TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC SƯ PHẠM TP HỒ CHÍ MINH
Tập 23, Số 7 (2026): 1616-1627
Website: https://journal.hcmue.edu.vn
ISSN: 2734-9918

HO CHI MINH CITY UNIVERSITY OF EDUCATION
JOURNAL OF SCIENCE
Vol. 23, No. 7 (2026): 1616-1627
https://doi.org/10.54607/hcmue.js.23.7.4935(2026)

Research Article

INEXACT BOOSTED DIFFERENCE OF CONVEX ALGORITHM

Bui Ho Kim Anh[1*], Pham Duy Khanh[1], Tran Trung Tin[1], Tran Ba Dat[2]

[1]Ho Chi Minh City University of Education, Vietnam

[2]Rowan University, The United States of America

[*]Corresponding author: Bui Ho Kim Anh – Email: 4801101004@student.hcmue.edu.vn

Received: April 28, 2025; Revised: May 20, 2025; Accepted: August 05, 2025

ABSTRACT

In this paper, we propose the Inexact Difference of Convex Algorithm (iBDCA), an inexact variant of the Boosted Difference of Convex Functions Algorithm (BDCA), to solve an optimization problem involving the difference of two strongly convex functions, where both components of the DC function are assumed to be differentiable, and the gradient of the first component is Lipschitz continuous. The proposed algorithm consists of two main steps: first, we employ a gradient descent-based scheme to find an approximate solution to a subproblem associated with the convex component of the DC problem; second, this approximate point is then used to compute a suitable step size for the next iteration. We establish the convergence and convergence rates of the iterate and objective-value sequences to the optimal solution and optimal value, respectively, as well as the complexity of the subproblem.

Keywords: boosted difference of convex functions algorithm; convergence analysis; DC functions; inexact gradient descent methods; line search; step size

1. Introduction

This paper focuses on minimizing *Difference of Convex* (DC) functions. We consider the problem

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\ \phi(x) := f_1(x) - f_2(x) \tag{1.1}$$

where both $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable and convex and to satisfy

$$\inf_{x\in\mathbb{R}^n}\ \phi(x) > -\infty. \tag{1.2}$$

To facilitate the subsequent analysis, we reformulate the problem (1.1) into an equivalent problem involving strongly convex functions. For any constant $\rho > 0$, define

Cite this article as: Bui, H. K. A., Pham, D. K., Tran, T. T., & Tran, B. D. (2026). Inexact boosted difference of convex algorithm. *Ho Chi Minh City University of Education Journal of Science, 23*(7), 1616-1627. https://doi.org/10.54607/hcmue.js.23.7.4935(2026)

$g(x) = f_1(x) + \frac{\rho}{2}\|x\|^2$ and $h(x) = f_2(x) + \frac{\rho}{2}\|x\|^2$. It follows immediately that both g and h are strongly convex functions with modulus ρ and $\phi(x) = f_1(x) - f_2(x) = g(x) - h(x)$, for all $x \in \mathbb{R}^n$. Consequently, the following equivalent problem is obtained

$$(\mathcal{P}) \quad \underset{x\in\mathbb{R}^n}{\text{minimize}}\ \phi(x) := g(x) - h(x).$$

DC programming has been successfully applied in various fields, including machine learning and image processing, biochemical modeling (Aragón-Artacho et al., 2024) and compressed sensing (Yin et al., 2015). The *Difference of Convex Algorithm* (DCA) was the first algorithm specifically designed for solving DC optimization problems (Tao, 1986; Tao & An, 1997). Because of its simplicity and computational efficiency, DCA, together with its variants, has been extensively applied to high-dimensional nonconvex optimization. Among those variants, *Boosted DC Algorithm* (BDCA) introduced in (Aragón-Artacho et al., 2018) is one of the most efficient methods because it combines DCA with a line search step inspired by Fukushima and Mine (1981). This combination enables BDCA to converge faster and require fewer iterations than DCA, particularly for large-scale problems.

