Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

An Interior Subgradient Method for DC Programming with Proximal Distance Regularization

Valdinês L. S. Júnior[1]
UFCA, Juazeiro do Norte, CE

Abstract. This paper introduces an interior subgradient algorithm designed to tackle a specific category of nonconvex minimization problems, specifically focusing on minimizing DC functions (difference of two convex functions). The algorithm was inspired by the interior gradient method proposed Auslender and Teboulle in [1].

Keywords. DC Programming, Proximal Distance, Nonconvex Optimization, Subgradient Methods

1 Introduction

Difference of convex functions (DC) programming is a powerful optimization framework gaining traction in recent decades. By exploiting the structure of convexity, DC programming enables the development of efficient algorithms for solving nonconvex optimization tasks. Its applications span various domains including engineering, economics, and machine learning. Notable advancements in theory and algorithms have propelled its growth, making it a promising avenue for tackling challenging optimization problems with practical relevance.

Recently, some authors have been proposed some algorithms and numerical experiments do study DC optimization problems in a lot of settings, see [5, 6, 8, 11–13]. To continue this path, we present an interior subgradient method for DC programming. Our method was inspired by the interior gradient methods presented in [1]. In our case, the method is applied for a DC function instead of a convex function. We prove that every accumulation point of its generated sequences, if any, is a critical point of a DC function over a nonempty, closed and convex set. Furthermore, with some additional assumptions, the whole sequence converges to a critical point of a DC function.

The organization of the paper is as follows. In Section 2, some notations and basic results used throughout the paper are presented. In Section 3, the two algorithms studied in this paper are presented and the main results are stated and proved. Some final remarks are made in Section 4.

2 Preliminaries

In this section, we present several concepts of non-smooth analysis that will be useful throughout this paper. The subdifferential of a convex lower semicontinuous (lsc) function $f : \mathbb{R}^n \to \mathbb{R}$ at x, is defined by $\partial f(x) = \{v \in \mathbb{R}^n : f(y) \geq f(x) + \langle v, y - x\rangle, \forall y \in \mathbb{R}^n\}$. If f is strongly convex with modulus $\rho > 0$, it is well known that, for all $v \in \partial f(x)$, $f(y) \geq f(x) + \langle v, y - x\rangle + \rho/2\|y - x\|^2$, for all $y \in \mathbb{R}^n$, and its subdifferential ∂f is strongly monotone with modulus ρ, i. e., for any $x, y \in \mathbb{R}^n, v \in \partial f(y)$ and $u \in \partial f(x)$, we have $\langle v - u, y - x\rangle \geq \rho\|y - x\|^2$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a

[1] valdines.leite@ufca.edu.br

locally Lipschitz function at $\bar{x} \in \mathbb{R}^n$ with constant $L > 0$ and $v \in \mathbb{R}^n$. The *Clarke's directional derivative* [4, page 25] of f at $\bar{x}$ in the direction v, denoted by $f^\circ(\bar{x}; v)$, is defined as

$$f^\circ(\bar{x}; v) := \limsup_{t \downarrow 0 \ y \to \bar{x}} \frac{f(y + tv) - f(y)}{t},$$

and *Clarke's subdifferential* [4, page 27] of f at $\bar{x}$, denoted by $\partial^\circ f(\bar{x})$, is defined as $\partial^\circ f(\bar{x}) := \{w \in \mathbb{R}^n \ : \ f^\circ(\bar{x}; v) \geq \langle w, v\rangle, \ \ v \in \mathbb{R}^n\}$. If f is convex, the Clarke's subdifferential coincides with the classical subdifferential ∂f.

A lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R}$, is called a DC function when there exist convex functions g and h such that,

$$f(x) = g(x) - h(x), \forall x \in \mathbb{R}^n. \tag{1}$$

The functions g and h are commonly called *components functions of* f. It is well known that a necessary condition for $x \in \mathbb{R}^n$ to be a local minimizer of a DC function f is $\partial h(x) \subset \partial g(x)$. In general, this condition is hard to be reached, often such condition is replaced by a relaxed one, namely points that satisfies $\partial g(x) \cap \partial h(x) \neq \emptyset$. Motivated by this condition and for the other definitions of critical points for constrained problems, we have the following definition.

