RAIRO-Oper. Res. 58 (2024) 3417–3438
https://doi.org/10.1051/ro/2024060

RAIRO Operations Research
www.rairo-ro.org

ABSTRACT GENERALIZED EPSILON-DESCENT ALGORITHM

Estefany Castillo Ventura[1] and Erik Alex Papa Quiroz[2,*]

Abstract. Given the problem of minimizing a possibly nonconvex and nonsmooth function in a real Hilbert space, we present a generalized epsilon-descent algorithm motivated from the abstract descent method introduced by Attouch *et al.* [*Math. Program.* **137** (2013) 91–129] with two essential additions, we consider scalar errors on the sufficient descent condition, as well as, on the relative inexact optimality condition. Under general conditions on the function to be minimized, we obtain that all accumulation points of the sequences generated by the algorithm, if they exist, are generalized critical limit points of the objective function.

Mathematics Subject Classification. 90C26, 65K05.

Received December 15, 2022. Accepted March 6, 2024.

1. Introduction

The theoretic analysis and develop of algorithms have been advanced in the study of convex optimization problems and even more if the objective functions are smooth. However, in the nonconvex case it is necessary to do more work to obtain new results.

The motivation of this paper comes from the study of optimization algorithms for lower semicontinuous functions that are possible nonconvex and nonsmooth, because that class of functions appear when modeling aspects of real life, see for example, for image restauration [12], image process, see [10, 16], machine learning [13], consumer theory [3, 8, 19, 31] and groups dynamic [9, 19].

Between several methods developed in Mathematical Optimization for this class of problems we are interested in the class of descent algorithms which produce, in each iteration, decreasing sequences with respect to the objective function such that the sequence generated by the algorithm converges to some minimum point or, at least, to a critical point when the function is not convex.

However, we known that when the objective function is nonconvex, descent methods might provide sequences with oscillatory behaviors. The first sighting of this phenomenon was possibly given by Curry, see [18]. Similar behaviors are found in the context of differential equations, see [29], as also, Absil *et al.* [1] for an adaptation to gradient methods. To avoid these behaviors several researchers have been motivated to work with functions that have appropriate structures, for example, functions satisfying the Kurdyka–Lojasiewicz (KL) inequality or coercivity, among others.

Keywords. Nonsmooth optimization, nonconvex optimization, coercive function, descent methods, relative error, scalar error.

[1] Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Lima, Peru.

[2] Pontificia Universidad Católica del Perú, Departamento Académico de Ciencias, San Miguel, Peru.

*Corresponding author: erikpapa@gmail.com

Attouch *et al.* [6] introduced an abstract descent algorithm, which consider a sufficient descent condition with respect to the objective function and a relative error condition, that includes several well known algorithms as gradient, proximal, forward-backward and Gauss–Seidel methods, and proved under the KL property the global convergence to some critical point of the objective function of the generated sequence of the algorithm.

The results of Attouch *et al.* [6] were extended to Hilbert space by Frankel *et al.* [21] where new results on rate of convergence were presented. Bonettini *et al.* [12] presented recently a variant of the abstract descent algorithm where the relative error condition is satisfied at a point that might be different from the actual iterate generated by the method. They obtained convergence results to a critical point using the KL inequality as a main assumption. Bolte *et al.* [11] also introduced a variant of the abstract descent algorithm proving the global convergence of the sequences using the KL inequality condition.

On the other hand, using a condition of coercivity, Papa Quiroz *et al.* [31] introduced an inexact proximal algorithm with quasi-distances where scalar errors for the sufficient descent condition and vectorial errors for the residual of the regularized critical point were considered. They showed that the algorithm converges weakly to a critical point for proper lower semicontinuous and coercive functions.

From the practical point of view, it becomes natural to consider scalar errors in the sufficient descent condition, as also, in the relative error condition. These errors could be attributed to the input data, the computation, truncated or rounded errors. So we believe that working with more computational errors will improve optimization techniques that use these methods such as machine learning algorithms. This is the main motivation of our approach.

The present paper studies a generalization in Hilbert spaces of the descent method introduced by Attouch *et al.* [6] and extended by Frankel *et al.* [21]. We follow the ideas of a recent paper by Papa Quiroz *et al.* [31] to analyze the convergence of the algorithm. The generalization consists for adding scalar errors in the sufficient descent condition and in the relative error condition, that is, given a function $g : H \to \mathbb{R} \cup \{+\infty\}$, where H is a Hilbert space, the proposed algorithm generates in H a sequence $\{x^k\}$ that satisfies the following conditions:

– For each $k \in \mathbb{N}$:

$$g\left(x^{k+1}\right) + \theta \left\|x^{k+1} - x^k\right\|^2 \leq g\left(x^k\right) + \epsilon_{k+1}. \tag{1}$$

– For each $k \in \mathbb{N}$, there exists $v^{k+1} \in \partial_{\epsilon_{k+1}} g(x^{k+1})$:

$$\left\|v^{k+1}\right\| \leq \kappa \left\|x^{k+1} - x^k\right\| + \beta_{k+1} \tag{2}$$

where θ, κ are positive real numbers and $\partial_{\epsilon_{k+1}} g(x^{k+1})$ denotes the ϵ_{k+1}-limiting Fréchet subdifferential (for details of this ϵ-subdifferential, see [24]). The first condition (1), called the ϵ-sufficient descent condition of $\{g(x^k)\}$, attributes a descent measure and an error criterion to the sequence $\{g(x^k)\}$. The second condition (2), reflects the relatively error condition for solving each minimization subproblem of an inexact or exact minimization process.

The contributions of the paper are the following:

(a) We present theoretic results on the convergence of the generalized ϵ-descent algorithm, (1) and (2), using general conditions on the objective function.
(b) The introduction of errors in the algorithm, permits that methods as gradient, proximal, forward-backward splitting and regularized Gauss–Seidel methods might work with more freedom to obtain the next point in each iteration.
(c) The proposed algorithm permits in the relatively error condition the use of an approximate subdifferential for g at the new point x^{k+1}.
(d) The theoretic results are given in Hilbert spaces, which are natural spaces to work with Partial Differential Equations problems when it is used the variational formulation to obtain optimization problems in these spaces, in particular in Sobolev spaces.
(e) We perform some first numerical results comparing descent algorithms with and without the introduction of the scalar error. We conclude that, considering the numerical experiments, the introduction of error improve the time and the total number of iteration of the algorithms.

The article is organized as follows: in Section 2, we present some concepts necessary for the development of the investigation, In Section 3, we present the abstract generalized epsilon-descent algorithm. In Section 4, we show specializations of the algorithm in already known descent methods (Sect. 4.1 gradient method, Sect. 4.2 proximal method, Sect. 4.3 the forward-backward algorithm). In Section 5, we show that all accumulation points (if they exist) of the sequences generated by the abstract algorithm satisfying appropriate conditions, are generalized critical points of the objective function. Then in Section 6 we study the convergence of the algorithm when the function is locally Lipschitz using the ϵ-Clarke-subdifferental. In Section 7, we analyze the convergence of the proposed algorithm when the function is convex using the ϵ-convex subdifferential. In Section 8, we present some numerical examples of the algorithm with and without the introduction of scalar errors to verify its performance. In the last section, we present our conclusions.

2. Basic concepts

The following definitions and results may be find in the book [14]. Throughout this article, we consider H as a real Hilbert space, $\langle \cdot, \cdot \rangle$ as its inner product which induces a norm $\| \cdot \|$, that is, $\|x\| := \sqrt{\langle x, x \rangle}$. Moreover, H^* denotes the dual space of H.

A sequence $\{x^k\}$ in H converges strongly to $\bar{x}$, denoted by $x^k \to \bar{x}$, when $\|x^k - \bar{x}\| \to 0$. The sequence $\{x^k\} \subset H$ converges weakly to $\bar{x}$, denoted by $x^k \rightharpoonup \bar{x}$, when it holds $\langle h, x^k \rangle \to \langle h, \bar{x} \rangle$, $\forall h \in H^*$.

Proposition 2.1 ([25], Sect. 4.8. problem 3, p. 261)**.** *If $\{x^k\}, \{y^k\} \subset H$ converges weakly to $\bar{x}$ and $\bar{y}$ respectively, then $\{x^k + y^k\}$ weakly converges to $\bar{x} + \bar{y}$.*

The set ℓ^1 is defined as following:

$$\ell^1 = \left\{ \{r^l\} \subset \mathbb{R} : \sum_{l=0}^{+\infty} |r^l| < +\infty \right\}. \tag{3}$$

Definition 2.1 ([14], p. 10)**.** Let V be a topological space and $g : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper function. We said that g is lower semicontinuous on V if for all $\beta \in \mathbb{R}$ the next set is closed:

$$\{x \in V : g(x) \leq \beta\}.$$

Remark 2.1. Taking $V = H$, since H is a metric space, the previous definition is equivalent to: If $\{x^k\} \subset H$ converges strongly to some $\bar{x} \in H$, then $\liminf_{k\to\infty} g(x^k) \geq g(\bar{x})$.

Theorem 2.1 ([14], Thm. 5.5, p. 135)**.** *Given any $\phi \in H^*$ there exists a unique $f \in H$ such that: $\phi(u) = \langle f, u \rangle$ $\forall u \in H$*

Definition 2.2 ([14], p. 57)**.** Given the collection $(\phi_f)_{f \in H^*}$, where each $\phi_f : H \longrightarrow \mathbb{R}$ is a linear functional defined by $\phi_f(x) = \langle f', x \rangle$, where $f' \in H$ is given by Theorem 2.1. The weak topology on H is the coarsest topology associated to the collection $(\phi_f)_{f \in H^*}$. This topology is denoted by $\sigma(H, H^*)$.

Definition 2.3 ([14], p. 61)**.** Let $g : H \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function with respect to the weak topology, that is, the set $\{x \in V : g(x) \leq \beta\}$ is closed in the weak topology, then we said that the function g is weakly lower semicontinuous.

Remark 2.2. Since the weak topology being non-metrizable, if g is a weakly lower semicontinuous function then, if $\{x^k\} \subset H$ converges weakly to some $x \in H$ we have

$$\liminf_{k\to\infty} g\left(x^k\right) \geq g(x),$$

the inverse in general is not true.