To further improve the *computational efficiency* of BDCA, we introduce an inexact variant of the *Boosted Difference of Convex Algorithm* (BDCA) to solve (1.1), referred to as the *Inexact Boosted Difference of Convex Algorithm* (iBDCA). This modified approach establishes a framework that allows inexactness in solving subproblems while still maintaining convergence properties. At each $k-$th iteration, unlike DCA and BDCA, which require solving the following subproblem $(\mathcal{P}_k)$ exactly

$$(\mathcal{P}_k) \ \underset{x\in\mathbb{R}^n}{\text{minimize}}\ g(x) - \langle \nabla h(x_k), x\rangle \tag{1.3}$$

we propose to compute an approximate solution y_k that satisfies the inexact condition

$$\|\nabla g(y_k) - \nabla h(x_k)\| \le \frac{\rho}{2}\|y_k - x_k\|. \tag{1.4}$$

This approach reduces the computational cost of each iteration while preserving the method's efficiency. We show that the convergence results of BDCA are preserved when we employ this inexact condition. We also provide a simple approach for finding y_k satisfying (1.4) using the *gradient descent method*. We also specify the number of iterations required for each subproblem to determine y_k.

The remainder of the paper is organized as follows. Section 2 presents the necessary definitions and preliminary results used in the subsequent analysis. The proposed iBDCA is then presented, and its convergence properties are analyzed in Section 3. Conclusions and directions for further research are discussed in Section 4.

2. Preliminaries and Notation

Throughout this paper, the *inner product* of two vectors $x, y \in \mathbb{R}^n$ is denoted by $\langle x, x \rangle$, while $\|\cdot\|$ denotes the *induced norm*, defined by $\|x\| = \sqrt{\langle x, x \rangle}$. We denote by $\mathbb{R}_+ = [0, \infty)$ the set of all nonnegative real numbers and $\mathbb{B}(x, r)$ the *closed ball* centered at x with radius $r > 0$. The *gradient* of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ at some point $x \in \mathbb{R}^n$ is denoted by $\nabla f(x) \in \mathbb{R}^n$.

We now recall some important definitions and results related to the convexity and monotonicity of a function. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* when

$$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y)$$

for all $\lambda \in (0,1)$. Furthermore, f is said to be *strongly convex* with modulus $\sigma > 0$ if

$$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y) - \frac{1}{2}\sigma\lambda(1-\lambda)\|x-y\|^2$$

for all $\lambda \in (0,1)$, or equivalently, when $f - \frac{\sigma}{2}\|\cdot\|^2$ is convex.

The following fundamental results on convex functions are useful in the next section.

Theorem 2.1. Let f be a differentiable function on $\mathbb{R}^n$. Then:

1. f is convex if and only if

$$f(x) \ge f(y) + \langle \nabla f(y), x - y \rangle \text{ for all } x, y \in \mathbb{R}^n; \tag{2.1}$$

2. f is strongly convex with modulus σ if and only if

$$f(x) \ge f(y) + \langle \nabla f(y), x - y \rangle + \frac{\sigma}{2}\|x-y\|^2 \text{ for all } x, y \in \mathbb{R}^n. \tag{2.2}$$

Definition 2.2. A mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be monotone if

$$\langle F(x) - F(y), x - y \rangle \ge 0 \text{ for all } x, y \in \mathbb{R}^n.$$

Furthermore, F is called strongly monotone with modulus σ if

$$\langle F(x) - F(y), x - y \rangle \ge \sigma\|x-y\|^2 \text{ for all } x, y \in \mathbb{R}^n.$$

The following proposition describes the relationship between the concepts of (strong) convexity and (strong) monotonicity when f is differentiable.

Proposition 2.3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then f is (strongly) convex if and only if ∇f is (strongly) monotone.

To establish our convergence results, we use the following definitions and theorems.

Definition 2.4. A mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is called Lipschitz continuous if there is some constant $L \ge 0$ such that

$$\left\|F\left(x\right)-F\left(y\right)\right\|\le L\left\|x-y\right\|,\quad \text{for all } x,y\in\mathbb{R}^n.$$

F is called locally Lipschitz continuous if for every $x\in\mathbb{R}^n$, there exists a neighborhood U of x such that F restricted to U is Lipschitz continuous.

The following result from Izmailov and Solodov (2014) provides a useful property of real-valued functions with Lipschitz continuous gradients.

Lemma 2.5. Let $f:\mathbb{R}^n\to\mathbb{R}$, let $x,y\in\mathbb{R}^n$, and let $L>0$. If f is differentiable on the line segment $\left[x,y\right]$ with its gradient being Lipschitz continuous with constant L on this segment, then

$$\left|f\left(y\right)-f\left(x\right)-\left\langle\nabla f\left(x\right),y-z\right\rangle\right|\le\frac{L}{2}\left\|y-x\right\|^2.$$

To establish our convergence results, we use the Łojasiewicz property (Lojasiewicz, 1965), defined next.

