



**HAL**  
open science

# Relaxed-Inertial Proximal Point Algorithms for Nonconvex Equilibrium Problems with Applications

Sorin-Mihai Grad, Felipe Lara, Raúl Marcavillaca

► **To cite this version:**

Sorin-Mihai Grad, Felipe Lara, Raúl Marcavillaca. Relaxed-Inertial Proximal Point Algorithms for Nonconvex Equilibrium Problems with Applications. *Journal of Optimization Theory and Applications*, In press. hal-04429671

**HAL Id: hal-04429671**

**<https://hal.science/hal-04429671v1>**

Submitted on 31 Jan 2024

**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.

Copyright

# Relaxed-Inertial Proximal Point Algorithms for Nonconvex Equilibrium Problems with Applications

Sorin-Mihai Grad\*   Felipe Lara †‡   Raúl Marcavillaca§

December 29, 2023

**Abstract** We propose a relaxed-inertial proximal point algorithm for solving equilibrium problems involving bifunctions which satisfy in the second variable a generalized convexity notion called strong quasiconvexity, introduced by Polyak in 1966. The method is suitable for solving mixed variational inequalities and inverse mixed variational inequalities involving strongly quasiconvex functions, as these can be written as special cases of equilibrium problems. Numerical experiments where the performance of the proposed algorithm outperforms the one of the standard proximal point methods are provided, too.

Communicated by Alexander Vladimirovich Gasnikov.

**Keywords:** Proximal point algorithms; inertial algorithms; equilibrium problems; nonconvex optimization; quasiconvexity

*In loving memory of Boris Teodorovich Polyak.*

## 1 Introduction

Let  $K \subseteq \mathbb{R}^n$  be a nonempty, closed and convex set and  $f : K \times K \rightarrow \mathbb{R}$  a bifunction. Then the *equilibrium problem* is defined by

$$\text{find } \bar{x} \in K : f(\bar{x}, y) \geq 0, \forall y \in K. \quad (\text{EP}) \quad \boxed{\text{EPO}}$$

Problem (EP) is a very general formulation which includes several mathematical models from continuous optimization and variational problems such as minimization problems, variational inequalities, minimax problems, fixed point

---

\*Unité de Mathématiques Appliquées, ENSTA Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France. E-mail: sorin-mihai.grad@ensta-paris.fr, ORCID-ID: 0000-0002-1139-7504

†Corresponding Author

‡Instituto de Alta investigación (IAI), Universidad de Tarapacá, Arica, Chile. E-mail: felipelaraobrequé@gmail.com; flarao@uta.cl. Web: felipelara.cl, ORCID-ID: 0000-0002-9965-0921

§Instituto de Alta investigación (IAI), Universidad de Tarapacá, Arica, Chile. E-mail: raultm.rt@gmail.com, ORCID-ID: 0000-0003-3748-0768

problems and complementarity problems among others (see the excellent seminal work [6] for more on this).

In particular, (EP) includes also two families of mixed variational inequalities in which we have special interest since applications in several fields of engineering and economics among others (see [1, 21, 22, 31, 15, 34, 38]) can be cast as such. Recall that given an operator  $T : K \rightarrow \mathbb{R}^n$  and a function  $\varphi : K \rightarrow \mathbb{R}$ , the *mixed variational inequality* (MVI) problem is defined as

$$\text{find } \bar{x} \in K : \langle T(\bar{x}), y - \bar{x} \rangle + \varphi(y) - \varphi(\bar{x}) \geq 0, \forall y \in K, \quad (\text{MVI}) \quad \boxed{\text{MVI}}$$

while the *inverse mixed variational inequality* (IMVI) problem is defined as

$$\text{find } \bar{x} \in K : \langle \bar{x}, T(y) - T(\bar{x}) \rangle + \varphi(y) - \varphi(\bar{x}) \geq 0, \forall y \in K. \quad (\text{IMVI}) \quad \boxed{\text{IMVI}}$$

Both problems are very general formulations, too, and have been intensively studied in the literature (see [1, 17, 21, 22, 36, 38, 51, 54] among others) since they encompass variational inequalities and inverse variational inequalities as a special case (when  $\varphi = 0$ , see [12, 13]); constrained optimization problems (when  $T = 0$ ) and even the minimization of the sum of two functions (when  $T$  is the gradient of a differentiable function defined on  $K$  in problem (MVI)). Furthermore, problem (IMVI) generalizes standard mathematical programming problems such as the extended linear-quadratic programming studied in [42] (see [21]). For all these reasons, problem (EP) became the object of study for many researchers in applied mathematics, engineering, economics and computer sciences, among others, especially, when the involved bifunction  $f$  is assumed to be convex on its second argument.

The convexity assumption in problems (EP), (MVI) and (IMVI) is standard and hard to be dropped since even usual results may fail without this hypothesis. Indeed, if  $\varphi$  is not convex, then problem (MVI) may have no solutions even on convex and compact sets (see the example in [26, page 127]) and the same happens with (IMVI). The reason behind this is the “conflict” between the properties of the function  $\varphi$  and those of the operator  $T$  (see [17, 26, 51]), that occurs because one is searching for a point where an inequality is satisfied for all possible values of a variable, so, if the function  $\varphi$  is monotonically decreasing on a certain subset of  $K$ , the operator needs to “go upwards” faster on the same subset, and vice-versa. As a consequence, an assumption regarding the relation between the involved function and operator needs to be imposed in order to ensure the existence of solutions to (MVI) and (IMVI) when  $\varphi$  is not convex even on convex and compact sets. Usually in the literature (see [17, 26, 27] and references therein), the so-called *mixed variational inequality property* introduced in [51, Definition 3.1]) is employed for this. Furthermore, similar obstacles or even higher are expected for algorithms for dealing with such problems, as indicated, for instance, by the projected gradient method for solving variational inequalities ( $\varphi = 0$  in problem (MVI)) where  $T$  is assumed  $\mu$ -strongly monotone and  $L$ -Lipschitz-continuous (with  $L, \mu > 0$ ) which converges when  $L < \sqrt{2\mu}$  (see for instance [13, page 1109]).

In this paper, we propose an iterative method for solving problem (EP) without the usual convexity assumption. In particular, we use the generalized convexity notion of *strong quasiconvexity*, a class of functions broader than the *strongly convex* ones.

Strongly quasiconvex functions were introduced by Polyak in his famous paper [39] and were investigated by different authors who provided various results involving them in works like [30, 50]. Moreover, as anticipated by Polyak in [39], there is also a significant literature on the calculus of variations involving strongly quasiconvex functions (see [32] for instance). New studies involving strongly quasiconvex functions in the field of optimization can be found, for instance, in [11], which provides asymptotic convergence properties of the saddle-point dynamics involving bifunctions which are strongly quasiconvex in one or both variables, and in [43], where one finds results on the asymptotic behavior of quasi-autonomous gradient expansive systems governed by a strongly quasiconvex function. The recent work [33] of one of the authors of this paper provides a fresh view on this class of functions, by recalling the most important existing results and providing a first investigation on minimizing its members by means of modern iterative methods. More precisely, there it is shown that the classical proximal point algorithm converges under standard assumptions when employed for minimizing a strongly quasiconvex function, too, while in [18] the authors of this study obtained a similar conclusion for a relaxed-inertial proximal point method.

Different proximal point type algorithms have been considered for solving equilibrium problems in the convex case, in some of them the employed bifunction being endowed with a generalized monotonicity property rather than the classical monotonicity. There are only a few (very recent) works where iterative methods for solving equilibrium problems with the underlying function (strongly) quasiconvex can be found, namely [28], where arguably the first proximal point method in such a framework is proposed; [52] with a subgradient method; [53] that contains a proximal subgradient algorithm and the recent contribution [4], which contains an extension of [28] to Bregman distances.

The main contribution of this work consists in showing that the relaxed-inertial proximal point algorithm for solving equilibrium problems (RIPPA-EP from now on) introduced in [2, 3] for solving monotone inclusion problems and extended to equilibrium problems in [24, 48] in the convex case remains convergent, under slightly different, but still standard assumptions, when the employed bifunction is strongly quasiconvex in its second variable and pseudomonotone, following thus in a certain sense our recent works [18, 24, 28, 33]. Therefore, the influence of B.T. Polyak's work on this paper is twofold, because the inertial proximal algorithms are based on his heavy ball method from [40]. Numerical experiments for solving problems (MVI) and (IMVI) where the involved function  $\varphi$  is strongly quasiconvex are presented, and the performance of the proposed algorithm is shown to be superior to the one of the standard proximal point method.

The paper is structured as follows. In Section 2, we introduce notations and preliminary results. In Section 3, the proposed algorithm is presented and its

convergence is investigated. Relevant special cases are discussed, and sufficient conditions that guarantee the fulfillment of the hypotheses of the convergence statements are provided, too. Section 4 contains computational results that illustrate the theoretical achievements.

## 2 Preliminaries and Basic Definitions

sec:2

We denote by  $\langle \cdot, \cdot \rangle$  the *inner product* of  $\mathbb{R}^n$ , while the corresponding *Euclidean norm* is  $\|\cdot\|$ . For  $x, y, z \in \mathbb{R}^n$  and  $\beta \in \mathbb{R}$ , one has

$$\langle x - z, y - x \rangle = \frac{1}{2}\|z - y\|^2 - \frac{1}{2}\|x - z\|^2 - \frac{1}{2}\|y - x\|^2, \quad (2.1)$$

3:points

$$\|\beta x + (1 - \beta)y\|^2 = \beta\|x\|^2 + (1 - \beta)\|y\|^2 - \beta(1 - \beta)\|x - y\|^2. \quad (2.2)$$

iden:1

Consider a function  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ . The *effective domain* of  $h$  is  $\text{dom } h := \{x \in \mathbb{R}^n : h(x) < +\infty\}$  and we call  $h$  *proper* when it has a nonempty domain and takes nowhere the value  $-\infty$ . The standard conventions  $\sup_{\emptyset} h := -\infty$ ,  $\inf_{\emptyset} h := +\infty$ , and  $0 \cdot (+\infty) = +\infty$ ,  $0 \cdot (-\infty) = 0$ , and  $+\infty + (-\infty) = +\infty$  are considered. The *sublevel set of  $h$  at height  $\lambda \in \mathbb{R}$*  is  $S_\lambda(h) := \{x \in \mathbb{R}^n : h(x) \leq \lambda\}$ , while its set *minimizers* is  $\arg \min_{\mathbb{R}^n} h$ . Denote by  $\text{Id} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  the identity operator on  $\mathbb{R}^n$ , and by  $K^* := \{y \in \mathbb{R}^n : \langle y, x \rangle \geq 0, \forall x \in K\}$  the positive *dual cone* associated to a set  $K \subseteq \mathbb{R}^n$ .

When  $\text{dom } h$  is convex, we say that  $h$  is

- (a) *convex*, when for every  $x, y \in \text{dom } h$

$$h(\lambda x + (1 - \lambda)y) \leq \lambda h(x) + (1 - \lambda)h(y), \forall \lambda \in [0, 1]; \quad (2.3)$$

def:convex

- (b) *semistrictly quasiconvex* if, given every  $x, y \in \text{dom } h$ , with  $h(x) \neq h(y)$ , then

$$h(\lambda x + (1 - \lambda)y) < \max\{h(x), h(y)\}, \forall x, y \in \text{dom } h, \forall \lambda \in ]0, 1[, \quad (2.4)$$

- (c) *quasiconvex*, when for every  $x, y \in \text{dom } h$

$$h(\lambda x + (1 - \lambda)y) \leq \max\{h(x), h(y)\}, \forall \lambda \in [0, 1]; \quad (2.5)$$

def:qcx

- (d) *strongly convex* [39] (on  $\text{dom } h$ ) with *modulus*  $\gamma > 0$ , when there is a  $\gamma \in ]0, +\infty[$  for which

$$h(\lambda y + (1 - \lambda)x) \leq \lambda h(y) + (1 - \lambda)h(x) - \lambda(1 - \lambda)\frac{\gamma}{2}\|x - y\|^2, \forall x, y \in \text{dom } h, \forall \lambda \in [0, 1];$$

- (e) *strongly quasiconvex* [39] (on  $\text{dom } h$ ) with *modulus*  $\gamma > 0$ , when there is a  $\gamma \in ]0, +\infty[$  for which

$$h(\lambda y + (1 - \lambda)x) \leq \max\{h(y), h(x)\} - \lambda(1 - \lambda)\frac{\gamma}{2}\|x - y\|^2, \forall x, y \in \text{dom } h, \forall \lambda \in [0, 1].$$

When (2.3) (or (2.5)) is strictly fulfilled whenever  $x \neq y$ , we call  $h$  *strictly (quasi)convex*.

