



**HAL**  
open science

## Coercivity and generalized proximal algorithms: application-traveling around the world

Erik Alex Papa Quiroz, Antoine Soubeyran, Paulo Roberto Oliveira

► **To cite this version:**

Erik Alex Papa Quiroz, Antoine Soubeyran, Paulo Roberto Oliveira. Coercivity and generalized proximal algorithms: application-traveling around the world. *Annals of Operations Research*, 2023, 321 (1), pp.451-467. 10.1007/s10479-022-04725-0 . hal-03665851

**HAL Id: hal-03665851**

**<https://amu.hal.science/hal-03665851>**

Submitted on 17 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Coercivity and generalized proximal algorithms: application—traveling around the world

Erik A. Papa Quiroz<sup>1,2,3</sup>  Antoine Soubeyran<sup>4</sup> Paulo R. Oliveira<sup>5</sup>

## Abstract

We present an inexact proximal point algorithm using quasi distances to solve a minimization problem in the Euclidean space. This algorithm is motivated by the proximal methods introduced by Attouch et al., section 4, (Math Program Ser A, 137: 91–129, 2013) and Solodov and Svaiter (Set Valued Anal 7:323–345, 1999). In contrast, in this paper we consider quasi distances, arbitrary (non necessary smooth) objective functions, scalar errors in each objective regularized approximation and vectorial errors on the residual of the regularized critical point, that is, we have an error on the optimality condition of the proximal subproblem at the new point. We obtain, under a coercivity assumption of the objective function, that all accumulation points of the sequence generated by the algorithm are critical points (minimizer points in the convex case) of the minimization problem. As an application we consider a human location problem: How to travel around the world and prepare the trip of a lifetime.

**Keywords** Proximal point methods · Inexact algorithms · Coercivity · Quasi distances · Variational rationality · Traveler problem

---

Some parts of the paper was developed during some Postdoctoral scholarship of the first author to PESC-COPPE Federal University of Rio de Janeiro.

---

✉ E. A. Papa Quiroz  
erikpapa@gmail.com

A. Soubeyran  
antoine.soubeyran@gmail.com

P. R. Oliveira  
poliveir@cos.ufrj.br

<sup>1</sup> Universidad Nacional Mayor de San Marcos, Lima , Peru

<sup>2</sup> Universidad Privada del Norte, Lima, Peru

<sup>3</sup> Universidade Federal de Goiás, Goiânia, Brazil

<sup>4</sup> AMSE, Aix-Marseille University, Marseille, France

<sup>5</sup> PESC-COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

# 1 Introduction

The proximal point method, introduced by Martinet (1970) to solve the minimization problem  $\min\{f(x) : x \in \mathbb{R}^n\}$  generates a sequence  $\{x^k\}$  given by  $x^0 \in \mathbb{R}^n$ , and

$$x^k \in \text{arg min}\{f(x) + (\lambda_k/2)\|x - x^{k-1}\|^2 : x \in \mathbb{R}^n\}, \quad (1.1)$$

where  $\lambda_k$  is a certain positive parameter. It is well known, see Güler (1992), that if  $f$  is convex in (1.1) and  $\{\lambda_k\}$  satisfies  $\sum_{k=1}^{+\infty} (1/\lambda_k) = +\infty$ , then  $\lim_{k \rightarrow \infty} f(x^k) = \inf f$ ; and if the optimal set is nonempty, we show that  $\{x^k\}$  converges to an optimal solution of the problem. Furthermore, if  $f$  is quasiconvex then the sequence converges to a critical point of the problem, see (Papa Quiroz & Oliveira, 2009, 2012).

To make possible an application of the proximal point method to solve problems in Computer Theory (Romaguera & Sanchis, 2003), Economics (consumer choice and utility functions) (Levin, 1991), and Behavioral Sciences (Soubeyran, 2010; Janeiro, 2010; Hirschi & Dauwalder, 2015) we need to consider in (1.1) a non convex function  $f$  and a quasi distance  $q(x, x^{k-1})$  instead of the Euclidan norm  $\|x - x^{k-1}\|$ . Thus, we consider the proximal point algorithm with quasi distances given by the following iteration:

$$x^k \in \text{arg min}\{f(x) + (\lambda_k/2)q^2(x, x^{k-1}) : x \in \mathbb{R}^n\}. \quad (1.2)$$

In 2011, Moreno et al. (2011) have been studied, for the first time, the convergence of the sequence  $\{x^k\}$  generated by the above iteration and proved that if  $f$  satisfies a Kurdyka-Lojasiewicz condition on the set of critical points and  $\{x^k\}$  is bounded then this sequence converges to a critical point of the problem. Moreover, the authors presented an application to habit's formation and observed that the classical proximal algorithm of Martinet (1970) and Rockafellar (1976) is unable to procure such an application in Economics and Social Sciences.

In 2015, Bento and Soubeyran (2015b), motivated by the papers of Attouch et al. (2013) and Moreno et al. (2011) presented an inexact version of the proximal point method using quasi distance given by the following main iteration: For  $k = 0, 1, 2, \dots$ , given  $x^k \in \mathbb{R}^n$ , find  $x^{k+1} \in \mathbb{R}^n$  and  $\varphi^{k+1} \in \mathbb{R}^n$  such that

$$f(x^{k+1}) \leq f(x^k) - \lambda_{k+1}(1 - \sigma)\Gamma(q(x^k, x^{k+1})) \quad (1.3)$$

$$\varphi^{k+1} \in \partial f(x^{k+1}) \quad (1.4)$$

$$\psi^{k+1} \in \partial q(x^k, \cdot)(x^{k+1}) \quad (1.5)$$

$$\|\varphi^{k+1}\| \leq b\Gamma'((q(x^k, x^{k+1})))\|\psi^{k+1}\| \quad (1.6)$$

where  $\Gamma(\cdot) : [0, +\infty) \rightarrow [0, +\infty)$  satisfies some adequate conditions as defined in (2) and (3) of Bento & Soubeyran (2015a),  $\delta \in (0, 1)$  and  $\partial f$  denotes the limiting subdifferential of  $f$ . Assuming that the sequence  $\{x^k\}$  is bounded and that  $f$  satisfies a Kurdyka-Lojasiewicz inequality condition, the authors proved that the sequence converges to a critical point of  $f$ , studied the rate of convergence of the sequence and introduced an application to change dynamics in behavioral sciences. The above algorithm has been applied to routine's formation with resistance to change, following worthwhile changes, see (Bento & Soubeyran, 2015a).

In this paper, motivated from the above papers and the work of Solodov and Svaiter (1999), we introduce other inexact iteration that consider two types of errors, scalar and vectorial, in the following way: For  $k = 0, 1, 2, \dots$ , given  $x^k \in \mathbb{R}^n$ , find  $x^{k+1} \in \mathbb{R}^n$  and  $\varphi^{k+1} \in \mathbb{R}^n$  such that

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2}q^2(x^k, x^{k+1}) + \epsilon_{k+1} \quad (1.7)$$

$$\varphi^{k+1} \in \partial_{\epsilon_{k+1}} f(x^{k+1}) \quad (1.8)$$

$$\|e^{k+1}\| \leq \sigma q(x^k, x^{k+1}), \quad (1.9)$$

where  $\sigma > 0$ ,

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1} \quad (1.10)$$

$$\psi^{k+1} \in \partial q(x^k, \cdot)(x^{k+1}), \quad (1.11)$$

and the scalar error  $\epsilon_{k+1}$  satisfies:

$$\sum_{k=0}^{\infty} \epsilon_{k+1} < +\infty. \quad (1.12)$$

The error criterion (1.7) is a  $\epsilon_{k+1}$  sufficient-decrease condition and conditions (1.8),(1.9) and (1.10) have been introduced by Solodov and Svaiter (1999) to find zeroes of maximal monotone operators but considering an extragradient condition.