Definition 2.6. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a differentiable function. f is said to have the Łojasiewicz property if for any critical point $\overline{x}$, there exist constants $M>0,\varepsilon>0$ and $\theta\in\left[0,1\right)$ such that

$$\left|f\left(x\right)-f\left(\overline{x}\right)\right|^{\theta}\le M\left\|\nabla f\left(x\right)\right\|,\quad \text{for all } x\in\mathbb{B}\left(\overline{x},\varepsilon\right),\tag{2.3}$$

where we adopt the convention $0^0=0$. The constant θ is called *Łojasiewicz exponent* of f at $\overline{x}$.

We also use the following proposition taken from (Khanh et al., 2024) to establish our convergence results.

Proposition 2.7. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a $\mathcal{C}^1$ – smooth function, and let the following conditions hold along a sequence of iterates $\left\{x^k\right\}\subset\mathbb{R}^n$ for the function f:

- *H1 (primary descent condition):* There is $\sigma>0$ such that we have
  $$f\left(x^k\right)-f\left(x^{k+1}\right)\ge\sigma\left\|\nabla f\left(x^k\right)\right\|\cdot\left\|x^{k+1}-x^k\right\|$$
  for sufficiently large $k\in\mathbb{N}$.
- *H2 (complementary descent condition):* The following implication holds:
  $$\left[f\left(x^{k+1}\right)=f\left(x^k\right)\right]\Rightarrow\left[x^{k+1}=x^k\right]$$
  for sufficiently large $k\in\mathbb{N}$.

If $\overline{x}$ is an accumulation point of $\left\{x^k\right\}$ and f satisfies *the Łojasiewicz property* at $\overline{x}$, then $x^k\to\overline{x}$ as $k\to\infty$.

Definition 2.8. Suppose that the sequence $\{x_k\}_{k=0}^{\infty} \subset \mathbb{R}^n$ converges to $x \in \mathbb{R}^n$. We say that the convergence is linear if there exists $\lambda \in (0,1)$ such that

$$\|x_{k+1} - x\| \leq \lambda \|x_k - x\|,$$

for all k sufficiently large.

3. Inexact Boosted Difference of Convex Algorithm

We now present the iBDCA algorithm:

Algorithm 1: iBDCA

Step 1. Fix $\alpha > 0, \bar{\lambda} > 0, 0 < \beta < 1$. Let x_0 be any initial point and set $k := 0$.

/*Initialization*/

Step 2. Find y_k such that

$$\|\nabla g(y_k) - \nabla h(x_k)\| \leq \frac{\rho}{2}\|y_k - x_k\|.$$

/*Find y_k */

Step 3. Set $d_k = y_k - x_k$.. If $d_k = 0$, STOP and RETURN x_k.Otherwise, go to Step 4.

/* Compute the direction d_k **/**

Step 4. Set $\lambda_k = \bar{\lambda}$. while $\phi(y_k + \lambda_k d_k) > \phi(y_k) - \frac{\alpha}{2}\lambda_k \|d_k\|^2$, do $\lambda_k := \beta\lambda_k$.

/* Line search */

Step 5. Set $x_{k+1} := y_k + \lambda_k d_k$, set $k := k+1$, and go to Step 2.

/* Update and iterate */

3.1. Analysis of Convergence

In this section, we establish the convergence results of Algorithm 1. We begin by proving that the inexact solution y_k yields a sufficient decrease in the objective function value.

Proposition 3.1. For any $k \in \mathbb{N}$, we can deduce that

$$\phi(y_k) \leq \phi(x_k) - \frac{\rho}{2}\|d_k\|^2. \tag{3.1}$$

Proof. As g and h are strongly convex and differentiable,

$$g(x_k) - g(y_k) \geq \langle \nabla g(y_k), x_k - y_k \rangle + \frac{\rho}{2}\|x_k - y_k\|^2$$

and

$$h(y_k) - h(x_k) \geq \langle \nabla h(x_k), y_k - x_k \rangle + \frac{\rho}{2}\|y_k - x_k\|^2.$$