Definition 2.1. *Let D be a closed and convex set and $f : \mathbb{R}^n \to \mathbb{R}$ be a DC function as in* (1)*. We say that a point $x^* \in D$ is a critical point of f in D if, there exist $v \in \partial g(x^*)$ and $u \in \partial h(x^*)$ such that $\langle v-u, y-x^*\rangle \geq 0$, for all $y \in D$. The set $\mathcal{S}^*_D(f)$ denotes the set of the critical points of f in D. Furthermore, we say that a point $x^* \in D$ is a Clarke-critical point of f in D if, $f^\circ(x^*; y - x^*) \geq 0$, for all $y \in D$. We denote by $\mathcal{S}^\circ_D(f)$, the set of the Clarke-critical points of f in D.*

In the case of $D = \mathbb{R}^n$, the definition of a critical point retrieves to the well-knwon condition $\partial g(x^*) \cap \partial h(x^*) \neq \emptyset$. Furthermore, if h is continuously differentiable, we have $\mathcal{S}^*_D(f) \subset \mathcal{S}^\circ_D(f)$, see [4, Corollary 1, page 39].

In our approach, we choose a proximal distance $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ as the regularization term. Such a well-known distance allows us to analyze the convergence of the algorithm under various settings. Following [1], let us recall the definition of the proximal and induced proximal distances.

Definition 2.2. *A function $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ is called a proximal distance with respect to a nonempty convex open set $C \subset \mathbb{R}^n$ if for each $y \in C$ it satisfies the following properties:*

(d1) *$d(\cdot, y)$ is proper, lsc, convex, and continuously differentiable on C;*

(d2) *dom $d(\cdot, y) \subset \bar{C}$ and dom $\partial_1 d(\cdot, y) = C$, where $\partial_1 d(\cdot, y)$ denotes the subgradient map of the function $d(\cdot, y)$ with respect to the first variable, where $\bar{C}$ denotes the closure of C;*

(d3) *$d(\cdot, y)$ is level bounded on $\mathbb{R}^n$, i.e., $\lim_{\|u\| \to +\infty} d(u, y) = +\infty$;*

(d4) *$d(y, y) = 0$.*

For each $y \in C$, let $\nabla_1 d(\cdot, y)$ denote the gradient map of the function $d(\cdot, y)$ with respect to the first variable. Note that by definition $d(\cdot, \cdot) \geq 0$, and from **(d3)** the global minimum of $d(\cdot, y))$ is obtained at y, which shows that $\nabla_1 d(y, y) = 0$. We denote by $\mathcal{D}(C)$ the family of functions d satisfying **(d1)-(d4)**.

Next, following the approach presented in [1], we associate to a given $d \in \mathcal{D}(C)$ a corresponding induced distance H that satisfies some desirable properties.

Definition 2.3. *Given $C \subset \mathbb{R}^n$, open and convex, and $d \in \mathcal{D}(C)$, a function $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ is called the induced proximal distance to d if H is finite valued on $C \times C$ and for each $x, y \in C$ satisfies the following properties:*

(H1) $H(x,x) = 0$;

(H2) $\langle z - y, \nabla_1 d(y,x)\rangle \leq H(z,x) - H(z,y)$, $z \in C$.

We write $(d,H) \in \Phi(C)$ to quantify the triple $[C,d,H]$ that satisfies the premises of Definition 2.3. Similarly, we write $(d,H) \in \Phi(\bar{C})$ for the triple $[\bar{C},d,H]$ whenever there exists H, which is finite valued on $\bar{C} \times C$, satisfies **(H1)-(H2)** for any $z \in C$, and is such that $z \in \bar{C}$ has $H(z,\cdot)$ level bounded on C. Clearly, one has $\Phi(\bar{C}) \subset \Phi(C)$. For examples and a thorough discussion about proximal and induced proximal distances see, for instance, [1, 3].