The following scheme exhibits the known relations among the mentioned (generalized) convexity concepts (where “qcx” stands for quasiconvex)

$$\begin{array}{ccccccc}
\text{strongly convex} & \implies & \text{strictly convex} & \implies & \text{convex} & \implies & \text{qcx} \\
\downarrow & & \downarrow & & \downarrow & & \\
\text{strongly qcx} & \implies & \text{strictly qcx} & \implies & \text{semistrictly qcx} & & \\
& & \downarrow & & & & \\
& & \text{qcx} & & & & 
\end{array}$$

In [30, Theorem 2] one can see that the Euclidean norm is strongly quasiconvex on any bounded convex set, but not strongly convex. Note also that a quasiconvex function can be seen as a strongly quasiconvex one with modulus  $\gamma = 0$ , so “ $h$  is strongly quasiconvex with modulus  $\gamma \geq 0$ ” covers both these notions.

Examples of strongly quasiconvex functions which are not convex can be found in [18, 28, 29, 33, 53, 30]. For instance, the function  $x \in \mathbb{R}^n \mapsto \sqrt{\|x\|}$ , which has been employed in studies on information protection (see [49]) or on Gram-Schmidt orthogonalization methods (see [47]), is (see [33, Theorem 17]) strongly quasiconvex on any bounded and convex subset of  $\mathbb{R}^n$  and, as we will see in Section 4, some classes of fractional functions (intensively studied for their economic applications, see [44, 46]) are strongly quasiconvex (see [29]), too.

A proper function  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is said to be *2-supercoercive* when

$$\liminf_{\|x\| \rightarrow +\infty} \frac{h(x)}{\|x\|^2} > 0.$$

According to [33, Theorem 1], any function that is strongly quasiconvex on a convex set  $K \subseteq \mathbb{R}^n$  is also 2-supercoercive. Hence, lower semicontinuous strongly quasiconvex functions have exactly one minimizer on closed convex finitely-dimensional subsets (see [33, Corollary 3]), an observation useful when investigating proximal point type methods for optimization problems involving such functions (see [18, 28, 33]).

Take  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  proper and  $K \subseteq \mathbb{R}^n$  closed and convex, such that  $K \subseteq \text{dom } h$ . The *proximal operator of  $h$  on  $K$  of parameter  $\beta > 0$  at  $x \in \mathbb{R}^n$*  is

$$\text{Prox}_{\beta h} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n, \quad \text{Prox}_{\beta h}(K, x) = \underset{y \in K}{\text{argmin}} \left\{ h(y) + \frac{1}{2\beta} \|y - x\|^2 \right\}.$$

When  $K = \mathbb{R}^n$ , we write  $\text{Prox}_{\beta h}(\mathbb{R}^n, \cdot)$  simpler as  $\text{Prox}_{\beta h}(\cdot)$ .

The next result from [33] will be necessary later.

**Lemma: HL**

**Lemma 1** (cf. [33, Proposition 6]) *Let  $K$  be a closed and convex set in  $\mathbb{R}^n$ ,  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be a proper, lower semicontinuous, strongly quasiconvex function with modulus  $\gamma \geq 0$  and such that  $K \subseteq \text{dom } h$ ,  $\beta > 0$  and  $x \in K$ . If  $\bar{x} \in \text{Prox}_{\beta h}(K, x)$ , then for all  $y \in K$  and all  $\lambda \in [0, 1]$  one has*

$$h(\bar{x}) - \max\{h(y), h(\bar{x})\} \leq \frac{\lambda}{\beta} \langle \bar{x} - x, y - \bar{x} \rangle + \frac{\lambda}{2} \left( \frac{\lambda}{\beta} - \gamma + \lambda\gamma \right) \|y - \bar{x}\|^2.$$

Different to the convex case, it has been shown in [33, (1.4)] that the sum of a quasiconvex and a half of a squared norm is not necessarily a strongly quasiconvex function. Additionally, [33, Remark 6] exhibits an example where not even the sum of a strongly quasiconvex function and a half of a squared norm is strongly quasiconvex. Consequently, (see also [33, Remark 19(ii)]) the proximal operator of a strongly quasiconvex function is usually set-valued, containing a singleton when the sum of the strongly quasiconvex function and a half of a squared norm is strongly quasiconvex.

A bifunction  $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  is *monotone on*  $K \subseteq \mathbb{R}^n$ , when

$$f(x, y) + f(y, x) \leq 0 \quad \forall x, y \in K.$$

When the following weaker condition (see [20] for counter-examples for the opposite implication) is fulfilled

$$f(x, y) \geq 0 \implies f(y, x) \leq 0 \quad \forall x, y \in K,$$

we say that  $f$  is *pseudomonotone on*  $K$ .

The next statement was essentially proven in [3, Theorem 2.1] and it will be useful for analyzing the convergence of the proposed algorithm.

lm:alv.att

**Lemma 2** *Let the sequences  $\{\varphi_k\}_k$ ,  $\{s_k\}_k$ ,  $\{\alpha_k\}_k$  and  $\{\delta_k\}_k$  in  $[0, +\infty[$  and let  $\alpha \in \mathbb{R}$  be such that  $0 \leq \alpha_k \leq \alpha < 1$  and*

$$\varphi_{k+1} - \varphi_k + s_{k+1} \leq \alpha_k(\varphi_k - \varphi_{k-1}) + \delta_k, \quad \forall k \geq 0. \quad (2.6)$$

eq:alv.att02

*Then the following assertions hold*

- (a) *for all  $k \geq 1$ , we have  $\varphi_k + \sum_{j=1}^k s_j \leq \varphi_0 + \frac{1}{1-\alpha} \sum_{j=0}^{k-1} \delta_j$ ;*
- (b) *if  $\sum_{k=0}^{\infty} \delta_k < +\infty$ , then  $\lim_{k \rightarrow \infty} \varphi_k$  exists, i.e., the sequence  $\{\varphi_k\}_k$  converges to some element in  $[0, +\infty[$ .*

More on generalized convexity (in particular (strong) quasiconvexity) and generalized monotonicity can be found, for instance, in [10, 20, 33, 30, 45, 39, 50].

### 3 Relaxed-Inertial Proximal Point Type Algorithm

sec:3

Let  $K$  be a nonempty set in  $\mathbb{R}^n$  and  $f : K \times K \rightarrow \mathbb{R}$  be a bifunction. Consider the *equilibrium problem*

$$\text{find } \bar{x} \in K : f(\bar{x}, y) \geq 0, \quad \forall y \in K, \quad (\text{EP})$$

EP

and denote by  $S(K, f)$  the set of solutions of (EP).

The following assumptions on the mathematical objects involved in the equilibrium problem are necessary for investigating the convergence of the algorithm we propose for solving (EP)

- (A1)  $y \mapsto f(\cdot, y)$  is upper semicontinuous for all  $y \in K$ ;
- (A2)  $f$  is pseudomonotone on  $K$ ;
- (A3)  $f$  is lower semicontinuous (jointly in both arguments);
- (A4)  $f(x, \cdot)$  is strongly quasiconvex on  $K$  with modulus  $\gamma > 0$  whenever  $x \in K$ ;
- (A5)  $f$  satisfies the following Lipschitz-continuity type condition: there exists  $\eta > 0$  for which

$$f(x, z) - f(x, y) - f(y, z) \leq \eta(\|x - y\|^2 + \|y - z\|^2), \quad \forall x, y, z \in K. \quad (3.1)$$

Lips:cond

As was noted in [28, Remark 3.1(i)], the following usual assumption when dealing with equilibrium problems is a direct consequence of (A2) and (A5). We mention it, as it will be employed in some intermediate results which do not require the satisfaction of (A5).

- (A0)  $f(x, x) = 0$  for all  $x \in K$ .

Moreover, one usually augments (A1) with the following hypothesis:

- (A1')  $f(x, \cdot)$  is lower semicontinuous for all  $x \in K$ ,

that is a consequence of (A3) and will be considered when the latter is not required.

While (A1), (A2), (A3) and (A5) are standard assumptions in the related literature, (A4) relaxes the standard convexity hypothesis usually imposed on  $f$  in its second argument in the literature.

Some comments on these hypotheses follow (see also [28, Remark 3.1]).

rem:assumptions

**Remark 3** (i) We know by [28, Proposition 3.1] that if  $f$  satisfies (Ai) with  $i = 0, 1, 1', 2, 4$ , then  $S(K, f) \neq \emptyset$  and is actually a singleton.

- (ii) When  $f(x, y) = \langle T(x), y - x \rangle$ , with  $L$ -Lipschitz continuous  $T : K \rightarrow \mathbb{R}^n$  ( $L > 0$ ), relation (3.1) is fulfilled for  $\eta = L/2$  (see [41]).

The relaxed-inertial proximal point type method we propose for solving (EP) is the following.



ppa:ep

---

**Algorithm 1** RIPPA-EP for Strongly Quasiconvex EP's

---

**Step 0.** (Initialization). Let  $x^0, x^{-1} \in K$ ,  $\alpha, \rho \in [0, 1[$  and  $\{\beta_k\}_{k \in \mathbb{N}} \subseteq ]0, +\infty[$ ,  $k = 0$ .

**Step 1.** Choose  $\alpha_k \in [0, \alpha]$ , set

$$y^k = x^k + \alpha_k(x^k - x^{k-1}), \quad [\text{extrapolation step}] \quad (3.2) \quad \text{step:extr01}$$

and compute

$$z^k \in \underset{x \in K}{\operatorname{argmin}} \left\{ f(y^k, x) + \frac{1}{2\beta_k} \|y^k - x\|^2 \right\}. \quad [\text{proximal step}] \quad (3.3) \quad \text{step:sqcx1}$$

**Step 2.** If  $y^k = z^k$ , then STOP. Otherwise, go to Step 3.

**Step 3.** Choose some relaxation parameter  $\rho_k \in [1 - \rho, 1 + \rho]$ , and update

$$x^{k+1} = (1 - \rho_k)y^k + \rho_k z^k. \quad [\text{relaxation step}] \quad (3.4) \quad \text{eq:relax.step}$$

**Step 4.** Let  $k \leftarrow k + 1$  and go to Step 1.

---

Next, we point out some observations regarding Algorithm 1.

rmk:EP1

**Remark 4** (i) *Step 1 in Algorithm 1 is actually a proximal point step with inertial effects applied to the function  $f(y^k, \cdot)$  at  $y^k$  over the set  $K$ , for every  $k \geq 0$ . Recall that inertial effects that improve the performance of proximal point algorithms were introduced by Alvarez and Attouch in [3] and are based on the heavy ball method introduced by Polyak in [40]. Even though the proximal operator is mostly considered in the literature with respect to the whole underlying space, in papers such as [7, 17, 18] it is taken on a certain set. The same algorithm has been proposed for solving (EP) in [24] (of which we became aware only when this paper was complete, hence the possible redundancies) as a relaxed inertial extragradient method and [48], via an explicit discretization of a second-order dynamical system, under different hypotheses that include the convexity of  $f$  in the second variable. Even if this is not explicitly acknowledged, in these works the proximal operator is considered with respect to the underlying set of (EP), too. Other relaxed inertial proximal point type algorithms for solving convex equilibrium problems were introduced in [23, 37].*

(ii) *In virtue of relations (3.2) and (3.4),  $K$  should be assumed to be an affine subspace and not only as a closed and convex set. We note that the authors in [2, 3, 16, 35] considered their algorithms on the whole space, while in [24, 48]  $K$  was taken closed and convex, however the issue of staying feasible after the extrapolation step does not seem to have been taken into consideration.*

(iii) If  $\alpha = 0$ , then Algorithm 1 reduces to the standard proximal point algorithm in [28] to which a relaxation step (that vanishes when  $\rho_k = 1$  for all  $k \geq 0$ ) is added.

### 3.1 Convergence Analysis

sec:3-1

As a first result, we validate the stopping criterion of Algorithm 1 whose proof is skipped as it is similar to the one of [28, Proposition 3.2].

stop:criteria

**Proposition 5** Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  sequences of positive numbers,  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  the sequences generated by Algorithm 1 and suppose that assumptions (A0), (A1), (A1') and (A5) are valid. If  $y^k = z^k$ , then  $x^{k+1} = y^k$  is solution of (EP), i.e.,  $\{y^k\} = S(K, f)$ .

