dots, p$ ),  $g_j$  ( $j = 1, \dots, q$ ) :  $\mathbb{R}^m \rightarrow \mathbb{R}$  are locally Lipschitz continuous functions and admit upper regular and upper semicontinuous convexifiers at the point of interest. We also choose  $\mathcal{C}$  as a nonempty closed convex subset of  $\mathbb{R}^m$ .

For a smooth optimization problem with inequality constraints, the classical Fritz John necessary optimality condition is well known, which has been proposed by Fritz John [14] in 1948. If  $\mathcal{C} = \mathbb{R}^m$ , then this classical Fritz John condition characterizes any local optimal solution  $x^*$  in terms of the existence of scalars  $\lambda_1^*, \dots, \lambda_p^*$  and

---

*Keywords.* Directional derivatives, convexifiers, optimality conditions, constraint qualifications.

Department of Mathematics and Statistics, Dr. Harisingh Gour Vishwavidyalaya, 470003 Sagar, Madhya Pradesh, India.

\*Corresponding author: [tnverma07@gmail.com](mailto:tnverma07@gmail.com)

$\mu_0^* \geq 0, \dots, \mu_q^* \geq 0$  not all zero such that

$$0 = \mu_0^* \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^q \mu_j^* \nabla g_j(x^*), \quad (1.1)$$

$$0 = \mu_j^* g_j(x^*); \quad \forall j = 1, \dots, q. \quad (1.2)$$

Here,  $\nabla$  represents the gradient of the corresponding functions and the equation (1.2) is called complementary slackness condition or CS condition. After Fritz John, many researchers have extended this condition for various problems such as Mangasarian and Fromovitz [17] for the smooth case having both equality and inequality constraints. Bector *et al.* [1] extended this for the nonsmooth problem without an abstract set constraint.

Enhanced Fritz John optimality condition was first given by Hestenes [11] and improved by Bertsekas [2] for the smooth case. Hestenes' enhanced Fritz John condition is a refinement of classical Fritz John condition, in which complementarity slackness (CS) conditions have been replaced by complementarity violation (CV) conditions as given in Proposition 2.1(iv) of [3]. The enhanced Fritz John condition states about a local optimal solution  $x^*$  for  $\mathcal{C} = \mathbb{R}^m$ , that there exist scalars  $\lambda_1^*, \dots, \lambda_p^*$  and  $\mu_0^* \geq 0, \dots, \mu_q^* \geq 0$  not all zero satisfying equation (1.1) and additionally, a sequential property, which says that: If the index set  $I \cup J$  is nonempty, where  $I = \{i | \lambda_i^* \neq 0\}$  and  $J = \{j \neq 0 | \mu_j^* > 0\}$ , then there is a sequence  $\{x^k\} \subset \mathcal{C}$  converging to  $x^*$  such that for all  $k$ ,

$$f(x^k) < f(x^*), \quad \lambda_i^* h_i(x^k) > 0, \quad \forall i \in I, \quad \mu_j^* g_j(x^k) > 0, \quad \forall j \in J. \quad (1.3)$$

Evidently, the CV conditions (1.3) proposed by Bertsekas are stronger than the classical CS conditions (1.2) because if  $\mu_j^* > 0$ , then according to condition (1.3), the corresponding  $j$ th inequality constraint must be violated arbitrarily close to  $x^*$ , which shows that  $g_j(x^*) = 0$ , and that is the reason to call them complementarity violation conditions or CV. Enhanced Fritz John condition further extended to the convex and nonconvex abstract set constraints in [2–4].

For nonsmooth problems with no abstract set constraints, enhanced Fritz John condition was first studied by Bector *et al.* [1] by using Clarke subdifferentials instead of the classical gradients. Later on, Ye and Zhang [27] sharpened this work with Lipschitz assumptions for nonsmooth problem having an abstract set constraint. The authors used limiting subdifferentials and limiting normal cone and provided some extra information about the proximal subdifferentiability ([27], Thm. 1(iv)). The significance of this information lies in the fact that a continuously differentiable function need not be proximal subdifferentiable, *e.g.*  $f(x) = -|x|^{2/3}$  is continuously differentiable at  $x = 0$  but admits no proximal subdifferential at  $x = 0$  (sufficient condition for a function to be proximal subdifferentiable is  $C^{1+}$ , *i.e.* the gradient of the function is locally Lipschitz). Hence, the condition of Theorem 1(iv) from [27] is much stronger than that of Proposition 2.1 of [3].

Here, we are dealing with the same optimization problem but the abstract set constraint is convex as well. We are generalizing the idea of Ye and Zhang [27] with a different substitute named convexificator, pioneered by Demyanov [7] and developed by Jeyakumar and Luc [13]. The notion of convexificators emerged in the development of various generalized subdifferentials and it provides both an upper convex and a lower concave approximation to the concerned functions, so we can use it as a tool to handle those problems which do not fulfill convexity requirements. Generally, convexificators are closed sets, but not necessarily bounded or convex. With the assumption of locally bounded convexificator at a point, a continuous function becomes locally Lipschitz at that point. For the calculus and more results on convexificators, which we have used in this paper, we refer to [9, 13]. A continuous function may admit several convexificators, so there is a need to know which will be the minimal, whether it will be unique or not. To answer these questions Jeyakumar and Luc [13] proposed some specific conditions in terms of the set of extreme points. The authors also showed that for a locally Lipschitz function, the Clarke subdifferential [6], Michel-Penot [19], Ioffe [12], Mordukhovich [22] and Treiman [25] are also convexificators and may contain the convex hull of some other convexificators. The authors [13] also introduced an optimality condition for a nonsmooth problem in terms of convexificators, which we use as an initial tool to establish our results.

In this paper, we obtain results in terms of the convex hull of convexificators of involved functions, which may be infinite whereas limiting subdifferential may be finite for the same function. Our result would be better for those cases also, in the sense that optimality condition depends on the convex hull of few points of limiting subdifferential, *i.e.* we do not need all the points of limiting subdifferential.

In addition to this, we provide a piece of new information about the differentiability at every point of a sequence converging towards the extremal point is given, where CV conditions hold. Further, we are generalizing the notion of pseudonormality and quasinormality given by Bertsekas and Ozdaglar [3] and sharpened by Ye and Zhang [27], as  $\partial^*$ -pseudonormality and  $\partial^*$ -quasinormality. We are using here convexificators instead of limiting subdifferentials, which may be smaller than the limiting subdifferentials (see *e.g.* Example 2.9) in Sect. 2), so our definitions are weaker than the classical notions and results are generalized. Later, by using these two definitions some characterizations for the feasible set is also improved.

We are organizing our paper as follows. Section 2 is devoted to the basic concepts of nonsmooth analysis, convexificators and the results that will be used in our paper. In Section 3, we will derive our main result *improved enhanced Fritz John condition* using convexificators. The notion of  $\partial^*$ -pseudonormality and  $\partial^*$ -quasinormality, the relationship between these two and some other results related to them are presented in Section 4.

## 2. PRELIMINARIES

In this section, some background material related to nonsmooth analysis and convexicator has recalled, which will be used later. Here, the definitions and results may seem to be concise but they are enough to fulfill our need. For details we refer [8, 9, 13, 16, 26, 27].

Some notations that will be used throughout the paper are as follows. Throughout the paper, we suppose  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  is the extended real line.  $\mathbb{B}[x^*, \epsilon]$  is the closed ball centered at  $x^*$  with radius  $\epsilon$  and  $\mathbb{B}$  is used to denote the closed unit ball. For any set  $S \subset \mathbb{R}^n$ ,  $\bar{S}$ ,  $\text{int}S$ ,  $\text{co}S$ ,  $\overline{\text{co}}S$  are the closure, interior, convex hull, closed convex hull of  $S$  respectively. The symbol  $\|\cdot\|$  denotes the Euclidean norm. For a function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g^+(x) = \max\{0, g(x)\}$ , here  $g^+$  is a vector if max function is defined componentwise. For a cone  $\mathcal{N}$ , we denote  $\mathcal{N}^0$  as its polar cone.  $\limsup_{x \rightarrow x_0} f(x)$ ,  $\liminf_{x \rightarrow x_0} f(x)$ ,  $\sup_{x \in \mathbb{R}^n} f(x)$   $\inf_{x \in \mathbb{R}^n} f(x)$  are the Kuratowski–Painlevé upper(outer), lower(inner), supremum, infimum value of  $f$  over  $\mathbb{R}^n$  respectively.

Here are some basic definitions of subdifferentials.

**Definition 2.1.** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be a lower semicontinuous (l.s.c.) function and  $x_0 \in \text{dom}f := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$ .

(i) The proximal subdifferential of  $f$  at  $x_0$  is the set

$$\partial^\pi f(x_0) := \{ \xi \in \mathbb{R}^n \mid \exists \sigma > 0, \eta > 0 \text{ s.t. } f(x) \geq f(x_0) + \langle \xi, x - x_0 \rangle - \sigma \|x - x_0\|^2, \forall x \in \mathbb{B}(x_0, \eta) \}.$$

If  $\partial^\pi f(x_0)$  is nonempty at  $x_0$ , then  $f$  is said to be proximal subdifferentiable at  $x_0$ .