We should compare the two algorithms: Observe that condition (1.3) in the algorithm of Bento and Soubeyran (2015b) has not a scalar error condition and so their inequality is not inexact compared with (1.7) of the proposed algorithm. Also, we consider  $\epsilon$ -subdifferentials in (1.8). The condition (1.6) of Bento & Soubeyran (2015a) is more general but it does not include condition (1.10) of the proposed algorithm. However, sufficient conditions such that equation (1.10) satisfies equation (1.6) are  $\lambda_k \|\psi^{k+1}\| \geq 1$ ,  $\sigma < 1$  and  $\{\lambda_k\}$  bounded. In fact, from (1.10) and (1.9) we have

$$\|e^{k+1}\|^2 \leq \sigma^2 q(x^k, x^{k+1})^2 \leq \sigma^2 (\|\varphi^{k+1}\|^2 + \|\lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}\|^2).$$

As  $\sigma < 1$ . Then, applying Lemma 4.1 of Attouch et al. (2013) we obtain that  $\|\varphi^{k+1}\| \leq M \lambda_{k+1} q(x^k, x^{k+1}) \|\psi^{k+1}\|$ . Due that  $\{\lambda_k\}$  is bounded then, we obtain (1.6). Thus, we may conclude that the inexact algorithm proposed by Bento & Soubeyran (2015a) and the inexact algorithm proposed in this paper are, in general, different.

In this paper we are interested in the analysis of convergence of the proposed algorithm when  $f$  is an arbitrary coercive function (may be nonconvex, nonsmooth), that is, that does not need to satisfy the Kurdyka–Lojasiewicz property studied by Bento and Soubeyran (2015b). Observe even that a convex function does not need to satisfy the Kurdyka–Lojasiewicz condition, see for example Theorem 36 of Bolte et al. (2010). On the other hand, we are also interested in the application of the proposed algorithm to Behavioral Sciences, in particular we study an application to traveler problem in the context of a recent theory of stay and changes human dynamics called variational rationality approach.

This paper examines the convergence of the proposed algorithm. In fact, assuming that  $f$  is a proper lower semicontinuous and coercive function we prove that the sequence generated by the proposed algorithm is well defined, bounded and if, furthermore,  $f$  is continuous on  $\text{dom}(f)$ , then all accumulation points of the sequence are generalized limiting critical points of the problem. If  $f$  is locally Lipschitz on  $\mathbb{R}^n$  we prove that the accumulation points are Clarke critical points of the problem and if  $f$  is convex then these accumulation points are minimal points of the minimization problem.

The paper is organized as follows. In Sect. 2 we present the basic aspects on subdifferential and quasi distances. In Sect. 3 we present the problem and introduced the assumptions. In Sect. 4 we analyze the convergence of the proposed algorithm for lower semicontinuous functions. In Sect. 5 we study the convergence of the algorithm for locally Lipschitz functions. In Sect. 6 we analyze the convergence when the function is convex. In Sect. 7 we consider, as an application of the proposed algorithm, a traveler problem, in the context of a recent

theory of stays and changes human dynamics, i.e., continue, stop and start doing things in behavioral sciences. That is, the variational rationality approach. Finally in Sect. 8 we give some conclusions.

## 2 Basic facts

### 2.1 Fréchet and limiting subdifferentials

**Definition 2.1** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be a proper function.

- (a) For each  $x \in \text{dom}(f)$ , the set of regular subgradients (also called Fréchet subdifferential) of  $f$  at  $x$ , denoted by  $\hat{\partial}f(x)$ , is the set of vectors  $v \in \mathbb{R}^n$  such that

$$f(y) \geq f(x) + \langle v, y - x \rangle + o(\|y - x\|), \text{ where } \lim_{y \rightarrow x} \frac{o(\|y - x\|)}{\|y - x\|} = 0.$$

Or equivalently,  $\hat{\partial}f(x) := \left\{ v \in \mathbb{R}^n : \liminf_{y \neq x, y \rightarrow x} \frac{f(y) - f(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$ . If  $x \notin \text{dom}(f)$  then  $\hat{\partial}f(x) = \emptyset$ .

- (b) The set of general subgradients (also called limiting subdifferential) of  $f$  at  $x \in \mathbb{R}^n$ , denoted by  $\partial f(x)$ , is defined as follows:

$$\partial f(x) := \left\{ v \in \mathbb{R}^n : \exists x^l \rightarrow x, f(x^l) \rightarrow f(x), v^l \in \hat{\partial}f(x^l) \text{ and } v^l \rightarrow v \right\}.$$

- (c) For each  $x \in \text{dom}(f)$ , the set of  $\epsilon$ -regular subgradients (also called  $\epsilon$ -Fréchet subdifferential) of  $f$  at  $x$ , denoted by  $\hat{\partial}_\epsilon f(x)$ , is the set of vectors  $v \in \mathbb{R}^n$  such that

$$f(y) \geq f(x) + \langle v, y - x \rangle - \epsilon + o(\|y - x\|), \text{ where } \lim_{y \rightarrow x} \frac{o(\|y - x\|)}{\|y - x\|} = 0.$$

Or equivalently,  $\hat{\partial}_\epsilon f(x) := \left\{ v \in \mathbb{R}^n : \liminf_{y \neq x, y \rightarrow x} \frac{f(y) - f(x) - \langle v, y - x \rangle}{\|y - x\|} \geq -\epsilon \right\}$ . If  $x \notin \text{dom}(f)$  then  $\hat{\partial}_\epsilon f(x) = \emptyset$ .

- (d) The set of  $\epsilon$ -general subgradients (also called  $\epsilon$ -limiting subdifferential) of  $f$  at  $x \in \mathbb{R}^n$ , denoted by  $\partial_\epsilon f(x)$ , is defined as follows:

$$\partial_\epsilon f(x) := \left\{ v \in \mathbb{R}^n : \exists x^l \rightarrow x, f(x^l) \rightarrow f(x), v^l \in \hat{\partial}_\epsilon f(x^l) \text{ and } v^l \rightarrow v \right\}.$$

- (e) The set of generalized limiting subdifferential of  $f$  at  $x \in \mathbb{R}^n$ , denoted by  $\bar{\partial}f(x)$ , is defined as follows:

$$\bar{\partial}f(x) := \left\{ v \in \mathbb{R}^n : \exists x^l \rightarrow x, f(x^l) \rightarrow f(x), v^l \in \partial_{\epsilon_l} f(x^l), v^l \rightarrow v, \text{ and } \epsilon_l \rightarrow 0 \right\}.$$

**Proposition 2.1** For a function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  and a point  $\bar{x} \in \text{dom}(f)$ , the subgradient sets  $\partial f(\bar{x})$  and  $\hat{\partial}f(\bar{x})$  are closed, with  $\hat{\partial}f(\bar{x})$  convex and  $\hat{\partial}f(\bar{x}) \subset \partial f(\bar{x})$ .

**Proof** See Rockafellar & Wets (1990), Theorem 8.6. □

**Proposition 2.2** If a proper function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  has a local minimum at the point  $\bar{x} \in \text{dom}(f)$ , then  $0 \in \hat{\partial}f(\bar{x})$ .

**Proof** See Rockafellar & Wets (1990), Theorem 10.1. □

**Proposition 2.3** Let  $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be proper functions such that  $f$  is locally Lipschitz at  $\bar{x} \in \text{dom}(f) \cap \text{dom}(g)$  and  $g$  is a lower semicontinuous function at  $\bar{x}$ . Then,

$$\partial(f + g)(\bar{x}) \subset \partial f(\bar{x}) + \partial g(\bar{x})$$

**Proof** See Mordukhovich (2006), Theorem 2.33. □

## 2.2 Quasi distances

In this section we give the definition of quasidistance and present some properties which will be used in the following section.