Proposition 2.2 ([14], p. 61)**.** *Given a lower semicontinuous function $g : H \to \mathbb{R} \cup \{+\infty\}$. If g is also convex then g is weakly lower semicontinuous.*

Proposition 2.3 ([14], Thm. 3.18, p. 69)**.** *Let $\{x^k\} \subset H$ be a bounded sequence, then it has a weak accumulation point, i.e., there exists $\{x^{k_j}\} \subset \{x^k\}$ a subsequence weakly convergent.*

Definition 2.4 ([22], p. 6)**.** A function $g : H \to \mathbb{R}$ is coercive on H if

$$\lim_{\|x\| \to +\infty} g(x) = +\infty.$$

Proposition 2.4 ([7], Thm. 3.2.5, p. 87)**.** *If $g : H \to \mathbb{R} \cup \{+\infty\}$ is an extended proper coercive and weakly lower semicontinuous function, then there exists some $u \in H$ such that $f(u) \leq f(v)$, $\forall v \in H$.*

Lemma 2.1 ([32], Lem. 9.2, p. 64)**.** *If $g : H \to \mathbb{R}$ is a C^1 function with L-Lipschitz gradient, then the following inequality holds for every x, y in H:*

$$g(y) \leq g(x) + \langle \nabla g(x), y - x \rangle + \frac{L}{2}\|y - x\|^2.$$

2.1. Fréchet subdifferentials

Definition 2.5. For a proper function $g : H \to \mathbb{R} \cup \{+\infty\}$ and a point x in its domain, we have that:

(a) ([26], (1.1)) The Fréchet subdifferential of g at $x \in dom(g)$ is the set:

$$\hat{\partial} g(x) := \left\{ w \in H : \lim_{z \to x} \inf_{z \neq x} \frac{g(z) - g(x) - \langle w, z - x \rangle}{\|z - x\|} \geq 0 \right\}.$$

$\hat{\partial} g(x) = \emptyset$, when $x \notin dom(g)$.

(b) ([26], (1.21)) The limiting Fréchet subdifferential of g at x is the set:

$$\partial g(x) := \{ w \in H : \exists x^l \to x,\ g(x^l) \to g(x),\ w^l \in \hat{\partial} g(x^l) \text{ and } w^l \rightharpoonup w \}.$$

(c) ([24], (2.3)) The Fréchet ϵ-subdifferential of g at x is the set:

$$\hat{\partial}_\epsilon g(x) := \left\{ w \in H : \liminf_{\|h\| \to 0} \frac{g(x + h) - g(x) - \langle w, h \rangle}{\|h\|} \geq -\epsilon \right\}.$$

(d) ([24], (2.8)) The limiting Fréchet ϵ-subdifferential of g at $x \in dom(g)$ is the set:

$$\partial_\epsilon g(x) := \{ w \in H : \exists x^l \to x,\ g(x^l) \to g(x),\ w^l \in \hat{\partial}_\epsilon g(x^l) \text{ and } w^l \rightharpoonup w \}.$$

(e) The generalized limit subdifferential of g at x is the set:

$$\bar{\partial} g(x) := \{ w \in H : \exists x^l \to x,\ g(x^l) \to g(x),\ w^l \in \partial_{\epsilon_l} g(x^l)\ w^l \rightharpoonup w \text{ and } \epsilon_l \to 0 \}.$$

Remark 2.3. The subdifferential defined in (e) was introduced, for the finite dimensional case, in Papa Quiroz *et al.* [31].

Remark 2.4. $\hat{\partial} g(x)$ and $\partial g(x)$ denote the sets in (c) and (d) when $\epsilon = 0$ respectively.

Remark 2.5. Given a function $g : H \to \mathbb{R} \cup \{+\infty\}$ and a point $\bar{x} \in dom(g)$, then it holds that $\hat{\partial} g(\bar{x}) \subset \partial g(\bar{x})$.

Proposition 2.5 ([26], Prop. 1.10, p. 3328)**.** *Given $g : H \to \mathbb{R} \cup \{+\infty\}$ a proper function which has a local minimum point, $\bar{x} \in dom(f)$, then $0 \in \hat{\partial} g(\bar{x})$.*

Proposition 2.6 ([26], Prop. 1.1, p. 3326)**.** *Given a proper function $g : H \to \mathbb{R} \cup \{+\infty\}$ and Fréchet differentiable at $\bar{x}$, with derivative $Dg(\bar{x})$, we obtain that $\hat{\partial} g(\bar{x}) = \{Dg(\bar{x})\}$.*

Proposition 2.7 ([27], Thm. 2.33, p. 216)**.** *Let $g, h : H \to \mathbb{R} \cup \{+\infty\}$ be subdifferentiable functions at $\bar{x}$, which means that $\hat{\partial} g(\bar{x})$ and $\hat{\partial} h(\bar{x})$ are non empty sets. If g is a lower semicontinuous function and h is locally Lipschitz, then $\hat{\partial}(g+h)(\bar{x}) \neq \emptyset$ and $\hat{\partial} g(\bar{x}) + \hat{\partial} h(\bar{x}) \supset \hat{\partial}(g+h)(\bar{x})$.*

Remark 2.6. In the original statement the Proposition 2.7 holds when the considered space is an Asplund space. Observe that Hilbert spaces are reflective Banach spaces and reflective Banach spaces are Asplund spaces, see [2, 37] for that proof.

Proposition 2.8 ([24], p. 73)**.** *Let $h, g : H \to \mathbb{R} \cup \{+\infty\}$ be proper functions then:*

(1) *$\hat{\partial}_\epsilon h(x)$ is empty if $x \notin dom(h)$.*
(2) *$\hat{\partial}_{\epsilon_1+\epsilon_2}(h+g)(x) \supset \hat{\partial}_{\epsilon_1} h(x) + \hat{\partial}_{\epsilon_2} g(x)$.*
(3) *If $\epsilon_1 \leq \epsilon_2$ then $\hat{\partial}_{\epsilon_2} h(x) \supset \hat{\partial}_{\epsilon_1} h(x)$.*

Proposition 2.9 ([24], Prop. 2.8, p. 77)**.** *Let $g : H \longrightarrow \mathbb{R}$ be continuously Fréchet differentiable at x, with derivative $Dg(x)$. Then,*

$$\partial_\epsilon g(x) = Dg(a) + \epsilon B^*,$$

where B^* is the ball in H centered at zero.

3. Abstract generalized epsilon-descent algorithm

In this section we introduce an abstract inexact algorithm, which is inspired by the paper of Attouch *et al.* [6], with the difference that in this work we consider scalar errors in the condition of sufficient descent (1), an ϵ-subdifferential and an error on the boundedness of the relative error in (2).

The motivation to introduce this inexact algorithm is to consider in descent algorithms computational errors due to approximations, compression or noise.

Algorithm 1. Abstract generalized ϵ-descent algorithm.

Step 0: Choose $x^0 \in H$, two positive real numbers θ, κ and two nonnegative real sequences $\{\epsilon_{k+1}\}$, $\{\beta_{k+1}\} \subset \mathbb{R}$ such that:

$$\{\epsilon_{k+1}\} \subset \ell^1, \ \beta_{k+1} \to 0, \tag{4}$$

where ℓ^1 is defined by (3).
Step 1: Given $x^k \in H$, find $x^{k+1} \in H$ such that:

$$g\left(x^{k+1}\right) + \theta \left\|x^{k+1} - x^k\right\|^2 \leq g\left(x^k\right) + \epsilon_{k+1} \tag{5}$$

$$v^{k+1} \in \partial_{\epsilon_{k+1}} g\left(x^{k+1}\right) \tag{6}$$

where

$$\left\|v^{k+1}\right\| \leq \kappa \left\|x^{k+1} - x^k\right\| + \beta_{k+1}. \tag{7}$$

Step 2: If $0 \in \partial g(x^{k+1})$ or $\|x^k - x^{k+1}\| = 0$, then stop. Otherwise, do $k \to k+1$ and return to Step 1.

The introduced abstract inexact algorithm becomes the algorithm proposed by Attouch *et al.* [6] when H is the Euclidean space, $\epsilon_k = 0$ and $\beta_k = 0$ for all $k \in \mathbb{N}$ and includes several inexact descent algorithms, as we will see in the next section.

4. Algorithms with epsilon-descent property

4.1. Inexact gradient method

Consider a function $g : H \to \mathbb{R}$ of class C^1 with L-Lipschitz gradient, that is, ∇g is L-Lipschitz.

Algorithm 2. Inexact gradient method.

Step 0: Take two real numbers θ, κ with $\theta > L$. Choose $x^0 \in H$ and two nonnegative real sequence $\{\epsilon_{k+1}\}$ and $\{\beta_{k+1}\}$ such that:

$$\{\epsilon_{k+1}\},\ \{\beta_{k+1}\} \in \ell^1. \tag{8}$$

Step 1: Given $x^k \in H$, find $x^{k+1} \in H$ such that:

$$\frac{\theta}{2}\left\|x^{k+1}-x^k\right\|^2 \leq \left\langle \nabla g\left(x^k\right), x^k - x^{k+1}\right\rangle + \epsilon_{k+1} \tag{9}$$

$$\left\|\nabla g\left(x^k\right)\right\| \leq \kappa\left\|x^{k+1}-x^k\right\| + \beta_{k+1}. \tag{10}$$

Step 2: If $\|x^k - x^{k+1}\| = 0$ or $\nabla g(x^{k+1}) = 0$, then stop. Otherwise, do $k \to k+1$ and return to Step 1.

Remark 4.1. When $\epsilon_{k+1} = 0$ and H is finite-dimensional, the Algorithm 2 is reduced to the algorithm introduced in Section 2 of Attouch *et al.* [6].