(ii) The Fréchet (regular) subdifferential of  $f$  at  $x_0$  is the set

$$\partial^F f(x_0) := \left\{ \xi \in \mathbb{R}^n \mid \liminf_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - \langle \xi, h \rangle}{\|h\|} \geq 0 \right\}.$$

(iii) The limiting (Mordukhovich or basic) subdifferential of  $f$  at  $x_0$  is the set

$$\begin{aligned} \partial f(x_0) &:= \{ \xi \in \mathbb{R}^n \mid \exists x_k \rightarrow x_0 \text{ and } \xi_k \rightarrow \xi \text{ with } \xi_k \in \partial^F f(x_k) \}, \\ &:= \{ \xi \in \mathbb{R}^n \mid \exists x_k \rightarrow x_0 \text{ and } \xi_k \rightarrow \xi \text{ with } \xi_k \in \partial^\pi f(x_k) \}. \end{aligned}$$

(iv) The singular limiting (Mordukhovich) subdifferential of  $f$  at  $x_0$  is the set

$$\begin{aligned} \partial^\infty f(x_0) &:= \{ \xi \in \mathbb{R}^n \mid \exists x_k \rightarrow x_0 \text{ and } t_k \xi_k \rightarrow \xi \text{ with } \xi_k \in \partial^F f(x_k), t_k \downarrow 0 \}, \\ &:= \{ \xi \in \mathbb{R}^n \mid \exists x_k \rightarrow x_0 \text{ and } t_k \xi_k \rightarrow \xi \text{ with } \xi_k \in \partial^\pi f(x_k), t_k \downarrow 0 \}. \end{aligned}$$

We have the following results for subdifferentials in the literature, see [21, 23].

- Proposition 2.2.** (i) A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is Lipschitz near  $x_0$  and  $\partial f(x_0) = \{\zeta\}$  if and only if  $f$  is strictly differentiable at  $x_0$  and the gradient of  $f$  at  $x_0$  is equal to  $\zeta$ .  
(ii) If a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is Lipschitz near  $x_0$  with positive constant  $L_f$ , then

$$\partial f(x_0) \subset L_f \mathbb{B}.$$

- (iii) A lower semicontinuous function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is Lipschitz near  $x_0$  if and only if  $\partial^\infty f(x_0) = \{0\}$ .

**Definition 2.3** (Tangent Cone). Let  $S$  be a nonempty subset of  $\mathbb{R}^n$  and  $x \in \bar{S}$ . Then Contingent Cone  $\mathcal{T}_S(x)$  to  $S$  at  $x$  is defined as

$$\mathcal{T}_S(x) = \{v \in \mathbb{R}^n \mid \exists t_k \rightarrow 0 \text{ and } v_k \rightarrow v \text{ such that } x + t_k v_k \in S\},$$

and Polar of this cone is

$$\mathcal{T}_S^0(x) := \{v \in \mathbb{R}^n \mid \langle y, v \rangle \leq 0, \forall y \in \mathcal{T}_S(x)\}.$$

**Definition 2.4** (Normal Cones). ([8], Def. 4.2) Let  $S$  be a nonempty subset of  $\mathbb{R}^n$  and  $\bar{x} \in S$ . Then the Fréchet (regular) normal cone to  $S$  at  $\bar{x}$  is given as

$$\mathcal{N}_S^F(\bar{x}) := \left\{ v \in \mathbb{R}^n \mid \limsup_{x \rightarrow \bar{x}, x \in S} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\},$$

and the nonempty cone

$$\mathcal{N}_S(\bar{x}) := \limsup_{x \rightarrow \bar{x}} \mathcal{N}_S^F(\bar{x}) \tag{2.1}$$

is called the limiting (Mordukhovich or basic) normal cone to  $S$  at  $\bar{x}$ .

The Fréchet normal cone to a set can also be expressed as the polar cone of the tangent cone, *i.e.*

$$\mathcal{N}_S^F(\bar{x}) = \mathcal{T}_S^0(\bar{x}),$$

and if  $\mathcal{N}_S^F(\bar{x}) = \mathcal{N}_S(\bar{x})$  for all  $x \in S$ , then set  $S$  is called regular.

Now we are presenting here some basic terms borrowed from Section 2 of [9] which are required to study convexificators.

**Definition 2.5** (Dini Directional Derivatives). Let  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  be an extended real valued function and let  $x \in \mathbb{R}^n$  at which  $f$  is finite. Then Lower and Upper Dini derivatives of  $f$  at  $x$  in the direction  $v \in \mathbb{R}^n$  are defined as

$$f^-(x; v) := \liminf_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t},$$

$$f^+(x; v) := \limsup_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t}.$$

**Definition 2.6** (Convexificators). Let  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  be an extended real valued function and let  $x \in \mathbb{R}^n$  at which  $f$  is finite. Then

- (i) The function  $f$  is said to admit an upper convexificator (UCF)  $\partial^* f(x)$  at  $x$  if  $\partial^* f(x) \subset \mathbb{R}^n$  is closed and for every  $v \in \mathbb{R}^n$ ,

$$f^-(x; v) \leq \sup_{\xi \in \partial^* f(x)} \langle \xi, v \rangle.$$

- (ii) The function  $f$  is said to admit a lower convexificator (LCF)  $\partial_* f(x)$  at  $x$  if  $\partial_* f(x) \subset \mathbb{R}^n$  is closed and for every  $v \in \mathbb{R}^n$ ,

$$f^+(x; v) \geq \inf_{\xi \in \partial_* f(x)} \langle \xi, v \rangle.$$

- (iii)  $f$  is said to admit a convexificator (CF)  $\partial_*^* f(x)$  at  $x$  if  $\partial_*^* f(x)$  is both an UCF and an LCF of  $f$  at  $x$ .
- (iv)  $f$  is said to admit an upper regular convexificator (URCF)  $\partial^r f(x)$  at  $x$  if  $\partial^r f(x)$  is an UCF of  $f$  at  $x$  and for every  $v \in \mathbb{R}^n$ ,

$$f^+(x; v) = \sup_{\xi \in \partial^r f(x)} \langle \xi, v \rangle.$$

- (v)  $f$  is said to admit a lower regular convexificator (LRCF)  $\partial_r f(x)$  at  $x$  if  $\partial_r f(x)$  is a LCF of  $f$  at  $x$  and for every  $v \in \mathbb{R}^n$ ,

$$f^-(x; v) = \inf_{\xi \in \partial_r f(x)} \langle \xi, v \rangle.$$

**Remark 2.7.** It is obvious from the definition that every URCF (LRCF) is UCF (LCF).

Now we need the following result from Propositions 2.2 and 2.3 of [26].

- Proposition 2.8.** (i) Let  $\partial_*^* f(x)$  be a convexificator of  $f$  at  $x$ , then  $\alpha \partial_*^* f(x)$  is a convexificator of  $\alpha f$  at  $x$  for every  $\alpha \geq 0$ .
- (ii) Let  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ , for  $i = 1, 2$ . Suppose that for each  $i = 1, 2$ ,  $\partial_*^* f_i(x)$  is a convexificator of  $f_i$  at  $x$ , then the set  $\partial_*^* f_1(x) + \partial_*^* f_2(x)$  is a convexificator of  $(f_1 + f_2)$  at  $x$ . Convexificator  $\partial_*^* f_1(x) + \partial_*^* f_2(x)$  may properly contain another convexificator of  $f_1 + f_2$ .

The following example is from [26] which shows that  $\partial_*^* f_1(x) + \partial_*^* f_2(x)$  may contain another convexificator of  $(f_1 + f_2)$ .

**Example 2.9.** Let  $f_1(x, y) = |x| - |y|$ ,  $f_2(x, y) = |y| - |x|$ . Then

$$f_1(x, y) + f_2(x, y) = 0,$$

it is easy to verify that

$$\partial_*^* f_1(0, 0) = \partial_*^* f_2(0, 0) = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}^T, \begin{pmatrix} -1 \\ 1 \end{pmatrix}^T \right\},$$

and

$$\partial_*^*(f_1 + f_2)(0, 0) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}^T \right\},$$

hence,

$$\partial_*^*(f_1 + f_2)(0, 0) \subset \partial_*^* f_1(0, 0) + \partial_*^* f_2(0, 0).$$

Also, the limiting subdifferential of  $f_1$  and  $f_2$  is

$$\partial f_1(0, 0) = \partial f_2(0, 0) = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}^T, \begin{pmatrix} -1 \\ 1 \end{pmatrix}^T, \begin{pmatrix} -1 \\ -1 \end{pmatrix}^T, \begin{pmatrix} 1 \\ 1 \end{pmatrix}^T \right\},$$

and we must note that  $\overline{\text{co}} \partial_*^* f_1(0, 0) \subset \overline{\text{co}} \partial f_1(0, 0)$  with strict inclusion and  $\partial f_1(0, 0) \not\subset \overline{\text{co}} \partial_*^* f_1(0, 0)$ .