**Definition 2.2** Stojmirović (2004) Let  $X$  be a set. A mapping  $q : X \times X \rightarrow \mathbb{R}_+$  is called a quasi distance if for all  $x, y, z \in X$ ,

1.  $q(x, y) = q(y, x) = 0 \Leftrightarrow x = y$ ,
2.  $q(x, z) \leq q(x, y) + q(y, z)$ .

From Definition 2.2, if  $q$  is also symmetric, that is, for all  $x, y \in X$ ,  $q(x, y) = q(y, x)$ , then  $q$  is a distance. Therefore the concept of quasi distance generalizes the concept of a distance.

**Example 2.1** Concilio & Gerla (2006) The function  $q : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$  defined by

$$q(x, y) = \begin{cases} 0 & \text{if } x \leq y, \\ 1 & \text{otherwise.} \end{cases}$$

generates a quasi distance on  $\mathbb{R}$ .

**Example 2.2** (Example 3.3 of Moreno et al. 2011) For each  $i = 1, \dots, n$ , we consider  $c_i^-, c_i^+ > 0$  and the function  $q_i : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$  defined by

$$q_i(x_i, y_i) = \begin{cases} c_i^+(y_i - x_i) & \text{if } y_i - x_i > 0, \\ c_i^-(x_i - y_i) & \text{if } y_i - x_i \leq 0, \end{cases}$$

which is a quasi distance on  $\mathbb{R}$ . Therefore,  $q(x, y) = \sum_{i=1}^n q_i(x_i, y_i)$  is a quasi distance on  $\mathbb{R}^n$ . On the other hand, for each  $\bar{z} \in \mathbb{R}^n$  it holds

$$q(x, \bar{z}) = \sum_{i=1}^n q_i(x_i, \bar{z}_i) = \sum_{i=1}^n \max \{c_i^+(\bar{z}_i - x_i), c_i^-(x_i - \bar{z}_i)\}, \quad x \in \mathbb{R}^n,$$

Thus  $q(\cdot, \bar{z})$  is a convex function. By the same reasoning,  $q(\bar{z}, \cdot)$  is convex.

Consider the following condition on the quasi distance  $q(\cdot, \cdot)$ : there is a positive constant  $\beta$  such that

$$q(x, y) \leq \beta \|x - y\|, \quad \forall x, y \in \mathbb{R}^n. \quad (2.1)$$

We note that the quasi distances of Example 2.2 verifies (2.1).

The next results are important for our study.

**Proposition 2.4** Let  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  be a quasi distance that verifies (2.1). Then for each  $\bar{z} \in \mathbb{R}^n$  the functions  $q(\bar{z}, \cdot)$  and  $q(\cdot, \bar{z})$  are Lipschitz continuous.

**Proof** See Proposition 3.6 of Moreno et al. (2011). □

**Proposition 2.5** Let  $\bar{z} \in \mathbb{R}^n$ . If  $q$  verifies (2.1) then  $q^2(\bar{z}, \cdot)$  and  $q^2(\cdot, \bar{z})$  are locally Lipschitz continuous functions on  $\mathbb{R}^n$ .

**Proof** See Proposition 3.7 of Moreno et al. (2011). □

### 3 The problem

Consider the optimization problem:

$$\min\{f(x) : x \in \mathbb{R}^n\} \quad (3.1)$$

where  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is an extended proper function satisfying the following assumptions:

- (H<sub>1</sub>)  $f$  is a proper lower semicontinuous function
- (H<sub>2</sub>)  $f$  is coercive, that is,  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$
- (H<sub>3</sub>)  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  is a quasi distance satisfying: there exists  $\beta > 0$  such that

$$q(x, y) \leq \beta \|x - y\|. \quad (3.2)$$

Observe that the main justification for using coercive functions, assumption (H<sub>2</sub>), instead of functions satisfying the Kurdyka-Lojasiewicz property is that in behavioral sciences we need global solutions and not only local solutions as for Kurdyka-Lojasiewicz functions. Furthermore, the Kurdyka-Lojasiewicz hypothesis can be very difficult to verify.

### 4 Inexact proximal algorithm with quasi distance (IPAQ) for lower semicontinuous functions

#### IPAQ

**Initialization:** Choose a quasi distance  $q(\cdot, \cdot)$  defined on  $\mathbb{R}^n$ , an initial point  $x^0 \in \mathbb{R}^n$ , a real number  $\sigma > 0$  and two positive real sequences  $\{\lambda_{k+1}\}$  and  $\{\epsilon_{k+1}\}$  such that

$$\sum_{k=0}^{+\infty} \epsilon_{k+1} < +\infty. \quad (4.1)$$

**Main Step:** For  $k = 0, 1, 2, \dots$ , given  $x^k \in \mathbb{R}^n$ , find  $x^{k+1} \in \mathbb{R}^n$  and  $\varphi^{k+1} \in \mathbb{R}^n$  such that

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) + \epsilon_{k+1} \quad (4.2)$$

$$\varphi^{k+1} \in \partial_{\epsilon_{k+1}} f(x^{k+1}) \quad (4.3)$$

$$\|e^{k+1}\| \leq \sigma q(x^k, x^{k+1}) \quad (4.4)$$

where

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}, \quad (4.5)$$

$\psi^{k+1} \in \partial q(x^k, \cdot)(x^{k+1})$ ,  $\partial_{\epsilon_{k+1}} f$  is given by Definition 2.1, (d).

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

**Theorem 4.1** Under assumptions (H<sub>1</sub>)-(H<sub>3</sub>), for each  $k$ , there exists  $(x^{k+1}, \varphi^{k+1})$  satisfying (4.2), (4.3) and (4.4), that is, the (IPAQ) is well defined.

**Proof** Given  $x^k \in \mathbb{R}^n$ , consider the problem  $\min\{f(x) + \frac{\lambda_{k+1}}{2}q^2(x^k, x) : x \in \mathbb{R}^n\}$ . From the lower semicontinuity of  $f$  and  $q^2(x^k, \cdot)$  and coercivity of  $f$ , there exists  $x^{k+1}$  such that

$$f(x^{k+1}) + \frac{\lambda_{k+1}}{2}q^2(x^k, x^{k+1}) \leq f(x) + \frac{\lambda_{k+1}}{2}q^2(x^k, x),$$

for all  $x \in \mathbb{R}^n$ . Taking  $x = x^k$  we get (4.2). On the other hand, from the optimality condition, see Proposition 2.2, we have

$$0 \in \hat{\partial} \left( f(\cdot) + \frac{\lambda_{k+1}}{2}q^2(x^k, \cdot) \right) (x^{k+1})$$

Due to Proposition 2.1 and Proposition 2.3 we have (4.3) and (4.5) with  $e^{k+1} = 0$ .  $\square$

**Remark 4.1** The result of the above theorem is also true if we substitute the condition (3.2) in assumption **(H<sub>3</sub>)** by the assumption that  $q(x, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}_+$  is lower semicontinuous for each  $x \in \mathbb{R}^n$ .

**Lemma 4.1** Let  $\{v_k\}, \{\gamma_k\}$ , and  $\{\beta_k\}$  be nonnegative sequences of real numbers satisfying  $v_{k+1} \leq (1 + \gamma_k)v_k + \beta_k$  and such that  $\sum_{k=1}^{\infty} \beta_k < \infty$ ,  $\sum_{k=1}^{\infty} \gamma_k < \infty$ . Then, the sequence  $\{v_k\}$  converges.