The following result is the starting point of our analysis.

lmm:auxR1

**Proposition 6** Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  be sequences of positive numbers,  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  be the sequences generated by Algorithm 1 and suppose that assumptions (Ai) with  $i = 1, 1', 2, 4, 5$  hold. Take  $\{\bar{x}\} = S(K, f)$ . Then for every  $k \geq 0$ , at least one of the following inequalities holds

$$\|x^{k+1} - \bar{x}\|^2 \leq \|y^k - \bar{x}\|^2 - \frac{2 - \rho_k}{\rho_k} \|x^{k+1} - y^k\|^2 - \frac{\rho_k(\gamma\beta_k - 1)}{2} \|z^k - \bar{x}\|^2, \quad (3.5)$$

conclusion:01

$$\begin{aligned} \|x^{k+1} - \bar{x}\|^2 \leq & \|y^k - \bar{x}\|^2 - \frac{2 - 4\eta\beta_k - \rho_k}{\rho_k} \|x^{k+1} - y^k\|^2 \\ & - \frac{\rho_k(\gamma\beta_k - 1 - 8\eta\beta_k)}{2} \|z^k - \bar{x}\|^2. \end{aligned} \quad (3.6)$$

conclusion:02

**Proof.** First, note that the relaxation step (3.4) of Algorithm 1 yields

$$x^{k+1} - \bar{x} = (1 - \rho_k)(y^k - \bar{x}) + \rho_k(z^k - \bar{x}).$$

Then using identity (2.2), we have

$$\|x^{k+1} - \bar{x}\|^2 = (1 - \rho_k)\|y^k - \bar{x}\|^2 + \rho_k\|z^k - \bar{x}\|^2 - \rho_k(1 - \rho_k)\|y^k - z^k\|^2,$$

or equivalently,

$$\|x^{k+1} - \bar{x}\|^2 - \|y^k - \bar{x}\|^2 = \rho_k(\|z^k - \bar{x}\|^2 - \|y^k - \bar{x}\|^2) - \rho_k(1 - \rho_k)\|y^k - z^k\|^2.$$

Furthermore, it follows from relation (2.1) that

$$\|z^k - \bar{x}\|^2 - \|y^k - \bar{x}\|^2 = \|z^k - y^k\|^2 + 2\langle z^k - y^k, y^k - \bar{x} \rangle.$$

Therefore, combining the last two inequalities we obtain

$$\|x^{k+1} - \bar{x}\|^2 - \|y^k - \bar{x}\|^2 = 2\rho_k\langle z^k - y^k, y^k - \bar{x} \rangle + \rho_k^2\|y^k - z^k\|^2. \quad (3.7)$$

eq:tm01

On the other hand, for every  $k \geq 0$ , step (3.3) and Lemma 1 yield

$$\begin{aligned} f(y^k, z^k) - \max\{f(y^k, z^k), f(y^k, \bar{x})\} &\leq \frac{\lambda}{\beta_k} \langle z^k - y^k, \bar{x} - z^k \rangle \\ &\quad + \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|z^k - \bar{x}\|^2, \forall \lambda \in [0, 1]. \end{aligned}$$

We consider two cases.

(i) If  $f(y^k, z^k) \geq f(y^k, \bar{x})$ , then

$$\begin{aligned} \frac{1}{\beta_k} \langle z^k - y^k, z^k - \bar{x} \rangle &\leq \frac{1}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|z^k - \bar{x}\|^2, \forall \lambda \in ]0, 1[ \\ \stackrel{\lambda=\frac{1}{2}}{\implies} \quad \langle z^k - y^k, z^k - \bar{x} \rangle &\leq \frac{1 - \gamma\beta_k}{4} \|z^k - \bar{x}\|^2, \\ \implies \langle z^k - y^k, y^k - \bar{x} \rangle &\leq -\|z^k - y^k\|^2 + \frac{1 - \gamma\beta_k}{4} \|z^k - \bar{x}\|^2. \end{aligned} \quad (3.8) \quad \boxed{\text{eq: TM1}}$$

Thus, combining relations (3.7) and (3.8), and then using the fact  $z^k - y^k = (1/\rho_k)(x^{k+1} - y^k)$ , we obtain (3.5).

(ii) If  $f(y^k, z^k) < f(y^k, \bar{x})$ , then

$$\begin{aligned} 0 &\leq \frac{\lambda}{\beta_k} \langle z^k - y^k, \bar{x} - z^k \rangle + \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|z^k - \bar{x}\|^2 \\ &\quad + f(y^k, \bar{x}) - f(y^k, z^k) \\ &\leq \frac{\lambda}{\beta_k} \langle z^k - y^k, \bar{x} - z^k \rangle + \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|z^k - \bar{x}\|^2 \\ &\quad + f(z^k, \bar{x}) + \eta (\|z^k - y^k\|^2 + \|z^k - \bar{x}\|^2), \end{aligned}$$

where the last inequality follows from assumption (A5). Furthermore, since  $\{\bar{x}\} = S(K, f)$  and  $f$  is pseudomonotone,  $f(z^k, \bar{x}) \leq 0$ , so that

$$\begin{aligned} 0 &\leq \frac{\lambda}{\beta_k} \langle z^k - y^k, \bar{x} - z^k \rangle + \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|z^k - \bar{x}\|^2 \\ &\quad + \eta (\|z^k - y^k\|^2 + \|z^k - \bar{x}\|^2), \forall \lambda \in [0, 1]. \end{aligned}$$

This implies

$$\begin{aligned} \frac{\lambda}{\beta_k} \langle z^k - y^k, z^k - \bar{x} \rangle &\leq \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|z^k - \bar{x}\|^2 + \eta (\|z^k - y^k\|^2 + \|z^k - \bar{x}\|^2) \\ &= \eta \|z^k - y^k\|^2 + \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma + \frac{2\eta}{\lambda} \right) \|z^k - \bar{x}\|^2 \\ \stackrel{\lambda=\frac{1}{2}}{\implies} \quad \langle z^k - y^k, y^k - \bar{x} \rangle &\leq (2\eta\beta_k - 1) \|z^k - y^k\|^2 \\ &\quad + \left( 2\eta\beta_k + \frac{1 - \gamma\beta_k}{4} \right) \|z^k - \bar{x}\|^2. \end{aligned} \quad (3.9) \quad \boxed{\text{eq: TM2}}$$

Finally, combining (3.9) with (3.7) and using  $z^k - y^k = (1/\rho_k)(x^{k+1} - y^k)$ , the desired inequality (3.6) holds.

Therefore, for every  $k \geq 0$ , we obtain (3.5) or (3.6). ■

**rmk:bound**

**Remark 7** (i) *Observe that*

$$\frac{\rho_k(\gamma\beta_k - 1 - 8\eta\beta_k)}{2} > 0 \iff \beta_k > \frac{1}{\gamma - 8\eta}.$$

By this and since  $\rho_k > 0$  for every  $k \geq 0$ , we have  $\rho_k(\gamma\beta_k - 1) \geq \rho_k(\gamma\beta_k - 1 - 8\eta\beta_k) > 0$ . Hence, the last terms in the right-hand side of inequalities (3.5) and (3.6) are nonnegative, i.e.,

$$\frac{\rho_k(\gamma\beta_k - 1)}{2} \|z^k - \bar{x}\|^2 \geq 0 \quad \text{and} \quad \frac{\rho_k(\gamma\beta_k - 1 - 8\eta\beta_k)}{2} \|z^k - \bar{x}\|^2 \geq 0.$$

(ii) *Since  $\{\rho_k\}_k \subseteq [1 - \rho, 1 + \rho]$ , we have*

$$\frac{2 - 4\eta\beta_k - \rho_k}{\rho_k} \geq \frac{2 - 4\eta\beta_k - (1 + \rho)}{1 + \rho} = \frac{1 - 4\eta\beta_k - \rho}{1 + \rho} > 0,$$

*if and only if there exists  $\epsilon > 0$  such that  $\beta_k < \epsilon \leq 1/(4\eta)$  and  $0 \leq \rho \leq 1 - 4\eta\epsilon$ . So,*

$$\frac{1 - 4\eta\beta_k - \rho}{1 + \rho} \geq \frac{1 - 4\eta\epsilon - \rho}{1 + \rho} \geq 0 \iff 0 \leq \rho \leq 1 - 4\eta\epsilon.$$

As a consequence of the previous remark, and in order to facilitate the convergence analysis of our proposed algorithm, we assume that the parameter sequences  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  satisfy the following hypotheses: there exists  $\epsilon > 0$  such that

$$(C1) \quad \frac{1}{\gamma - 8\eta} < \beta_k < \epsilon \leq \frac{1}{4\eta} \text{ for every } k \geq 0;$$

$$(C2) \quad 0 < 1 - \rho \leq \rho_k \leq 1 + \rho \text{ with } 0 \leq \rho \leq 1 - 4\eta\epsilon.$$

We immediately note the following connections of these hypotheses with the conditions required for the convergence of the standard proximal point algorithm for solving (EP) proposed in [28].

**rem:link**

**Remark 8** (i) *If we consider  $\epsilon = \frac{1}{4\eta}$ , then  $\rho = 0$ ,  $\rho_k = 1$  for all  $k \geq 0$  and (C1) becomes  $\frac{1}{\gamma - 8\eta} < \beta_k < \frac{1}{4\eta}$  for every  $k \geq 0$ , which is actually [28, Assumption (A6)] for the usual proximal point algorithm for pseudomonotone equilibrium problems. Note also that the hypotheses considered in [48, Section 5] or [24, Section 3] for ensuring the convergence of Algorithm 1 in the convex case are not significantly less complicated than the ones imposed in this paper.*

(ii) Assumption (C1) (in particular, [28, Assumption (A6)]) is related to (MVI) and (IMVI) and their variants, when the involved function  $\varphi$  is nonconvex and, as we explained before, this assumption mediates the “conflict” between the operator  $T$  and the function  $\varphi$  (e.g. in Example 25), which in general could lead to equilibrium problems without solutions, even when they are defined over convex compact sets.

As a consequence of Proposition 6 and Remark 7(i), we have the following statement.

prop:auxR1

**Corollary 9** Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $f$  such that assumptions (Ai) with  $i = 1, 1', 2, 4, 5$  hold,  $\{\alpha_k\}_k$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  sequences of positive numbers such that assumption (C1) holds and  $\{x^k\}_k$  and  $\{y^k\}_k$  the sequences generated by Algorithm 1. Taking  $\{\bar{x}\} = S(K, f)$ , then for every  $k \geq 0$ , at least one of the following inequalities holds

$$\|x^{k+1} - \bar{x}\|^2 \leq \|y^k - \bar{x}\|^2 - \frac{2 - \rho_k}{\rho_k} \|x^{k+1} - y^k\|^2, \quad (3.10)$$

conclusion:T1

$$\|x^{k+1} - \bar{x}\|^2 \leq \|y^k - \bar{x}\|^2 - \frac{2 - 4\eta\beta_k - \rho_k}{\rho_k} \|x^{k+1} - y^k\|^2. \quad (3.11)$$

conclusion:T2

The following result will be useful later.

prop:01

**Proposition 10** Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $f$  such that assumptions (Ai) with  $i = 1, 1', 2, 4, 5$  hold,  $\{\alpha_k\}_k$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  sequences of positive numbers such that assumption (C1) holds and  $\{x^k\}_k$  and  $\{y^k\}_k$  the sequences generated by Algorithm 1. Take  $\{\bar{x}\} = S(K, f)$  and  $\varphi_k := \|x^k - \bar{x}\|^2$  for every  $k \geq 0$ . Then, for every  $k \geq 0$ , at least one of the following inequalities holds

$$\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) + \frac{2 - \rho_k}{\rho_k} \|x^{k+1} - y^k\|^2 \leq (\alpha_k^2 + \alpha_k) \|x^k - x^{k-1}\|^2, \quad (3.12)$$

eq:5

$$\begin{aligned} \varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) + \frac{2 - 4\eta\beta_k - \rho_k}{\rho_k} \|x^{k+1} - y^k\|^2 \\ \leq (\alpha_k^2 + \alpha_k) \|x^k - x^{k-1}\|^2. \end{aligned} \quad (3.13)$$

eq:6

**Proof.** Observe that using relation (3.2), identity (2.1) and the definition of  $\varphi_k$ , we have

$$\begin{aligned} \|y^k - \bar{x}\|^2 &= \|x^k + \alpha_k(x^k - x^{k-1}) - \bar{x}\|^2 \\ &= \|x^k - \bar{x}\|^2 + \alpha_k^2 \|x^k - x^{k-1}\|^2 + 2\alpha_k \langle x^k - \bar{x}, x^k - x^{k-1} \rangle \\ &= \varphi_k + \alpha_k^2 \|x^k - x^{k-1}\|^2 + \alpha_k(\varphi_k + \|x^k - x^{k-1}\|^2 - \varphi_{k-1}) \\ &= \varphi_k + \alpha_k(\varphi_k - \varphi_{k-1}) + (\alpha_k^2 + \alpha_k) \|x^k - x^{k-1}\|^2. \end{aligned} \quad (3.14)$$

for:first-eq

Hence, the desired inequalities (3.12) and (3.13) follow directly by replacing (3.14) in (3.10) and (3.11), respectively. ■

Our first main result, which shows that the sequence generated by Algorithm 1 converges to a solution of problem (EP) is given below.

conver:solution

**Theorem 11** Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $f$  be such that assumptions (Ai) with  $i = 1, 2, 3, 4, 5$  hold,  $\{\alpha_k\}_k$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  be sequences of positive numbers such that assumptions (Ci) with  $i = 1, 2$  hold,  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  be the sequences generated by Algorithm 1. If

$$\sum_{k=0}^{\infty} \alpha_k \|x^k - x^{k-1}\|^2 < +\infty, \quad (3.15) \quad \text{eq:th:main.01}$$

then the following assertions hold

(a) for  $\{\bar{x}\} = S(K, f)$ , the limit  $\lim_{k \rightarrow \infty} \|x^k - \bar{x}\|$  exists and

$$\lim_{k \rightarrow +\infty} \|x^{k+1} - y^k\|^2 = \lim_{k \rightarrow +\infty} \|z^k - y^k\|^2 = 0; \quad (3.16) \quad \text{eq:16}$$

(b) the sequences  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  converge to the same limit, that is a solution to (EP).