Now, some calculus rules and results for convexificators ([13], Rule 4.1, 4.2), which we use in this paper are as follows

**Lemma 2.10.** Let  $\partial^* f(x)$ ,  $\partial_* f(x)$  and  $\partial_*^* f(x)$  be UCF, LCF and CF of  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  at  $x$  respectively and  $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  admits an URCF  $\partial^r f_1(x)$  at  $x$ . Then

- (i)  $\forall \lambda > 0$ ,  $\lambda \partial^* f(x)$  and  $\lambda \partial_* f(x)$  respectively are UCF and LCF of  $\lambda f$  at  $x$ .
- (ii)  $\forall \lambda < 0$ ,  $\lambda \partial^* f(x)$  and  $\lambda \partial_* f(x)$  respectively are LCF and UCF of  $\lambda f$  at  $x$ .
- (iii)  $\forall \lambda \in \mathbb{R}$ ,  $\lambda \partial^* f(x)$  is a CF of  $\lambda f$  at  $x$ .
- (iv)  $\partial^* f(x) + \partial^r f_1(x)$  is an UCF of  $f + f_1$  at  $x$ .

**Remark 2.11.** Since limiting subdifferential is also a convexificator, so there may be some convexificators which are smaller than limiting subdifferential (see e.g. Example 2.9). For smaller convexificators we have  $\partial^* f(x_0) \subset L_f \mathbb{B}$  directly by the use of Proposition 2.2(ii), otherwise we can have the following result.

**Lemma 2.12.** *If a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is Lipschitz near  $x_0$  with positive constant  $L_f$  and admits an upper regular convexificator  $\partial^r f(x_0)$ , then  $\partial^r f(x_0) \subset L_f \mathbb{B}$ .*

*Proof.* Since function  $f$  is Lipschitz, we have

$$|f(x_1) - f(x_2)| \leq L_f \|x_1 - x_2\|,$$

$$L_f \mathbb{B} = \{x \in \mathbb{R}^n \mid x = L_f u \text{ and } \|u\| \leq 1 \text{ i.e. } \|x\| \leq L_f\}.$$

To achieve our goal, we have to show that  $\|\xi\| \leq L_f$  for all  $\xi \in \partial^r f(x_0)$ .

Now, due to the admittance of upper regular convexificator by  $f$ , we have for every  $v \in \mathbb{R}^n$

$$f^+(x_0, v) = \sup_{\xi \in \partial^r f(x)} \langle \xi, v \rangle, \quad \forall v \in \mathbb{R}^n$$

i.e.

$$\limsup_{t \downarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \sup_{\xi \in \partial^r f(x)} \langle \xi, v \rangle, \quad \forall v \in \mathbb{R}^n. \tag{2.2}$$

Since we have by the properties of modulus function

$$\begin{aligned} \left| \limsup_{t \downarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} \right| &\leq \limsup_{t \downarrow 0} \frac{|f(x_0 + tv) - f(x_0)|}{|t|} \\ &= \limsup_{t \downarrow 0} \frac{1}{|t|} L_f \|x_0 + tv - x_0\| \\ &= \limsup_{t \downarrow 0} \frac{1}{|t|} L_f |t| |v| \\ &= L_f |v|, \end{aligned}$$

then involving (2.2) gives us

$$\begin{aligned} \left| \sup_{\xi \in \partial^r f(x)} \langle \xi, v \rangle \right| &\leq L_f \|v\|, \quad \forall v \in \mathbb{R}^n \\ \langle \xi, v \rangle &\leq L_f \|v\|, \quad \forall v \in \mathbb{R}^n, \quad \forall \xi \in \partial^r f(x^*), \end{aligned} \tag{2.3}$$

which implies that  $\|\xi\| \leq L_f$ . For, let  $\xi \in \partial^r f(x)$  is such that  $\|\xi\| > L_f$ , then taking  $v = \xi$  in equation (2.3), we get

$$\begin{aligned} \langle \xi, \xi \rangle &\leq L_f \|\xi\| \\ \implies \|\xi\| &\leq L_f, \end{aligned}$$

hence

$$\partial^r f(x) \subset L_f \mathbb{B}.$$

□

Since we know that most of the subdifferentials are convexificators and they often contain convex hull of convexificator [13], therefore classical necessary optimality condition has also been introduced for nonsmooth problems in terms of convexificators by replacing the classical gradient with the convex hull of convexificators of concerned functions. Which is as follows.

**Theorem 2.13** ([9], Thm. 3.6). *Let  $S$  be a convex set and let  $x_0 \in S$  be a minimum of a locally Lipschitz function  $f$  over  $S$ . Let  $f$  admits an upper convexificator  $\partial^* f(x_0)$  at  $x_0$ . Then*

$$0 \in \overline{\text{co}(\partial^* f(x_0)) + \mathcal{N}_S(x_0)},$$

where  $\mathcal{N}_S(x_0)$  denotes the normal cone to the convex set  $S$  at  $x_0 \in S$ .

### 3. IMPROVED ENHANCED FRITZ JOHN NECESSARY OPTIMALITY CONDITION

The Fritz John condition is a necessary optimality condition, which has been established for the smooth case and many analogous extensions have been made for cases of nonsmooth optimization problems. For nonsmooth case gradient of the involved functions has been replaced by the generalized gradient, see Clarke [5]. Then by using this generalized subdifferential concept, a lot of work has been done using several different types of subdifferential objects ([6], Thm. 6.1.1 and [20]). Here we have focused on nonsmooth optimization problem having equality, inequality and set constraint to establishing some of the results, which have already been established in the literature, see [27].

The first theorem provides the enhanced Fritz John necessary optimality conditions in terms of convexificators. Here, we replace the limiting subdifferentials of involved functions in Theorem 1 of [27] by their convexificators. Although we have assumed all the involved functions to be locally Lipschitz and for this case limiting subdifferential also becomes convexificators, but we must note that they are not the only convexificators, there may be some other which can be smaller than limiting subdifferential, see Example 2.9. So for these convexificators, the provided result is sharper than Theorem 1 of [27] in the sense that our result will rely on the convex hull of fewer points than the limiting subdifferential. Therefore, to make our results fruitful and sharper than the existing one, we are assuming that all the involved functions admit convexificators, which are smaller than their limiting subdifferentials. For the otherwise case, applying the proposed result is not a smart way, because for that situation using limiting subdifferentials will give more perfect optimality condition.

Here, Theorem 3.1(v) is an additional outcome in contrast to Theorem 1 of [27], it ensures the differentiability of involved functions on a sequence converging at the optimal point and satisfying CV conditions. Hence, our is an improved enhanced Fritz John condition.

**Theorem 3.1.** *Let  $\bar{x}$  be a local minimum of problem (P) and set  $\mathcal{C}$  is closed and convex, where all the involved functions admit upper regular and upper semicontinuous convexificators. Then there exist scalars  $\bar{\mu}_0, \bar{\lambda}_1, \dots, \bar{\lambda}_p, \bar{\mu}_1, \dots, \bar{\mu}_q$  satisfying the following conditions:*

- (i)  $0 \in \bar{\mu}_0 \overline{\text{co}} \partial^r f(\bar{x}) + \sum_{i=1}^p \bar{\lambda}_i \overline{\text{co}} \partial^r (\bar{\lambda}_i h_i)(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \overline{\text{co}} \partial^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x})$ .
- (ii)  $\bar{\mu}_j \geq 0$  for all  $j = 0, 1, \dots, q$ .
- (iii)  $\bar{\mu}_0, \bar{\lambda}_1, \dots, \bar{\lambda}_p, \bar{\mu}_1, \dots, \bar{\mu}_q$  are not all equal to zero.
- (iv) *The complementary violation condition holds with proximal subdifferentiability: If the index set  $I \cup J$  is nonempty, where  $I = \{i | \bar{\lambda}_i \neq 0\}$ ,  $J = \{j \neq 0 | \bar{\mu}_j > 0\}$ , then there exist a sequence  $\{x^k\} \subset \mathcal{C}$  converging to  $\bar{x}$  such that for all  $k$ ,  $f(x^k) < f(\bar{x})$ ,  $\bar{\lambda}_i h_i(x^k) > 0, \forall i \in I$ ,  $\bar{\mu}_j g_j(x^k) > 0, \forall j \in J$ , and  $f, h_i (i \in I), g_j (j \in J)$  are all proximal subdifferentiable at  $x^k$ .*
- (v) *The complementary violation condition holds with differentiability also i.e. there exist an another sequence  $\{x^n\} \subset \mathcal{C}$  converging to  $\bar{x}$  and satisfying the same inequalities as by  $\{x^k\}$  in (iv) and  $f, h_i (i \in I), g_j (j \in J)$  are all differentiable at  $x^n$ .*

*Proof.* Let  $\bar{x}$  be a local minimum of problem  $(P)$ . Then we can adopt a quadratic penalty function approach originally given by McShane [18] and used by Bertsekas [3] and later by Ye and Zhang [27]. Firstly, we consider a penalized problem for each  $k = 1, 2, \dots$

$$(P_k) \quad \min F^k(x) = f(x) + \frac{k}{2} \sum_{i=1}^p (h_i(x))^2 + \frac{k}{2} \sum_{j=1}^q g_j^+(x)^2 + \frac{1}{2} \|x - \bar{x}\|^2$$

s.t.  $x \in \mathcal{C} \cap \mathbb{B}[\bar{x}, \epsilon],$

where  $\epsilon > 0$  is such that  $f(\bar{x}) \leq f(x)$  for all feasible  $x \in \mathbb{B}[\bar{x}, \epsilon]$ . Due to the compactness of  $\mathcal{C} \cap \mathbb{B}[\bar{x}, \epsilon]$  and by the Weierstrass theorem, problem  $(P_k)$  admits an optimal solution, say  $x^k$ . Then

$$F^k(x^k) \leq F^k(\bar{x}), \tag{3.1}$$

$$f(x^k) + \frac{k}{2} \sum_{i=1}^p (h_i(x^k))^2 + \frac{k}{2} \sum_{j=1}^q (g_j^+(x^k))^2 + \frac{1}{2} \|x^k - \bar{x}\|^2 \leq F^k(\bar{x}) = f(\bar{x}). \tag{3.2}$$