**Proof** See Lemma 2, pp. 44, of Polyak (1987).  $\square$

**Proposition 4.1** Under assumptions **(H<sub>1</sub>)**, **(H<sub>2</sub>)**, **(H<sub>3</sub>)** and the sequence  $\{x^k\}$ , generated by the (IPAQ), satisfies:

- (i)  $\{f(x^k)\}$  converges;
- (ii) If  $\lambda_- < \lambda_{k+1}$ , for some  $\lambda_- > 0$ , then  $q(x^k, x^{k+1})$  converges to zero;
- (iii)  $\{x^k\}$  is bounded;
- (iv) If  $\lambda_- < \lambda_{k+1}$ , for some  $\lambda_- > 0$ , and  $\{x^{k_j}\}$  converges to  $\bar{x}$ , then  $\{x^{k_j+1}\}$  converges to  $\bar{x}$ .

**Proof** (i) From (4.2) we have

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2}q^2(x^k, x^{k+1}) + \epsilon_{k+1} \quad (4.6)$$

Lemma 4.1 and (4.1) imply that  $\{f(x^k)\}$  converges.

(ii) From (4.6) and  $\lambda_- < \lambda_{k+1}$ , we have  $q^2(x^k, x^{k+1}) \leq \frac{2}{\lambda_-} (f(x^k) - f(x^{k+1}) + \epsilon_{k+1})$ .

Given that  $\{f(x^k)\}$  converges and  $\{\epsilon_{k+1}\}$  converges to zero, then taking  $k \rightarrow +\infty$  we obtain the result.

(iii) From (4.6) we have  $f(x^k) \leq f(x^0) + \sum_{i=0}^{k-1} \epsilon_{k+1} \leq f(x^0) + \sum_{i=0}^{+\infty} \epsilon_{k+1} < +\infty$ . This implies that  $x^k \in L_f(\alpha_0) := \{x \in \mathbb{R}^n : f(x) \leq \alpha_0\}$  where  $\alpha_0 = f(x^0) + \sum_{i=0}^{+\infty} \epsilon_{k+1}$ . As  $L_f(\alpha_0)$  is bounded (because  $f$  is coercive) then  $\{x^k\}$  is bounded.

(iv) From the triangular inequality property we obtain

$$\|x^{k_j+1} - \bar{x}\| \leq \|x^{k_j+1} - x^{k_j}\| + \|x^{k_j} - \bar{x}\| \leq \beta q(x^{k_j}, x^{k_j+1}) + \|x^{k_j} - \bar{x}\|,$$

where the last inequality is due to (3.2). As  $q(x^{k_j+1}, x^{k_j})$  converges to zero the result follows.  $\square$

**Remark 4.2** The results (i), (ii) and (iii) of the above proposition are also true if we substitute the condition (3.2) in assumption **(H<sub>3</sub>)** by the assumption that  $q(x, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}_+$  is lower semicontinuous for each  $x \in \mathbb{R}^n$ . However, to obtain the item (iv) is necessary the condition (3.2).

Now, define the foollowing set

$$U := \{x \in \mathbb{R}^n : f(x) \leq \lim_{k \rightarrow +\infty} f(x^k)\}.$$

Observe that due to assumptions **(H<sub>1</sub>)**-**(H<sub>2</sub>)** this set is nonempty closed and bounded.

**Theorem 4.2** *Under assumptions **(H<sub>1</sub>)**-**(H<sub>3</sub>)**, each accumulation point of the sequence  $\{x^k\}$ , generated by the (IPAQ), belongs to  $U$ .*

**Proof** Let  $\bar{x} \in \mathbb{R}^n$  be an accumulation point of  $\{x^k\}$ , then there exists  $\{x^{k_j}\}$  such that  $x^{k_j} \rightarrow \bar{x}$ . As  $f$  is lower semicontinuous and  $\{f(x^k)\}$  converges then  $f(\bar{x}) \leq \lim_{j \rightarrow \infty} f(x^{k_j}) = \lim_{k \rightarrow \infty} f(x^k)$ . That is,  $\bar{x} \in U$ .  $\square$

**Remark 4.3** The result of the above theorem is also true if we substitute the condition (3.2) in assumption **(H<sub>3</sub>)** by the assumption that  $q(x, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}_+$  is lower semicontinuous for each  $x \in \mathbb{R}^n$ .

Consider the following assumption:

**(H<sub>1</sub>)'**  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is proper lower semicontinuous and continuous on  $\text{dom}(f)$ .

Observe that there are functions that are continuous on their domain but are not lower semicontinuous on  $\mathbb{R}^n$ , consider for example the function  $f(x) = 3$ , for  $x \in (1, 2]$  and  $+\infty$ , otherwise.

**Theorem 4.3** *Under assumptions **(H<sub>1</sub>)'**, **(H<sub>2</sub>)**, **(H<sub>3</sub>)**, and that  $0 < \lambda_- < \lambda_k < \lambda_+$ , each accumulation point of the sequence  $\{x^k\}$ , generated by the (IPAQ), is a generalized limiting critical point of the problem (3.1).*

**Proof** Let  $\bar{x} \in \mathbb{R}^n$  be an accumulation point of  $\{x^k\}$ , then there exists  $\{x^{k_j}\}$  such that  $x^{k_j} \rightarrow \bar{x}$  and due to Proposition 4.1, (iv),  $x^{k_j+1} \rightarrow \bar{x}$ . From (4.5) and the triangular inequality we have

$$\|\varphi^{k_j+1}\| \leq \|e^{k_j+1}\| + \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1}) \|\psi^{k_j+1}\| \leq (\sigma + \lambda_{k_j+1} \|\psi^{k_j+1}\|) q(x^{k_j}, x^{k_j+1}) \quad (4.7)$$

where the last inequality comes from (4.4). As  $q(x^k, \cdot)$  is locally Lipschitz in  $\mathbb{R}^n$  then  $\partial q(x^k, \cdot)$  is locally bounded, see Proposition 9.13 of Rockafellar & Wets (1990), then from Lemma 5.1 of Moreno et al. (2011), there exists  $M > 0$  such that  $\|\psi^{k_j+1}\| \leq M, \forall j$ . Thus, the above inequality (4.7) implies that  $\|\varphi^{k_j+1}\| \leq (\sigma + \lambda_+ M) q(x^{k_j}, x^{k_j+1})$ . Therefore, there exists  $x^{k_j+1} \rightarrow \bar{x}$ , with  $f(x^{k_j+1}) \rightarrow f(\bar{x})$  and  $\varphi^{k_j+1} \rightarrow 0$  with  $\varphi^{k_j+1} \in \partial_{\epsilon_{k_j+1}} f(x^{k_j+1})$ , then from Definition 2.1,  $0 \in \bar{\partial} f(\bar{x})$ .  $\square$

## 5 Inexact proximal algorithm with quasi distance (IPAQ) for lower semicontinuous and locally lipschitz functions

In this section we assume that the function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is locally Lipschitz on  $\mathbb{R}^n$ , thus  $\text{dom}(f)$  is open. We consider the  $\epsilon$ -Clarke subdifferential

$$\partial_\epsilon^\circ f(x) := \{w \in \mathbb{R}^n : f^\circ(x, v) \geq \langle w, v \rangle - \epsilon, \forall v \in \mathbb{R}^n\}, \quad (5.1)$$

where

$$f^\circ(x, v) = \limsup_{t \downarrow 0, y \rightarrow x} \frac{f(y + tv) - f(y)}{t}.$$

Now, we consider a variant of the (IPAQ). We will only present the main step and the stop criterion because the initialization step is the same as in the (IPAQ).