**Proof.** (a): Defining  $\delta_k = (\alpha_k^2 + \alpha_k) \|x^k - x^{k-1}\|^2$ ,  $k \geq 0$ , and using Proposition 10 we observe that condition (2.6) in Lemma 2 is fulfilled with  $s_{k+1} = ((2 - \rho_k)/\rho_k) \|x^{k+1} - y^k\|^2$  or  $s_{k+1} = ((2 - 4\eta\beta_k - \rho_k)/\rho_k) \|x^{k+1} - y^k\|^2$ . In addition, from  $\{\alpha_k\}_k \subset [0, 1[$  and condition (3.15) follows

$$\sum_{k=0}^{+\infty} \delta_k = \sum_{k=0}^{+\infty} (\alpha_k^2 + \alpha_k) \|x^k - x^{k-1}\|^2 \leq 2 \sum_{k=0}^{+\infty} \alpha_k \|x^k - x^{k-1}\|^2 < +\infty.$$

Hence, we infer from this and Lemma 2(b) that  $\lim_{k \rightarrow +\infty} \varphi_k$  exists. Thus for  $\{\bar{x}\} = S(K, f)$ , we have that  $\lim_{k \rightarrow \infty} \|x^k - \bar{x}\|$  exists, which, in particular, guarantees the boundedness of the sequence  $\{x^k\}_k$ . Likewise, condition (3.15) and Lemma 2(a) yield  $\sum_{k=0}^{\infty} s_{k+1} < +\infty$ , and so,  $s_{k+1} \rightarrow 0$  as  $k \rightarrow +\infty$ . From this and (Ci),  $i = 1, 2$ , we conclude that  $\lim_{k \rightarrow +\infty} \|x^{k+1} - y^k\|^2 = 0$ , and since  $x^{k+1} - y^k = \rho_k(z^k - y^k)$  for all  $k \geq 0$  by relation (3.4), it follows that  $\lim_{k \rightarrow +\infty} \|z^k - y^k\|^2 = 0$ .

(b): By part (a), the sequence  $\{x^k\}_k$  is bounded, hence it has cluster points. Let  $\hat{x} \in K$  be one of them. Then, there exists a subsequence  $\{x^{k_l}\}_l$  of  $\{x^k\}_k$  such that  $x^{k_l} \rightarrow \hat{x}$  as  $l \rightarrow +\infty$ . Applying (3.16) to  $\{x^{k_l}\}_l$  and using that  $\lim_{l \rightarrow +\infty} x^{k_l} = \hat{x}$ , we conclude that

$$\lim_{l \rightarrow +\infty} y^{k_l} = \lim_{l \rightarrow +\infty} z^{k_l} = \hat{x}. \quad (3.17) \quad \text{eq:sub1}$$

It follows from (3.3) that for any  $k \geq 0$  one has

$$\begin{aligned} f(y^k, z^k) + \frac{1}{2\beta_k} \|y^k - z^k\|^2 &\leq f(y^k, x) + \frac{1}{2\beta_k} \|y^k - x\|^2, \quad \forall x \in K \\ \implies f(y^k, z^k) &\leq f(y^k, x) + \frac{1}{2\beta_k} \|y^k - x\|^2, \quad \forall x \in K. \end{aligned}$$

Take  $x = \lambda y + (1 - \lambda)\hat{x}$  with  $y \in K$  and  $\lambda \in [0, 1]$ . Then,

$$\begin{aligned}
f(y^k, z^k) &\leq f(y^k, \lambda y + (1 - \lambda)\hat{x}) + \frac{1}{2\beta_k} \|\lambda(y^k - y) + (1 - \lambda)(y^k - \hat{x})\|^2 \\
&\leq \max\{f(y^k, y), f(y^k, \hat{x})\} - \frac{\lambda(1 - \lambda)\gamma}{2} \|y - \hat{x}\|^2 \\
&\quad + \frac{\lambda}{2\beta_k} \|y^k - y\|^2 + \frac{(1 - \lambda)}{2\beta_k} \|y^k - \hat{x}\|^2 - \frac{\lambda(1 - \lambda)}{2\beta_k} \|y - \hat{x}\|^2 \\
&= \max\{f(y^k, y), f(y^k, \hat{x})\} + \frac{\lambda}{\beta_k} \langle \hat{x} - y^k, y - \hat{x} \rangle + \frac{1}{2\beta_k} \|y^k - \hat{x}\|^2 \\
&\quad + \frac{\lambda}{2} \left( \frac{\lambda}{\beta_k} - \gamma + \lambda\gamma \right) \|y - \hat{x}\|^2, \quad \forall y \in K, \quad \forall \lambda \in [0, 1].
\end{aligned}$$

Take  $\hat{\beta} := 1/(\gamma - 8\eta)$ . Replace  $k$  by  $k_\ell$  in the previous inequalities. Then, taking  $\limsup_{\ell \rightarrow +\infty}$  and using (3.17) and assumption (A3), we have

$$\begin{aligned}
0 &= f(\hat{x}, \hat{x}) \leq \liminf_{\ell \rightarrow +\infty} f(y^{k_\ell}, z^{k_\ell}) \leq \limsup_{\ell \rightarrow +\infty} f(y^{k_\ell}, z^{k_\ell}) \\
&\leq \max\{f(\hat{x}, y), 0\} + \frac{\lambda}{2} \left( \frac{\lambda}{\hat{\beta}} - \gamma + \lambda\gamma \right) \|y - \hat{x}\|^2, \quad \forall y \in K, \quad \forall \lambda \in [0, 1].
\end{aligned}$$

Since  $\gamma > 0$ , we can take  $\lambda < (\hat{\beta}\gamma)/(1 + \hat{\beta}\gamma)$  in such a way that  $\lambda/\hat{\beta} - \gamma + \lambda\gamma < 0$ . Thus

$$0 \leq \max\{f(\hat{x}, y), 0\} = f(\hat{x}, y), \quad \forall y \in K \setminus \{\hat{x}\}.$$

Therefore,  $\hat{x} \in S(K, f)$ .

Finally, since  $f$  is strongly quasiconvex in its second argument,  $S(K, f)$  is a singleton, invoking Opial's Lemma (cf. [5, Lemma 2.39]), the whole sequence  $\{x^k\}_k$  converges to  $\{\bar{x}\} = S(K, f)$ , which completes the proof. ■

choose:alphak

**Remark 12** A sufficient condition for the fulfillment of (3.15) is to choose the inertial parameter  $\alpha_k$  as follows: for some  $\theta \in ]0, 1[$ ,

$$\alpha_k = \begin{cases} \min \left\{ \frac{\theta^k}{\|x^k - x^{k-1}\|^2}, \alpha \right\}, & \text{if } x^k \neq x^{k-1}, \\ \alpha, & \text{otherwise.} \end{cases}$$

In the next section, we provide another sufficient condition that connects the relaxation and Lipschitz condition parameters.

As consequences of Theorem 11, we obtain some particular cases. The first one discusses the convergence of Algorithm 1 when  $\alpha = 0$  (i.e. without the inertial step). Note that its convergence statement does not require  $K$  to be an affine subspace when  $\rho_k \in [0, 1]$ .

coro:rppa-EP

**Corollary 13** Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $f$  such that assumptions (Ai) with  $i = 1, 2, 3, 4, 5$  hold, and  $\{\alpha_k\}_k$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  sequences of positive numbers such that assumptions (Ci) with  $i = 1, 2$  hold. Then, for the sequence  $\{x^k\}_k$  generated as follows: choose  $x^0 \in K$  and for  $k \geq 0$  do

$$(RPPA-EP) \quad \begin{cases} z^k \in \operatorname{argmin} \left\{ f(x^k, x) + \frac{1}{2\beta_k} \|x^k - x\|^2 : x \in K \right\} \\ x^{k+1} = (1 - \rho_k)x^k + \rho_k z^k, \end{cases} \quad (3.18) \quad \boxed{\text{rppa}}$$

we have

- (a)  $\sum_{k=0}^{+\infty} \frac{2 - \rho_k}{\rho_k} \|x^{k+1} - x^k\|^2 < +\infty$  and  $\sum_{k=0}^{+\infty} \frac{2 - 4\eta\beta_k - \rho_k}{\rho_k} \|x^{k+1} - x^k\|^2 < +\infty$ ;
- (b) when  $\{\bar{x}\} = S(K, f)$ ,  $\lim_{k \rightarrow +\infty} \|x^k - \bar{x}\|$  exists, and hence  $\{x^k\}_k$  is bounded;
- (c) the sequence  $\{x^k\}_k$  generated by (3.18) converges to a solution of (EP) as  $k \rightarrow +\infty$ .

Let us now assume that  $\rho = \alpha = 0$ , that yields  $\rho_k = 1$  and  $\alpha_k = 0$  for every  $k \geq 0$ . In this case we obtain a variant of [28, Theorem 3.1].

**Corollary 14** Let  $K$  be a closed and convex set in  $\mathbb{R}^n$ ,  $f$  such that assumptions (Ai) with  $i = 1, 2, 3, 4, 5$  hold and the sequence  $\{\beta_k\}_k$  such that assumption (C1) holds. Let  $\{x^k\}_k$  be the sequence generated by the following iterative rule, for  $k \geq 0$ ,

$$(PPA-EP) \quad \begin{cases} x^0 \in K, \\ x^{k+1} \in \operatorname{argmin} \left\{ f(x^k, x) + \frac{1}{2\beta_k} \|x^k - x\|^2 : x \in K \right\}. \end{cases}$$

Then  $\{x^k\}_k$  converges to  $\{\bar{x}\} = S(K, f)$  as  $k \rightarrow +\infty$ .

**Proof.** Take  $\epsilon = 1/(4\eta)$  and use Remark 8. ■

Another consequence of Theorem 11 is [18, Theorem 3.1], which establishes the convergence of the inertial relaxed proximal point algorithm for minimizing strongly quasiconvex functions.

cor:ppa

**Corollary 15** Let  $K$  be a convex and closed set in  $\mathbb{R}^n$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  a lower semicontinuous and strongly quasiconvex function on  $K$  with modulus  $\gamma > 0$ ,  $\{\alpha_k\}_k$  and  $\{\beta_k\}_k$  sequences of positive numbers bounded away from 0 and  $\{\rho_k\}_k$  a sequence of positive numbers such that assumption (C2) holds. Let us consider the problem

$$\min_{x \in K} h(x). \quad (3.19) \quad \boxed{\text{min:problem}}$$

If  $\sum_{k=0}^{\infty} \alpha_k \|x^k - x^{k-1}\|^2 < +\infty$ , then the sequence  $\{x^k\}_k$ , generated by Algorithm 1 applied to the bifunction  $f_h : K \times K \rightarrow \mathbb{R}$ ,  $f_h(x, y) := h(y) - h(x)$ ,  $x, y \in K$ , converges to the unique minimizer of  $h$  on  $K$ . Moreover, the sequences  $\{y^k\}_k$  and  $\{z^k\}_k$  converge both to the same limit.