Taking the sum of the two previous inequalities leads to

$$g(x_k)-g(y_k)+h(y_k)-h(x_k)\geq \langle \nabla g(y_k)-\nabla h(x_k), x_k-y_k\rangle+\rho\|x_k-y_k\|^2,$$

or equivalently

$$\phi(y_k)-\phi(x_k)\leq \langle \nabla g(y_k)-\nabla h(x_k), d_k\rangle-\rho\|d_k\|^2,$$

where d_k is given in Step 3 of the algorithm. It follows from Step 2 and the Cauchy-Schwarz inequality that

$$\langle \nabla g(y_k)-\nabla h(x_k), d_k\rangle\leq \|\nabla g(y_k)-\nabla h(x_k)\|\cdot\|d_k\|\leq \frac{\rho}{2}\|d_k\|^2,$$

which in turn yields $\phi(y_k)-\phi(x_k)\leq -\frac{\rho}{2}\|d_k\|^2$. □

We now show that $d_k=y_k-x_k$ serves as a descent direction for ϕ at y_k.

Proposition 3.2. For all $k\in\mathbb{N}$, we have $\langle \nabla\phi(y_k), d_k\rangle\leq -\frac{\rho}{2}\|d_k\|^2$.

Proof. The strong convexity of the function h with constant ρ implies the strong monotonicity of ∇h with constant ρ. Applying this property at x_k, y_k, we have,

$$\langle \nabla h(x_k)-\nabla h(y_k), x_k-y_k\rangle\geq \rho\|x_k-y_k\|^2.$$

Combining this with Step 2 and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\langle \nabla\phi(y_k), d_k\rangle &= \langle \nabla g(y_k)-\nabla h(y_k), d_k\rangle\\ &= \langle \nabla h(x_k)-\nabla h(y_k), d_k\rangle+\langle \nabla g(y_k)-\nabla h(x_k), d_k\rangle\\ &\leq \langle \nabla h(x_k)-\nabla h(y_k), d_k\rangle+\|\nabla g(y_k)-\nabla h(x_k)\|\|d_k\|\\ &\leq -\rho\|d_k\|^2+\frac{\rho}{2}\|d_k\|^2\leq -\frac{\rho}{2}\|d_k\|^2,\end{aligned}$$

which completes the proof. □

The following corollary ensures that the backtracking procedure in Step 4 of Algorithm 1 terminates after finitely many steps whenever $\rho>\alpha$.

Corollary 3.3. Let $\rho>\alpha>0$. Then, for every $k\in\mathbb{N}$, there exists $\delta_k>0$ satisfying

$$\phi(y_k+\lambda d_k)\leq \phi(y_k)-\frac{\alpha}{2}\lambda\|d_k\|^2, \text{ for all } \lambda\in[0,\delta_k]. \tag{3.2}$$

Proof. The inequality trivially holds for all positive numbers λ when $d_k=0$. For $d_k\neq 0$, since $\rho>\alpha>0$, one has

$$\lim_{\lambda\to 0^+}\frac{\phi(y_k+\lambda d_k)-\phi(y_k)}{\lambda}=\langle\nabla\phi(y_k), d_k\rangle\leq -\frac{\rho}{2}\|d_k\|^2<-\frac{\alpha}{2}\|d_k\|^2.$$

Hence, there is some $\delta_k>0$ such that, for all $\lambda\in(0,\delta_k)$, the following holds

$$\frac{\phi(y_k+\lambda d_k)-\phi(y_k)}{\lambda} < -\frac{\alpha}{2}\|d_k\|^2,$$

or equivalently

$$\phi(y_k+\lambda d_k) < \phi(y_k) - \frac{\alpha}{2}\lambda\|d_k\|^2. \square$$

The following result shows the monotonicity of the sequence $\{\phi(x_k)\}$. The proof *follows immediately from Proposition 3.1 and Step 4.*

Proposition 3.4. For all $k \in \mathbb{N}$, we have the following inequality

$$\left(\frac{\rho}{2}+\frac{\alpha}{2}\lambda_k\right)\|y_k - x_k\|^2 \le \phi(x_k) - \phi(x_{k+1}).$$

We now provide several results on the convergence of Algorithm 1. The proof is analogous to that in (Aragón-Artacho et al., 2018).

Proposition 3.5. For any initial point $x_0 \in \mathbb{R}^n$, either Algorithm 1 terminates by returning a stationary point of $(\mathcal{P})$ or it generates an infinite sequence that satisfies the following conditions:

i. The sequence $\{\phi(x_k)\}$ is monotonically decreasing and converges to a limit ϕ^*;

ii. Every limit point of $\{x_k\}$ is a stationary point of $(\mathcal{P})$.

iii. $\sum_{k=0}^{\infty}\|d_k\|^2 < \infty$ and $\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2 < \infty$.