Here,  $f(x^k)$  is bounded over  $x^k \in \mathcal{C} \cap \mathbb{B}[\bar{x}, \epsilon]$ , therefore we obtained from (3.2)

$$\lim_{k \rightarrow \infty} |h_i(x^k)| = 0, \quad i = 1, \dots, p,$$

and  $\lim_{k \rightarrow \infty} |g_j^+(x^k)| = 0, \quad j = 1, \dots, q,$

hence, every limit point  $x_0$  of  $\{x^k\}$  is feasible and lie in  $\Omega$ . Further, from (3.2) it follows that

$$f(x^k) + \frac{1}{2} \|x^k - \bar{x}\|^2 \leq f(\bar{x}); \quad \forall k \tag{3.3}$$

by taking limit as  $k \rightarrow \infty$ , we get

$$f(x_0) + \frac{1}{2} \|x_0 - \bar{x}\|^2 \leq f(\bar{x}). \tag{3.4}$$

Here, since  $x_0 \in \Omega$  and lies in an  $\epsilon$ -neighbourhood of  $\bar{x}$ , therefore we have

$$f(\bar{x}) \leq f(x_0). \tag{3.5}$$

Now, combining these two equations (3.4), and (3.5) we have

$$\|x_0 - \bar{x}\| = 0 \Rightarrow x_0 = \bar{x}.$$

Thus, the sequence  $\{x^k\}$  converges to  $\bar{x}$ , it follows that  $x^k$  is an interior point of the closed ball  $\mathbb{B}[\bar{x}, \epsilon]$  for all  $k$  greater than some  $\bar{k}$ .

Since  $x^k$  is an optimal solution of  $(P_k)$  and also is an interior point of  $\mathbb{B}[\bar{x}, \epsilon]$ , we have a necessary optimality condition in terms of convexifiers and normal cone at the optimal point, see Theorem 2.13. Although, this condition is true for any convexifiers, we are assuming admittance of upper regular, upper semicontinuous convexifier at the concerned point for sharp results. Now by using Theorem 2.13 we get

$$0 \in \overline{\overline{c\partial} \partial^r F^k(x^k) + \mathcal{N}_{\mathcal{C}}(x^k)},$$

$$0 \in \overline{\overline{c\partial} \left[ \partial^r f(x^k) + \frac{k}{2} \sum_{i=1}^p \partial^r (h_i(x^k))^2 + \frac{k}{2} \sum_{j=1}^q \partial^r (g_j^+(x^k))^2 + \frac{1}{2} \partial^r \|x^k - \bar{x}\|^2 \right] + \mathcal{N}_{\mathcal{C}}(x^k)}.$$

Using basic calculus rules of closure and convex hull of a set gives

$$0 \in \overline{c\partial} \partial^r f(x^k) + \overline{c\partial} \left[ \frac{k}{2} \sum_{i=1}^p \partial^r (h_i(x^k))^2 \right] + \overline{c\partial} \left[ \frac{k}{2} \sum_{j=1}^q \partial^r (g_j^+(x^k))^2 \right] + \frac{1}{2} \partial^r \|x^k - \bar{x}\|^2 + \mathcal{N}_{\mathcal{C}}(x^k).$$



Here, we have the existence of multipliers

$$\alpha_i^k = \frac{k}{2}h_i(x^k), \quad \beta_j^k = \frac{k}{2}g_j^+(x^k). \tag{3.6}$$

Then,

$$\begin{aligned} 0 &\in \overline{co}\partial^r f(x^k) + \overline{co} \left[ \sum_{i=1}^p \partial^r(\alpha_i^k h_i(x^k)) \right] + \overline{co} \left[ \sum_{j=1}^q \partial^r(\beta_j^k g_j^+(x^k)) \right] + (x^k - \bar{x}) + \mathcal{N}_C(x^k), \\ 0 &\in \overline{co}\partial^r f(x^k) + \overline{co} \left[ \sum_{i=1}^p \partial^r(\alpha_i^k h_i(x^k)) \right] + \overline{co} \left[ \sum_{j \in J'} \partial^r(\beta_j^k g_j^+(x^k)) \right] + (x^k - \bar{x}) + \mathcal{N}_C(x^k), \end{aligned}$$

where,

$$J' = \{j \in (1, \dots, q) \mid g_j(x^k) > 0\}.$$

By (3.6),  $\beta_j^k$  is a non-negative scalar, therefore using Lemma 2.10(i)

$$0 \in \overline{co}\partial^r f(x^k) + \overline{co} \left[ \sum_{i=1}^p \partial^r(\alpha_i^k h_i(x^k)) \right] + \overline{co} \left[ \sum_{j \in J'} \beta_j^k \partial^r(g_j^+(x^k)) \right] + (x^k - \bar{x}) + \mathcal{N}_C(x^k). \tag{3.7}$$

Let

$$\begin{aligned} \delta^k &= \sqrt{1 + \sum_{i=1}^p (\alpha_i^k)^2 + \sum_{j \in J'} (\beta_j^k)^2}, \\ \mu_0^k &= \frac{1}{\delta^k}, \quad \lambda_i^k = \frac{\alpha_i^k}{\delta^k}, \quad i = 1, \dots, p, \quad \mu_j^k = \frac{\beta_j^k}{\delta^k}, \quad j = 1, \dots, q. \end{aligned} \tag{3.8}$$

Due to positivity of  $\delta^k$ , we can divide (3.7) by  $\delta^k$  and also using some calculus, we have for all  $k > \bar{k}$

$$0 \in \overline{co} [\mu_0^k \partial^r f(x^k)] + \sum_{i=1}^p \overline{co} \partial^r(\lambda_i^k h_i(x^k)) + \sum_{j \in J'} \mu_j^k \overline{co} \partial^r(g_j^+(x^k)) + \frac{1}{\delta^k}(x^k - \bar{x}) + \mathcal{N}_C(x^k), \tag{3.9}$$

since

$$(\mu_0^k)^2 + \sum_{i=1}^p (\lambda_i^k)^2 + \sum_{j=1}^q (\mu_j^k)^2 = 1, \tag{3.10}$$

therefore this sequence  $\{\mu_0^k, \lambda_1^k, \dots, \lambda_p^k, \mu_1^k, \dots, \mu_q^k\}$  is bounded, hence contains a subsequence converging to some limit  $\{\bar{\mu}_0, \bar{\lambda}_1, \dots, \bar{\lambda}_p, \bar{\mu}_1, \dots, \bar{\mu}_q\}$ . Since  $h_i$  is locally Lipschitz and by the Proposition 2.8, we have

$$\begin{aligned} \partial^r(\lambda_i^k h_i)(x^k) &= \partial^r(\lambda_i^k h_i + \bar{\lambda}_i h_i - \bar{\lambda}_i h_i)(x^k) \\ &\subseteq \mathcal{L}_{h_i} |\lambda_i^k - \bar{\lambda}_i| \mathbb{B} + \partial^r[\bar{\lambda}_i h_i(x^k)] \end{aligned} \tag{by Lem. 2.12}$$

and therefore,

$$\overline{co}\partial^r(\lambda_i^k h_i)(x^k) \subseteq \mathcal{L}_{h_i} |\lambda_i^k - \bar{\lambda}_i| \mathbb{B} + \overline{co}\partial^r[\bar{\lambda}_i h_i(x^k)].$$

Similarly for  $f$  and  $g_j$ , we have

$$\begin{aligned} \overline{co}[\mu_0^k \partial^r f(x^k)] &\subseteq \mathcal{L}_f |\mu_0^k - \bar{\mu}_0| \mathbb{B} + \overline{co}\bar{\mu}_0 \partial^r[f(x^k)], \\ \overline{co}[\mu_j^k \partial^r(g_j)(x^k)] &\subseteq \mathcal{L}_{g_j} |\mu_j^k - \bar{\mu}_j| \mathbb{B} + \overline{co}[\bar{\mu}_j \partial^r g_j(x^k)]. \end{aligned}$$

Using these results, equation (3.9) yields

$$0 \in \bar{\mu}_0 \bar{c} \bar{o} \partial^r f(x^k) + \sum_{i=1}^p \bar{c} \bar{o} \partial^r (\bar{\lambda}_i h_i)(x^k) + \sum_{j \in J'} \bar{\mu}_j \bar{c} \bar{o} \partial^r g_j(x^k) + (\mathcal{L}_{h_i} |\lambda_i^k - \bar{\lambda}_i| + \mathcal{L}_f |\mu_0^k - \bar{\mu}_0| + \mathcal{L}_{g_j} |\mu_j^k - \bar{\mu}_j|) \mathbb{B} + \frac{1}{\delta_k} (x^k - \bar{x}) + \mathcal{N}_{\mathcal{C}}(x^k).$$

Since all the involved convexificators are upper semi continuous and by the definition of normal cone, taking limit as  $k \rightarrow \infty$  and the existence of  $\bar{\mu}_0, \bar{\lambda}_i, \bar{\mu}_j$  yields result (i)

$$0 \in \bar{\mu}_0 \bar{c} \bar{o} \partial^r f(\bar{x}) + \sum_{i=1}^p \bar{c} \bar{o} \partial^r (\bar{\lambda}_i h_i)(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \bar{c} \bar{o} \partial^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}). \tag{3.11}$$

By the construction of  $\mu_j^k$  in equation (3.8), we have  $\mu_j^k \geq 0$  for all  $k$  and  $j$ . Therefore  $\bar{\mu}_j \geq 0$  for all  $j = 1, \dots, q$ , which is (ii). Similarly (3.10) ensures condition (iii).