**Proof.** It can be immediately verified that assumptions (A2) and (A4) hold for  $f_h$ . Furthermore, since  $h$  is lower semicontinuous,  $f_h$  is lower semicontinuous on its first argument and upper semicontinuous on its second argument, so assumptions (A1) and (A1') hold, which easily entails the validity of (A3). Finally, by definition of  $f_h$ , we have

$$f_h(x, z) - f_h(x, y) - f_h(y, z) = h(z) - h(x) - (h(y) - h(x)) - (h(z) - h(y)) = 0,$$

i.e., assumption (A5) holds with  $\eta = 0$ , and hence assumption (C1) is valid as well since  $\gamma > 0$ . Hence, all assumptions needed for our previous results are valid in this case. Therefore, it follows from Theorem 11 that  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  converge to the unique optimal solution of (3.19). ■

**Remark 16** *In particular, considering  $\rho = \alpha = 0$ , i.e.,  $\rho_k = 1$  and  $\alpha_k = 0$  for every  $k \geq 0$  in Algorithm 1, Corollary 15 collapses to [33, Theorem 15] (by taking  $\epsilon = 1/(4\eta)$  and using Remark 8).*

## 3.2 Sufficient Conditions

sec:3-2

In order to provide other sufficient hypotheses for the convergence of the sequences generated by Algorithm 1 to the solution of (EP) by ensuring the fulfilment of (3.15), additionally to assumptions (Ai) with  $i = 1, 2, 3, 4, 5$  and (Ci) with  $i = 1, 2$ , we consider the following condition

(C3) the sequence  $\{\alpha_k\}_k$  is nondecreasing and there exists  $\bar{\alpha} = \xi/(2 + \xi)$  such that

$$0 \leq \alpha_k \leq \alpha_{k+1} \leq \alpha < \bar{\alpha} < 1, \forall k \geq 0, \quad (3.20) \quad \text{eq:alpha}$$

where

$$\xi := \frac{1 - 4\eta\epsilon - \rho}{1 + \rho} > 0. \quad (3.21) \quad \text{lower.b}$$

To this end, we prove the following auxiliary result. Note that in its proof we “unify” the statements given in terms of “at least one inequality holds” that begin with Proposition 6.

prop:02

**Proposition 17** *Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $f$  such that assumptions (Ai) with  $i = 1, 1', 2, 4, 5$  hold,  $\{\alpha_k\}_k$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  sequences of positive numbers such that assumptions (Ci) with  $i = 1, 2$  hold, and  $\{x^k\}$  the sequence generated by Algorithm 1. Take  $\{\bar{x}\} = S(K, f)$  and  $\varphi_k := \|x^k - \bar{x}\|^2$  for every  $k \geq 0$ . Then, for every  $k \geq 0$ , we have*

$$\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) - \nu_k \|x^k - x^{k-1}\|^2 \leq \xi(\alpha_k - 1) \|x^{k+1} - x^k\|^2, \quad (3.22) \quad \text{eq:20}$$

where  $\nu_k := (1 + \xi)\alpha_k^2 + (1 - \xi)\alpha_k$  and  $\xi$  is defined by (3.21).

**Proof.** To show (3.22) it is enough to bound  $\{\|x^{k+1} - y^k\|^2\}_k$  from below in relation (3.12) (and (3.13)) of Proposition 10. For  $k \geq 0$ , a direct calculation using definition of  $y^k$  and the Cauchy-Schwartz inequality yield

$$\begin{aligned} \|x^{k+1} - y^k\|^2 &= \|x^{k+1} - x^k\|^2 + \alpha_k^2 \|x^k - x^{k-1}\|^2 - 2\alpha_k \langle x^{k+1} - x^k, x^k - x^{k-1} \rangle \\ &\geq \|x^{k+1} - x^k\|^2 + \alpha_k^2 \|x^k - x^{k-1}\|^2 - 2\alpha_k \|x^{k+1} - x^k\| \|x^k - x^{k-1}\|. \end{aligned}$$

Using the well-known identity  $2pq \leq p^2 + q^2$  with  $p = \|x^{k+1} - x^k\|$  and  $q = \|x^k - x^{k-1}\|$ , we conclude

$$\|x^{k+1} - y^k\|^2 \geq (1 - \alpha_k) \|x^{k+1} - x^k\|^2 + (\alpha_k^2 - \alpha_k) \|x^k - x^{k-1}\|^2. \quad (3.23) \quad \boxed{\text{eq:d1}}$$

Moreover, observe that assumptions (C1) and (C2) yield

$$\frac{2 - \rho_k}{\rho_k} \geq \frac{2 - 4\eta\beta_k - \rho_k}{\rho_k} > \frac{1 - 4\eta\epsilon - \rho}{1 + \rho} = \xi > 0.$$

Finally, replacing (3.23) in (3.12), and taking into account the latter bound,

$$\begin{aligned} \varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) &\leq [(1 + \xi)\alpha_k^2 + (1 - \xi)\alpha_k] \|x^k - x^{k-1}\|^2 \\ &\quad + \xi(\alpha_k - 1) \|x^{k+1} - x^k\|^2, \end{aligned}$$

which proves relation (3.22). ■

Our second main result, which provides another sufficient condition for the convergence of Algorithm 1, is given below.

**thM:main3**

**Theorem 18** *Let  $K$  be an affine subspace in  $\mathbb{R}^n$ ,  $f$  such that assumptions (Ai) with  $i = 1, 2, 3, 4, 5$  hold,  $\{\alpha_k\}_k$ ,  $\{\beta_k\}_k$  and  $\{\rho_k\}_k$  sequences of positive numbers such that assumptions (Ci) with  $i = 1, 2, 3$  hold, and  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  the sequences generated by Algorithm 1. Then,*

$$\sum_{k=1}^{\infty} \|x^k - x^{k-1}\|^2 < +\infty. \quad (3.24) \quad \boxed{\text{eq:sum}}$$

As a consequence, the sequences  $\{x^k\}_k$ ,  $\{y^k\}_k$  and  $\{z^k\}_k$  converge to  $\{\bar{x}\} = S(K, f)$ .

**Proof.** Define  $\mu_0 := (1 - \alpha_0)\varphi_0 \geq 0$  and

$$\mu_k := \varphi_k - \alpha_k \varphi_{k-1} + \nu_k \|x^k - x^{k-1}\|^2, \quad \forall k \geq 1. \quad (3.25) \quad \boxed{\text{eq:mu}_k}$$

Since  $\varphi_k \geq 0$  and  $\alpha_k \leq \alpha_{k+1}$  for all  $k \geq 0$ , it follows from (3.22) and (3.25) that

$$\begin{aligned} \mu_{k+1} - \mu_k &= (\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) - \nu_k \|x^k - x^{k-1}\|^2) + \nu_{k+1} \|x^{k+1} - x^k\|^2 \\ &\leq (\xi(\alpha_k - 1) + \nu_{k+1}) \|x^{k+1} - x^k\|^2 \\ &= ((1 + \xi)\alpha_{k+1}^2 + (1 - \xi)\alpha_{k+1} + \xi(\alpha_k - 1)) \|x^{k+1} - x^k\|^2 \\ &\leq ((1 + \xi)\alpha_{k+1}^2 + \alpha_{k+1} - \xi) \|x^{k+1} - x^k\|^2. \end{aligned} \quad (3.26) \quad \boxed{\text{eq:23}}$$

Now, as  $\alpha_k \in [0, 1[$  (hence  $\alpha_k^2 \leq \alpha_k$ ) for every  $k \geq 0$ , using (3.20), we have

$$(1 + \xi)\alpha_{k+1}^2 + \alpha_{k+1} - \xi \leq (2 + \xi)\alpha_{k+1} - \xi \leq (2 + \xi)\alpha - \xi < 0.$$

This implies that  $\mu_{k+1} \leq \mu_k$  for all  $k \geq 0$ , consequently, the sequence  $\{\mu_k\}_k$  is non increasing and bounded above by  $\mu_0 = (1 - \alpha_0)\varphi_0$ . So, equation (3.25) and the monotonicity of  $\{\mu_k\}_k$  yield

$$\mu_0 \geq \mu_k = \varphi_k - \alpha_k \varphi_{k-1} + \nu_k \|x^k - x^{k-1}\|^2 \geq \varphi_k - \alpha_k \varphi_{k-1}.$$

From this and (3.20), we recursively obtain

$$\varphi_k \leq \alpha \varphi_{k-1} + \mu_0 \leq \dots \leq \alpha^k \varphi_0 + \mu_0 \sum_{j=0}^{k-1} \alpha^j \leq \alpha^k \varphi_0 + \frac{\mu_0}{1 - \alpha}.$$

By using relation (3.26), the above inequality, and taking into account that  $\mu_{k+1} \geq -\alpha \varphi_k$ , we have

$$\begin{aligned} \sum_{j=0}^k \|x^{j+1} - x^j\|^2 &\leq \frac{1}{\xi - (2 + \xi)\alpha} (\mu_0 - \mu_{k+1}) \leq \frac{1}{\xi - (2 + \xi)\alpha} (\mu_0 + \alpha \varphi_k) \\ &\leq \frac{1}{\xi - (2 + \xi)\alpha} \left( \mu_0 + \alpha \left( \alpha^k \varphi_0 + \frac{\mu_0}{1 - \alpha} \right) \right) \\ &= \frac{1}{\xi - (2 + \xi)\alpha} \left( \frac{\mu_0}{1 - \alpha} + \alpha^{k+1} \varphi_0 \right). \end{aligned} \quad (3.27) \quad \boxed{\text{eq:rate}}$$

Hence, taking  $k \rightarrow +\infty$ , we obtain

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 \leq \frac{1}{\xi - (2 + \xi)\alpha} \frac{\mu_0}{1 - \alpha} < +\infty,$$

which proves (3.24). The other conclusions follow directly from (3.24) and Theorem 11. ■

plans **Remark 19** *The above result extends [35, Theorem 3.2] and [18, Theorem 4.1] from minimization problems to pseudomonotone equilibrium problems involving strongly quasiconvex functions.*

As a consequence of Theorem 18, we obtain the following result, which provides an estimate for the number of iterations necessary for Algorithm 1 to reach a solution of (EP).

numer:iter **Corollary 20** *Consider the hypotheses of Theorem 18. For  $\{\bar{x}\} = S(K, f)$  and any positive integer  $k$ , there exists  $i \in \{0, \dots, k\}$  such that*

$$\|x^{i+1} - x^i\| \leq \frac{1}{\sqrt{k}} C, \quad (3.28) \quad \boxed{\text{eq:MT1}}$$

where  $C := \sqrt{\frac{1 + \alpha - \alpha^2}{(1 - \alpha)[\xi - (2 + \xi)\alpha]}} \|x^0 - \bar{x}\|$ . Moreover, given a tolerance  $\|z^k - y^k\| < \varepsilon$ , Algorithm 1 finds a solution of (EP) after performing at most

$$\frac{(1 + \alpha)(1 + \alpha - \alpha^2)}{(1 - \alpha)[\xi - (2 + \xi)\alpha]} \frac{\|x^0 - \bar{x}\|^2}{\varepsilon^2},$$

iterations.

**Proof.** Indeed, since  $\alpha \in ]0, 1[$ , from (3.27) we have

$$\begin{aligned} \sum_{i=0}^k \|x^i - x^{i-1}\|^2 &\leq \frac{1}{\xi - (2 + \xi)\alpha} \left( \frac{\mu_0}{1 - \alpha} + \alpha^{k+1} \varphi_0 \right) \\ &\leq \frac{1 + \alpha - \alpha^2}{(1 - \alpha)[\xi - (2 + \xi)\alpha]} \|x^0 - \bar{x}\|^2. \end{aligned} \quad (3.29) \quad \boxed{\text{eq:130}}$$

Choosing some  $i \in \{0, 1, \dots, k\}$  in (3.29), we deduce (3.28).

Since  $\|x^{k+1} - y^k\| \leq \|x^{k+1} - x^k\| + \alpha_k \|x^k - x^{k-1}\|$  and  $\rho_k \|z^k - y^k\| = \|x^{k+1} - y^k\|$  by (3.2), the last statement follows directly from (3.28). ■

We conclude this section by discussing the choice of the parameters involved in our analysis.