Proof. If Algorithm 1 stops at Step 3 and returns x_k, then from Step 2, we deduce that x_k must be a stationary point of $(\mathcal{P})$. Otherwise, it follows from Proposition 3.4 that the sequence $\{\phi(x_k)\}$ decreases monotonically. Given that it is bounded from below by (1.2), it converges to a limit ϕ^*, which proves (i). As a result, we deduce that $\phi(x_{k+1})-\phi(x_k)\to 0$. From Proposition 3.4, one has

$$\frac{\rho}{2}\|d_k\|^2 \le \left(\frac{\rho}{2}+\frac{\alpha}{2}\lambda_k\right)\|d_k\|^2 \le \phi(x_k)-\phi(x_{k+1}).$$

Letting $k\to\infty$ gives $\|y_k - x_k\| = \|d_k\| \to 0$. Let $\bar{x}$ be an arbitrary limit point of the sequence $\{x_k\}$. Then, there exists a subsequence $\{x_{k_i}\}$ that converges to $\bar{x}$. From this, it follows that

$$\|y_{k_i} - \bar{x}\| \le \|y_{k_i} - x_{k_i}\| + \|x_{k_i} - \bar{x}\| \to 0 \text{ as } i\to\infty.$$

Using Step 2, we have

$$0 \le \|\nabla g(y_{k_i}) - \nabla h(x_{k_i})\| \le \frac{\rho}{2}\|y_{k_i} - x_{k_i}\| \to 0 \text{ as } i\to\infty,$$

which gives $\nabla h(\bar{x}) = \nabla g(\bar{x})$, i.e. $\nabla\phi(\bar{x}) = 0$ or $\bar{x}$ is a stationary point of $(\mathcal{P})$.

To prove (iii), by Proposition 3.4, we obtain

$$\left(\frac{\rho}{2}+\frac{\alpha}{2}\lambda_k\right)\|d_k\|^2 \le \phi(x_k)-\phi(x_{k+1}). \tag{3.3}$$

Taking the sum of this inequality over $k=0$ up to $\mathbb{N}$, we obtain

$$\sum_{k=0}^{N}\left(\frac{\rho}{2}+\frac{\alpha}{2}\lambda_k\right)\|d_k\|^2 \le \phi(x_0)-\phi(x_{N+1}) \le \phi(x_0)-\inf_{x\in\mathbb{R}^n}\phi(x). \tag{3.4}$$

Therefore, letting N tend to infinity, we deduce that

$$\sum_{k=0}^{\infty}\frac{\rho}{2}\|d_k\|^2 \le \sum_{k=0}^{\infty}\left(\frac{\rho}{2}+\frac{\alpha}{2}\lambda_k\right)\|d_k\|^2 \le \phi(x_0)-\inf_{x\in\mathbb{R}^n}\phi(x)<\infty,$$

Thus, we obtain $\sum_{k=0}^{\infty}\|d_k\|^2<\infty$. As

$$x_{k+1}-x_k = y_k-x_k+\lambda_k d_k = (1+\lambda_k)d_k,$$

there holds

$$\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2 = \sum_{k=0}^{\infty}(1+\lambda_k)^2\|d_k\|^2 \le (1+\bar{\lambda})^2\sum_{k=0}^{\infty}\|d_k\|^2<\infty,$$

and the proof is complete. □

The following theorem states the convergence results and convergence rate of the algorithm.

Theorem 3.6. Let ∇g be Lipschitz continuous and ϕ satisfy the Łojasiewicz property with an exponent $\theta\in[0,1)$. For any initial point $x_0\in\mathbb{R}^n$, consider the sequence $\{x_k\}$ generated by Algorithm 1. If $\{x_k\}$ has a cluster point x^*, then the entire sequence converges to x^*, which is a stationary point of $(\mathcal{P})$. Moreover, the following convergence rate estimates hold:

(i) if $\theta=0$, then the sequences $\{x_k\}$ and $\{\phi(x_k)\}$ converge in a finite number of iterations to x^* and $\phi(x^*)$, respectively;

(ii) if $\theta\in\left(0,\frac{1}{2}\right]$, then the sequences $\{x_k\}$ and $\{\phi(x_k)\}$ converge linearly to x^* and $\phi(x^*)$, respectively;

(iii) if $\theta\in\left(\frac{1}{2},1\right)$, then there are some positive constants η_1 and η_2 satisfying.

$$\|x_k-x^*\| \le \eta_1 k^{-\frac{1-\theta}{2\theta-1}},$$

$$\phi(x_k)-\phi(x^*) \le \eta_2 k^{-\frac{1}{2\theta-1}},$$

for all large k.