Remark 4.2. Let a and b two positive numbers and $\{\gamma^k\}$ such that $0 < a \leq \gamma^k \leq b < \frac{2}{L}$, where L is a Lipschitz constant of the gradient of g. Consider a sequence generated by the inexact gradient method

$$x^{k+1} = x^k - \gamma^k \nabla g\left(x^k\right) + e^k. \tag{11}$$

Observe that in applications e^k model measurement errors, noises, quantization errors due to compression, among others. We suppose that

$$\sum_{k=1}^{+\infty}\left\|e^k\right\| < +\infty. \tag{12}$$

We affirm that the gradient method (11) and (12) is a particular case of Algorithm 2. In fact, from (11) we have

$$\begin{aligned}\left\|x^{k+1}-x^k\right\|^2 &= \left\|-\gamma^k\nabla g\left(x^k\right)+e^k\right\|^2\\ &= \left\langle -\gamma^k\nabla g\left(x^k\right)+e^k, x^{k+1}-x^k\right\rangle = \gamma^k\left\langle \nabla g\left(x^k\right), x^k-x^{k+1}\right\rangle + \left\langle e^k, x^{k+1}-x^k\right\rangle\\ &\leq \gamma^k\left\langle \nabla g\left(x^k\right), x^k-x^{k+1}\right\rangle + \left\|e^k\right\|\left\|x^{k+1}-x^k\right\|. \end{aligned}\tag{13}$$

Nothing that $ab \leq \frac{a^2+b^2}{2}$, for all $a, b \in \mathbb{R}$ we have

$$\left\|e^k\right\|\left\|x^{k+1}-x^k\right\| \leq \frac{\left\|e^k\right\|^2 + \left\|x^{k+1}-x^k\right\|^2}{2}.$$

Using the above inequality in (13)

$$\left\|x^{k+1}-x^k\right\|^2 \leq \gamma^k\left\langle \nabla g\left(x^k\right), x^k-x^{k+1}\right\rangle + \frac{\left\|e^k\right\|^2}{2} + \frac{\left\|x^{k+1}-x^k\right\|^2}{2}.$$

It gives

$$\frac{1}{2}\left\|x^{k+1}-x^k\right\|^2 \leq \gamma^k\left\langle \nabla g(x^k), x^k-x^{k+1}\right\rangle + \frac{\left\|e^k\right\|^2}{2}.$$

Using the fact $0 < a \leq \gamma^k \leq b < \frac{2}{L}$ we have

$$\frac{1}{2b}\left\|x^{k+1}-x^k\right\|^2 \leq \left\langle \nabla g\left(x^k\right), x^k - x^{k+1}\right\rangle + \frac{\left\|e^k\right\|^2}{2a}.$$

Taking $\epsilon_{k+1} = \frac{\|e^k\|^2}{2a}$, condition (9) is satisfied.

Now, we show that condition (10) is also satisfied. From (11)

$$\left\|\nabla g\left(x^k\right)\right\| = \frac{1}{\gamma^k}\left\|x^{k+1}-x^k-e^k\right\| \leq \frac{1}{a}\left\|x^{k+1}-x^k\right\| + \frac{1}{a}\left\|e^k\right\|.$$

Taking $\beta_{k+1} = \frac{1}{a}\|e^k\|$ we have condition (10).

Now, we prove that Algorithm 2 falls in the form of the abstract generalized epsilon-descent algorithm, that is, Algorithm 1.

Condition (4) of Algorithm 1 is obtained above. To prove that condition (5) is satisfied, we apply the Lemma 2.1 at the points $x = x^k$ and $y = x^{k+1}$ and add the term $\frac{\theta}{2}\|x^{k+1}-x^k\|^2$ in both members of inequality, then we get:

$$\left\langle \nabla g\left(x^k\right), x^{k+1}-x^k\right\rangle + \frac{L}{2}\left\|x^{k+1}-x^k\right\|^2 + \frac{\theta}{2}\left\|x^{k+1}-x^k\right\|^2 \geq g\left(x^{k+1}\right) - g\left(x^k\right) + \frac{\theta}{2}\left\|x^{k+1}-x^k\right\|^2.$$

By (9):

$$g\left(x^{k+1}\right) + \frac{\theta}{2}\left\|x^{k+1}-x^k\right\|^2 \leq g\left(x^k\right) + \frac{L}{2}\left\|x^{k+1}-x^k\right\|^2 + \epsilon_{k+1}.$$

Then,

$$g\left(x^{k+1}\right) + \left(\frac{\theta-L}{2}\right)\left\|x^{k+1}-x^k\right\|^2 \leq g\left(x^k\right) + \epsilon_{k+1}.$$

Since $\theta > L$ we have (5).

By Proposition 2.9 we get:

$$\partial_{\epsilon_{k+1}} g\left(x^{k+1}\right) = \left\{\nabla g\left(x^{k+1}\right) + \epsilon_{k+1} B^*\right\}.$$

Then, for some $z \in B^*$, we have that $v^{k+1} = \nabla g(x^{k+1}) + \epsilon_{k+1} z \in \partial_{\epsilon_{k+1}} g(x^{k+1})$, and since ∇g is L-Lipschiz, applying the triangular inequality we get that

$$\begin{aligned}\left\|v^{k+1}\right\| &= \left\|\nabla g\left(x^{k+1}\right) + \epsilon_{k+1} z\right\| \leq \left\|\nabla g\left(x^{k+1}\right)\right\| + \left\|\epsilon_{k+1} z\right\| \\ &\leq \left\|\nabla g\left(x^{k+1}\right) - \nabla g\left(x^k\right)\right\| + \left\|\nabla g\left(x^k\right)\right\| + \epsilon_{k+1}\|z\| \\ &\leq (L+\kappa)\left\|x^{k+1}-x^k\right\| + \epsilon_{k+1} + \beta_{k+1}.\end{aligned}$$

Therefore, condition (7) of Algorithm 1 is fulfilled.

4.2. Inexact proximal method

The proximal point method to find a minimum point of a proper lower semicontinuous function $g : H \to \mathbb{R} \cup \{+\infty\}$, starting at x^0, generates a sequence $\{x^k\}$ given by the following iteration

$$x^{k+1} \in \arg\min \left\{ g(x) + \frac{\lambda_k}{2}\left\|x - x^k\right\|^2 : x \in H \right\} \tag{14}$$

where λ_k is a positive parameter. It is easy to verify that a sufficient condition to assure the existence of each x^k is the lower bounded and lower semicontinuity of g.

Assuming the existence of x^{k+1} we have from the optimality condition that

$$g\left(x^{k+1}\right) \leq g\left(x^{k}\right) - \frac{\lambda_k}{2}\left\|x^{k+1} - x^{k}\right\|^2,$$

and

$$0 \in \hat{\partial}\left(g(.) + \frac{\lambda_k}{2}\left\|\cdot - x^{k}\right\|^2\right)\left(x^{k+1}\right).$$

The above two conditions will be exploited to define an inexact proximal point algorithm.

4.2.1. An inexact proximal algorithm type

The following algorithm extends the algorithm introduced by Attouch *et al.* [6], Section 4, from $\mathbb{R}^n$ to H considering the scalar error in the descent condition and relative error. This algorithm can be applied to a proper function $g : H \to \mathbb{R} \cup \{+\infty\}$, which is weakly lower semicontinuous and bounded from below.

Algorithm 3. Inexact proximal method.

Step 0: Choose $x^0 \in H$, $0 \leq \omega < 1$, $0 < \underline{\tau} \leq \overline{\tau} < \infty$, $0 < \delta \leq 1$, and a non negative real sequence $\{\epsilon_{k+1}\}$ such that:

$$\{\epsilon_{k+1}\} \subset \ell^1. \tag{15}$$

Step 1: Take $\tau_k \in [\underline{\tau}, \overline{\tau}]$, given $x^k \in H$, find $x^{k+1}, v^{k+1} \in H$ such that:

$$g\left(x^{k+1}\right) \leq g\left(x^{k}\right) - \frac{\delta}{2\tau_k}\left\|x^{k} - x^{k+1}\right\|^2 + \epsilon_{k+1}, \tag{16}$$

$$v^{k+1} \in \partial_{\epsilon_{k+1}} g\left(x^{k+1}\right), \tag{17}$$

$$\left\|\tau_k v^{k+1} + x^{k+1} - x^{k}\right\|^2 \leq \omega\left(\left\|\tau_k v^{k+1}\right\|^2 + \left\|x^{k+1} - x^{k}\right\|^2\right). \tag{18}$$

Stop Criterion: If $0 \in \partial g(x^{k+1})$ or $\|x^k - x^{k+1}\| = 0$, then stop. Otherwise, do $k \to k+1$ and return to Step 1.

The proximal point algorithm (14), for $\lambda_k = \frac{\delta}{\tau_k}$, is a particular case of the above algorithm. In fact, given $x^k \in H$, we note that

- It is clear that there exists x^{k+1} solution of the problem $\min\{g(x) + \frac{\delta}{2\tau_k}\|x^k - x\|^2 : x \in H\}$. This is because the function to minimize is bounded from below and weakly lower semicontinuous function (see Prop. 2.4). Then, for all $x \in H$, x^{k+1} holds:

$$g(x) + \frac{\delta}{2\tau_k}\left\|x^{k} - x\right\|^2 \geq g\left(x^{k+1}\right) + \frac{\delta}{2\tau_k}\left\|x^{k} - x^{k+1}\right\|^2.$$

  Taking $x = x^k$ we obtain:

$$g\left(x^{k}\right) - \frac{\delta}{2\tau_k}\left\|x^{k} - x^{k+1}\right\| \geq g\left(x^{k+1}\right).$$

- By Proposition 2.5, $0 \in \partial(g(\cdot) + \frac{\delta}{2\tau_k}\|x^k - \cdot\|^2)(x^{k+1})$ and using Proposition 2.7, we obtain the conditions (17) and (18) of the above algorithm.

From (15), (16) and (17) we directly obtain conditions (4), (5) and (6) of the Algorithm 1. To prove condition (7) is satisfied, we need to use the next lemma, which is an extension to Hilbert spaces from Lemma 4.1, page 113 of [6].

Lemma 4.1. *Given $x, y \in H$ and $\rho \in (0,1]$ such that $\rho(\|x\|^2+\|y\|^2) \geq \|x+y\|^2$. The next inequality is satisfied*

$$2\langle x, y\rangle \leq (\rho-1)\left(\|x\|^2+\|y\|^2\right).$$

Moreover, if also $\rho \in (0,1)$, then:

$$\alpha^{-1}\left(1+\sqrt{1-\alpha^2}\right)\|y\| \geq \|x\| \geq \alpha^{-1}\left(1-\sqrt{1-\alpha^2}\right)\|y\| \tag{19}$$

where $\alpha = 1-\rho$.