**rem:disc**

**Remark 21** (i) Recall that the idea behind the proximal point algorithm is the regularization of a possibly ill-conditioned problem  $0 \in T(x)$  (where  $T$  is a nonlinear operator, for instance, the subdifferential of a convex function) by replacing it with a sequence of well-conditioned subproblems of the form  $0 \in \beta_k T(x) + x - x^k$ ,  $k \geq 0$ . For any  $k \geq 0$  the choice of  $\beta_k$  in each specific instance must be a trade-off between numerical stability and fast progress of the algorithm, because if  $\beta_k$  is very small, the subproblem is very well-conditioned, but the progress of the algorithm slow because  $x^k$  remains very close to  $x^{k-1}$ . On the other hand, when  $\beta_k$  is large, the algorithm advances fast (in the limit, when  $\beta_k$  approaches  $+\infty$ , the regularization term becomes negligible, and we solve the problem in one iteration), but when the problem is ill-conditioned, this entails high numerical instability. Therefore, in the proximal type methods we consider for solving equilibrium problems the choice of  $\beta_k$  must adhere to the trade-off between numerical stability and fast progress. This is ensured by (C1).

On the other hand, the choice of the parameters of the inertial and relaxation parameters  $\{\alpha_k\}_k$  and  $\{\rho_k\}_k$  is controlled by the input parameters  $\rho$ ,  $\alpha$ , and the Lipschitz-type constant  $\eta$ , as can be observed in (C2) and (C3).

(ii) Note that the sequence  $\{\beta_k\}_k$  can be updated by using in each iteration a line-search procedure or a self-adaptive strategy. When employing them, there is no need to know the value of the Lipschitz-type condition parameter (which can be challenging to calculate in many situations), but the price

to pay is that an extra gradient step is needed. For further details on this approach, see, for instance, [24, 41] in the convex case and [53] in the quasiconvex setting (without inertial and relaxation effects).

- (iii) Relaxation techniques are another ingredient in the formulation of algorithms in continuous optimization because they provide more flexibility to the output of an iterative scheme (see [5]) and have the property of speeding up the algorithm in a similar manner to the effect of inertial steps. Without using inertia, over-relaxation (i.e.  $\rho_k > 1$  when  $k \geq 0$ ) provides a natural way of speeding up algorithms. The interplay of the inertial and relaxation parameters plays a crucial role in the convergence analysis of our proposed algorithm. The choice of the values of  $\alpha$  and  $\rho$  (given as inputs to the algorithm) is thus crucial, as they need to satisfy condition (C3).

## 4 Applications and Numerical Experiments

sec:4

In this section, we discuss some equilibrium problems involving bifunctions that are strongly quasiconvex in the second variable and fulfill the hypotheses guaranteeing the convergence of the proposed algorithm.

### 4.1 Strongly quasiconvex ratios

Given a subset  $K \subseteq \mathbb{R}^n$ , and functions  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  and  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $K \cap \text{dom } h \neq \emptyset$  and  $0 \notin g(K)$ , consider the fractional minimization problem given by

$$\min_{x \in K} \varphi(x) = \min_{x \in K} \frac{h(x)}{g(x)}. \quad (\text{FMP}) \quad \boxed{\text{FMP}}$$

This problem has been intensively studied in the literature (see [10, 44, 46] among others) due to its concrete applications in several fields of mathematical sciences as well as in economics, for instance, in the theory of productivity as maximization of return/risk or profit/cost and minimization of cost/time among others.

In general, problem (FMP) is not convex. Indeed, if, for instance,  $h$  is convex and  $g$  is affine, then  $\varphi : K \rightarrow \overline{\mathbb{R}}$ ,  $\varphi = h/g$ , is semistrictly quasiconvex, and the same happens when  $h$  is nonnegative and convex, and  $g$  is positive and concave, see [10, Theorem 2.3.8]. Sufficient conditions for guaranteeing that  $\varphi$  is strongly quasiconvex are given below.

prop:frac

**Proposition 22** (cf. [29]) *When  $h$  is strongly convex with modulus  $\gamma > 0$ ,  $g$  is positive and bounded from above by  $M > 0$ , and one of the following hypotheses holds*

- (a)  $g$  is affine;
- (b)  $h$  is nonnegative on  $\text{dom } h$  and  $g$  is concave;

(c)  $h$  is nonpositive on  $\text{dom } h$  and  $g$  is convex,

then  $\varphi : K \rightarrow \overline{\mathbb{R}}$ ,  $\varphi = h/g$ , is strongly quasiconvex with modulus  $\gamma' = \gamma/M > 0$ .

A consequence of this statement follows.

quad:frac

**Corollary 23** (cf. [29]) Let  $A, B \in \mathbb{R}^{n \times n}$ ,  $a, b \in \mathbb{R}^n$  and  $\alpha, \beta \in \mathbb{R}$ ,  $h(x) = \frac{1}{2}\langle Ax, x \rangle + \langle a, x \rangle + \alpha$  and  $g(x) = \frac{1}{2}\langle Bx, x \rangle + \langle b, x \rangle + \beta$ . Take  $K = \{x \in \mathbb{R}^n : m \leq g(x) \leq M\}$ , with  $0 < m < M$ , and  $\varphi(x) = h(x)/g(x)$ , with  $x \in K$ . Suppose that  $A$  is a positive definite matrix, and  $\lambda_{\min}(A)$  is its smallest eigenvalue. If any of the following conditions holds

- (a)  $B = 0$  (the null matrix);
- (b)  $h$  is nonnegative on  $K$  and  $B$  is negative semidefinite;
- (c)  $h$  is nonpositive on  $K$  and  $B$  is positive semidefinite,

then  $\varphi$  is strongly quasiconvex with modulus  $\gamma' = \lambda_{\min}(A)/M > 0$ .

In the following example, we present families of problems of type (IMVI) for which assumptions (Ai) with  $i = 0, 1, 2, 3, 4, 5$  are satisfied. This example already appears in [29], but we include it here for completeness.

ex:IMVI

**Example 24** (cf. [29]) Let  $A, B \in \mathbb{R}^{n \times n}$  be matrices,  $a, b, c \in \mathbb{R}^n$  and  $\alpha, \beta \in \mathbb{R}$ . Let us consider  $K := \{x \in \mathbb{R}^n : m \leq \langle c, x \rangle + \alpha \leq M\}$  with  $0 < m < M$ , and  $F : K \rightarrow \mathbb{R}^n$  and  $\varphi : K \rightarrow \mathbb{R}$  be the operator and function defined by

$$F(x) := \frac{1}{\langle c, x \rangle + \alpha}(Ax + a), \quad \varphi(x) := \frac{\frac{1}{2}\langle Bx, x \rangle + \langle b, x \rangle + \beta}{\langle c, x \rangle + \alpha}. \quad (4.1)$$

ex:operators

By taking  $f : K \times K \rightarrow \mathbb{R}$  given by

$$f(x, y) := \left\langle x, \frac{Ay + a}{\langle c, y \rangle + \alpha} - \frac{Ax + a}{\langle c, x \rangle + \alpha} \right\rangle + \frac{\frac{1}{2}\langle By, y \rangle + \langle b, y \rangle + \beta}{\langle c, y \rangle + \alpha} - \frac{\frac{1}{2}\langle Bx, x \rangle + \langle b, x \rangle + \beta}{\langle c, x \rangle + \alpha}, \quad (4.2)$$

equi:bifunction

(IMVI) reduces to an equilibrium problem (EP).

Let us check that the bifunction  $f$  defined in (4.2) satisfies assumptions (Ai) with  $i = 0, 1, 2, 3, 4, 5$  under mild hypotheses. Indeed, (Ai) with  $i = 0, 1, 3$  are straightforward while (A2) follows from the monotonicity of  $F$ .

For (A4): note that  $f$  can be rewritten as

$$\begin{aligned} f(x, y) &= \langle x, F(y) - F(x) \rangle + \varphi(y) - \varphi(x) \\ &= \frac{\frac{1}{2}\langle By, y \rangle + (x^\top A + b^\top)y + a^\top x + \beta}{\langle c, y \rangle + \alpha} - \frac{x^\top (Ax + a)}{\langle c, x \rangle + \alpha} - \frac{\frac{1}{2}\langle Bx, x \rangle + \langle b, x \rangle + \beta}{\langle c, x \rangle + \alpha}. \end{aligned}$$

Hence, in virtue of Corollary 23, for any  $x \in K$  the function  $f(x, \cdot)$  is strongly quasiconvex on  $K$  when the matrix  $B$  is positive definite, and its modulus of strong quasiconvexity is  $\gamma = \lambda_{\min}(B)/M > 0$ .

For (A5): Since  $F$  is continuously differentiable on  $K$ , it is Lipschitz-continuous on  $K$ . Let us denote by  $L > 0$  its Lipschitz constant. Then, for every  $x, y, z \in K$ ,

$$\begin{aligned} f(x, y) + f(y, z) - f(x, z) &= \langle x - y, F(y) - F(z) \rangle \\ &\geq -\frac{\sqrt{L}}{2} \|x - y\|^2 - \frac{1}{2\sqrt{L}} \|F(y) - F(z)\|^2 \\ &\geq -\frac{\sqrt{L}}{2} (\|x - y\|^2 + \|y - z\|^2), \end{aligned}$$

i.e., (A5) holds.

Finally, for (C1) and (C2), we simply take the matrix  $B$  having its smallest eigenvalue large enough to satisfy (A6). This can be done because  $\eta$  depends on  $A, a, c, \alpha$ , but not on  $B$ .

In the following example, we present families of problems that satisfy our assumptions  $(A_i)$  with  $i = 0, 1, 2, 3, 4, 5$ . This new example will be used for the numerical experiments below.

ex:MVI

**Example 25** Let  $B \in \mathbb{R}^{n \times n}$ ,  $b, c \in \mathbb{R}^n$  and  $\alpha, \beta \in \mathbb{R}$ ,  $K := \{x \in \mathbb{R}^n : m \leq \langle c, x \rangle + \alpha \leq M\}$  where  $0 < m < M$ ,  $T : K \rightarrow K$  be an operator and the function  $\varphi : K \rightarrow \mathbb{R}$  defined as in (4.1). Taking  $f : K \times K \rightarrow \mathbb{R}$  given by

$$f(x, y) := \langle T(x), y - x \rangle + \frac{\frac{1}{2} \langle B y, y \rangle + \langle b, y \rangle + \beta}{\langle c, y \rangle + \alpha} - \frac{\frac{1}{2} \langle B x, x \rangle + \langle b, x \rangle + \beta}{\langle c, x \rangle + \alpha}, \quad (4.3)$$

equi:bifunction2

problem (MVI) reduces to an equilibrium problem.

Let us check that bifunction  $f$  defined in (4.3) satisfies assumptions  $(A_i)$  with  $i = 0, 1, 2, 3$  under mild hypotheses. Indeed, again the verification of assumptions  $(A_i)$  with  $i = 0, 1, 3$ , is straightforward; if  $T$  is monotone, then (A2) holds and if  $T$  is  $L$ -Lipschitz-continuous, then (A5) follows from Remark 3(ii).

For (A4): note that  $f$  can be rewritten as

$$\begin{aligned} f(x, y) &= \langle T(x), y - x \rangle + \varphi(y) - \varphi(x) \\ &= \langle T(x), y \rangle + \frac{\frac{1}{2} \langle B y, y \rangle + \langle b, y \rangle + \beta}{\langle c, y \rangle + \alpha} - \langle T(x), x \rangle - \frac{\frac{1}{2} \langle B x, x \rangle + \langle b, x \rangle + \beta}{\langle c, x \rangle + \alpha} \\ &= \frac{y^\top (T(x)c^\top + \frac{1}{2} B) y + (b + \alpha T(x))^\top y + \beta}{\langle c, y \rangle + \alpha} - \langle T(x), x \rangle - \frac{\frac{1}{2} \langle B x, x \rangle + \langle b, x \rangle + \beta}{\langle c, x \rangle + \alpha}. \end{aligned}$$

Take an  $x \in K$  such that  $T(x)^\top c \neq 0$ . Then, by Brauer's Eigenvalue Theorem (see [8]), the matrix  $T(x)c^\top$  has rank 1 and its only nonzero eigenvalue is  $\lambda_{\neq 0}(x) = T(x)^\top c$ . Hence, in virtue of Proposition 22, for any  $x \in K$  the function  $f(x, \cdot)$  is strongly quasiconvex on  $K$  when the matrix  $T(x)c^\top + (1/2)B$  is positive definite for all  $x \in K$ .