Proof. First, we prove that the conditions (H1) and (H2) of Proposition 2.7 with the sequence $\{x_k\}$ generated by Algorithm 1 for the function ϕ.

H2: Assume that for some $k \in \mathbb{N}$ one has $\phi(x_k) = \phi(x_{k+1})$. From Proposition 2.7 and $x_{k+1} - x_k = (1+\lambda_k)(y_k - x_k)$, we obtain the inequality

$$\frac{\rho + \alpha\lambda_k}{2(1+\lambda_k)^2}\|x_{k+1} - x_k\|^2 \le \phi(x_k) - \phi(x_{k+1}) = 0,$$

which implies that $x_{k+1} = x_k$, since $0 \le \lambda_k \le \bar{\lambda}$.

H1: Combining Step 2 and the Lipschitz continuity of ∇g, we deduce that

$$\begin{aligned}\|\nabla\phi(x_k)\| &= \|\nabla g(x_k) - \nabla h(x_k)\| \\ &\le \|\nabla g(x_k) - \nabla g(y_k)\| + \|\nabla g(y_k) - \nabla h(x_k)\| \\ &\le L\|x_k - y_k\| + 0.5\rho\|x_k - y_k\| \\ &= (L + 0.5\rho)\|d_k\|.\end{aligned} \tag{3.5}$$

Combining (3.5) with Proposition 3.4, we have

$$\begin{aligned}\phi(x_k) - \phi(x_{k+1}) &\ge \left(\frac{\rho}{2} + \frac{\alpha}{2}\lambda_k\right)\|d_k\|^2 \ge \frac{\rho}{2}\|d_k\|^2 \ge \frac{\rho}{2}\|d_k\| \cdot \frac{\|\nabla\phi(x_k)\|}{L + 0.5\rho} \\ &\ge \frac{\rho}{(1+\lambda_k)(2L+\rho)}\|\nabla\phi(x_k)\| . \|x_{k+1} - x_k\| \\ &\ge \frac{\rho}{(1+\bar{\lambda})(2L+\rho)}\|\nabla\phi(x_k)\| . \|x_{k+1} - x_k\|.\end{aligned}$$

Suppose that x^* is a cluster point of $\{x_k\}$. Since ϕ satisfies the Łojasiewicz property at x^*, applying Proposition 2.7, we obtain $x_k \to x^*$ as $k \to \infty$.

The proof of the convergence rates based on θ is similar to the proof of Theorem 1 in Aragón-Artacho et al. (2018).

3.2. Solving subproblems by Gradient Descent

For each $x_k \in \mathbb{R}^n$, recall that $f_k(x) = g(x) - \langle \nabla h(x_k), x \rangle$. Since f_k is strongly convex with modulus ρ, the optimization problem

$$\underset{x \in \mathbb{R}^n}{\text{minimise}} \; f_k(x)$$

has a unique minimiser x_k^*. Then,

$$\nabla f_k(x_k^*) = \nabla g(x_k^*) - \nabla h(x_k) = 0.$$

To obtain an approximate solution to the problem, we construct the gradient descent sequence $\{z_j\}_{j\in\mathbb{N}}$ given by $z_{j+1} = z_j - \frac{1}{L}\nabla f_k(z_j),$, where $z_0 = x_k$. The following convergence result is standard.

Proposition 3.7. Let f_k be strongly convex with modulus ρ and suppose ∇f_k is Lipschitz continuous with constant L. From an initial point $z_0 \in \mathbb{R}^n$, the sequence $\{z_j\}_{j\in\mathbb{N}}$ above satisfies the following estimate

$$\|z_{j+1} - x_k^*\|^2 \le \left(1 - \frac{\rho}{L}\right)^{j+1} \|z_0 - x_k^*\|^2.$$

Depending on the relationship between ρ and L, the appropriate elements of the gradient descent sequence solve the subproblem at each $k-$th iteration.