Proof. Similar to the proof of Lemma 4.1 of Attouch *et al.* [6]. □

Since $\omega < 1$, considering $\rho = \omega$, the inequality (18) with $x = \tau_k v^{k+1}$ and $y = x^{k+1} - x^k$ and using Lemma 4.1, left inequality of (19), we obtain for some $\kappa > 0$:

$$\kappa\left\|x^{k+1}-x^k\right\| \geq \left\|\tau_k v^{k+1}\right\|,$$

then, since $\tau_k \in [\underline{\tau}, \overline{\tau}]$:

$$\left\|v^{k+1}\right\| \leq \frac{\kappa}{\tau_k}\left\|x^{k+1}-x^k\right\| \leq \frac{\kappa}{\underline{\tau}}\left\|x^{k+1}-x^k\right\|,$$

and thus the condition (7) holds.

4.2.2. Inexact proximal algorithm with quasi-distances

We extend the algorithm developed in Section 4 by Papa Quiroz *et al.* [31] to Hilbert spaces. We first consider some extra definitions and propositions.

Definition 4.1. Given a set V. A quasi-distance in V is a mapping $\phi : V \times V \to \mathbb{R}_+$, such that for every $r, s, t \in V$,

(i) $\phi(r,s) = \phi(s,r) = 0$ whenever $s = r$,
(ii) $\phi(r,s) \leq \phi(r,t) + \phi(t,s)$.

Let $\phi : V \times V \to \mathbb{R}_+$ be a quasi-distance. The quasidistance is called not so asymmetric if for every $r, s \in V$ and some $\alpha, \beta > 0$, it holds:

$$\alpha\|r-s\| \leq \phi(r,s) \leq \beta\|r-s\|. \tag{20}$$

The main properties of quasi-distances can be find in Moreno *et al.* [28]. Here we sketch some of them when the quasi-distance satisfies the condition (20).

- For every $\overline{\omega} \in H$ the functions $\phi(\overline{\omega}, \cdot)$ and $\phi(\cdot, \overline{\omega})$ are Lipschitz.
- For every $\overline{\omega} \in H$ $\phi^2(\overline{\omega}, \cdot)$ and $\phi^2(\cdot, \overline{\omega})$ are locally Lipschitz functions and coercives.

For an extended proper lower weakly semicontinuous coercive function $g : H \to \mathbb{R} \cup \{+\infty\}$, we choose a quasi-distance $\phi : H \times H \to \mathbb{R}_+$ such that it satisfies the condition (20). We consider the following algorithm:

Given $x^k \in H$, we have that the above algorithm is well defined (x^{k+1} and w^{k+1} exist satisfying all the condition of Algorithm 4) because:

- Equation (22) holds because the problem $\min\{f(x) + \frac{\mu_{k+1}}{2}\phi^2(x^k, x) : x \in H\}$ has a solution, this is because the function to minimize is a coercive and weakly lower semicontinuous function, and we apply the Proposition 2.4.
- Now, by Proposition 2.5, $0 \in \hat{\partial}(f(\cdot) + \frac{\mu_{k+1}}{2}\phi^2(x^k, \cdot))(x^{k+1})$ and by Proposition 2.7, (23)–(25) holds when $e^{k+1} = 0$.

Algorithm 4.

Step 0: Choose ϕ a quasi-distance on H, $x^0 \in H$, $\rho \in \mathbb{R}_{++}$ and a pair of real positive sequences $\{\mu_{k+1}\}$ and $\{\epsilon_{k+1}\}$ such that:

$$\{\epsilon_{k+1}\} \subset \ell^1. \tag{21}$$

Step 1: Given $x^k \in H$, find $x^{k+1}, w^{k+1} \in H$ such that:

$$g\left(x^{k+1}\right) + \frac{\mu_{k+1}}{2}\phi^2\left(x^k, x^{k+1}\right) \leq g\left(x^k\right) + \epsilon_{k+1}, \tag{22}$$

$$w^{k+1} \in \partial_{\epsilon_{k+1}} f\left(x^{k+1}\right), \tag{23}$$

$$\left\|e^{k+1}\right\| \leq \rho\phi\left(x^k, x^{k+1}\right), \tag{24}$$

where

$$\begin{aligned} e^{k+1} &= w^{k+1} + \mu_{k+1}\ \phi\left(x^k, x^{k+1}\right)\psi^{k+1} \\ \psi^{k+1} &\in \partial\phi\left(x^k, \cdot\right)\left(x^{k+1}\right). \end{aligned} \tag{25}$$

Step 2: If $x^k = x^{k+1}$ or $0 \in \partial g(x^{k+1})$, then stop. Otherwise, do $k \to k+1$ and return to Step 1.

Under some extra conditions we will verify that the algorithm falls into our abstract inexact algorithm. From (25) and the triangular inequality we have:

$$\begin{aligned} \left\|w^{k+1}\right\| &\leq \left\|e^{k+1} - \mu_{k+1}\ \phi\left(x^k, x^{k+1}\right)\psi^{k+1}\right\| \leq \left\|e^{k+1}\right\| + \left\|\mu_{k+1}\ \phi\left(x^k, x^{k+1}\right)\psi^{k+1}\right\|, \\ &\leq \rho\phi\left(x^k, x^{k+1}\right) + \mu_{k+1}\ \phi\left(x^k, x^{k+1}\right)\left\|\psi^{k+1}\right\| = \left(\rho + \mu_{k+1}\left\|\psi^{k+1}\right\|\right)\ \phi\left(x^k, x^{k+1}\right). \end{aligned}$$

Thus we obtain:

$$\left\|w^{k+1}\right\| \leq \left(\rho + \mu_{k+1}\left\|\psi^{k+1}\right\|\right)\ \phi\left(x^k, x^{k+1}\right).$$

Now, whenever $\mu_{k+1}\|\psi^{k+1}\| \leq M$, where M is a constant, we have the inequality:

$$\left\|w^{k+1}\right\| \leq (\rho + M)\ \phi\left(x^k, x^{k+1}\right),$$

and since $\epsilon_k \geq 0,\ \forall k \in \mathbb{N}$ we get:

$$\left\|w^{k+1}\right\| \leq (\rho + M)\ \phi\left(x^k, x^{k+1}\right) + \epsilon_{k+1}.$$

If we consider the quasi-distance as the Euclidean distance, we have that Algorithm 4 satisfies condition (7) of our abstract inexact algorithm, the conditions (4), (5) and (6) are obtained directly from (21), (22) and (23).

4.3. Inexact forward–backward algorithm

Let us consider a weakly lower semicontinuous and lower bounded proper function $g : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ which can be expressed in the form

$$g = \phi + \varphi,$$

where $\phi : H \longrightarrow \mathbb{R}$ is a nonsmooth and nonconvex function and $\varphi : H \longrightarrow \mathbb{R}$ is a C^1 function such that $\nabla\varphi$ is L-Lipschitz. This type of structured problem is frequently founded (see for example [5] for problems of systems of coupled monotone inclusions in Hilbert spaces and [17] for signal recovery problems).

Consider the following algorithm:

Algorithm 5.

Step 0: Take $\theta, \kappa > 0$ with $a > L$. Set an initial point $x^0 \in dom(\phi)$ and a nonnegative real sequence $\{\epsilon_{k+1}\}$ such that:

$$\{\epsilon_{k+1}\} \subset \ell^1. \tag{26}$$

Step 1: Given $x^k \in H$, find $x^{k+1}, v^{k+1} \in H$ such that:

$$\phi\left(x^{k+1}\right) + \left\langle x^{k+1} - x^k, \nabla\varphi\left(x^k\right)\right\rangle + \frac{\theta}{2}\left\|x^{k+1} - x^k\right\|^2 \leq \phi\left(x^k\right) + \epsilon_{k+1} \tag{27}$$

$$v^{k+1} \in \partial_{\epsilon_{k+1}}\phi\left(x^{k+1}\right) \tag{28}$$

$$\left\|v^{k+1} + \nabla\varphi\left(x^k\right)\right\| \leq \kappa\left\|x^{k+1} - x^k\right\|. \tag{29}$$

Step 2: If $\|x^k - x^{k+1}\| = 0$ or $0 \in \partial g(x^{k+1})$, then stop. Otherwise, do $k \to k+1$ and return to Step 1.

Remark 4.3. The forward-backward splitting method

$$x^{k+1} \in prox_{\lambda_k\phi}\left(x^k - \lambda_k\nabla\varphi\left(x^k\right)\right),$$

where $prox_{\lambda f}(x) := \arg\min\{f(y) + \frac{1}{2\lambda}||y - x||^2 : y \in H\}$ and $\lambda_k > 0$, is an example of Algorithm 5 considering in each iteration $\epsilon_{k+1} = 0$. In fact, that affirmation was proved by Section 5.1 from Attouch *et al.* [6], under the introduction of two positive numbers $\underline{\lambda}$, $\overline{\lambda}$ such that: $0 < \underline{\lambda} < \lambda_k < \overline{\lambda} < \frac{1}{L}$.

Now, we prove that Algorithm 5 satisfies the conditions of our abstract inexact algorithm (Algorithm 1). Applying Lemma 2.1 to the function φ with $y = x^{k+1}$ and $x = x^k$:

$$\varphi\left(x^{k+1}\right) \leq \varphi\left(x^k\right) + \left\langle x^{k+1} - x^k, \nabla\varphi\left(x^k\right)\right\rangle + \frac{L}{2}\left\|x^{k+1} - x^k\right\|^2.$$

Adding in both members $\phi(x^{k+1}) + \frac{\theta}{2}\|x^{k+1} - x^k\|^2$ and by (27):

$$\begin{aligned}
&\varphi\left(x^{k+1}\right) + \phi\left(x^{k+1}\right) + \frac{\theta}{2}\left\|x^{k+1} - x^k\right\|^2 \\
&\quad\leq \varphi\left(x^k\right) + \left\langle x^{k+1} - x^k, \nabla\varphi\left(x^k\right)\right\rangle + \frac{L}{2}\left\|x^{k+1} - x^k\right\|^2 + \phi\left(x^{k+1}\right) + \frac{\theta}{2}\left\|x^{k+1} - x^k\right\|^2 \\
&\quad\leq \phi\left(x^k\right) + \epsilon_{k+1} + \varphi\left(x^k\right) + \frac{L}{2}\left\|x^{k+1} - x^k\right\|^2.
\end{aligned}$$

After some algebraic operations, we have:

$$\varphi\left(x^{k+1}\right) + \phi\left(x^{k+1}\right) + \left(\frac{\theta - L}{2}\right)\left\|x^{k+1} - x^k\right\|^2 \leq \phi\left(x^k\right) + \varphi\left(x^k\right) + \epsilon_{k+1}.$$

Since $g = \varphi + \phi$, it holds:

$$g\left(x^{k+1}\right) + \left(\frac{\theta - L}{2}\right)\left\|x^{k+1} - x^k\right\|^2 \leq g\left(x^k\right) + \epsilon_{k+1}.$$

That is, condition (5) of Algorithm 1 is satisfied.