### IPAQ2

**Main Step:** For  $k = 0, 1, 2, \dots$ , given  $x^k \in \mathbb{R}^n$ , find  $x^{k+1} \in \mathbb{R}^n$  and  $\varphi^{k+1} \in \mathbb{R}^n$  such that

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) + \epsilon_{k+1} \quad (5.2)$$

$$\varphi^{k+1} \in \partial_{\epsilon_{k+1}}^\circ f(x^{k+1}) \quad (5.3)$$

$$\|e^{k+1}\| \leq \sigma q(x^k, x^{k+1}), \quad (5.4)$$

where

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1} \quad (5.5)$$

where  $\psi^{k+1} \in \partial^\circ q(x^k, \cdot)(x^{k+1})$ ,  $\partial_{\epsilon_{k+1}}^\circ f$  is given by (5.1).

**Stop Criterion:** If  $x^{k+1} = x^k$  or  $0 \in \partial^\circ f(x^{k+1})$ , then stop. Otherwise, do  $k \leftarrow k + 1$  and return to Main Step.

We assume the following assumption:

**(H<sub>1</sub>)''**  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is proper lower semicontinuous and locally Lipschitz on  $\text{dom}(f)$ .

Observe that under the assumptions **(H<sub>1</sub>)''**, **(H<sub>2</sub>)**, **(H<sub>3</sub>)**, the results of Theorem 4.1, Proposition 4.1 and Theorem 4.2 are maintained.

**Theorem 5.1** *Under assumptions **(H<sub>1</sub>)''**, **(H<sub>2</sub>)**, **(H<sub>3</sub>)**, and  $0 < \lambda_- < \lambda_k < \lambda_+$ , each accumulation point of the sequence  $\{x^k\}$ , generated by the (IPAQ2), is a Clarke critical point of the problem (3.1).*

**Proof** From (5.5),  $e^{k+1} - \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1} \in \partial_{\epsilon_{k+1}}^\circ f(x^{k+1})$ , so

$$f^\circ(x^{k+1}, v) \geq \langle e^{k+1}, v \rangle - \lambda_{k+1} q(x^k, x^{k+1}) \langle \psi^{k+1}, v \rangle - \epsilon_{k+1}, \quad \forall v \in \mathbb{R}^n.$$

Let  $\bar{x}$  be a limit of  $\{x^{k_j}\}$ , which implies that  $\{x^{k_j+1}\}$  converges to  $\bar{x}$ . Then,

$$f^\circ(x^{k_j+1}, v) \geq \langle e^{k_j+1}, v \rangle - \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1}) M \|v\| - \epsilon_{k_j+1},$$

where  $M$  satisfies  $\|\psi^{k_j+1}\| \leq M$ . Taking  $\limsup$  in the above inequality and using the property that  $f^\circ(\cdot, \cdot)$  is upper semicontinuous we obtain that  $f^\circ(\bar{x}, v) \geq \limsup_{j \rightarrow +\infty} f^\circ(x^{k_j+1}, v) \geq 0$ . It follows that  $0 \in \partial^\circ f(\bar{x})$ .  $\square$

**Remark 5.1** If  $f$  is a proper, lower semicontinuous and convex function and  $\text{dom}(f)$  is open, then  $f$  satisfies the assumption **(H<sub>1</sub>)''** and thus, from Theorem 5.1, we obtain that each accumulation point of the sequence  $\{x^k\}$ , generated by the (IPAQ2), is a minimum point of problem (3.1). However, if  $\text{dom}(f)$  is not open then we should search another way to obtain that result.

**Remark 5.2** We should observe that if  $f$  is convex and differentiable then,  $\partial_\epsilon^\circ f(x) \subset \partial_\epsilon^F f(x) = \nabla f(x) + \epsilon B(x, 0)$ , where  $\partial_\epsilon^F f$  is the Fenchel  $\epsilon$ -subdifferential:

$$\partial_\epsilon^F f(x) = \{s \in \mathbb{R}^n : f(y) \geq f(x) + \langle s, y - x \rangle - \epsilon, \text{ for all } y \in \mathbb{R}^n\}. \quad (5.6)$$

That is, for the convex and differentiable case, the Fenchel  $\epsilon$ -subdifferential includes the Clarke  $\epsilon$ -subdifferential defined in (5.1).

The above remarks motivate us to study the (IPAQ) considering the Fenchel  $\epsilon$ -subdifferential when  $f$  is a proper lower semicontinuous convex function.

## 6 Inexact proximal algorithm with quasi distance (IPAQ) for lower semicontinuous convex functions

In this section, we complete the convergence of the method for arbitrary proper lower semicontinuous convex functions. For that, we assume the following condition:

(H<sub>1</sub>)'''  $f$  is a proper lower semicontinuous and convex function.

As in the previous section, we will only present the main step and the stop criterion because the initialization step is the same as in the (IPAQ).

### IPAQ3

**Main Step:** For  $k = 0, 1, 2, \dots$ , given  $x^k \in \mathbb{R}^n$ , find  $x^{k+1} \in \mathbb{R}^n$  and  $\varphi^{k+1} \in \mathbb{R}^n$  such that

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) + \epsilon_{k+1} \quad (6.1)$$

$$\varphi^{k+1} \in \partial_{\epsilon_{k+1}}^F f(x^{k+1}) \quad (6.2)$$

$$\|e^{k+1}\| \leq \sigma q(x^k, x^{k+1}) \quad (6.3)$$

where

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1} \quad (6.4)$$

$\psi^{k+1} \in \partial^\circ q(x^k, \cdot)(x^{k+1})$ ,  $\partial_{\epsilon_{k+1}}^F f$  is given in (5.6).

**Stop Criterion:** If  $x^{k+1} = x^k$  or  $0 \in \partial^F f(x^{k+1})$ , then stop. Otherwise, do  $k \leftarrow k + 1$  and return to Main Step.

**Theorem 6.1** *Under assumptions (H<sub>1</sub>)''', (H<sub>2</sub>), (H<sub>3</sub>) and that  $\{\lambda_k\}$  is bounded, each accumulation point of the sequence  $\{x^k\}$ , generated by the (IPAQ3), is an optimal solution of the problem (3.1).*

**Proof** Let  $\bar{x} \in \mathbb{R}^n$  be an accumulation point of  $\{x^k\}$ , then there exists  $\{x^{k_j}\}$  such that  $x^{k_j} \rightarrow \bar{x}$ . As  $f$  is lower semicontinuous and  $\{f(x^k)\}$  converges then  $f(\bar{x}) \leq \lim_{k \rightarrow \infty} f(x^k)$ . On the other hand, as  $f$  is convex, then for each  $x \in \mathbb{R}^n$  we have

$$f(x) \geq f(x^{k+1}) + \langle s^{k+1}, x - x^{k+1} \rangle - \epsilon_{k+1} \quad (6.5)$$

for all  $s^{k+1} \in \partial_{\epsilon_{k+1}}^F f(x^{k+1})$ . From (6.4)

$$e^{k+1} - \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1} \in \partial_{\epsilon_{k+1}}^F f(x^{k+1}) \quad (6.6)$$

where  $\psi^{k_j+1} \in \partial^\circ q(x^k, \cdot)(x^{k+1})$ . From (6.6) and (6.5) we have

$$f(x) \geq f(x^{k+1}) + \langle e^{k+1}, x - x^{k+1} \rangle - \lambda_{k+1} q(x^k, x^{k+1}) \langle \psi^{k+1}, x - x^{k+1} \rangle - \epsilon_{k+1}. \quad (6.7)$$

As  $q(x^k, \cdot)$  is locally Lipschitz in  $\mathbb{R}^n$  then  $\partial q(x^k, \cdot)$  is locally bounded, see Proposition 5.15 from Rockafellar & Wets (1990), then from Lemma 5.1 of Moreno et al. (2011), there exists  $M > 0$  such that  $\|\psi^{k+1}\| \leq M, \forall k$ . Using the above fact in (6.7) and considering that  $\{x^k\}$  and  $\{\lambda_k\}$  are bounded and  $q(x^k, x^{k+1}) \rightarrow 0$ , we have  $f(x) \geq \lim_{k \rightarrow \infty} f(x^{k+1}) \geq f(\bar{x})$ .  $\square$

## 7 Application to location theory: traveling around the world

This section provides an application to location theory in behavioral sciences (economics, decision theory, management sciences, psychology, sociology,...).