Now, the existence of a sequence  $\{x^k\} \rightarrow \bar{x}$  at which all the involved functions are proximal subdifferentiable, i.e. property (iv) can be shown similarly as in Ye ([27], Thm. 1), and finally to prove (v) we use Lipschitz continuity of involved functions. Firstly, it is easy to find a sequence  $\{z^n\}$  in a similar manner which satisfies following inequalities:

$$\begin{aligned} f(z^n) &< f(\bar{x}), \\ \bar{\lambda}_i h_i(z^n) &> 0, \quad \forall i \in I, \\ \bar{\mu}_j g_j(z^n) &> 0, \quad \forall j \in J. \end{aligned}$$

Since  $f, h_i, g_j$  are locally Lipschitz and therefore differentiable almost everywhere (thanks to Rademacher’s theorem). Hence, for each  $n \geq N_0$  we can find a sequence  $\{z^{n,l}\} \subset \mathcal{C}$  at which  $f, h_i, g_j$  are differentiable and

$$z^{n,l} \rightarrow z^n.$$

Due to continuity, we have for sufficiently large  $l$ ,

$$\begin{aligned} f(z^{n,l}) &< f(\bar{x}), \\ \bar{\lambda}_i h_i(z^{n,l}) &> 0, \quad \forall i \in I, \\ \bar{\mu}_j g_j(z^{n,l}) &> 0, \quad \forall j \in J. \end{aligned}$$

For  $n \geq N_0$ , choose  $l_n$  such that  $l_1 < \dots < l_{n-1} < l_n$  and  $\lim_{n \rightarrow \infty} z^{n,l_n} = \bar{x}$ . Now we construct a sequence  $\{x^n\}$  as

$$x^n = z^{(n+N_0), (l_n+N_0)}, \quad N_0 = 1, 2, \dots$$

Then by combining preceding results for such sequence, we have  $\{x^n\} \subset \mathcal{C}$  and

$$\lim_{n \rightarrow \infty} x^n = \bar{x}, \quad f(x^n) < f(\bar{x}), \quad \bar{\lambda}_i h_i(x^n) > 0, \quad \forall i \in I, \quad \bar{\mu}_j g_j(x^n) > 0, \quad \forall j \in J,$$

and  $f, h_i (i \in I), g_j (j \in J)$  are differentiable at  $x^n$ . □

**Remark 3.2.** (1) Since convexificators may be smaller than limiting subdifferentials, our result is sharper than Theorem 1 of [27] because it relies on the convex hull of potentially fewer points.

(2) It is important to note that the proof is similar to the proof of Theorem 1 in [27], but without employing the Lemma 2.12, we can not conclude the condition (i) in Theorem 3.1.

Since a locally Lipschitz continuous function need not be differentiable, the condition (v) is important as it provides information about a sequence at which all the involved functions are differentiable and where CV conditions hold, which may be useful in further analysis.

**Remark 3.3.** In the obtained equation (3.11), it should be noted that multipliers associated to the equality constraints are free, so we can not write  $\bar{\lambda}_i \partial^r h_i(\bar{x})$  instead of  $\partial^r(\bar{\lambda}_i h_i)(\bar{x})$  directly. As already pointed out in Proposition 2(ii) of [27], if we fix  $f$  to be the zero function, then we get  $\partial(-g)(\bar{x}) \subset -\partial g(\bar{x})$  for the limiting subdifferential case, but under the restriction that the Fréchet subdifferential of  $g$  is nonempty near the point  $\bar{x}$ . Since limiting subdifferentials are also convexifiers, so we can not say that  $\partial^r(\bar{\lambda}_i h_i)(\bar{x})$  is contained in  $\bar{\lambda}_i \partial^r h_i(\bar{x})$ .

Now, we illustrate the above Theorem 3.1 in the following example.

**Example 3.4.** Consider the mathematical program

$$\begin{aligned} & \min (x_1 + 1)^2 + |x_2| \\ \text{s.t. } & -2x_1 + |x_1| + |x_2| \leq 0. \end{aligned}$$

Clearly,  $(0, 0)$  is the optimal solution of the above problem. Here

$$\begin{aligned} f(x_1, x_2) &= (x_1 + 1)^2 + |x_2| \\ g(x_1, x_2) &= -2x_1 + |x_1| + |x_2| \end{aligned}$$

and

$$C = \mathbb{R}^2.$$

Then, both the functions admit upper regular, upper semicontinuous convexifiers at  $(0, 0)$

$$\begin{aligned} \partial^r f(0, 0) &= \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}^T, \begin{pmatrix} 2 \\ -1 \end{pmatrix}^T \right\}, \\ \partial^r g(0, 0) &= \left\{ \begin{pmatrix} -3 \\ -1 \end{pmatrix}^T, \begin{pmatrix} -1 \\ 1 \end{pmatrix}^T \right\}. \end{aligned}$$

We can easily see that

$$\begin{aligned} \mu_0 \bar{c} \partial^r f(0, 0) &= \left\{ \begin{pmatrix} 2\mu_0 \\ (2s - 1)\mu_0 \end{pmatrix}^T \mid 0 \leq s \leq 1 \right\} \\ \bar{c} \partial^r(\mu g)(0, 0) &= \left\{ \begin{pmatrix} (-2t - 1)\mu \\ (-2t + 1)\mu \end{pmatrix}^T \mid 0 \leq t \leq 1 \right\}. \end{aligned}$$

If  $t = 0$  then Theorem 3.1(i) holds for  $\mu_0 = 1$  and  $\mu = \frac{3}{2}$ . Infact, if we fix  $\mu_0 = 1$ , value of  $\mu$  depends upon value of  $t$ ,  $\mu$  varies from  $1/2 \leq \mu \leq 3/2$  whenever  $0 \leq t \leq 1$ .

We also have a sequence  $\{(\frac{-1}{n}, \frac{1}{n})\}$  where condition (iv) and (v) holds.

For the above example, we can see that the limiting subdifferentials of involved functions are

$$\partial f(0, 0) = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}^T, \begin{pmatrix} 2 \\ -1 \end{pmatrix}^T \right\},$$

and

$$\partial g(0, 0) = \left\{ \begin{pmatrix} -3 \\ -1 \end{pmatrix}^T, \begin{pmatrix} -1 \\ 1 \end{pmatrix}^T, \begin{pmatrix} -3 \\ 1 \end{pmatrix}^T, \begin{pmatrix} -1 \\ -1 \end{pmatrix}^T \right\}.$$

It is clear that, the limiting subdifferential may be finite, whereas convex hull of convexifier is of course need not be a finite set. In optimization, considering convex hull is natural. So when limiting subdifferential is bigger than convexifier, our result would be better in the sense that the optimality condition rely on the convex hull of few points compare to limiting subdifferential.

On the other hand, dropping the locally Lipschitz continuity we may lose condition (iv) and (v) in Theorem 3.1.

## 4. IMPROVED ENHANCED KKT CONDITION

Now by the use of above Fritz John condition, we can derive enhanced KKT-condition by using the classical approach in the literature, *i.e.* by making coefficient nonzero, associated to objective function in enhanced Fritz John condition.

**Definition 4.1.** Let  $\bar{x}$  be a feasible point of the problem  $(P)$ . Then improved enhanced KKT condition holds at  $\bar{x}$  if the improved Fritz John condition holds with  $\bar{\mu}_0 = 1$ .

**Theorem 4.2.** Let  $\bar{x}$  be a local minimum of the problem  $(P)$ . Suppose there is no nonzero vector  $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}_+^q$  such that

$$0 \in \sum_{i=1}^p \bar{c}_0 \partial^r (\bar{\lambda}_i h_i)(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \bar{c}_0 \partial^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}), \quad (4.1)$$

and the CV conditions defined in Theorem 3.1(iv) and (v) hold. Then the improved enhanced KKT condition holds at  $\bar{x}$ .

*Proof.* If  $\bar{\mu}_0 = 0$ , then according to Theorem 3.1 there exists a sequence  $(\lambda, \mu)$  which satisfies (4.1) and hence it violates assumption. Therefore  $\bar{\mu}_0$  must be nonzero and after a proper scaling of  $\bar{\lambda}, \bar{\mu}$  the enhanced KKT conditions holds.  $\square$

**Remark 4.3.** In Example 3.4, we can see that only zero vector satisfies equation (4.1), *i.e.* improved enhanced KKT condition also holds for Example 3.4.

Now with the help of Theorem 4.2, we are in a position to define some new terms and constraint qualifications (CQ) associated with convexifiers, similar to the *pseudonormality* and *quasinormality* concepts [3] of the smooth case. These CQs will play a significant role in exploring the behavior of concerned problem. It is necessary to mention here that condition in Theorem 4.2 helps in defining these CQ but it may not be a constraint qualification itself as it involves the objective function.