Corollary 3.8. At each $k-$th iteration, with $z_0 = x_k$, the following assertions hold for the sequence $\{z_j\}_{j\in\mathbb{N}}$:

- If $\rho = L,$ then

$$\|\nabla g(z_1) - \nabla h(x_k)\| \le \frac{\rho}{2}\|z_1 - x_k\|. \qquad (3.6)$$

- If $p < L,$ then for $j_k \in \mathbb{N}$ and $j_k \ge 2\log_{1-\frac{\rho}{L}} \frac{\rho}{\rho + 2L},$ one has

$$\|\nabla g(z_{j_k}) - \nabla h(x_k)\| \le \frac{\rho}{2}\|z_{j_k} - x_k\|.$$

Proof. If $\rho = L,$ then by Proposition 3.7, we deduce that $z_{j+1} = x_k^*$ for all $j \in \mathbb{N}$. Hence, with $j = 0,$ we have

$$\|\nabla g(z_1) - \nabla h(x_k)\| = \|\nabla g(x_k^*) - \nabla h(x_k)\| = 0 \le \frac{\rho}{2}\|z_1 - x_k\|.$$

Now suppose that $\rho < L.$ By Proposition 3.7, one has

$$\|z_j - x_k^*\| \le d^{\frac{j}{2}} \|z_0 - x_k^*\|, \quad \forall j \in \mathbb{N}, \qquad (3.7)$$

where $d = 1 - \frac{\rho}{L} \in (0,1)$. If $j_k \ge 2\log_{1-\frac{\rho}{L}} \frac{\rho}{\rho + 2L},$ then we have

$$\|z_{j_k} - x_k^*\| \le d^{\frac{j_k}{2}} \|z_0 - x_k^*\| \le d^{\frac{j_k}{2}} \left(\|z_0 - z_{j_k}\| + \|z_{j_k} - x_k^*\|\right),$$

and hence,

$$\|z_{j_k} - x_k^*\| \le \frac{d^{\frac{j_k}{2}}}{1 - d^{\frac{j_k}{2}}} \|z_{j_k} - z_0\|. \qquad (3.8)$$

Combining the Lipschitz continuity with constant L of ∇f_k and $\nabla f_k\left(x_k^*\right)=0$ with (3.8), one has

$$\left\|\nabla f_k\left(z_{j_k}\right)\right\| \le \left\|\nabla f_k\left(z_{j_k}\right)-\nabla f_k\left(x_k^*\right)\right\| \le L\left\|z_{j_k}-x_k^*\right\| \le \frac{Ld^{\frac{j_k}{2}}}{1-d^{\frac{j_k}{2}}}\left\|z_{j_k}-z_0\right\| \le \frac{\rho}{2}\left\|z_{j_k}-z_0\right\|,$$

where the last inequality follows from the choice of j_k. Equivalently, we have

$$\left\|\nabla g\left(z_{j_k}\right)-\nabla h\left(x_k\right)\right\| \le \frac{\rho}{2}\left\|z_{j_k}-x_k\right\|. \square$$

4. Conclusion

In this paper, we have introduced the *Inexact Boosted Difference of Convex Algorithm* (iBDCA), a computationally efficient variant of the *Boosted Difference of Convex Algorithm* (BDCA), designed to solve optimization problems involving the difference of two strongly convex and differentiable functions, where the first component has a Lipschitz continuous gradient. By allowing inexact solutions to the subproblems, iBDCA significantly reduces the computational burden while preserving the convergence properties of BDCA.

The paper provides a rigorous convergence analysis of the iBDCA algorithm. Key results include:

- Proof of monotonic decrease in the objective function values,
- Guarantee that the sequence of iterates either reaches a stationary point or converges to one,
- Establishment of convergence rate results using the Łojasiewicz inequality, showing that convergence can be finite, linear, or sublinear depending on the Łojasiewicz exponent.

Additionally, we show that the number of inner gradient descent iterations required to obtain a sufficiently accurate solution can be bounded, thereby characterizing the method's overall complexity.