*Location theory* This discipline makes a clear distinction between, i) transportation problems where goods are transported from a given location (source) to another one (target) to satisfy human needs in the target location with, ii) translocation problems where individuals (not goods) move from one location to another one to better satisfy their needs.

In mathematics and management sciences, a transportation problem is a special type of linear programming problem where the objective is to minimize the cost of distributing a product from a number of sources or origins to a number of destinations.

In economics a translocation problem can be a migration problem where an individual moves from living in one country to live in an other one, or a changing job problem where a worker moves from doing a given job in a given city to do the same or a different job in an other city, or a tourism problem when an individual or a family visits a region of his country, a city, a landscape, a foreign country....

*The variational rationality approach* Our application will be examined in the context of the (VR) variational rationality approach of stop, continue and go human dynamics where, each period, individuals constantly navigate (life and move) between stays, i.e., continue to do, and changes, i.e., stop and start to do different situated activities (Soubeyran, 2009, 2010, 2016, 2021a,b,c,d). A situated activity, i.e., an action, being an activity done in a given context = environment. The VR approach is a spatial way to model stay and change human dynamics.

*The traveler model* The central picture of the VR approach is the psychological traveler model that makes a clear distinction between two aspects of human dynamics: weak and strong resistance to move. It offers a very general solution to the Lewin's celebrated tentative, i.e., to build a topological approach of psychology (Lewin, 1935, 1936, 1938, 1951). Lewin's model of human dynamics is a psychological translocation model of how to move and live in a psychological life space. This life space includes, for each individual, his internal environment (needs, goals that help to choose how much of each need to satisfy each period, thoughts and emotions) and his external environment (landscape, objects, and persons that help or forbid to satisfy his needs).

*Our application: traveling around the world* Our application is a simple translocation model. It poses the question: "How to travel around the world and prepare the trip of a lifetime". It represents a special instance of the traveler model in the context of weak resistance to move (cf, the proximal algorithm). See, for instance, Moreno et al. (2011), Bento and Soubeyran (2015a; 2015b). This application can be contrasted with other applications of the VR approach (changing jobs, see (Bao & Soubeyran, 2019)), improving the work-life balance (Alfuraidan et al., 2020), transit in migration problems (Fakhar et al., 2020) in the context of strong resistance to move (i.e., the Ekeland variational principle).

### 7.1 A variational rationality "world tour" model

For simplification our application concerns a tourist who wants to make a "world tour", visiting successively different countries until ending his "world tour" in a given country and decides, eventually, to live there. The problem is: when does he will stop to visit a new country.

*A simple world tour model* For simplification, consider an individual that wants to visit successively different locations  $x, y, \dots \in X$ , one location each period. To economize notations, the space of locations will be a straight line  $X = \mathbb{R}$ . Consider two successive periods. Suppose that, in the previous period, this individual has visited location  $x \in X$ . Then, to become able to visit location  $y \in X$  in the current period, that is, to get access to what he wants to visit, he must travel from  $x$  to  $y$  at the beginning of the current period. That is, he must do the move  $m = x \curvearrowright y$ . Suppose also that, as does each simplified version of the VR model,

- (i) the utility of visiting a location  $z$  for the first time is  $g(z) \in \mathbb{R}$ . To economize notations, this utility is net of costs to visit location  $z$ . For simplification we suppose that the utility of visiting a location for the first and second times are equal. The opposite case, which is much more realistic, has been examined in Soubeyran (2021b, c, d).
- (ii) the physical distance between location  $x$  and location  $y$  is  $d(x, y) = \|y - x\|$ .
- (iii) travel costs of moving from having visited location  $x \in X$  in the previous period to become able to visit location  $y \in X$  in the current period are  $T(x, y) \in \mathbb{R}_+$ . We suppose that these travel costs can be modeled as a quasi distance  $T(x, y) = \begin{cases} t_r(y - x), & y \geq x \\ t_l(x - y), & y < x \end{cases}$ , where  $t_r, t_l > 0$  model per unit of distance costs to travel from the right and from the left, depending of the relative position of  $x$  and  $y$ . If  $y = x$ ; travelling costs are zero:  $T(x, x) = 0$ .
- (iv) the disutility  $\Delta(x, y) = T(x, y)^2$  of bearing travel costs increases strongly (quadratically for simplification) with travel costs.

*Payoffs to move (change, or stay).* A payoff to change  $P(y/x) = g(y) - \xi T(x, y)^2$ ,  $\xi > 0$  models a traveler who,

- (i) starts from having visited for the first time a location  $x$  in the previous period;
- (ii) travels, in the current period, from location  $x$  to reach a different location  $y \neq x$  and,
- (iii) visits for the first time location  $y$  at the end of the current period.

A payoff to stay  $P(x/x) = g(x) - \xi T(x, x)^2 = g(x)$  models a traveler who visits location  $x$  for the second time in the current period. The term  $\xi > 0$  measures the importance given to the disutility of travel costs. It depends strongly of the personality of the traveler.

## 7.2 When to start, continue or stop visiting: the VR approach

The main question we pose is: should the traveler stays, or should he changes. Thus, the following definition.

*A worthwhile balance* The VR approach models the difference between a payoff to change and a payoff to stay. That is,  $B(y/x) = P(y/x) - P(x/x) = g(y) - g(x) - \xi T(x, y)^2$ .

In this simple context,

- (i) the difference  $A(y/x) = g(y) - g(x)$  provides a simple formulation of advantages to move (change rather than stay).
- (ii) travel costs define inconveniences to move,  $I(y/x) = T(x, y)$ .

Motivation to move refers to the utility of advantages to move  $M(y/x) = U[A(y/x)]$ , where  $U[\cdot] : A \in \mathbb{R}_+ \mapsto U[A] \in \mathbb{R}_+$  is strictly increasing.

Resistance to move is the disutility of inconveniences to move  $R(y/x) = D[I(y/x)]$ , where  $D[\cdot] : I \in \mathbb{R}_+ \mapsto D[I] \in \mathbb{R}_+$  is strictly increasing.

In a simple linear quadratic translocation model with weak resistance to move,  $U[A] = A$  and  $D[I] = I^2$ .

Then, in this setting, a worthwhile balance is the difference between advantages to move and the disutility of inconveniences to move. That is,  $B(y/x) = A(y/x) - \xi I(y/x)^2$ .

We come back, in this simple context, to the general VR formulation of a worthwhile balance, as the weighted difference between motivation and resistance to move:

$$B(y/x) = M(y/x) - \xi R(y/x).$$

*Stop and go dynamics* They represent a succession of moves over a succession of periods (stays or changes). Each period, a move  $m = x \rightsquigarrow y$  can be,

- (1) improving, if  $A(y/x) \geq 0$ ;
- (2) improving enough (satisficing), if  $A(y/x) \geq \gamma > 0$ ;
- (3) worthwhile if  $B(y/x) \geq 0$ , worthwhile enough if  $B(y/x) \geq \gamma > 0$ ;
- (4) not advantageous, if  $A(y/x) \leq 0$ ;

A position  $x$  is,

- (5) a desired end  $x = x^*$ , if  $A(y/x^*) = g(y) - g(x^*) \leq 0$  for all  $y \in X$ . That is, if  $x^*$  is a maximum of  $g(\cdot)$ .
- (6) a critical point  $x^\bullet$  of a differentiable utility function  $g(\cdot)$ , if  $\nabla g(x^\bullet) = 0$ ;
- (7) an approximate critical point if  $\|\nabla g(x^\bullet)\| \leq \delta$ ,  $\delta > 0$ .