3 The Algorithm and Convergence Analysis

Let $C \subset \mathbb{R}^n$ be an open nonempty convex set. From now on, $f : \mathbb{R}^n \to \mathbb{R}$ is a lower semicontinuous bounded below DC function and $g, h : \mathbb{R}^n \to \mathbb{R}$ are lower semicontinuous and convex functions such that $f(x) = g(x) - h(x)$. In addition, in all further results, assume that $(d,H) \in \Phi_+(C)$. To solve the problem of finding a critical point of f on $\bar{C}$, we will study the following algorithm:

Algorithm 1

Let $\lambda_k > 0$, $k \in \mathbb{N}$. Start from a point $x^0 \in C$ and generates a sequence $\{x^k\} \subset C$ satisfying

$$v^k \in \partial g(x^k), \quad w^k \in \partial h(x^k),$$

$$x^{k+1} \in \operatorname{argmin}\left\{\lambda_k \langle v^k - w^k, z\rangle + d(z,x^k) \mid z \in C\right\}.$$

The existence of $\{x^k\} \subset C$ of Algorithm 1 is guaranteed by using similar arguments as in the proof of [1, Proposition 2.1]. Then, from optimality conditions, we obtain

$$\lambda_k(v^k - w^k) + \nabla_1 d(x^{k+1}, x^k) = 0, \quad k \in \mathbb{N}. \tag{2}$$

Next we present an important result to our convergence analysis.

Proposition 3.1. *Set $\beta_k := \langle \nabla_1 d(x^{k+1}, x^k), x^{k+1} - x^k\rangle$ and assume that h is strongly convex with modulus ρ. Then for all $k \in \mathbb{N}$, $\beta_k \geq 0$. Furthermore, assume that there exists a positive constant κ satisfying*

$$\partial g(x) \subset \partial g(y) + \kappa\|x - y\|\mathbb{B} \quad \forall x, y \in C, \tag{3}$$

where $\mathbb{B}$ denotes the closed unit ball in $\mathbb{R}^n$. Then,

$$f(x^k) - f(x^{k+1}) \geq (\rho - \kappa)\left\|x^k - x^{k+1}\right\|^2 + \frac{\beta_k}{\lambda_k}, \quad k \in \mathbb{N}. \tag{4}$$

Besides, if $\lambda_k \leq \lambda^+$, $k \in \mathbb{N}$ and $\rho > \kappa$, we have $\sum_k \beta_k < \infty$.

Proof. From **(H2)**, with $z = x = x^k$, $y = x^{k+1}$, and taking into account that $H(x^k, x^k) = 0$, we obtain

$$H(x^k, x^{k+1}) \leq \langle x^{k+1} - x^k, \nabla_1 d(x^{k+1}, x^k)\rangle, \quad k \in \mathbb{N}. \tag{5}$$

Since $H(x^k, x^{k+1}) \geq 0$, we have that $\beta_k \geq 0$, $k \in \mathbb{N}$. Now, let us prove (4). First, in view of (2), we have

$$\lambda_k(v^k - w^k) + \nabla_1 d(x^{k+1}, x^k) = 0, \quad k \in \mathbb{N}. \tag{6}$$

Since $x^k \in C$ for all $k \geq 0$, we can use (3), to obtain

$$\partial g(x^k) \subset \partial g(x^{k+1}) + \kappa \left\|x^k - x^{k+1}\right\| \mathbb{B}, \quad k \in \mathbb{N}.$$

Taking into account that $v^k \in \partial g(x^k)$, last inclusion implies that there exist $u^k \in \partial g(x^{k+1})$ and $b^k \in \mathbb{B}$ satisfying

$$v^k = u^k + \kappa \left\|x^k - x^{k+1}\right\| b^k, \quad k \in \mathbb{N}. \tag{7}$$

From convexity of g,

$$g(x^k) \geq g(x^{k+1}) + \langle u^k, x^k - x^{k+1}\rangle, \quad k \in \mathbb{N}.$$

Now, combining last inequality with (7), we obtain

$$g(x^k) \geq g(x^{k+1}) + \langle v^k, x^k - x^{k+1}\rangle - \kappa \left\|x^k - x^{k+1}\right\| \langle b^k, x^k - x^{k+1}\rangle, \quad k \in \mathbb{N}.$$