Finally, for (C1) and (C2), we simply take  $b \in \text{int } K^*$  and the matrix  $B$  such that its smallest eigenvalue is large enough to satisfy (C1) and (C2).

## 4.2 Numerical Experiments

In order to stress the applicability of the algorithms proposed in this paper to equilibrium problems for which until recently there have been no reliable iterative solving methods we present some computational results obtained in MATLAB 2019b-Win64 on a Dell Latitude 5420 Laptop with Windows 11 Pro and an Intel Core i7 1165G7 CPU with 2.80 GHz and 32GB RAM by implementing the relaxed-inertial proximal point method investigated in this work Algorithm 1 (RIPPA-EP) and its non-inertial version (3.18) (RPPA-EP), and, for comparison, the “pure” proximal point method [28, Algorithm 1] (PPA-EP). The examples treated below exhibit situations where the methods proposed in this paper have a superior performance (in terms of the number of necessary iterations and CPU time until reaching the solution to (EP)) to their “pure” proximal point counterpart. As stopping criteria of the proposed algorithms we considered the situations when the norm, in particular the absolute value, of the difference between  $y^k$  and  $z^k$ , respectively  $x^k$  and  $z^k$  in the non-inertial case, and between two consecutive members of  $\{x^k\}_k$  in the “pure” proximal point case is not larger than an a priori given error  $\varepsilon > 0$ , i.e., we stop when  $\|z^k - y^k\| < \varepsilon$  for (RIPPA-EP),  $\|z^k - x^k\| < \varepsilon$  for (RPPA-EP), and  $\|x^k - x^{k-1}\| < \varepsilon$  for (PPA-EP).

We begin with two simpler equilibrium problems, followed by the fractional one considered in Example 25.

**ex1** **Example 26** Take  $p, q \in \mathbb{N}$ ,  $p, q > 1$ , and the bifunction  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ ,

$$f(x, y) := p \left( \max\{\sqrt{|y|}, (y - q)^2 - q\} \right) - p \left( \max\{\sqrt{|x|}, (x - q)^2 - q\} \right) + x(y - x). \quad (4.4)$$

**bifunction**

We solve (EP) with Algorithm 1 and [28, Algorithm 1]. Its solution set  $S(\mathbb{R}, f)$  depends on the values of  $p, q > 1$ .

Assumptions (A1) and (A3) trivially hold by the continuity of  $f$ , while (A5) holds when, for instance,  $\eta = 1/2$ . Taking  $x, y \in K$  one has

$$f(x, y) + f(y, x) = x(y - x) + y(x - y) = -(y - x)^2 \leq 0,$$

hence  $f$  is monotone on  $K$ , i.e. (A2) holds as well.

Regarding (A4), for all  $x \in K$ , since the functions  $y \mapsto x(y - x)$  and  $y \mapsto p \left( \max\{\sqrt{|y|}, (y - q)^2 - q\} \right)$  are increasing on  $[0, +\infty[$ , their sum is increasing too, so that  $f$  is quasiconvex in its second argument by, for instance, [19, Proposition 4.9]. Furthermore,  $f$  is strongly quasiconvex in its second argument as the maximum of two strongly quasiconvex functions. Indeed, for every  $q \in \mathbb{N}$ , taking  $\delta > 0$  to be larger than the positive solution of the equation  $p(y - q)^2 - \sqrt{|y|} = pq$ , the function  $y \mapsto p\sqrt{|y|} + x(y - x)$  is strongly quasiconvex on  $[0, \delta]$  with modulus  $\gamma = 1/(2\sqrt{\delta^3}) > 0$  by [28, Example 4.2], while the function  $y \mapsto (y - q)^2 - q + x(y - x)$  is strongly convex, hence also strongly quasiconvex on  $[\delta, +\infty[$ . Therefore,  $f$  is strongly quasiconvex on  $[0, +\infty[$  and assumption (A4) holds. Also, as  $q > 1$ , the functions  $\sqrt{|y|}$  and  $(y - q)^2 - q$  take the same value twice on  $[0, +\infty[$ , so

$$y \mapsto p \max\{\sqrt{|y|}, (y - q)^2 - q\} + x(y - x) = p((y - q)^2 - q) + x(y - x),$$



is strongly convex on  $]-\infty, 0]$ , hence also strongly quasiconvex. Therefore, (A4) holds when  $K = \mathbb{R}$ . Finally,  $p, q > 1$  have to be chosen such that  $\gamma$  satisfies conditions (C1) and (C2), while  $\{\alpha_k\}_k$  needs to fulfill (C3). Note that one can multiply the last term of (4.4) by a parameter  $t > 0$  without affecting the mentioned properties of the considered problem.

We performed several experiments in MATLAB for various constellations of the involved elements, that confirm the conclusions reached in [2, 3, 24] for the convex case, namely that a clever choice of the parameters makes the relaxed-inertial proximal point algorithms faster and cheaper to use than their “pure” proximal point counterparts.

For instance, when  $q = 99$ ,  $\epsilon = 1/4$ ,  $\eta = 1/2$ ,  $\rho = 2/5$ ,  $\varepsilon = 10^{-7}$ ,  $\alpha = 1/29 - \varepsilon$  (as  $\bar{\alpha} = 1/29$  by (3.21)),  $\alpha_k = \alpha - 1/(10 + 7k)$ ,  $\beta_k = 1/(8k + 4)$ ,  $k \geq 0$ , and the starting points  $x^0 = 8297$  and  $x^{-1} = 3210$ , taking  $\rho_k = (1 - 1/k)(1 - \rho) + (1/k)(1 + \rho)$ ,  $k \geq 0$ , made Algorithm 1 (RIPPA-EP) deliver the solution  $\bar{x} = 88.5879$  to (EP) in 20 iterations that required 0.6232 seconds, while [28, Algorithm 1] (PPA-EP) provided  $\bar{x}$  in 123 steps, after 1.3024 seconds. Moreover, when  $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$ ,  $k \geq 0$ , Algorithm 1 (RIPPA-EP) remained faster and cheaper than the standard proximal point method as it delivered  $\bar{x}$  after 38 iterations, in 0.9610 seconds. Improving the accuracy to  $\varepsilon = 10^{-9}$ , Algorithm 1 (RIPPA-EP) reached  $\bar{x}$  in 30 iterations and 0.8728 seconds with the first choice of the relaxation parameters, and in 53 iterations and 1.3601 seconds with the second one, while [28, Algorithm 1] (PPA-EP) required 124 iterations and 1.9781 seconds. These results are summarized in Table 1.

Algorithm	RIPPA-EP(a)	RIPPA-EP(b)	[28, PPA-EP]
	$\varepsilon = 10^{-7}$		
Time (s)	<b>0.6232</b>	0.961	0.8727
Iterations	<b>20</b>	38	123
	$\varepsilon = 10^{-9}$		
Time (s)	<b>1.1303</b>	1.3601	1.9781
Iterations	<b>30</b>	53	124

**t1** Table 1: Comparison of the performance of RIPPA-EP(a) (Algorithm 1 with  $\rho_k = (1 - 1/k)(1 - \rho) + (1/k)(1 + \rho)$ ), RIPPA-EP(b) (Algorithm 1 with  $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$ ), and [28, Algorithm 1] with starting points  $x^0 = 8297$  and  $x^{-1} = 3210$  while reaching the stopping criterion.

Taking the starting points closer to the obtained solution,  $x^0 = 200$  and  $x^{-1} = -300$ , the tendency observed above remained valid, as Algorithm 1 (RIPPA-EP) reached  $\bar{x}$  in 13 iterations and 0.4052 seconds, while [28, Algorithm 1] (PPA-EP) required 83 iterations and 0.8381 seconds.

We can also report on the curious fact that when we mistakenly implemented  $\beta_k = 1/(8k)+4$  (instead of  $\beta_k = 1/(8k+4)$ ),  $k \geq 0$ , in the last parameter constellation presented above, Algorithm 1 converged and delivered the same solution

$\bar{x}$  to (EP), requiring 25 iterations and 0.8981 seconds for this. In this situation, however, (C1) is violated and thus Theorem 11 does not apply, hence the convergence of the proposed algorithm cannot be theoretically guaranteed. As this seems to work in practice, at least in this case, the legitimate (at the moment open) question of alternative hypotheses that keep Algorithm 1 convergent beyond the setting of Theorem 11 is raised.

Next we consider the problem from (4.4) in higher dimensions. The relaxed-inertial proximal point algorithm turns out to outperform its standard proximal point counterpart in this case, too.

**ex3** **Example 27** Take  $p, q \in \mathbb{N}$ ,  $p, q > 1$ , and the bifunction  $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$f(x, y) := p \left( \max\{\sqrt{\|y\|}, (y - qe)^2 - q\} \right) - p \left( \max\{\sqrt{\|x\|}, (x - qe)^2 - q\} \right) + x^\top (y - x), \quad (4.5) \quad \text{high}$$

where  $e = (1, \dots, 1)^\top \in \mathbb{R}^n$ . We solve (EP) with Algorithm 1 and [28, Algorithm 1]. Its solution set  $S(\mathbb{R}^n, f)$  depends on the value of  $q > 1$ . One can verify the fulfillment of the hypotheses in a similar manner to Example 26.

For  $n = 60$ ,  $\varepsilon = 10^{-8}$ ,  $\alpha_k = \alpha - 1/(120 + 3k)$ ,  $\beta_k = 1/(28k + 4)$ ,  $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$ ,  $k \geq 0$ , the other parameters as above, and randomly generated starting points, Algorithm 1 (RIPPA-EP) delivered the solution to (EP) in 84 iterations that required 4.8055 seconds, while [28, Algorithm 1] (PPA-EP) provided it in 137 steps, after 7.8315 seconds. Increasing the dimension to  $n = 99$  and relaxing the admissible error to  $\varepsilon = 10^{-6}$ , for  $\alpha_k = \alpha - 1/(120 + 7k)$ ,  $\rho_k = (1 - 1/k)(1 - \rho) + (1/k)(1 + \rho)$ ,  $k \geq 0$ , the other parameters as above, and randomly generated starting points, Algorithm 1 (RIPPA-EP) needed 71 iterations that required 5.6575 seconds to solve (EP), while [28, Algorithm 1] (PPA-EP) did the job in 228 steps, after 22.3490 seconds. The results are summarized in Table 2. The observed tendency was confirmed by testing the algorithms for higher dimensions like  $n = 500$  and  $n = 1000$ , too.

Algorithm	$n = 60, \varepsilon = 10^{-8}$	
	RIPPA-EP(a)	[28, PPA-EP]
Time (s)	<b>4.8055</b>	7.8315
Iterations	<b>84</b>	123
Algorithm	$n = 99, \varepsilon = 10^{-6}$	
	RIPPA-EP(b)	[28, PPA-EP]
Time (s)	<b>5.6575</b>	22.3490
Iterations	<b>71</b>	228

**t2** Table 2: Comparison of the performance of RIPPA-EP(a) (Algorithm 1 with  $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$ ), RIPPA-EP(b) (Algorithm 1 with  $\rho_k = (1 - 1/k)(1 - \rho) + (1/k)(1 + \rho)$ ), and PPA-EP [28, Algorithm 1].

Next, we examine an equilibrium problem related to the one discussed in Example 26. In this scenario, the governing bifunction is strongly quasiconvex in its second variable, representing a special case of the strongly quasiconvex equilibrium problem outlined in [28, Example 4.2]. This specific case has been addressed in [53, Example 1] using a proximal subgradient algorithm.

ex:22

**Example 28** Consider the equilibrium problem from Example 25 with  $n = 2$ ,  $\alpha = 0.3$ ,  $c = (2, 3)^\top$ ,  $B = 10 \cdot \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ,  $b = (3, 8)^\top$ ,  $\beta = 0$ ,  $T = (1/13) \cdot \text{Id}$ ,  $m = 1$  and  $M = 10$ . Hence  $\gamma = 1$ . Because of the structure of the considered problem, as the feasible set is convex and closed, but not an affine subspace, we solved it with the non-inertial version (3.18) (RPPA-EP) of Algorithm 1 (as the extrapolation steps could lead to infeasible points). We performed experiments in MATLAB for various constellations of the involved parameters. For instance, when  $\varepsilon = 10^{-6}$ ,  $\epsilon = 1/3$ ,  $\eta = 1/13$ ,  $\rho = 0.15$ ,  $\beta_k = 1/(k + 20)$ ,  $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$ ,  $k \geq 0$ , with the starting point  $x^0 = (5, 5)^\top$ , (3.18) (RPPA-EP) delivered the solution  $(0.1948, 0.1035)^\top$  to the considered problem after 66 iterations and 0.6249 seconds. This performance was improved when we took larger relaxation intervals, the best result being achieved for  $\rho = 0.7$ , namely 41 iterations and 0.3625 seconds. Interestingly, for higher values of  $\rho$  (up to its maximal value of 35/39 allowed by (C2)) the performance of the algorithm resembles the one of the standard proximal point method [28, Algorithm 1], which reached the solution after 78 iterations, in 0.7149 seconds. Lowering the acceptable error to  $\varepsilon = 10^{-8}$ , (3.18) (RPPA-EP) delivered an approximate solution to the equilibrium problem from Example 25 after 75 iterations and 0.6980 seconds when  $\rho = 0.7$ , that is improved to 69 iterations and 0.6211 seconds for  $\rho = 0.45$ , while [28, Algorithm 1] required 127 iterations and 1.4180 seconds. These findings are summarized in Table 3.