Now, we define $w^{k+1} = v^{k+1} + \nabla\varphi(x^{k+1})$. We will show that:

$$w^{k+1} \in \partial_{\epsilon_{k+1}}g\left(x^{k+1}\right). \tag{30}$$

In fact, from Algorithm 5 we know that $v^{k+1} \in \partial_{\epsilon_{k+1}}\phi(x^{k+1})$. Therefore, by Definition 2.5 item (d):

$$\exists x^l \longrightarrow x^{k+1}, \phi(x^l) \longrightarrow \phi\left(x^{k+1}\right), v^l \in \hat{\partial}_{\epsilon_{k+1}}\phi\left(x^l\right) \text{ and } v^l \rightharpoonup v^{k+1}.$$

Since φ is continuous, we have:

$$\exists x^l \longrightarrow x^{k+1} : g\left(x^l\right) = \phi\left(x^l\right) + \varphi\left(x^l\right) \longrightarrow \phi\left(x^{k+1}\right) + \varphi\left(x^{k+1}\right) = g\left(x^{k+1}\right). \tag{31}$$

By Proposition 2.8, item (2), it follows that

$$\hat{\partial}_{\epsilon_{k+1}}\phi\left(x^l\right) + \hat{\partial}_0\varphi\left(x^l\right) \subset \hat{\partial}_{\epsilon_{k+1}}(\phi+\varphi)\left(x^l\right) = \hat{\partial}_{\epsilon_{k+1}}g\left(x^l\right). \tag{32}$$

Then, by Proposition 2.6, we get $\nabla\varphi(x^l) \in \hat{\partial}_0\varphi(x^l) = \hat{\partial}\varphi(x^l)$ moreover $v^l \in \hat{\partial}_{\epsilon_{k+1}}\phi(x^l)$, then from (32) we obtain:

$$v^l + \nabla\varphi\left(x^l\right) \in \hat{\partial}_{\epsilon_{k+1}}g\left(x^l\right). \tag{33}$$

Since $\nabla\varphi$ is continuous and $v^l \rightharpoonup v^{k+1}$, by Proposition 2.1:

$$v^l + \nabla\varphi\left(x^l\right) \rightharpoonup v^{k+1} + \nabla\varphi\left(x^{k+1}\right) = w^{k+1}. \tag{34}$$

From (31), (33) and (34) we conclude that:

$$w^{k+1} \in \partial_{\epsilon_{k+1}}g\left(x^{k+1}\right).$$

By inequality (29) of Algorithm 5, and the triangular inequality we obtain the condition (7) of Algorithm 1, as follows:

$$\begin{aligned}\left\|w^{k+1}\right\| &\leq \left\|v^{k+1} + \nabla\varphi\left(x^k\right)\right\| + \left\|\nabla\varphi\left(x^{k+1}\right) - \nabla\varphi\left(x^k\right)\right\| \\ &\leq \kappa\left\|x^{k+1} - x^k\right\| + L\left\|x^{k+1} - x^k\right\| \\ &= (\kappa + L)\left\|x^{k+1} - x^k\right\|.\end{aligned}$$

Then, since $\epsilon_{k+1} \geq 0$ we have:

$$\left\|w^{k+1}\right\| \leq (\kappa + L)\left\|x^{k+1} - x^k\right\| + \epsilon_{k+1}.$$

Remark 4.4. In the article introduced by Frankel *et al.* [21], Section 4, they developed a forward-backward and Gauss–Seidel type algorithm, which thanks to its nature, they fall into the abstract inexact algorithm introduced in this paper.

Before continuing with the above algorithm we need to introduce the proximal operator given metric presented in Section 2.3 of [21].

We considered a space of bounded, uniformly elliptic and self-adjoint operators on H denoted by $S_{++}(H)$. Each $A \in S_{++}(H)$ induces a metric on H by the norm $\|x\|_A = \sqrt{\langle x, x\rangle_A}$ where the inner product is defined by $\langle x, y\rangle_A = \langle A(x), y\rangle$. Moreover, we denoted $\alpha(A)$ as the infimum of the spectral values of A, such that $\|x^2\|_A \geq \alpha(A)\|x\|^2\ \forall x \in H$. Given $f : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ the proximal operator of f in the metric induced by A is the set-valued mapping $prox_f^A : H \Longrightarrow H$, defined as:

$$prox_f^A(x) = \arg\min\left\{f(y) + \frac{1}{2}\|y - x\|_A^2 : y \in H\right\}.$$

We choose an integer $p > 0$ and consider $H_1, \ldots, H_p$ Hilbert spaces with $\langle\cdot,\cdot\rangle_{H_i}$ as inner product and $\|\cdot\|_{H_i}$ and norm respectively. Moreover $H = \prod_{i=1}^p H_i$ with $\langle\cdot,\cdot\rangle = \sum_{i=1}^p \langle\cdot,\cdot\rangle_{H_i}$ and norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$.

Let's consider the problem:

$$\min_{x_i \in H_i}\ f\left(x_1, \ldots, x_p\right) = h\left(x_1, \ldots, x_p\right) + \sum_{i=1}^p g_i\left(x_i\right),$$

where $h : H \to \mathbb{R}$ is continuously differentiable and each $g_i : H_i \to \mathbb{R}$ is lower semicontinuous function. Take $L \geq 0$ such that the functions $x \in H_i \longmapsto h(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_p)$ are L-Lipschitz continuous gradient. For the sequence generated for the next algorithm, we denote $X^k = (x_1^k, \ldots, x_p^k)$ and $X_i^k = (x_1^{k+1}, \ldots, x_{i-1}^{k+1}, x_i^k, \ldots, x_p^k)$.

Algorithm 6.

Step 0: Take $\alpha > L$, an integer $p > 0$, and for each $i \in \{1, \dots, p\}$ choose a sequence $\{A_{i,k}\} \subset S_{++}(H_i)$, and α_k and β_k as the infimum and maximum of the eigenvalues of $\{A_{i,k}\}_{i\in\{1,\dots,p\}}$ such that, $\alpha_k > \alpha$, $\frac{1}{\beta_k} \notin \ell^1$ and $\sup_{k\in\mathbb{N}} \frac{\beta_k}{\alpha_{k+1}} < +\infty$. Take $X^0 \in H$. And choose two sequences $\{S^k = (s_1^k, \dots, s_p^k)\}, \{R^k = (r_1^k, \dots, r_p^k)\} \subset H$ such that, there exists $\eta \geq 0$, $\mu \in]0, 1]$ with $\frac{\eta+1}{\mu} < \frac{\alpha}{L}$:

- $\|S_i^k\| \leq \frac{\eta}{2}\|y_i^{k+1} - y_i^k\|$, where $S_i^k = (s_1^{k+1}, \dots, s_{i-1}^{k+1}, s_i^k, \dots, s_p^k)$,
- $\|r_i^k\| \leq \frac{\eta}{2}\|y_i^{k+1} - y_i^k\| + \delta^k$, with $\delta_k \geq 0$ and $\delta_k \in \ell^1$,
- $\langle r_i^k + s_i^k, y_i^{k+1} - y_i^k\rangle_{A_{i,k}} \leq \frac{1-\mu}{2}\|y_i^{k+1} - y - i^k\|^2_{A_{i,k}}$.

Step 1: Given X^k, for each $i = 1, \dots, p$, do:

$$y_i^{k+1} \in prox_{g_i}^{A_{i,k}} \left(x_i^k - A_{i,k}^{-1}\nabla_i h \left(X_i^k \right) + r_i^k \right)$$
$$x_i^{k+1} = y_i^{k+1} + s_i^{k+1}.$$

Step 2: If $\|X^k - X^{k+1}\| = 0$ or $0 \in \partial f(X^{k+1})$, stop. Otherwise, do $k \to k + 1$ and return to Step 1.

5. Convergence analysis

In this section, we analyze the convergence of the sequence generated by Algorithm 1, when the function under study is proper and lower semicontinuous.

Lemma 5.1 ([33], Lem. 2, p. 44)**.** *Let $\{\mu_k\}$, $\{\chi_k\}$, $\{\alpha_k\} \subset \mathbb{R}_+$, such that the sequences satisfy the inequality $\mu_{k+1} \leq (1 + \chi_k)\mu_k + \alpha_k$, with $\sum_{k=1}^{+\infty} \alpha_k < +\infty$ and $\sum_{k=1}^{+\infty} \chi_k < +\infty$. Then, $\{\mu_k\}$ is convergent.*

Proposition 5.1. *Given a lower semicontinuous proper function $g : H \to \mathbb{R} \cup \{+\infty\}$. The sequence $\{x^k\}$ generated by Algorithm 1 satisfies:*

(i) *$\{g(x^k)\}$ converges.*
(ii) $\|x^{k+1} - x^k\| \to 0$.
(iii) *If $\{x^{k_l}\}$ is convergent to x^* then $\{x^{k_l+1}\}$ converges to the same point x^*.*

Proof. (i) By condition (5) we get:

$$g\left(x^{k+1}\right) \leq g\left(x^k\right) + \epsilon_{k+1}. \tag{35}$$

Denoting $\omega(x^k) = g(x^k) - \inf g(x)$ and using (35) we have:

$$\omega\left(x^{k+1}\right) \leq \omega\left(x^k\right) + \epsilon_{k+1}. \tag{36}$$

Since $\omega(x^k) \geq 0$, we can apply Lemma 5.1, (4) and (36) to obtain that $\{\omega(x^k)\}$ converges and thus we obtain (i).

(ii) From (5) we get:

$$\left\|x^{k+1} - x^k\right\|^2 \leq \frac{1}{\theta}\left(g\left(x^k\right) - g\left(x^{k+1}\right)\right) + \epsilon_{k+1}. \tag{37}$$

Since $\{g(x^k)\}$ converges by item (i) and $\{\epsilon_{k+1}\}$ converges to 0 by (4), we get:

$$\frac{1}{\theta}\left(g\left(x^k\right) - g\left(x^{k+1}\right)\right) + \epsilon_{k+1} \to 0.$$

Then, from (37), $\|x^{k+1} - x^k\|$ converges to zero.