❖ ***Conflicts of Interest:*** *Authors have no conflict of interest to declare.*

❖ ***Aknowledgments:*** *This research is funded by Ho Chi Minh City University of Education Foundation for Science and Technology under grant number CS.2025.19.04.TĐ.*

REFERENCES

Aragón Artacho, F. J., & Vuong, P. T. (2020). The boosted difference of convex functions algorithm for nonsmooth functions. *SIAM Journal on Optimization, 30*(1), 980-1006. https://doi.org/10.1137/18M123339X

Aragón-Artacho, F. J., Campoy, R., & Vuong, P. T. (2022). The boosted DC algorithm for linearly constrained DC programming. *Set-Valued and Variational Analysis, 30*(4), 1265-1289. https://doi.org/10.1007/s11228-022-00656-x

Aragón-Artacho, F. J., Fleming, R. M., & Vuong, P. T. (2018). Accelerating the DC algorithm for smooth functions. *Mathematical programming, 169*, 95-118. https://doi.org/10.1007/s10107-017-1180-1

Aragón-Artacho, F. J., Mordukhovich, B. S., & Pérez-Aros, P. (2024). Coderivative-based semi-Newton method in nonsmooth difference programming. *Mathematical Programming*, 1-48. https://doi.org/10.1007/s10107-024-02142-8

Fukushima, M., & Mine, H. (1981). A generalized proximal point algorithm for certain non-convex minimization problems. *International Journal of Systems Science, 12*(8), 989-1000. https://doi.org/10.1080/00207728108963798

Izmailov, A. F., & Solodov, M. V. (2014). Newton-type methods for optimization and variational problems . *Springer, 1*. https://doi.org/10.1007/978-3-319-04247-3

Khanh, P. D., Mordukhovich, B. S., & Tran, D. B. (2024). A new inexact gradient descent method with applications to nonsmooth convex optimization. *Optimization Methods and Software*, 1-29. https://doi.org/10.1080/10556788.2024.2322700

Lojasiewicz, S. (1965). Ensembles semi-analytiques. *Lectures Notes IHES (Bures-sur-Yvette)*.

Tao. (1986). Algorithms for solving a class of nonconvex optimization problems. Methods of subgradients. *North-Holland Mathematics Studies, 129*, 249-271. https://doi.org/10.1016/S0304-0208(08)72402-2

Tao, P. D., & An, L. T. (1997). Convex analysis approach to DC programming: theory, algorithms and applications. *Acta mathematica vietnamica, 22*(1), 289-355.

Yin, P., Lou, Y., He, Q., & Xin, J. (2015). Minimization of 1-2 for compressed sensing. *SIAM Journal on Scientific Computing, 37*(1), A536-A563. https://doi.org/10.1137/140952363

THUẬT TOÁN DCA GIẢI BÀI TOÁN XẤP XỈ HIỆU HAI HÀM LỒI

Bùi Hồ Kim Ánh[1*], Phạm Duy Khánh[1], Trần Trung Tín[1], Trần Bá Đạt[2]

[1]Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Việt Nam

[2]Trường Đại học Rowan, Mĩ

[*]Tác giả liên hệ: Bùi Hồ Kim Ánh – Email: 4801101004@student.hcmue.edu.vn

Ngày nhận bài: 28-4-2025; Ngày nhận bài sửa: 20-5-2025; Ngày duyệt đăng: 05-8-2025

TÓM TẮT

Trong bài báo này, chúng tôi đề xuất Thuật toán hiệu hai hàm lồi xấp xỉ (iBDCA), một biến thể xấp xỉ của Thuật toán hiệu hai hàm lồi tăng cường (BDCA), để giải một bài toán tối ưu liên quan đến hiệu của hai hàm lồi mạnh, trong đó cả hai thành phần của hàm DC đều được giả thiết là khả vi và gradient của thành phần đầu tiên là liên tục Lipschitz. Thuật toán đề xuất bao gồm hai bước chính: đầu tiên, chúng tôi sử dụng một lược đồ dựa trên phương pháp hạ gradient để tìm nghiệm xấp xỉ cho một bài toán con liên quan đến phần lồi của hàm DC; tiếp theo, nghiệm xấp xỉ này được sử dụng để tính toán kích thước bước phù hợp cho lần lặp tiếp theo. Chúng tôi thiết lập sự hội tụ và tốc độ hội tụ của dãy lặp và dãy giá trị hàm mục tiêu tương ứng đến nghiệm tối ưu và giá trị tối ưu.

Từ khóa: thuật toán DCA tăng cường; phân tích sự hội tụ; hàm DC; phương pháp hạ gradient xấp xỉ; tìm kiếm theo đường thẳng; kích thước bước