**Definition 4.4.** Let  $\bar{x}$  be a feasible solution of problem  $(P)$ , Then

- (i)  $\bar{x}$  is said to satisfy NNAMCQ if there is no nonzero vector  $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}_+^q$  such that condition (4.1) and the CS holds:  $\mu_j g_j(\bar{x}) = 0, \forall j = 1, \dots, q$ .
- (ii)  $\bar{x}$  is said to satisfy  $\partial^*$ -pseudonormality (for the region  $\Omega$ ), if there is no vector  $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}_+^q$  and no infeasible sequence  $\{x^t\} \subset \mathcal{C}$  ( $t = k, n$ ) converging to  $\bar{x}$  such that condition (4.1) and the pseudo-complementarity slackness condition (pseudo-CS) holds: if the index set  $I \cup J$  is nonempty for  $I = \{i | \lambda_i \neq 0\}, J = \{j | \mu_j > 0\}$ , then for each  $t$  ( $t = k, n$ )

$$\sum_{i=1}^p \lambda_i h_i(x^t) + \sum_{j=1}^q \mu_j g_j(x^t) > 0,$$

and  $h_i (i \in I), g_j (j \in J)$  are all proximal subdifferentiable (differentiable) at  $x^t$  for each  $t$  ( $t = k, n$ ).

- (iii)  $\bar{x}$  is said to satisfy  $\partial^*$ -quasinormality (for the region  $\Omega$ ), if there is no vector  $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}_+^q$  and no infeasible sequence  $\{x^t\} \subset \mathcal{C}$  ( $t = k, n$ ) converging to  $\bar{x}$  such that condition (4.1) and the quasi-complementarity slackness condition (quasi-CS) holds: if the index set  $I \cup J$  is nonempty for  $I = \{i | \lambda_i \neq 0\}, J = \{j | \mu_j > 0\}$ , then for all  $i \in I, j \in J, \lambda_i h_i(x^t) > 0$  and  $\mu_j g_j(x^t) > 0$  and  $h_i (i \in I), g_j (j \in J)$  are all proximal subdifferentiable (differentiable) at  $x^t$  for each  $t = k$  ( $t = n$ ).

**Remark 4.5.** Due to the involvement of convexifiers in equation (4.1), above definitions are different from the Definition 5 of [27].

In Proposition 3.1 of [3] Bertsekas and Ozdaglar employed some restrictions on constraints for a point to be pseudonormal. They proved that any feasible point of a constraint region must be pseudonormal if equality functions are linear and inequality functions are concave and smooth and there is no abstract set constraint. Ye ([27], Prop. 3) extended it to the nonsmooth case, where all involved functions have limiting subdifferential. Here, this paper is about such problems, where functions may admit more general objects called convexificators, which includes limiting and many other subdifferentials frequently consisting in nonsmooth optimization, see [9, 13, 16, 26]. Hence our definition of  $\partial^*$ -pseudonormality and  $\partial^*$ -quasinormality is much weaker than that of Ye and Zhang [27] or Bertsekas and Ozdaglar [3]. Here we have an enhanced result for any feasible point of the problems including functions which admit convexificators, to be  $\partial^*$ -pseudonormal. We are presenting only statement here, because the proof is similar to Proposition 3 of [27].

**Proposition 4.6.** *Any feasible point of the problem (P) is  $\partial^*$ -pseudonormal under the restrictions that  $C = \mathbb{R}^m$ ,  $h_i$  are linear and  $g_j$  is concave.*

**Definition 4.7.** Let  $\bar{x}$  be a point in the feasible region  $\Omega$ . A vector  $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}_+^q$  is said to be  $\partial^*$ -quasinormal multiplier, if it satisfies the following weaker version of the improved enhanced KKT conditions

- (i)  $0 \in \bar{\mu}_0 \bar{c} \partial^r f(\bar{x}) + \sum_{i=1}^p \bar{c} \partial^r (\bar{\lambda}_i h_i)(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \bar{c} \partial^r g_j(\bar{x}) + \mathcal{N}_C(\bar{x})$ .
- (ii) There exist sequences  $\{x^k\}$  and  $\{x^n\} \subset C$  converging to  $\bar{x}$  such that the quasi-CS in Definition 4.4 holds.

**Corollary 4.8.** *Let  $\bar{x}$  be a local minimum of the problem (P). Then the weaker version of the improved enhanced KKT condition holds at  $\bar{x}$ , provided that at  $\bar{x}$ , NNAMCQ, or  $\partial^*$ -pseudonormality or  $\partial^*$ -quasinormality is satisfied.*

*Proof.* Suppose  $\partial^*$ -quasinormality holds at  $\bar{x}$ . Also, if  $\bar{x}$  is a local minimum, then according to Theorem 3.1, there is a vector  $(\mu_0, \lambda, \mu)$  for which condition (i) and (ii) of Definition 4.7 easily holds. Now, since quasi-CS conditions hold for  $(\lambda, \mu)$ , therefore according to  $\partial^*$ -quasinormality, equation (4.1) does not holds for  $(\lambda, \mu)$ , it means  $\mu_0$  can not be zero. Hence, weaker version of improved enhanced KKT condition holds at  $\bar{x}$ .

Since NNAMCQ and  $\partial^*$ -pseudonormality is stronger than  $\partial^*$ -quasinormality, *i.e.* same result holds under these CQs. □

Now it has already been established in [15] that NNAMCQ implies the boundedness of the set of all normal multipliers. Later, Ye and Zhang [27] presented a similar result for the nonsmooth problem. The author has shown that at any feasible point, the set of all quasinormal multiplier is bounded under the assumption that the concerned point is quasinormal.

Here, we are dealing with the convexificators as a tool for generalized subdifferential, we extend this result for the set of all  $\partial^*$ -quasinormal multipliers in terms of the convexificators.

**Theorem 4.9.** *Let  $\bar{x}$  be a feasible point for the problem (P). If  $\partial^*$ -quasinormality holds at  $\bar{x}$ , then the set of all  $\partial^*$ -quasinormal multipliers  $M(\bar{x})$  is bounded.*

*Proof.* Suppose, on contrary that  $M(\bar{x})$  is unbounded. Then there is a sequence  $\{(\lambda^n, \mu^n)\} \in M(\bar{x})$  such that  $\|(\lambda^n, \mu^n)\| \rightarrow \infty$  as  $n \rightarrow \infty$ . Then by the definition of  $\partial^*$ -quasinormal multiplier, for all  $n$ , there is a sequence  $\{x_n^k\}$  and  $\{x_n^l\} \subset C$  converging to  $\bar{x}$ , such that

$$0 \in \bar{c} \partial^r f(\bar{x}) + \sum_{i=1}^p \bar{c} \partial^r (\lambda_i^n h_i)(\bar{x}) + \sum_{j=1}^q \mu_j^n \bar{c} \partial^r g_j(\bar{x}) + \mathcal{N}_C(\bar{x}), \tag{4.2}$$

$$\mu_j^n \geq 0, \quad \forall j = 1, \dots, q, \tag{4.3}$$

$$\lambda_i^n h_i(x_n^m) > 0, \quad \forall i \in I^n, \quad (m = k, l) \tag{4.4}$$

$$\mu_j^n g_j(x_n^m) > 0, \quad \forall j \in J^n, \quad (m = k, l) \tag{4.5}$$

$h_i(i \in I^n), g_j(j \in J^n)$  are proximal subdifferentiable at  $x_n^k$  (and differentiable at  $x_n^l$ ), where  $I^n = \{i \mid \lambda_i^n \neq 0\}$  and  $J^n = \{j \mid \mu_j^n > 0\}$ .

Let  $\xi^n = \frac{\lambda^n}{\|(\lambda^n, \mu^n)\|}$  and  $\zeta^n = \frac{\mu^n}{\|(\lambda^n, \mu^n)\|}$  and we can assume that  $(\xi^n, \mu^n) \rightarrow (\bar{\xi}, \bar{\mu})$ , divide both sides of (4.2) by  $\|(\lambda^n, \mu^n)\|$ , we obtain

$$0 \in \sum_{i=1}^p \bar{c}_i \bar{\partial}^r(\xi_i^n h_i)(\bar{x}) + \sum_{j=1}^q \zeta_j^n \bar{c}_j \bar{\partial}^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}).$$

Now, taking limit as  $n \rightarrow \infty$  yields that

$$0 \in \sum_{i=1}^p \bar{c}_i \bar{\partial}^r(\bar{\xi}_i h_i)(\bar{x}) + \sum_{j=1}^q \bar{\zeta}_j \bar{c}_j \bar{\partial}^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}).$$

Since, we have

$$\mu_j^n \geq 0, \quad \forall j = 1, \dots, q,$$

therefore,

$$\bar{\zeta}_j \geq 0, \quad \forall j = 1, \dots, q.$$

Now we can construct index sets as

$$I = \{i \mid \bar{\xi}_i \neq 0\}; \quad J = \{j \mid \bar{\zeta}_j > 0\},$$

then  $I \cup J$  is nonempty and by equations (4.4) and (4.5), there is some  $N$  such that for all  $n > N$  we must have

$$\frac{\lambda_i^n(h_i(x_n^m))}{\|(\lambda_i^n, \mu_j^n)\|} > 0, \quad \forall i \in I^n, \quad (m = k, l),$$

or

$$\bar{\xi}_i h_i(x_n^m) > 0, \quad \forall i \in I, \quad (m = k, l),$$

similarly

$$\bar{\zeta}_j g_j(x_n^m) > 0, \quad \forall j \in J, \quad (m = k, l).$$

Since  $h_i(i \in I^n), g_j(j \in J^n)$  are proximal subdifferentiable at  $x_n^k$  and differentiable at  $x_n^l$ , there exist some scalars  $(\bar{\xi}_1, \dots, \bar{\xi}_p, \bar{\zeta}_1, \dots, \bar{\zeta}_q)$  not all zero and sequences  $\{x_n^k\}, \{x_n^l\} \subset \mathcal{C}$  that satisfies these relations, but it is violation of  $\partial^*$ -quasinormality at  $\bar{x}$ . Hence, our assumption is wrong, *i.e.* set of all  $\partial^*$ -quasinormal multipliers  $M(\bar{x})$  must be bounded. This completes the proof.  $\square$

By using the proof of our Theorem 3.1, we can also generalize a well-established result of Lemma 2 from [24] to our problem. The following result characterizes a  $\partial^*$ -quasinormal vector in its neighbourhood.