(iii) Assume that $\{x^{k_l}\}$ converges to x^*, then using the triangular inequality property we get:

$$\left\|x^{k_l+1} - x^*\right\| \leq \left\|x^{k_l+1} - x^{k_l}\right\| + \left\|x^{k_l} - x^*\right\|.$$

Since $x^{k_l} \to x^*$ and from (ii) we get $\|x^{k_l+1} - x^*\| \to 0$. Therefore, $\{x^{k_l+1}\}$ converges to x^*.

□

Remark 5.1. If $\{x^{k_l}\}$ converges weakly to x^* then $\{x^{k_l+1}\}$ converges weakly to x^*. Indeed, using the triangular inequality property we get:

$$\left|\left\langle h, x^{k_l+1}\right\rangle - \langle h, x^*\rangle\right| \leq \left|\left\langle h, x^{k_l+1}\right\rangle - \left\langle h, x^{k_l}\right\rangle\right| + \left|\left\langle h, x^{k_l}\right\rangle - \langle h, x^*\rangle\right|, \quad \forall h \in H^*.$$

Since $x^{k_l} \rightharpoonup x^*$ implies $|\langle h, x^{k_l}\rangle - \langle h, x^*\rangle| \to 0,\ \forall h \in H^*$. Moreover, since $|\langle h, x^{k_l+1}\rangle - \langle h, x^{k_l}\rangle| \leq \|h\|\|x^{k_l+1} - x^{k_l}\|,\ \forall h \in H^*$, and from (ii) we get that $|\langle h, x^{k_l+1}\rangle - \langle h, x^*\rangle| \to 0$. That is, $\{x^{k_l+1}\}$ converges weakly to x^*.

We define the set:

$$U = \left\{x \in H : g(x) \leq \lim_{k\to+\infty} g\left(x^k\right)\right\}.$$

Proposition 5.2. *Given a lower semicontinuous proper function $g : H \to \mathbb{R}\cup\{+\infty\}$. Each accumulation point of the sequence $\{x^k\}$ generated by Algorithm 1, if exists, it belongs to U.*

Proof. Suppose there exists $\bar{x}$ an accumulation point of the sequence $\{x^k\}$, then there is $\{x^{k_l}\} \subset \{x^k\}$ that converges to $\bar{x}$. By the lower semicontinuity of g and since $\{g(x^k)\}$ converges by Proposition 5.1, item (i), we obtain:

$$g(\bar{x}) \leq \lim_{l\to+\infty} g\left(x^{k_l}\right) = \lim_{k\to+\infty} g\left(x^k\right).$$

Then $\bar{x} \in U$. □

Theorem 5.1. *Given a lower semicontinuous proper function $g : H \to \mathbb{R}\cup\{+\infty\}$ and moreover g is continuous in its domain. Each accumulation point of the sequence $\{x^k\}$ generated by Algorithm 1, if exists, it is a generalized limiting critical point of g.*

Proof. Suppose that there exists $\bar{x}$ an accumulation point of the sequence $\{x^k\}$, then there is $\{x^{k_l}\} \subset \{x^k\}$ that converges to $\bar{x}$. By Proposition 5.1, (iii), $x^{k_l+1} \to \bar{x}$ and furthermore $g(x^{k_l+1}) \to g(\bar{x})$ since g is continuous in its domain. From (6) and (7) we have:

$$v^{k_l+1} \in \partial_{\epsilon_{k_l+1}} g\left(x^{k_l+1}\right) \text{ and } \left\|v^{k_l+1}\right\| \leq \kappa \left\|x^{k_l+1} - x^{k_l}\right\| + \beta_{k_l+1}.$$

By Proposition 5.1, (ii), we know that $\|x^{k_l+1} - x^{k_l}\| \to 0$, and thanks to (4) we have that $\beta_{k+1} \to 0$, then

$$\left\|v^{k_l+1}\right\| \to 0.$$

Thus, we have that there is $x^{k_l+1} \to \bar{x}$ with $g(x^{k_l+1}) \to g(\bar{x})$, $v^{k_l+1} \in \partial_{\epsilon_{k_l+1}} g(x^{k_l+1})$ with $v^{k_l+1} \to 0$ and $\epsilon_{k_l+1} \to 0$. Then, $0 \in \overline{\partial} g(\bar{x})$. □

Remark 5.2. If we have that $\bar{x} \in H$ is a weak accumulation point, *i.e.*, there exists $\{x^{k_j}\}$ converging weakly to $\bar{x}$, to obtain a similar result as Theorem 5.1, thanks to Remark 5.1, we need that g let be weakly lower semicontinuous and weakly continuous in its domain.

Remark 5.3. It should be noted that we use the generalized of generalized subdifferential, $\overline{\partial} g(\cdot)$, to obtain that the point $\bar{x}$ in the proof of Theorem 5.1 is in a certain sense a generalized limiting critical point.

In the next sections we study the convergence properties of the algorithm for two ϵ-subdifferentials: ϵ-Clarke subdifferential and ϵ-convex subdifferential. The motivation to do this study is to find better properties than the general case.

6. Abstract generalized epsilon-descent algorithm for locally Lipschitz functions

We will use the ϵ-Clarke subdifferential, introduced in Section 5 of Papa Quiroz *et al.* [31].

Definition 6.1. The ϵ-Clarke subdifferential of g at $x \in dom(g)$, denoted by $\partial_\epsilon^\circ g(x)$, is the next set.

$$\partial_\epsilon^\circ g(x) = \{v \in H : g^\circ(x,z) \geq \langle v, u\rangle - \epsilon, \forall u \in H\}, \tag{38}$$

where

$$g^\circ(x,u) = \limsup_{\substack{t\downarrow 0 \\ y\to x}} \frac{g(y+tu)-g(y)}{t}.$$

In the following, we present a variant of our abstract inexact algorithm when $g : H \to \mathbb{R}\cup\{+\infty\}$ is a locally Lipschitz function.

Algorithm 7. Abstract generalized epsilon-descent algorithm for locally Lipschitz functions.

Step 0: Choose $x^0 \in H$, θ, $\kappa \in \mathbb{R}$ positive constants and two nonnegative sequences $\{\epsilon_{k+1}\}$, $\{\beta_{k+1}\} \subset \mathbb{R}$ such that:

$$\{\epsilon_{k+1}\} \subset \ell^1,\ \beta_{k+1} \to 0. \tag{39}$$

Step 1: Given x^k find x^{k+1} that holds:

$$g\left(x^{k+1}\right) + \theta\left\|x^{k+1}-x^k\right\|^2 \leq g\left(x^k\right) + \epsilon_{k+1} \tag{40}$$

$$v^{k+1} \in \partial^\circ_{\epsilon_{k+1}} g\left(x^{k+1}\right) \tag{41}$$

where

$$\left\|v^{k+1}\right\| \leq \kappa\left\|x^{k+1}-x^k\right\| + \beta_{k+1}. \tag{42}$$

Step 2: If $\|x^k - x^{k+1}\| = 0$ or $0 \in \partial^\circ g(x^{k+1})$, stop. Otherwise, do $k \to k+1$ and return to Step 1.

Proposition 6.1. *Given a lower semicontinuous proper and locally Lipschitz function $g : H \longrightarrow \mathbb{R}\cup\{+\infty\}$ and $\{x^k\}$ be a sequence generated by Algorithm 7. Each accumulation point of $\{x^k\}$, if it exists, is a Clarke critical point of g.*

Proof. Assume that $\bar{x} \in H$ is an accumulation point of $\{x^k\}$, then exists $\{x^{k_j}\}$ a subsequence of $\{x^k\}$ that converges to $\bar{x}$. By Proposition 5.1, (iii), $x^{k_j+1} \to \bar{x}$. From (41), $v^{k_j+1} \in \partial^\circ_{\epsilon_{k_j+1}} g(x^{k_j+1})$ and thus

$$g^\circ\left(x^{k_j+1}, z\right) \geq \left\langle v^{k_j+1}, z\right\rangle - \epsilon_{k_j+1}, \qquad \forall z \in H. \tag{43}$$

By Cauchy–Schwarz inequality and (42):

$$\left\langle v^{k_j+1}, z\right\rangle \leq \left\|v^{k_j+1}\right\| \|z\| \leq \kappa\left\|x^{k_j+1}-x^{k_j}\right\| \|z\| + \beta_{k+1}\|z\|.$$

Analogous to the reasoning used in the demonstration of Theorem 5.1, it follows that $\|v^{k_j+1}\|\|z\| \longrightarrow 0$, $\forall z \in H$. Therefore, for all $z \in H$, $\langle v^{k_j+1}, z\rangle \longrightarrow 0$. Finally, taking lim sup in both members of the inequality (43) and by the upper semicontinuity of $g^\circ(\cdot,\cdot)$, we obtain that: $g^\circ(\bar{x}, z) \geq \limsup_{j\to+\infty} g^\circ(x^{k_j+1}, z) \geq 0$. This implies that $0 \in \partial^\circ g(\bar{x})$. □

7. Abstract generalized epsilon-descend algorithm for convex functions

Definition 7.1 ([24], Rem. 2.4, p. 75)**.** The ϵ-Fenchel subdifferential of g at $x \in dom(g)$, denoted by $\partial_\epsilon^F g(x)$, is the next set:

$$\partial_\epsilon^F g(x) = \{v \in H : g(z) \geq g(x) + \langle v, z - x\rangle - \epsilon, \ \forall z \in H\}.$$

We present a variant of our abstract inexact algorithm for the convex case.

Algorithm 8. Abstract generalized epsilon-descend algorithm for convex functions.

Step 0: Choose $x^0 \in H$, θ, $\kappa \in \mathbb{R}$ positive constants and two nonnegative sequences $\{\epsilon_{k+1}\}$,$\{\beta_{k+1}\} \subset \mathbb{R}$ such that:

$$\{\epsilon_{k+1}\} \subset \ell^1, \ \beta_{k+1} \to 0. \tag{44}$$

Step 1: Given x^k, calculate x^{k+1} such that:

$$g\left(x^{k+1}\right) \leq g\left(x^k\right) - \theta \left\|x^{k+1} - x^k\right\|^2 + \epsilon_{k+1} \tag{45}$$

$$v^{k+1} \in \partial_{\epsilon_{k+1}}^F g\left(x^{k+1}\right) \tag{46}$$

where

$$\left\|v^{k+1}\right\| \leq \kappa \left\|x^{k+1} - x^k\right\| + \beta_{k+1}. \tag{47}$$

Step 2: If $0 \in \partial^F g(x^{k+1})$ or $\|x^k - x^{k+1}\| = 0$, stop. Otherwise, do $k \to k+1$ and return to Step 1.