Consequently, from (6), we have

$$\begin{aligned} g(x^k) \geq g(x^{k+1}) &- \frac{1}{\lambda_k}\langle \nabla_1 d(x^{k+1}, x^k), x^k - x^{k+1}\rangle \\ &+ \langle w^k, x^k - x^{k+1}\rangle - \kappa \left\|x^k - x^{k+1}\right\| \langle b^k, x^k - x^{k+1}\rangle, \quad k \in \mathbb{N}. \end{aligned}$$

Since $\beta_k = \langle \nabla_1 d(x^{k+1}, x^k), x^{k+1} - x^k\rangle$, we have

$$\begin{aligned} g(x^k) \geq g(x^{k+1}) &+ \langle w^k, x^k - x^{k+1}\rangle \\ &- \kappa \left\|x^k - x^{k+1}\right\| \langle b^k, x^k - x^{k+1}\rangle + \frac{\beta_k}{\lambda_k}, \quad k \in \mathbb{N}. \end{aligned}$$

On the other hand, as h is strongly convex with modulus $\rho > 0$, we have

$$h(x^{k+1}) \geq h(x^k) + \langle w^k, x^{k+1} - x^k\rangle + \rho \left\|x^k - x^{k+1}\right\|^2, \quad k \in \mathbb{N}. \tag{8}$$

Then, we obtain

$$\begin{aligned} g(x^k) \geq g(x^{k+1}) &+ h(x^k) - h(x^{k+1}) - \kappa \left\|x^k - x^{k+1}\right\| \langle b^k, x^k - x^{k+1}\rangle \\ &+ \rho\|x^k - x^{k+1}\|^2 + \frac{\beta_k}{\lambda_k}, \quad k \in \mathbb{N}. \end{aligned}$$

Using Cauchy–Schwartz inequality, we have

$$g(x^k) - h(x^k) \geq g(x^{k+1}) - h(x^{k+1}) + (\rho - \kappa) \left\|x^k - x^{k+1}\right\|^2 + \frac{\beta_k}{\lambda_k}, \quad k \in \mathbb{N}.$$

Finally, since $f(x) = g(x) - h(x)$ we obtain (4). The fact that $\sum \beta_k < \infty$, follows immediately from (4), and from the fact that f is bounded below. □

Remark 3.1. *Assumption* (3) *refers to the Lipschitz property of subdifferential* ∂g. *This is a commonly used condition in gradient-type algorithms.*

The forthcoming result is crucial to our work.

Lemma 3.1. *Let* $\{\lambda_k\}$ *be a sequence of positive numbers,* $\{a_k\}$ *a sequence of real numbers, and* $b_n := \sigma_n^{-1} \sum_{k=1}^n \lambda_k a_k$, *where* $\sigma_n := \sum_{k=1}^n \lambda_k$. *If* $\sigma_n \to \infty$, $\liminf a_n \leq \liminf b_n \leq \limsup b_n \leq \limsup a_n$.

Proof. See Lemaire [7], Lemma 3.5, and also [2], Lemma 2.3. □

DOI: 10.5540/03.2025.011.01.0340

Theorem 3.1. *Under all the assumptions of Proposition 3.1, suppose that $\lambda_k \geq \lambda_- > 0$, $k \in \mathbb{N}$. If $\{x^k\}$ is generated by Algorithm 1, its accumulation points, if any, are critical points of f in $\bar{C}$.*

Proof. Let $\bar{x}$ be an accumulation point of $\{x^k\}$ and let $\{x^{k_j}\}$ a subsequence of $\{x^k\}$ such that $\lim_{j\to\infty} x^{k_j} = \bar{x}$. From Algorithm 1, $v^{k_j} \in \partial g(x^{k_j})$ and $w^{k_j} \in \partial h(x^{k_j})$. Now, due do (4), we obtain that $\lim_{j\to\infty} x^{k_j+1} = \bar{x}$. Then, we can use [10, Theorem 9.13] and, without loss of generality, we can assume that $\{v^{k_j}\}$ and $\{w^{k_j}\}$ converge to $\bar{v}$ and $\bar{w}$ respectively. Now, consider any $y_0 \in \bar{C}$ fixed. Based on (2),

$$\langle v^k - w^k, y_0 - x^k \rangle = -\frac{1}{\lambda_k} \langle \nabla_1 d(x^{k+1}, x^k), y_0 - x^k \rangle, \quad k \in \mathbb{N}.$$