**Lemma 4.10.** *If a vector  $\bar{x} \in \Omega$  is  $\partial^*$ -quasinormal, then all the feasible vectors in its neighbourhood are  $\partial^*$ -quasinormal.*

We have studied yet  $\partial^*$ -quasinormality for a point and some characterization of the set of  $\partial^*$ -quasimultipliers, so we may extend now an inclusion result of Proposition 1 from [24] to the concerned optimization problem. Bertsekas and Ozdaglar represented a Fréchet normal cone in terms of a set of quasinormal multipliers for the smooth case.

Although this result also has been extended by Ye and Zhang ([27], Prop. 4) to the nonsmooth case, where involved functions have limiting subdifferentials. The author modified this result for limiting normal cone in terms of quasinormal multipliers. But the case where the function may admit generalized objects still remained. Since we are using convexificators, so our result enhances that of Ye and Zhang [27], Ozdaglar and Bertsekas [24] and Henrion ([10], Thm. 4.1). Here, we are giving a specific representation of the polar of the tangent cone (or Fréchet normal cone) to the constraint region in terms of the set of  $\partial^*$ -quasinormal multipliers.

**Theorem 4.11.** *Let  $\bar{x}$  be  $\partial^*$ -quasinormal for  $\Omega$ , then*

$$\mathcal{T}_\Omega^0(\bar{x}) \subset \left\{ \sum_{i=1}^p \overline{co}\partial^r(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^q \mu_j \overline{co}\partial^r g_j(\bar{x}) + \mathcal{N}_C(\bar{x}) \mid (\lambda, \mu) \in M(\bar{x}) \right\}.$$

*Proof. Step I.* Let  $y$  be any point of  $\mathcal{T}_\Omega^0(\bar{x})$ , then by Ozdaglar and Bertsekas ([24], Lem. 3), there exist a smooth function  $F$  with the property

$$-\nabla F(\bar{x}) = y, \tag{4.6}$$

which achieves a strict global minimum over  $\Omega$  at  $\bar{x}$ . This result can be reduced to  $-\overline{co}\partial^*F(\bar{x}) = y$  due to the smoothness of  $F$ .

Now similar to the proof of Theorem 3.1, we can use a quadratic penalty function approach for each  $k = 1, 2, \dots$

$$\min F^k(x) \quad \text{s.t. } x \in C \cap \mathbb{B}[\bar{x}, \epsilon],$$

where

$$F^k(x) = F(x) + \frac{k}{2} \sum_{i=1}^p (h_i(x))^2 + \frac{k}{2} \sum_{j=1}^q (g_j^+(x))^2.$$

Since  $C \cap \mathbb{B}$  is compact, Therefore due to Weierstrass theorem, there is an optimal solution  $x^k$  for the above problem. *i.e.*

$$\begin{aligned} F^k(x^k) &\leq F^k(\bar{x}) = F(\bar{x}), \\ F(x^k) + \frac{k}{2} \sum_{i=1}^p (h_i(x^k))^2 + \frac{k}{2} \sum_{j=1}^q (g_j^+(x^k))^2 &\leq F(\bar{x}). \end{aligned}$$

Now due to the boundedness of  $F(x^k)$  and by using the same approach as in Theorem 3.1, we conclude that this sequence  $\{x^k\}$  converges to  $\bar{x}$  and  $x^k$  is an interior point of  $\mathbb{B}[\bar{x}, \epsilon]$  for all  $k$  greater than some  $\bar{k}$ .

Now, by the Theorem 3.1,

$$\begin{aligned} 0 &\in \overline{co}\partial^r F^k(x^k) + \mathcal{N}_C(x^k), \\ 0 &\in \overline{co}\partial^r F(x^k) + \sum_{i=1}^p \overline{co}\partial^r(\lambda_i^k h_i)(x^k) + \sum_{j=1}^q \mu_j^k \overline{co}\partial^r g_j(x^k) + \mathcal{N}_C(x^k), \end{aligned}$$

hence,

$$-\overline{co}\partial^r F(x^k) \in \sum_{i=1}^p \overline{co}\partial^r(\lambda_i^k h_i)(x^k) + \sum_{j=1}^q \mu_j^k \overline{co}\partial^r g_j(x^k) + \mathcal{N}_C(x^k). \tag{4.7}$$

There also exist sequences  $\{x^{k,r}\}$  and  $\{x^{k,s}\} \subset C$  converges to  $\bar{x}$  such that , for all  $k$

$$\begin{aligned} \lambda_i^k h_i(x^{k,m}) &> 0, \quad \forall i \in I^k, \quad (m = r, s), \\ \mu_j^k g_j(x^{k,m}) &> 0, \quad \forall j \in J^k, \quad (m = r, s), \end{aligned}$$

and  $h_i$  ( $i \in I^k$ ),  $g_j$  ( $j \in J^k$ ) are proximal subdifferentiable (differentiable) at  $x^{k,r}$  (at  $x^{k,s}$ ), where

$$I^k = \{i \mid \lambda_i^k \neq 0\},$$

$$J^k = \{j \mid \mu_j^k > 0\}.$$

**Step II.** Now to show that the sequence  $\{\lambda_i^k\}$  and  $\{\mu_j^k\}$  is bounded, on contrary assume that this sequence is unbounded. Then denote

$$\delta^k = \sqrt{1 + \sum_{i=1}^k (\lambda_i^k)^2 + \sum_{j=1}^q (\mu_j^k)^2},$$

$$\xi_i^k = \frac{\lambda_i^k}{\delta^k}, \quad \forall i = 1, \dots, p,$$

and

$$\zeta_j^k = \frac{\mu_j^k}{\delta^k}, \quad \forall j = 1, \dots, q.$$

Then by equation (4.7), we have

$$\frac{-\overline{co}\partial^r F(x^k)}{\delta^k} \in \sum_{i=1}^p \overline{co}\partial^r \left(\frac{\lambda_i^k}{\delta^k} h_i\right)(x^k) + \sum_{j=1}^q \frac{\mu_j^k}{\delta^k} \overline{co}\partial^r (g_j)(x^k) + \mathcal{N}_{\mathcal{C}}(x^k),$$

$$\frac{-\overline{co}\partial^r F(x^k)}{\delta^k} \in \sum_{i=1}^p \overline{co}\partial^r (\xi_i^k h_i)(x^k) + \sum_{j=1}^q \zeta_j^k \overline{co}\partial^r g_j(x^k) + \mathcal{N}_{\mathcal{C}}(x^k).$$

Since the sequence  $\{\xi_1^k, \dots, \xi_p^k, \zeta_1^k, \dots, \zeta_q^k\}$  is bounded, we may assume that it converges to  $\{\bar{\xi}_1, \dots, \bar{\xi}_p, \bar{\zeta}_1, \dots, \bar{\zeta}_q\} \neq 0$  as  $k \rightarrow \infty$ . Then taking limit in the last inclusion yields that

$$0 \in \sum_{i=1}^p \overline{co}\partial^r (\bar{\xi}_i h_i)(x^k) + \sum_{j=1}^q \bar{\zeta}_j \overline{co}\partial^r g_j(x^k) + \mathcal{N}_{\mathcal{C}}(x^k),$$

where  $\bar{\zeta}_j \geq 0$  for all  $j = 1, 2, \dots, q$  and  $\bar{\xi}_1, \dots, \bar{\xi}_p, \bar{\zeta}_1, \dots, \bar{\zeta}_q$  are not all zero.