Proposition 7.1. *Let $g : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous and convex function and $\{x^k\}$ is a sequence generated by Algorithm 1. If $\{x^k\}$ is bounded, each weak accumulation point of $\{x^k\}$ is a global minimum point of g.*

Proof. Since $\{x^k\}$ is bounded and from Proposition 2.3, there exists a weak accumulation point, which we denote by $\bar{x}$. Thus, there exists a subsequence $\{x^{k_j}\}$ converging weakly to $\bar{x}$. Since g is weakly lower semicontinuous because g is lower semicontinuous and convex (see Prop. 2.2), and since $\{g(x^k)\}$ converges by Proposition 5.1, (i), we get that:

$$g(\bar{x}) \leq \lim_{k\to+\infty} g\left(x^{k_j}\right) = \lim_{k\to+\infty} g\left(x^k\right). \tag{48}$$

Now, since $v^{k+1} \in \partial_{\epsilon_{k+1}}^F g(x^{k+1})$, by Definition 7.1, for all $z \in H$ the next inequality is satisfied

$$g(z) \geq g\left(x^{k+1}\right) + \left\langle v^{k+1}, z - x^{k+1}\right\rangle - \epsilon_{k+1}. \tag{49}$$

Thanks to conditions (44), (47), Proposition 5.1, (ii), we have that $\|v^{k+1}\| \to 0$. Furthermore, as $\{x^k\}$ is bounded, we get $\langle v^{k+1}, z - x^{k+1}\rangle \to 0$. Using this fact and (48) in (49), jointly with the fact $\epsilon_{k+1} \to 0$, we obtain that

$$g\left(\bar{x}\right) \leq \lim_{k\longrightarrow\infty} g\left(x^k\right) \leq g(z), \qquad \forall z \in H. \tag{50}$$

Therefore, $\bar{x} \in H$ is a minimum point of g. □

Remark 7.1. A sufficient condition to obtain that the sequence $\{x^k\}$ is bounded and thus to obtain the existence of weak accumulation points of $\{x^k\}$ is that the function g let lower semicontinuous and coercive. In fact, from (35), we have

$$g\left(x^k\right) < \alpha_0 := g\left(x^0\right) + \sum_{i=1}^{+\infty} \epsilon_i.$$

Thus, $x^k \in L_g(\alpha_0)$ and as g is lower semicontinuous and coercive, then $\{x^k\}$ is bounded. Now, from Proposition 2.3, there exists $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ that converges weakly to some $\bar{x} \in H$ *i.e.* $\bar{x}$ is a weak accumulation point.

8. Numerical experiments

In this section we present a numerical example that shows the advantages of considering scalar errors. For that we use a personal computer (Intel Core i5-1155G7 @ 2.50 GHz, 12.00 GB RAM) and implement our code using MATLAB R2023a.

We implemented the Algorithm 2 of Section 4.1 for two examples where calculate the inequalities (9) and (10) based on the gradient method with Armijo's rule. The Matlab code is the following

```
function X=xax1(x0,F,e,t,N)
    syms x y;
    f=inline(F);
    G=gradient(F);
    gf=inline(G);
    s=1;
    sigma=0.1;
    beta=0.5;
    test=0;
    v=gf(x0(1),x0(2))';
    m=0;
    x1=x0-v;
    des=dot(v,x1-x0)+t/2*norm(x1-x0)^2-e;
    lambda=1;
    test=test+1;
    while des>0
         m=m+1;
         test=test+1;
         lambda=beta^m*s;
         x1=x0-lambda*v;
         des=dot(v,x1-x0)+t/2*norm(x1-x0)^2-e;
    end
    X=x0-lambda*v;
end
```

It is necessary to remember that in order to apply Algorithm 2 we need that the function to be minimized let be C^1 with L-Lipschitz gradient.

Example 8.1. Consider the problem:

$$\min\left\{f(x,y) = 2x^2 + y^2 : (x,y) \in \mathbb{R}^2\right\}. \tag{51}$$

It is not difficult to notice that the minimum values of the function $f(x,y)$ is reached at the points $(0,0)$ as shown in Figure 1. Now we use the Algorithm 2.

The objective function is C^1 with gradient 4-Lipschitz continuous and we choose: $x^0 = (1,1)$, $\theta = 4.5$, $\kappa = 4$, $\epsilon_{k+1} = \frac{1}{2^{k+1}}$, $\beta_{k+1} = \frac{1}{k(k+1)}$ and the stopping criterion $\|x^{k+1} - x^k\| <= 10^{-4}$.

Table 1 shows the results of the Algorithm 2 to solve the problem (51). Iter is the number of iterations of the Algorithm 2, InnIt represents the number of inner iterations to solve the subproblems in each iteration of

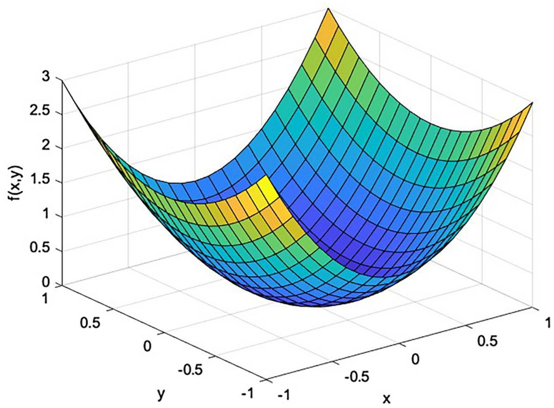

Figure 1. Objective function.

Table 1. Results of inexact gradient method with errors.

Iter	$\nabla f(x^k)$	$\|x^{k+1} - x^k\|$	$x^k = (x_1^k, x_2^k)$	$f(x^k)$	ϵ_k
1	1.51624	2.64947	(1.75565, 1.01572)	0.135759	0.1
2	0.173517	0.165591	(1.90323, 0.940606)	0.00984928	0.1
3	0.0418656	0.0864268	(1.97663, 0.986233)	0.000563453	0.1
4	0.0135019	0.0187443	(1.99307, 0.995244)	5	0.1
5	0.00388978	0.00602352	(1.99804, 0.998637)	0.216411	0.1
6	0.00130969	0.00160636	(1.99937, 0.999547)	0.216411	0.1
7	0.000389387	0.000527001	(1.99978, 0.999872)	0.216411	0.1
8	0.000134748	0.000165438	(1.99993, 0.999957)	0.216411	0.1
9	3.46405e−05	6.44615e−05	(1.99998, 0.999989)	0.216411	0.1
10	9.36943e−06	1.66832e−05	(2, 0.999997)	0.216411	0.1
11	2.82063e−06	3.7222e−06	(2, 0.999999)	0.216411	0.1
12	9.35557e−07	1.22327e−06	(2, 1)	0.216411	0.1

the Algorithm 2, that is, to obtain the inequalities (9) and (10). We observe that 3 iterations are necessary to reach the specified tolerance with one inner iteration for each one, 0.125900 s were used.

Now. with the intention of comparing, we use the Algorithm 2 with $\epsilon_{k+1} = 0$ and $\beta_{k+1} = 0$. The initial conditions and the stopping criterion are the same as the above inexact case.

Table 2 shows the results of the Algorithm 2 without errors. We observe that 17 iterations are necessary to reach the specified tolerance with a time of 0.626055 s.

When comparing the results in both tables and with the time information, we conclude that with Algorithm 2, by including the errors, we need less iterations with less inner iterations and therefore in less time while complying with a given tolerance.

Example 8.2. In this example we will consider the problem:

$$\min\left\{f(x,y) = 4x^2 + 2x + 2y^2 : (x,y) \in \mathbb{R}^2\right\}. \tag{52}$$

It is not difficult to notice that the minimum values of the function $f(x,y)$ is reached at the points $(0,0)$ as shown in Figure 2.

TABLE 2. Results of inexact gradient method without errors.

Iter	InnIt	$\nabla f(x^k)$	$\|x^{k+1} - x^k\|$	$x^k = (x_1^k, x_2^k)$	$f(x^k)$
1	1	4.472136	1.118034	(0.000000, 0.500000)	0.250000
2	1	1.000000	0.250000	(0.000000, 0.250000)	0.062500
3	1	0.500000	0.125000	(0.000000, 0.125000)	0.015625
4	1	0.250000	0.062500	(0.000000, 0.062500)	0.003906
5	1	0.125000	0.031250	(0.000000, 0.031250)	0.0009766
...	...	...	...	...	...
13	1	0.000488	0.000122	(0.000000, 0.000122)	$1.49012 \cdot 10^{-8}$
14	1	0.000244	0.000061	(0.000000, $6.10352 \cdot 10^{-5}$)	$3.72529 \cdot 10^{-9}$
15	1	0.000122	0.000031	(0.000000, $3.05176 \cdot 10^{-5}$)	$9.31323 \cdot 10^{-10}$
16	1	0.000061	0.000015	(0.000000, $1.52588 \cdot 10^{-5}$)	$2.32831 \cdot 10^{-10}$
17	1	0.000031	0.000008	(0.000000, $7.62939 \cdot 10^{-6}$)	$5.82077 \cdot 10^{-11}$

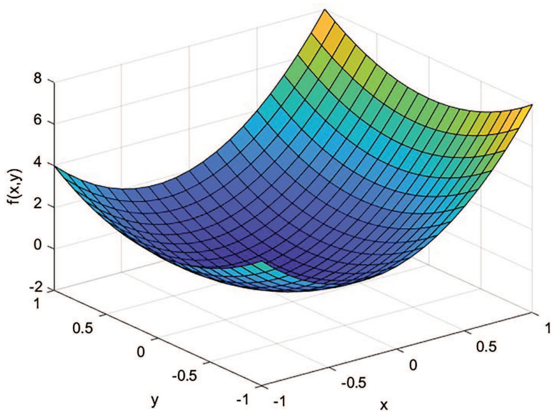

FIGURE 2. Objective function.

TABLE 3. Results of inexact gradient method with errors.