Then, tanking into account that $\beta_k := \langle \nabla_1 d(x^{k+1}, x^k), x^{k+1} - x^k \rangle$, we have

$$\langle v^k - w^k, y_0 - x^k \rangle = -\frac{1}{\lambda_k} \langle \nabla_1 d(x^{k+1}, x^k), y_0 - x^{k+1} \rangle - \frac{\beta_k}{\lambda_k}, \quad k \in \mathbb{N}. \tag{9}$$

From **(H2)**, with $z = y_0, y = x^{k+1}, x = x^k$, we obtain

$$\langle y_0 - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle \leq H(y_0, x^k) - H(y_0, x^{k+1}), \quad k \in \mathbb{N}. \tag{10}$$

Combining the last inequality with (9), for all $k \geq 0$ we obtain

$$H(y_0, x^k) - H(y_0, x^{k+1}) + \beta_k \geq -\lambda_k \langle v^k - w^k, y_0 - x^k \rangle.$$

Summing the last inequality over $k = 1, \ldots, n$, for all $k \geq 0$ we have

$$H(y_0, x^1) - H(y_0, x^{n+1}) + \sum_{k=1}^{n} \beta_k \geq \sum_{k=1}^{n} \lambda_k \left(-\langle v^k - w^k, y_0 - x^k \rangle \right).$$

Since $H(\cdot, \cdot) \geq 0$, for all $k \geq 0$ we obtain

$$\sigma_n^{-1} H(y_0, x^1) + \sigma_n^{-1} \sum_{k=1}^{n} \beta_k \geq \sigma_n^{-1} \sum_{k=1}^{n} \lambda_k \left(-\langle v^k - w^k, y_0 - x^k \rangle \right),$$

where $\sigma_n := \sum_{k=1}^{n} \lambda_k$. As $\lambda_k \geq \lambda_-$, then $\sigma_n \to \infty$, and considering that $\sum_{k=1}^{\infty} \epsilon_k < \infty$, we can use Lemma 3.1 to obtain $\limsup_{k\to+\infty} \langle v^k - w^k, y_0 - x^k \rangle \geq 0$. Since ∂g and ∂h are closed, we obtain $\bar{v} \in \partial g(\bar{x})$ and $\bar{w} \in \partial h(\bar{x})$. Thus, last inequality imples that $\langle \bar{v} - \bar{w}, y_0 - \bar{x} \rangle \geq 0$, for all $y_0 \in \bar{C}$. Therefore $\bar{x} \in \mathcal{S}^*_{\bar{C}}(f)$. □

Lemma 3.2. *Under all the assumptions of Proposition 3.1, suppose furthermore that g is strongly convex with modulus $\gamma > 0$, h is continuously differentiable and ∇h is L–Lipschitz continuous on C. Consider any $\bar{x} \in \mathcal{S}^*_{\bar{C}}(f)$. Then the following hold:*

$$H(\bar{x}, x^{k+1}) + \lambda_k(\gamma - L)\|x^k - \bar{x}\|^2 \leq H(\bar{x}, x^k) + \beta_k, \quad k \in \mathbb{N}. \tag{11}$$

Proof. Take any $\bar{x} \in \mathcal{S}^*_{\bar{C}}(f)$ and let $v \in \partial g(\bar{x})$ be such that, for all $y \in C$,

$$\langle v - \nabla h(\bar{x}), y - \bar{x} \rangle \geq 0.$$

Let us prove (11). Since $\{x^k\} \subset C$, we obtain $\langle v, x^k - \bar{x} \rangle \geq \langle \nabla h(\bar{x}), x^k - \bar{x} \rangle$, $k \in \mathbb{N}$. Since g is strongly convex with modulus γ, we have $\gamma \|x^k - \bar{x}\|^2 \leq \langle x^k - \bar{x}, v^k - v \rangle$, $k \in \mathbb{N}$. Consequently, $\gamma \|x^k - \bar{x}\|^2 \leq \langle x^k - \bar{x}, v^k - \nabla h(\bar{x}) \rangle$, $k \in \mathbb{N}$. From (2),