Now, let  $i \in I^* = \{i : \bar{\xi}_i \neq 0\}$ , since  $\xi_i^k \rightarrow \bar{\xi}_i \neq 0$  as  $k \rightarrow \infty$ , therefore  $\xi_i^k \neq 0$ , and has the same sign as  $\bar{\xi}_i$  for sufficiently large  $l$ . Consequently, we have

$$\xi_i^l h_i(x^{k,m}) > 0, \quad (m = r, s)$$

$$\Rightarrow \bar{\xi}_i h_i(x^{k,m}) > 0, \quad \forall \text{ sufficiently large } k \text{ and } m, (m = r, s),$$

similarly for  $j \in J^* = \{j \mid \bar{\zeta}_j > 0\}$ , we have

$$\bar{\zeta}_j g_j(x^{k,m}) > 0, \quad (m = r, s),$$

also similar to the proof of Theorem 3.1, using density theorem we can find subsequences

$$\{x^{k,r_k}\} \subset \{x^{k,r}\} \subset \mathcal{C}$$

and  $(\{x^{k,s_k}\} \subset \{x^{k,s}\} \subset \mathcal{C})$  converging to  $\bar{x}$  as  $k \rightarrow \infty$  such that for all sufficiently large  $k$ , we have

$$\bar{\xi}_i h_i(x^{k,m_k}) > 0, \quad \forall i \in I^*, \quad (m = r, s),$$

$$\bar{\zeta}_j g_j(x^{k,m_k}) > 0, \quad \forall j \in J^*, \quad (m = r, s),$$



and  $h_i(x^{k,r_k})$ , ( $i \in I^*$ ),  $g_j(x^{k,r_k})$ , ( $j \in J^*$ ) ( $h_i(x^{k,s_k})$ , ( $i \in I^*$ ),  $g_j(x^{k,s_k})$ , ( $j \in J^*$ )) are proximal subdifferentiable at  $x^{k,r_k}$  (differentiable at  $x^{k,s_k}$ ). But this violates the  $\partial^*$ -quasinormality at  $\bar{x}$ , hence the sequence  $\{\lambda_1^k, \dots, \lambda_p^k, \mu_1^k, \dots, \mu_q^k\}$  must be bounded.

**Step III.** Now without loss of generality, we assume that  $(\lambda_1^k, \dots, \lambda_p^k, \mu_1^k, \dots, \mu_q^k)$  converges to  $(\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q)$  as  $k \rightarrow \infty$  and since  $F$  is smooth, it admits a continuous convexificator, then taking limit in (4.7) we have

$$\begin{aligned} -\overline{co}\partial^r F(\bar{x}) &\in \sum_{i=1}^p \overline{co}\partial^r(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^q \mu_j \overline{co}\partial^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}) \\ \Rightarrow y &\in \sum_{i=1}^p \overline{co}\partial^r(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^q \mu_j \overline{co}\partial^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}), \end{aligned}$$

hence,

$$\mathcal{T}_{\Omega}^0(\bar{x}) \subset \left\{ \sum_{i=1}^p \overline{co}\partial^r(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^q \mu_j \overline{co}\partial^r g_j(\bar{x}) + \mathcal{N}_{\mathcal{C}}(\bar{x}) \mid (\lambda, \mu) \in M(\bar{x}) \right\}.$$

Similar to step II, for sets  $I = \{i \mid \lambda_i \neq 0\}$  and  $J = \{j \mid \mu_j > 0\}$ , we can find subsequences

$$\{x^{k,r_k}\} \subset \{x^{k,r}\} \subset \mathcal{C},$$

and

$$\{x^{k,s_k}\} \subset \{x^{k,s}\} \subset \mathcal{C},$$

converging to  $\bar{x}$  and satisfying the inequalities same in step II and proximal subdifferentiability and differentiability notions hold respectively. □

**Remark 4.12.** If  $\Omega$  is regular, then the result is true for limiting normal cone instead of Fréchet normal cone and again we can say that this result is sharper than Proposition 4 of [27], because it provides a set for limiting normal cone to be contained in, which is sharper than established in the literature.

From Theorems 4.11 and 2.13, The improved enhanced KKT condition can be sharpened for the case, where the objective function is Fréchet differentiable, since the minimal convexificator in this case is the gradient of the function.

**Corollary 4.13.** *Suppose  $\bar{x}$  is a local minimizer of the problem (P) with Fréchet differentiable objective function at  $\bar{x}$ . If  $\bar{x}$  satisfies one of the NNAMCQ,  $\partial^*$ -Pseudonormality or  $\partial^*$ -Quasinormality, then the improved version of enhanced KKT condition holds with*

$$\overline{co} \partial^r f(\bar{x}) = \{\nabla f(\bar{x})\}.$$

### 5. CONCLUSION

In this paper, we have used convexificators in place of limiting subdifferentials to study a mathematical program involving equality, inequality and an abstract set constraint under locally Lipschitz assumption. Although limiting subdifferential of a locally Lipschitz continuous function becomes convexificator, but there may be some other convexificators smaller than limiting subdifferential. Significance of our results can be seen for those convexificators.

Due to the involvement of convexificators, the obtained result becomes sharper, because our result does not depend on the whole limiting subdifferential, but the convex hull of some few points of it to check the necessary optimality condition. we have also analyzed the pseudonormality and the quasinormality in terms of convexificators as  $\partial^*$ -pseudonormality and  $\partial^*$ -quasinormality, which are sharper than the previously established in the literature. Later we have generalized some results from [27] by using  $\partial^*$ -pseudonormality and  $\partial^*$ -quasinormality, which are stronger than [27].

## REFERENCES

- [1] C.S. Bector, S. Chandra and J. Dutta, Principle of Optimization Theory. Narosa Publishers, India and Alpha Science Publishers (2004).
- [2] D.P. Bertsekas, Nonlinear Programming. Athena Scientific Publishers (1999).
- [3] D.P. Bertsekas and A.E. Ozdaglar, Pseudonormality and Lagrange multiplier theory for constrained optimization. *J. Optim. Theory App.* **114** (2002) 287–343.
- [4] D.P. Bertsekas, A. Nedić and A.E. Ozdaglar, Convex Analysis and Optimization. Athena Scientific, Belmont, MA (2003).
- [5] F.H. Clarke, A new approach to Lagrange multipliers. *Math. Oper. Res.* **1** (1976) 165–174.
- [6] F.H. Clarke, Optimization and Nonsmooth Analysis. Wiley Interscience (1983); reprinted as Vol. 5 of *Classics Appl. Math.* SIAM J. Optim., Philadelphia, PA (1990).
- [7] V.F. Demyanov, Convexification and Concavification of Positively Homogeneous Functions by the Same Family of Linear Functions. Technical Report, University of Pisa (1994) 1–11.
- [8] J. Dutta, Generalized derivatives and nonsmooth optimization, a finite dimensional tour. *Soc. Estadística Invest. Oper.* **13** (2005) 185–314.
- [9] J. Dutta and S. Chandra, Convexifactors, generalized convexity and optimality conditions. *J. Optim. Theory App.* **113** (2002) 41–64.
- [10] R. Henrion, A. Jourani and J. Outrata, On the calmness of a class of multifunctions. *SIAM J. Optim.* **13** (2002) 603–618.
- [11] M.R. Hestenes, Optimization Theory: The Finite Dimensional Case. Wiley, New York, NY (1975).
- [12] A.D. Ioffe, Approximate subdifferentials and applications II. *Mathematika* **33** (1986) 111–128.
- [13] V. Jeyakumar and D.T. Luc, Nonsmooth calculus, minimality, and monotonicity of convexificators. *J. Optim. Theory App.* **101** (1999) 599–621.
- [14] F. John, Extremum problems with inequalities as side constraints. In: Studies and Essays. Courant Anniversary Volume. Wiley (interscience) (1948) 187–204.
- [15] A. Jourani, Constraint qualification and Lagrange multipliers in nondifferentiable programming problems. *J. Optim. Theory App.* **81** (1994) 533–548.
- [16] X.F. Li and J.Z. Zhang, Necessary optimality conditions in terms of convexificators in lipschitz optimization. *J. Optim. Theory App.* **131** (2006) 429–452.
- [17] O.L. Mangasarian and S. Fromovitz, The Fritz john necessary optimality conditions in the presence of equality and inequality constraints. *J. Math. Anal. Appl.* **17** (1967) 37–47.
- [18] E.J. McShane, The Lagrange multiplier rule. *Am. Math. Monthly.* **80** (1973) 922–925.
- [19] P. Michel and J.P. Penot, A generalized derivative for calm and stable functions. *Differ. Integral Equ.* **5** (1992) 433–454.
- [20] B.S. Mordukhovich, Metric approximation and necessary optimality condition for general classes of nonsmooth extremal problems. *Sov. Math. Dokl.* **22** (1980) 526–530.
- [21] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I: Basic Theory. Springer, New York (2006).
- [22] B.S. Mordukhovich and Y. Shao, On non-convex subdifferential calculus in Banach spaces. *J. Convex Anal.* **2** (1995) 211–228.
- [23] B.S. Mordukhovich, N.M. Nam and N.D. Yen, Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming. *Optimization* **55** (2006) 685–708.
- [24] A.E. Ozdaglar and D.P. Bertsekas, The relation between pseudonormality and quasiregularity in constrained optimization. *Optim. Methods Softw.* **19** (2004) 493–506.
- [25] J.S. Treiman, The Linear nonconvex generalized gradient and Lagrange multipliers. *SIAM J. Optim.* **5** (1995) 670–680.
- [26] X. Wang and V. Jeyakumar, A sharp langrage multiplier rule for nonsmooth mathematical programming problems involving equality constraints. *SIAM J. Optim.* **10** (2000) 1136–1148.
- [27] J.J. Ye and J. Zhang, Enhanced Karush–Kuhn–Tucker condition and weaker constraint qualifications. *Math. Program. Ser. B* **139** (2013) 353–381.