Two other concepts played a major role in the VR approach:

- (8) a stationary trap  $x = x_*$ , if  $B(y/x_*) \leq 0$  for all  $y \in X$ ;
- (9) a variational trap  $x_*$  relative to  $x_0$ , if, i)  $B(x_*/x_0) \geq 0$  and, ii)  $B(z/x_*) \leq 0$  for all  $z \in X$ . That is, if, i) it is worthwhile to reach  $x_*$  from  $x_0$  and if, ii) it is not worthwhile to leave  $x_*$ .

To save space, these two concepts will be examined elsewhere in the context of coercivity hypothesis.

**Remark 1** *The status of travel costs.* They represent, within the VR approach, capability costs to do something. In the traveler example, they model capability costs to visit. They are spend ex ante (before visiting, as fixed costs). Being able to visit  $y$  requires to know what to visit at location  $y$  (what are the most beautiful landscapes to contemplate, what are the more interesting museums to visit) and how to visit  $y$  (what kind of quick and cheap locomotion to use to visit  $y$ , where to eat tasty food,...).

**Remark 2** *The status of execution costs.* They represent, in this paper, costs to visit a location. They have been included in the utility function.

## 7.3 A VR formulation of an exact/inexact proximal algorithm

### 7.3.1 Exact proximal algorithms

An exact proximal algorithm (Martinet, 1970) defined proximal payoffs  $\Pi(y/x) = f(y) + \xi \|y - x\|^2$ , where  $f(\cdot) : y \in X = \mathbb{R}^n \mapsto f(y) \in \mathbb{R}$  is an objective function to be minimized and  $\xi = \lambda/2$ . The algorithm follows a succession of stages  $k, k + 1, \dots$ . The current stage  $k + 1$  solves the problem  $\arg \min \{ \Pi_{k+1}(y/x) = f(y) + \xi_{k+1} \|y - x\|^2 : y \in X \}$ , where  $x = x^k \in X$  is the given status quo in the current stage  $k + 1$ .

The VR approach (Soubeyran, 2021a, b, c, d) makes easily a direct link between a payoff to move and a proximal payoff as follow. Given an utility function  $g(\cdot) : y \in X \mapsto g(y) \in \mathbb{R}$ ,

$g^* = \sup \{g(y) : y \in X\} < +\infty$  defines an aspiration utility level. It is the maximum utility level an individual can hope to reach. In contrast,  $f(y) = g^* - g(y) \geq 0$  represents the disutility of the frustration level that an individual feels if he does not reach his best when he performs  $y \in X$ . Then, in the simple context of the present paper, given that inconveniences to move  $I(y/x)$  can be identified to travel costs  $T(x, y)$  (if costs to move are symmetric, equal to the Euclidean distance) a proximal payoff models a payoff to move, i.e.,  $Q(y/x) = \Pi(y/x) = f(y) + \xi T(x, y)^2$ .

Then, each current period  $k + 1$ , a proximal algorithm minimizes the total cost of moving, i.e., the sum of the disutility of travel costs  $T(x, y)^2$  plus the disutility of frustration feelings  $f(y) = g^* - g(y) \geq 0$  coming from not having done his best at the end of the move.

## 7.4 Interpretation of the results of this paper

### 7.4.1 Interpretation of our new inexact proximal algorithm

Our new (IPAQ) poses two main conditions, each period  $k + 1$  :

(i) *Weak worthwhile moves.* Inequality (4.18) can be written

$$f(x^k) - f(x^{k+1}) \geq (\lambda_{k+1}/2)q(x^k, x^{k+1})^2 - \epsilon_{k+1},$$

with  $\epsilon_{k+1} > 0$  and  $\sum_{k=0}^{+\infty} \epsilon_{k+1} < +\infty$ . This means that each move  $m^{k+1} = x^k \rightsquigarrow x^{k+1}$  must be, a) almost worthwhile ( $\epsilon_{k+1} > 0$ ), b) and more and more worthwhile ( $\sum_{k=0}^{+\infty} \epsilon_{k+1} < +\infty$  implies  $\epsilon_{k+1} \rightarrow 0$ ).

(ii) *Vanishing marginal total disutility to move.* That is, given that

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1}q(x^k, x^{k+1})\psi^{k+1},$$

given in (4.5), represents the marginal total disutility to move, such marginal total disutility must be lower enough relative to costs to move, i.e.,  $\|e^{k+1}\| \leq \sigma q(x^k, x^{k+1})$ , (4.4).

*Importance of coercivity. An almost finite world.* The coercivity hypothesis (H2) (infinity at infinity) done in this paper ( $f(\cdot) = g^* - g(\cdot)$  coercive) means that the utility  $g(y) = g^* - f(y)$  to visit a location  $y \in X$  far away of the origin (initial location) in the space  $X = \mathbb{R}^n$  ( $n = 1$ , in the example) becomes strongly negative. This models a finite world with boundaries. This is a very realistic hypothesis in behavioral sciences. The quasi distance hypothesis models, within the VR approach, asymmetric costs of moving (see also the present example with traveling costs). Hypothesis (H3) means that costs to move are not too asymmetric. Hypothesis (H1)', i.e., upper semicontinuity, prevents downward jumps of the utility function  $g(\cdot)$ , but allows upward jumps.

*Results: Ending in "almost zero" or zero motivation to move.* This paper shows that, under the list of hypothesis given in Theorem 4.3, (H1)', (H2), (H3), each accumulation point of the sequence  $\{x^k\}$ , generated by the (IPAQ) is a generalized limiting critical point of the problem (3.1). That is, in behavioral terms, in the context of the VR approach, a succession of almost worthwhile moves converges to a critical point where the marginal utility to move  $\nabla g(x^*)$  (when  $g$  is differentiable) is zero. Thus, ending in a zero, or almost zero, motivation to move.

Interpretations of theorems 5.1 and 6.1 are similar. However, in theorem 6.1, the end point  $x^*$  maximizes the utility function, given its concavity.

## 8 Conclusions

We present a general inexact proximal algorithm using quasi distances, called (IPAQ), to find critical points and minimizer points of convergent subsequences generated by the algorithm. Under the assumption of coercivity of the objective function, we analyze the cases of lower semicontinuous, locally lipschitz and convex functions and show that all accumulation points of the sequence generated by the (IPAQ) algorithm are critical points (minimizer points for the convex case) of the minimization problem.

If we consider (IPAQ) with  $\epsilon_{k+1} = 0$ , for all  $k = 0, 1, \dots$  and the algorithm does not finish (that is,  $x^{k+1} \neq x^k$  for all  $k = 1, 2, 3, \dots$ ) then  $x^{k+j} \neq x^k$  for all  $j = 1, 2, 3, \dots$  and this mean that, in our model, for the asymptotic case the tourist cannot visit the same place twice. On the other hand, if the scalar error  $\epsilon_{k+1} > 0$ , for some  $k = 0, 1, \dots$  then, for the asymptotic case, it is possible to have  $x^{k+j_0} = x^k$ , for some  $j_0$  then, it is easy to prove that the point  $x^k$  is an accumulation point of the sequence  $\{x^k\}$  and therefore a critical point (minimum in the convex case) of the problem (3.1). If the algorithm finishes in a finite number of iterations, the actual point is a critical point (minimum in the convex case) of the function or an approximate critical point (approximate minimum in the convex case), that is,  $e^{k+1} \in \partial_{\epsilon_k} f(x^{k+1})$ .