$$\gamma \|x^k - \bar{x}\|^2 \leq -\frac{1}{\lambda_k} \langle x^k - \bar{x}, \nabla_1 d(x^{k+1}, x^k) \rangle + \langle x^k - \bar{x}, \nabla h(x^k) - \nabla h(x) \rangle, \quad k \in \mathbb{N}.$$

Taking into account that $\beta_k = \langle \nabla_1 d(x^{k+1}, x^k), x^{k+1} - x^k \rangle$, for all $k \in \mathbb{N}$,

$$\gamma \|x^k - \bar{x}\|^2 \leq \frac{\beta_k}{\lambda_k} - \frac{1}{\lambda_k} \langle x^{k+1} - \bar{x}, \nabla_1 d(x^{k+1}, x^k) \rangle + \langle x^k - \bar{x}, \nabla h(x^k) - \nabla h(\bar{x}) \rangle.$$

As ∇h is L–Lipschitz continuous on C and using Cauchy-Schwarz, we obtain

$$\gamma \|x^k - \bar{x}\|^2 \leq \frac{\beta_k}{\lambda_k} + \frac{1}{\lambda_k} \langle \bar{x} - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle + L\|x^k - \bar{x}\|^2, \quad k \in \mathbb{N}.$$

Again using **(H2)**, with $z = \bar{x}, y = x^{k+1}, x = x^k$, we have

$$\langle \bar{x} - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle \leq H(\bar{x}, x^k) - H(\bar{x}, x^{k+1}), \quad k \in \mathbb{N}. \tag{12}$$

Hence, we can combine the last two inequalities to obtain (11). □

To set the convergence of any sequence generated by Algorithm 1, we need to make further assumptions on the induced proximal distance H, which were also considered in [1]. Let $(d, H) \in \Phi_+(\bar{C}) \subset \Phi(\bar{C})$ be such that the function H satisfies the following two additional properties: For $y \in \bar{C}$ and $\{y^k\} \subset C$,

(Ha) $\lim_{k\to+\infty} y^k = y$, whenever $\{y^k\}$ is bounded and $\lim_{k\to+\infty} H(y, y^k) = 0$;

(Hb) $\lim_{k\to+\infty} H(y, y^k) = 0$, whenever $\lim_{k\to+\infty} y^k = y$.

We also make the following assumption:

$$\mathcal{S}^*_{\bar{C}}(f) \neq \emptyset. \tag{13}$$

Under these assumptions, we prove that Algorithm 1 converges to a Clarke critical of f.

Before we introduce the main result of the present work, we recall the following well-known result of nonnegative sequences.

Lemma 3.3. *Let $\{u^k\}$, $\{\alpha_k\}$, and $\{\beta_k\}$ be nonnegative sequences of real numbers satisfying $u^{k+1} \leq (1+\alpha_k)u^k + \beta_k$ such that $\sum_k \alpha_k < \infty$ and $\sum_k \beta_k < \infty$. Then, the sequence $\{u^k\}$ converges.*

Proof. See Lemma 2 on page 44 by Polyak [9]. □

Theorem 3.2. *Under all assumptions of Lemma 3.2, suppose that $0 < \lambda_- \leq \lambda_k \leq \lambda^+$, $k \in \mathbb{N}$ and $\gamma > L + 1/2$. If $\{x^k\}$ is generated by Algorithm 1, then it converges to a Clarke critical point of f in $\bar{C}$.*