Iter	InnIt	$\nabla f(x^k)$	$\|x^{k+1} - x^k\|$	$x^k = (x_1^k, x_2^k)$	$f(x^k)$	ϵ_k	β_k
1	1	10.7703	1.34629	(0.25, 0.5)	0.25	0.5	1
2	1	2	0.5	$(-0.25, 0)$	-0.25	0.25	0.166667
3	1	0	0	$(-0.25, 0)$	-0.25	0.125	0.083333

It is easy to see that the problem (51) satisfies the necessary conditions of Algorithm 2 because $\nabla f(x, y)$ is 8-Lipschitz continuous. The initial conditions that we choose are: $x^0 = (1, 1)$, $\theta = 8.5$, $\kappa = 8$, $\epsilon_{k+1} = \frac{1}{2^{k+1}}$, $\beta_{k+1} = \frac{1}{k(k+1)}$ and the stopping criterion $\|x^{k+1} - x^k\| <= 10^{-4}$.

We observe that 3 iterations are necessary to reach the specified tolerance with 1 inner iteration for each point with a time of 0.116492 s (Tab. 3).

Now we use the Algorithm 2 without errors. We observe that 17 iterations are necessary to reach the specified tolerance with a time of 0.620852 s (Tab. 4).

TABLE 4. Results of inexact gradient method without errors.

Iter	InnIt	$\nabla f(x^k)$	$\|x^{k+1} - x^k\|$	$x^k = (x_1^k, x_2^k)$	$f(x^k)$
1	1	10.770330	1.346291	$(-0.250000, 0.500000)$	0.250000
2	1	2.000000	0.250000	$(-0.250000,\ 0.250000)$	-0.125000
3	1	1.000000	0.125000	$(-0.250000,\ 0.125000)$	-0.218750
4	1	0.500000	0.062500	$(-0.250000,\ 0.062500)$	-0.242188
5	1	0.250000	0.031250	$(-0.250000,\ 0.031250)$	-0.242047
...	...	...	...	...	...
13	1	0.000977	0.000122	$(-0.250000,\ 0.000122)$	-0.250000
14	1	0.000488	0.000061	$(-0.250000,\ 6.10352 \cdot 10^{-5})$	-0.250000
15	1	0.000244	0.000031	$(-0.250000,\ 3.05176 \cdot 10^{-5})$	-0.250000
16	1	0.000122	0.000015	$(-0.250000,\ 1.52588 \cdot 10^{-5})$	-0.250000
17	1	0.000061	0.000008	$(-0.250000,\ 7.62939 \cdot 10^{-6})$	-0.250000

When comparing the results in both tables and with the time information, we conclude that with Algorithm 1, by including the errors, we need less iterations with less inner iterations and therefore in less time while complying with a given tolerance.

9. Conclusions

We present an abstract inexact algorithm that covers inexact descent algorithms for finding critical points of an extended proper and lower semicontinuous function, which takes into account scalar errors in each iteration. We show that if there exists an accumulation point of the sequence generated by Algorithm 1, this is a generalized limit critical point.

An important contribution of our work is evidenced in the computational application of the methods worked in Section 4, because many algorithms, as for example, inexact gradient, proximal, forward-backward and Gauss–Seidel regularized methods, already include in each approximation scalar errors.

We note that coercivity implies that the sequence generated by Algorithm 1 is bounded, thus guaranteeing the existence of weak accumulation points (or accumulation points in the finite dimensional case). As a consequence of the above we have that when the function is also coercive, the results proved in Section 5 can be applied directly, but when we have weak accumulation point it is needed to require additional assumptions as weakly lower semicontinuity of the objective function (see Rem. 5.2).

This paper presents weak convergence results. The analyze of global convergence of Algorithm 1 under KL property as it was studied by Attouch *et al.* [6] and Frankel *et al.* [21] is ongoing in the working paper of Castillo Ventura and Papa Quiroz [15] where we also study the rate of convergence of the algorithm to know if the introduction of scalar errors increase the speed of convergence of the algorithm.

Acknowledgements

The authors are grateful to UNMSM for funding the PCONFIGI-B23141311 project.

References

[1] P.A. Absil, R. Mahony and B. Andrews, Convergence of the iterates of descent methods for analytic cost functions. *SIAM J. Optim.* **16** (2005) 531–547.

[2] E. Asplund, Fréchet differentiability of convex functions. *Acta Math.* **121** (1968) 31–47.

[3] H. Attouch and A. Soubeyran, Local search proximal algorithms as decision dynamics with costs to move. *Set-Valued Var. Anal.* **19** (2011) 157–177.

[4] H. Attouch, J. Bolte, P. Redont and A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka–Lojasiewicz inequality. *Math. Oper. Res.* **35** (2010) 438–457.

[5] H. Attouch, L.M. Briceño-Arias and P.L. Combettes, A parallel splitting method for coupled monotone inclusions. *SIAM J. Control Optim.* **48** (2010) 3246–3270.

[6] H. Attouch, J. Bolte and B. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. *Math. Program.* **137** (2013) 91–129.

[7] H. Attouch, G. Buttazzo and G. Michaillen, Variational Analysis in Sobolev and BV Spaces. *MPS-SIAM Series on Optimization*. Springer, New York (2014).

[8] T.Q. Bao, B.S. Mordukhovich, A. Soubeyran and C. Tammer, Vector optimization with domination structures: variational principles and applications. *Set-Valued Var. Anal.* **30** (2022) 695–729.

[9] G.C. Bento, O.P. Ferreira, A. Soubeyran and J.V.L. Sousa, Inexact multi-objective local search proximal algorithms: application to group dynamic and distributive justice problems. *J. Optim. Theory Appl.* **177** (2018) 181–200.

[10] M. Bertero, P. Boccacci and V. Ruggiero, Inverse Imaging with Poisson Data. IOP Publishing, Bristol (2018).

[11] J. Bolte, S. Sabach, M. Teboulle and Y. Vaisbourd, First order methods beyond convexity and Lipschitz gradient continuity with applications to quadratic inverse problems. *SIAM J. Optim.* **28** (2018) 2131–2151.

[12] S. Bonettini, P. Ochs, M. Prato and S. Rebegoldi, An abstract convergence framework with application to inertial inexact forward–backward methods. *Comput. Optim. Appl.* **84** (2023) 319–362.

[13] L. Bottou, F.E. Curtis and J. Nocedal, Optimization methods for large-scale machine learning. *SIAM Rev.* **60** (2018) 223–311.

[14] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer, New York (2011).

[15] E. Castillo Ventura and E.A. Papa Quiroz, Abstract algorithm of generalized epsilon descent: global convergence and rate of convergence. Working Paper (2022).

[16] A. Chambolle and T. Pock, An introduction to continuous optimization for imaging. *Acta Numer.* **25** (2016) 161–319.

[17] P.L. Combettes and V.R. Wajs, Signal recovery by proximal forward–backward splitting. *Multiscale Model. Simul.* **4** (2005) 1168–1200.

[18] H.B. Curry, The method of steepest descent for non-linear minimization problems. *Q. Appl. Math.* **1944** (2013) 258–261.

[19] B.G. de Carvalho, S.D.B. Bitar, N.J.X. da Cruz, A. Soubeyran and S.J.C. de Oliveira, A proximal point method for difference of convex functions in multi-objective optimization with application to group dynamic problems. *Comput. Optim. Appl.* **75** (2020) 263–290.

[20] O.P. Ferreira and P.R. Oliveira, Subgradient algorithm on Riemannian manifolds. *J. Optim. Theory Appl.* **97** (1998) 93–104.

[21] P. Frankel, G. Garrigos and J. Peypouquet, Splitting methods with variable metric for Kurdyka–Lojasiewicz functions and general convergence rates. *J. Optim. Theory Appl.* **165** (2015) 874–900.

[22] P. Gajardo, Introducción al Análisis Convexo. Instituto de Matemáticas. Pontificie Universidad Catóolica de Valparaíso (2006).

[23] Y. Hu, X. Yang and C.K. Sim, Inexact subgradient methods for quasi-convex optimization problems. *Eur. J. Oper. Res.* **240** (2015) 315–327.

[24] A. Jofre, D.T. Luc and M. Thera, ϵ-subdiferential and ϵ-monotonicity. *Nonlinear Anal. Theory Methods Appl.* **33** (1998) 71–90.

[25] E. Kreyszig, Introductory Functional Analysis with Applications. Wiley Classics Library (1989).

[26] A.Y. Kruger, On Fréchet subdifferentials. *J. Math. Sci.* **116** (2003) 3325–3358.

[27] B.S. Mordukhovich, Principio Extremal en Análisis Variacional, En: Análisis variacional y diferenciación generalizada I. Grundlehren der mathematischen Wissenschaften. Vol. 330. Springer, Berlin (2006).

[28] F.A.G. Moreno, P.R. Oliveira and A. Soubeyran, A proximal algorithm with quasi distance. Application to Habit's formation. *Optimization* **61** (2012) 1383–1403.

[29] J. Palis and W. Melo, Geometric Theory of Dynamical Systems. Springer, New York (1982).

[30] E.A. Papa Quiroz, M. Ramirez and P.R. Oliveira, An inexact proximal method for quasiconvex minimization. *Eur. J. Oper. Res.* **246** (2015) 721–729.

[31] E.A. Papa Quiroz, A. Soubeyran and P.R. Oliveira, Coercivity and generalized proximal algorithms: application – traveling around the world. *Ann. Oper. Res.* **321** (2022) 451–467.

[32] J. Peypouquet, Optimización y Sistemas Dinámicos, Instituto Venezolano de Investigaciones Científicas (2013).

[33] B.T. Polyak, Introduction to Optimization. Optimization Software, New York (1987).

[34] B.E. Rojas, *Existencia de solución débil de un problema semilineal elíptico*. Tesis de Licenciatura inédita, Universidad Nacional Mayor de San Marcos (2016).

[35] A. Soubeyran and J.C.D.O. Souza, General descent method using w-distance. Application to emergence of habits following worthwhile moves. *J. Nonlinear Var. Anal.* **4** (2020) 285–300.

[36] T. Sun, K. Tang and D. Li, Gradient descent learning with floats. *IEEE Trans. Cybern.* **52** (2022) 1763–1771.

[37] S.L. Trojanski, On locally uniformly convex and differentiable norm in certain non-separable Banach spaces. *Stud. Math.* **37** (1971) 73–180.