Coercivity is very important for two points of view, theoretical and applied: At the mathematical level, a coercivity assumption implies that for each  $c \in \mathbb{R}$ , the level set  $\{x : f(x) \leq c\}$  is bounded and, therefore, the sequence generated by any descent algorithm (in particular the proximal point algorithm) is bounded. On the other hand, to obtain convergence results using a Kurdyka-Lojasiewicz (KL) inequality it is needed, see for example (Attouch et al., 2013; Bento & Soubeyran, 2015a, b; Moreno et al., 2011), to take the sequence generated by the algorithm to be bounded. Then, coercivity is a condition that helps the KL condition to obtain global convergence results.

At the moment, it is not known if in the absence of (3.2), but under the assumption that the function satisfies the Kurdyka-Lojasiewicz property on the set of accumulation points, is possible to obtain the same results of convergence. This may be a future work for our research. However, a result in this direction was given in the paper authored by Soubeyran and Souza (2020). The authors, using a  $w$ -distance (and in the absence of the hypotheses (3.2)) as regularized term, proved the convergence of bounded sequences generated by a general descent method proposed by Attouch et al. (2013).

In this paper, we introduce two new subdifferentials, the generalized limiting subdifferential  $\bar{\partial} f(x)$ , see Definition 2.1,(e), and the  $\epsilon$ -Clarke subdifferential in (5.1) which is different between all the  $\epsilon$ -Clarke subdifferential introduced by other authors. In this sense, the study of the properties of these subdifferentials are welcome.

**Acknowledgements** The first author is grateful to FAPERJ-CAPES by the economic support on the Pos-Doctoral Project PAPD-2011, Process *no E - 26/100.057/2013 - BBP*. The work of the second author was supported by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A\*MIDEX. The research of the third author was partially support by CNPq (Grant 302678/2017-4). The last version of the paper was written when the first author was working as a visiting professor at IME-Federal University of Goias.

## References

- Alfuraidan, M. R., Khamsi, M. A., & Soubeyran, A. (2020). Ekeland variational principle on quasi-weighted graphs. Improving the work- family balance. *Accepted in the Journal of Fixed Point Theory and Applications*.
- Attouch, H., Bolte, J., & Svaiter, B. (2013). Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Mathematical Programming Series A*, 137, 91–129.
- Bao, T. Q., & Soubeyran, A. (2019). Variational principles in set optimization with domination structures and application to changing jobs. *Journal of Applied and Numerical Optimization*, 1(3), 217–241.
- Bolte, J., Daniilidis, A., Ley, O., & Mazet, L. (2010). Characterizations of Lojasiewicz inequalities: Subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362(6), 3319–3363.
- Bento, G. C., & Soubeyran, A. (2015). Generalized inexact proximal algorithms: Routine's formation with resistance to change, following worthwhile changes. *Journal of Optimization Theory and Applications*, 166(1), 172–187.
- Bento, G. C., & Soubeyran, A. (2015). A Generalized inexact proximal point method for nonsmooth functions that satisfies Kurdyka Lojasiewicz inequality. *Set Valued and Variational Analysis*, 23(3), 501–517.
- Concilio, A., & Gerla, G. (2006). Quasi-metric spaces and point-free geometry. *Mathematical Structures in Computer Science*, 16, 115–137.
- Fakhar, M., Khodakhah, M., Mazyaki, A., Soubeyran, A., & Zafarani, J. (2020). Variational rationality, variational principles and the existence of traps in a changing environment. *Accepted in JOGO*.
- Güler, O. (1992). New proximal point proximal algorithms for convex minimization. *SIAM Journal Control and Optimization*, 2, 649–664.
- Hirschi, A. & Dauwalder, J. P. (2015). Dynamics in Career development. Personal and organizational perspective. In *Book: Handbook of the life design paradigm: From practice to theory, from theory to practice*. (pp. 27–39) Boston, MA: Hogrefe.
- Janeiro, I. N. (2010). Motivational dynamics in the development of career attitudes among adolescents. *Journal of Vocational Behavior*, 76(2), 170–177.
- Levin, V. L. (1991). Some applications of set-valued mappings in mathematical economics. *Journal of Mathematical Economics*, 20, 69–87.
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw Hill.
- Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw Hill.
- Lewin, K. (1938). *The conceptual representation and measurement of psychological forces*. Durham, NC: Duke University Press.
- Lewin, K. (1951). Intention, will and need. In D. Rapaport (Ed.), *Organization and pathology of thought: Selected sources* (pp. 95–153). New York, NY: Columbia University Press.
- Martinet, B. (1970). Régularisation, d'inéquations variationnelles par approximations successives. *Revue Française d'informatique et de Recherche opérationnelle*, 154–159.
- Mordukhovich, B. S. (2006). *Variational analysis and generalized differentiation I: Basic theory* (Vol. 330). Berlin: Springer Science and Business Media.
- Moreno, F. G., Oliveira, P. R., & Soubeyran, A. (2011). A proximal algorithm with quasi distance. *Application to Habit's Formation. Optimization*, 61(12), 1383–1403.
- Papa Quiroz, E. A., & Oliveira, P. R. (2009). Proximal point methods for quasiconvex and convex functions with Bregman distances on Hadamard manifolds. *Journal of Convex Analysis*, 16(1), 49–69.
- Papa Quiroz, E. A., & Oliveira, P. R. (2012). Full convergence of the proximal point method for quasiconvex function on Hadamard manifolds. *ESAIM: Control, Optimisation and Calculus of Variations*, 18(2), 483–500.
- Polyak, B. T. (1987). *Introduction to optimization*. Optimization Software, New York. Optimization Software.
- Romaguera, S., & Sanchis, M. (2003). Applications of utility functions defined on quasi-metric spaces. *Journal of Mathematical Analysis and Applications*, 283, 219–235.
- Rockafellar, R. T. (1976). Monotone operations and the proximal point method. *SIAM Journal of Control Optimization*, 14, 877–898.
- Rockafellar, R. T., & Wets, R. (1990). *Variational analysis*. Grundlehren der Mathematischen, Wissenschaften (Vol. 317). Springer.
- Solodov, M. V., & Svaiter, B. F. (1999). A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued and Variational Analysis*, 7, 323–345.
- Stojmirović, A. (2004). Quasi-metric spaces with measure. In *Proceedings of the 18th summer conference on topology and its applications* (Vol. 28, pp. 655-671).
- Soubeyran, A. (2009). *Variational rationality, a theory of individual stability and change: worthwhile and ambidexterity behaviors*. GREQAM, Aix Marseille University. Preprint.

- Soubeyran, A. (2010). *Variational rationality and the unsatisfied man: Routines and the course pursuit between aspirations, capabilities and beliefs*. GREQAM, Aix Marseille University. Preprint.
- Soubeyran, A. (2016). *Variational rationality. A theory of worthwhile stay and change approach-avoidance transitions ending in traps*. GREQAM-AMSE, Aix Marseille University. Preprint.
- Soubeyran, A. (2021). *Variational rationality: towards a grand theory of motivation driven by worthwhile moves*. AMSE, Aix-Marseille University. Preprint.
- Soubeyran, A. (2021b). Variational rationality: the concepts of motivation and motivational force. AMSE, Aix-Marseille University. Preprint.
- Soubeyran, A. (2021c). Variational rationality. The resolution of goal conflicts via stop and go approach-avoidance dynamics. AMSE, Aix-Marseille University. Preprint.
- Soubeyran, A. (2021d). Variational rationality. A general theory of moving goals and intentions as satisficing worthwhile moves. AMSE, Aix-Marseille University. Preprint.
- Soubeyran, A., & Souza, J. C. (2020). Generalized descent method using w-distance. Application to emergence of habits following worthwhile moves. *Journal of Nonlinear and Variational Analysis*, 4(2), 285–300.