Proof. In view of (13), take any $x \in \mathcal{S}^*_{\bar{C}}(f)$. As $\gamma > L+1/2$, Lemma 3.2 implies that, $H(x, x^{k+1}) \leq H(x, x^k) + \beta_k$, and $H(x, x^{k+1}) \leq H(x, x^k) + \alpha_k$, for all $k \in \mathbb{N}$. As $\sum_k \alpha_k < \infty$, and thanks to Proposition 3.1 (ii), in both cases, we can apply Lemma 3.3 we conclude that $\{H(x, x^k)\}$ converges to some point $\beta(x)$. Let x^* be an accumulation point of $\{x^k\}$. From Theorem 3.1, $x^* \in \mathcal{S}^*_{\bar{C}}(f)$. Based on **(Ha)**, we obtain $\lim_{\ell\to+\infty} H(x^*, x^{k_\ell}) = 0$. Considering that $\{H(x, x^k)\}$ converges, we conclude that $\lim_{k\to+\infty} H(x^*, x^k) = 0$. Now, by **(Hb)** it follows that $\{x^k\}$ converges to x^*. Therefore, from Theorem 3.1, x^* is a Clarke critical point of f in $\bar{C}$, which proves the theorem. □

4 Final Remarks

In this paper, we present an interior subgradient and a proximal linearized method for DC programming, whose regularization term is a proximal distance. Based on the methods presented in [1], we prove that any accumulation point of a sequences generated by the algorithm is a critical point, where the strong convexity of one of the components of the main function played a vital role in this analysis. Finally, with some additional assumptions, we prove that the whole sequence generated by the method converges to a Clarke-critical point. Our motivation in this paper is to increase the interest in this research field, intending that these results will inspire others to seek knowledge and make science stronger. In future research, we intend to investigate this kind of problem in more general settings as in Riemann Manifolds and Multi-objective Optimization. We foresee further progress in this topic shortly.

References

[1] A. Auslender and M. Teboulle. "Interior gradient and proximal methods for convex and conic optimization". In: **SIAM Journal on Optimization** 16.3 (2006), pp. 697–725.

[2] M. A. Bahraoui. "Diagonal Bundle Methods for Convex Minimization, Theoritical Aspect". In: **Int. J. Math. And Appl** 6.4 (2018).

[3] R. Burachik and J. Dutta. "Inexact proximal point methods for variational inequality problems". In: **SIAM Journal on Optimization** 20.5 (2010), pp. 2653–2678.

[4] F. H. Clarke. **Optimization and nonsmooth analysis**. SIAM, 1990.

[5] J. X. Cruz Neto, J. O. Lopes, P. S. M. Santos, and J. C. O. Souza. "An interior proximal linearized method for DC programming based on Bregman distance or second-order homogeneous kernels". In: **Optimization** 68.7 (2019), pp. 1305–1319.

[6] J. X. Cruz Neto, P. R. Oliveira, A. Soubeyran, and J. C. O. Souza. "A generalized proximal linearized algorithm for DC functions with application to the optimal size of the firm problem". In: **Annals of Operations Research** 289 (2020), pp. 313–339.

[7] B. Lemaire. "On the convergence of some iterative methods for convex minimization". In: **Recent Developments in Optimization: Seventh French-German Conference on Optimization**. Springer. 1995, pp. 252–268.

[8] A. Moudafi and P.-E. Maingé. "On the convergence of an approximate proximal method for DC functions". In: **Journal of computational Mathematics** (2006), pp. 475–480.

[9] B. T. Polyak. "Introduction to optimization. Optimization software". In: **Inc., Publications Division, New York** 1.32 (1987), p. 1.

[10] R. Tyrrell Rockafellar and R. J.-B. Wets. **Variational analysis**. Vol. 317. Springer Science & Business Media, 2009.

[11] J. C. O. Souza and P. R. Oliveira. "A proximal point algorithm for DC functions on Hadamard manifolds". In: **Journal of Global Optimization** 63 (2015), pp. 797–810.

[12] J. C. O. Souza, Paulo Roberto Oliveira, and Antoine Soubeyran. "Global convergence of a proximal linearized algorithm for difference of convex functions". In: **Optimization Letters** 10.7 (2016), pp. 1529–1539.

[13] W. Sun, Raimundo J. B. Sampaio, and M. A. B. Candido. "Proximal point algorithm for minimization of DC function". In: **Journal of computational Mathematics** (2003), pp. 451–462.

DOI: 10.5540/03.2025.011.01.0340