Algorithm	RPPA-EP ( $\rho = 0.15$ )	RPPA-EP ( $\rho = 0.7$ )	[28, PPA-EP]
		$\varepsilon = 10^{-6}$	
Time (s)	0.6244	<b>0.3649</b>	0.7149
Iterations	66	<b>41</b>	78
	RPPA-EP ( $\rho = 0.45$ )	RPPA-EP ( $\rho = 0.7$ )	[28, PPA-EP]
		$\varepsilon = 10^{-8}$	
Time (s)	<b>0.6211</b>	0.6980	1.4180
Iterations	<b>69</b>	75	127

t3

Table 3: Comparison of the performance of (3.18) RPPA-EP with  $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$  and PPA-EP [28, Algorithm 1].

All these experiments confirm, again, that a suitable choice of the parameters makes the relaxed-inertial proximal point algorithms faster and cheaper to use than their “pure” proximal point counterparts.

**Remark 29** When implementing all three considered algorithms the proximal step (3.3) and its counterpart in [28, Algorithm 1] were computed by employing

the MATLAB function `fmincon` because, as it is unfortunately often the case when dealing with nonconvex functions, a closed form of the proximal operator of the involved function (even when  $K = \mathbb{R}^n$ ) is not available for the moment. Depending on the involved function, it might be derived by direct calculation or possibly by adapting the method considered in [9] for determining the proximal operator of the root function.

**Remark 30** As noted in [18], the optimal choices of the involved relaxation and inertial parameters derived in the convex case do not always accelerate the proximal point algorithm in the strongly quasiconvex case. It remains thus an open question to determine choices of these parameters that deliver a theoretically guaranteed acceleration in the strongly quasiconvex setting.

## 5 Conclusions

sec:5

We show that the relaxed-inertial proximal point algorithm for solving equilibrium problems, recently introduced in the convex case, remains convergent, under standard hypotheses, when the employed bifunction is strongly quasiconvex in its second variable, and pseudomonotone. Various sufficient conditions for guaranteeing the fulfillment of these hypotheses are provided. Numerical experiments illustrate the theoretical achievements by showing that the proposed algorithm outperforms the standard proximal point one.

In subsequent work, we plan to provide similar results achieved under a more general quasiconvexity assumption on the involved bifunction in its second variable. Moreover, we aim to investigate the extension to the quasiconvex framework considered in this work of the modified relaxed inertial extragradient method proposed in [24, 48] in the convex case.

## 6 Declarations

### 6.1 Availability of supporting data

No data sets were generated during the current study. The used MATLAB codes are available from all authors on reasonable request.

### 6.2 Competing interests

There are no conflicts of interest or competing interests related to this paper.

### 6.3 Funding

The authors would like to thank the MATH AmSud cooperation program (Project AMSUD-220020) for its support. This research was partially supported by Anid-Chile under project Fondecyt Regular 1220379 (Lara), and by a CIAS Senior Research Fellow Grant of the Corvinus Institute for Advanced Studies, by

the Hi! PARIS Center and by a public grant from the Fondation Mathématique Jacques Hadamard as part of the Investissement d’avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH (Grad).

## 6.4 Authors’ contributions

All authors contributed equally to the study conception, design and implementation, and wrote and corrected the manuscript.

## References

- AAGS [1] K. ADDI, S. ADLY, G. GOELEVEN, H. SAOUD, A sensitivity analysis of a class of semi-coercive variational inequalities using recession tools, *J. Global Optim.*, **40**, 7–27, (2008).
- ALZ [2] F. ALVAREZ, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space, *SIAM J. Optim.*, **14**, 773–782, (2003).
- ALA [3] F. ALVAREZ, H. ATTOUCH, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, *Set-Valued Var. Anal.*, **9**, 3–11, (2001).
- ABR [4] Q.H. ANSARI, F. BABU, M.S. RAJU, Proximal point method with Bregman distance for quasiconvex pseudomonotone equilibrium problems, *Optimization*, DOI: 10.1080/02331934.2023.2252430, (2023).
- BCO [5] H.H. BAUSCHKE, P.L. COMBETTES, “Convex Analysis and Monotone Operators Theory in Hilbert Spaces”. *CMS Books in Mathematics*. Springer-Verlag, second edition, (2017).
- BLO [6] E. BLUM, W. OETTLI, From optimization and variational inequalities to equilibrium problems, *Math. Student.*, **63**, 123–145, (1994).
- BCF [7] R.I. BOT, E.R. CSETNEK, Proximal-gradient algorithms for fractional programming, *Optimization*, **66**, 1383–1396, (2017).
- BRA [8] A. BRAUER, Limits for the characteristic roots of matrices IV: applications to stochastic matrices, *Duke Math J*, **19**, 75–91, (1952).
- BRL [9] K. BREDIES, D. LORENZ, Iterated hard shrinkage for minimization problems with sparsity constraints, *SIAM J Sci Comput*, **30**, 657–683, (2008).
- CMM [10] A. CAMBINI, L. MARTEIN. “Generalized Convexity and Optimization: Theory and Applications”. Springer, 2009.
- CGC [11] A. CHERUKURI, B. GHARESIFARD, J. CORTÉS, Saddle-point dynamics: conditions for asymptotic stability of saddle points, *SIAM J Control Optim*, **55**, 486–511, (2017).

- [FP1] [12] F. FACCHINEI, J.-S. PANG. “Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. I”. Springer, New York, (2003).
- [FP2] [13] F. FACCHINEI, J.-S. PANG. “Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. II”. Springer, New York, (2003).
- [KYF] [14] K. FAN, A Minimax Inequality and Applications. In O. Shisha (ed.), “Inequality III”, pp. 103–113. Academic Press, New York, (1972).
- [Goe-book] [15] D. GOELEVELN. “Complementarity and Variational Inequalities in Electronics”. Academic Press, London, (2017).
- [GOM] [16] X. GOUDOU, J. MUNIER, The gradient and heavy ball with friction dynamical systems: the quasiconvex case, *Math. Programm.*, **116**, 173–191, (2007).
- [GL2] [17] S.-M. GRAD, F. LARA, Solving mixed variational inequalities beyond convexity, *J. Optim. Theory Appl.*, **190**, 565–580, (2021).
- [GLM] [18] S.-M. GRAD, F. LARA, R. T. MARCAVILLACA, Relaxed-inertial proximal point type algorithms for quasiconvex minimization, *J. Global Optim.*, **85**, 615–635, (2023).
- [HAD] [19] N. HADJISAVVAS, Convexity, generalized convexity and applications, In: S. Al-Mezel et al. (eds.): “Fixed Point Theory, Variational Analysis and Optimization”. pp. 139–169, Taylor & Francis, Boca Raton, Florida, (2014).
- [HKS] [20] N. HADJISAVVAS, S. KOMLOSI, S. SCHAIBLE, “Handbook of Generalized Convexity and Generalized Monotonicity”. Springer-Verlag, Boston, (2005).
- [HE] [21] B. HE, Algorithm for a class of generalized linear variational inequality and its application, *Sci China, A*, **25**, 939–945, (1995).
- [HHL] [22] B. HE, X.-Z. HE, H.K. LIU, Solving a class of constrained “black-box” inverse variational inequalities, *Eur. J. Oper. Res.*, **204**, 391–401, (2010).
- [HIE] [23] D. HIEU, An inertial-like proximal algorithm for equilibrium problems, *Math. Meth. Oper. Res.*, **88**, 399–415, (2018).
- [HDT] [24] D. HIEU, H.N. DUONG, B.H. THAI, Convergence of relaxed inertial methods for equilibrium problems, *J. Appl. Numer. Optim.*, **3**, 215–229, (2021).
- [IL3] [25] A. IUSEM, F. LARA, Optimality conditions for vector equilibrium problems with applications. *J. Optim. Theory Appl.*, **180**, 187–206, (2019).
- [IL5] [26] A. IUSEM, F. LARA, Existence results for noncoercive mixed variational inequalities in finite dimensional spaces, *J. Optim. Theory Appl.*, **183**, 122–138, (2019).

- [IL6] [27] A. IUSEM, F. LARA, A note on “Existence results for noncoercive mixed variational inequalities in finite dimensional spaces”, *J. Optim. Theory Appl.*, **187**, 607–608, (2020).
- [ILP] [28] A. IUSEM, F. LARA, Proximal point algorithms for quasiconvex pseudomonotone equilibrium problems, *J. Optim. Theory Appl.*, **193**, 443–461, (2022).
- [ILMY] [29] A. IUSEM, F. LARA, R.T. MARCAVILLACA, L.H. YEN, A two-step PPA for nonconvex equilibrium problems with applications to fractional programming, *Submitted*, (2023).
- [JOV] [30] M. JOVANOVIĆ, A note on strongly convex and quasiconvex functions, *Math. Notes*, **60**, 584–585, (1996).
- [KAR] [31] G. KASSAY, V. RĂDULESCU, “Equilibrium Problems and Applications”, Elsevier, (2018).
- [KOS] [32] K. KOUMATOS, S. SPIRITO, Quasiconvex elastodynamics: weak-strong uniqueness for measure-valued solutions, *Commun. Pure Appl. Math.*, **72**, 1288–1320, (2019).
- [LAP] [33] F. LARA, On strongly quasiconvex functions: existence results and proximal point algorithms, *J. Optim. Theory Appl.*, **192**, 891–911, (2022).
- [LAE] [34] F. LARA, On nonconvex pseudomonotone equilibrium problems with applications, *Set-Valued Var. Anal.*, **30**, 355–372, (2022).
- [MAI] [35] M. MAINGUÉ, Asymptotic convergence of an inertial proximal method for unconstrained quasiconvex minimization, *J. Global Optim.*, **45**, 631–644, (2009).
- [MAL] [36] Y. MALITSKY, Golden ratio algorithms for variational inequalities, *Math. Programm.*, **184**, 383–410, (2020).
- [MOU] [37] A. MOUDAFI, Second-order differential proximal methods for equilibrium problems, *J. Ineq. Pure Appl. Math.*, **4**, Art. 18, (2003).
- [OG] [38] N. OVCHAROVA, J. GWINNER, Semicoercive variational inequalities: From existence to numerical solutions of nonmonotone contact problems, *J. Optim. Theory Appl.*, **171**, 422–439, (2016).
- [POL] [39] B.T. POLYAK, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Sov Math Dokl.*, **7**, 72–75, (1966) (translation from *Dokl Akad Nauk SSSR*, **166**, 287–290, (1966)).
- [PO2] [40] B.T. POLYAK, Some methods of speeding up the convergence of iteration methods, *U.S.S.R. Comput. Math. Math. Phys.*, **4**, 1–17, (1967) (translation from *Zh. Vychisl. Mat. Mat. Fiz.*, **4**, 791–803, (1964)).

- [QMH] [41] T.D. QUOC, L.D. MUU, N.V. HIEN, Extragradient algorithms extended to equilibrium problems, *Optimization*, **57**, 749–766, (2008).
- [RWE] [42] R.T. ROCKAFELLAR, R. WETS, Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time, *SIAM J. Control Optim.*, **28**, 810–820, (1990).
- [RPI] [43] B.D. ROUHANI, M.R. PIRANFAR, Asymptotic behavior for a quasi-autonomous gradient system of expansive type governed by a quasiconvex function, *Electron. J. Differ. Equ.*, **2021**, paper no. 15, (2021).
- [SCH] [44] S. SCHAIBLE, Fractional programming, In: R. Horst and P. Pardalos (eds.), “Handbook of Global Optimization”, pp. 495–608. Kluwer Academic Publishers, Dordrecht, (1995).
- [SZI] [45] S. SCHAIBLE, W.T. ZIEMBA, “Generalized Concavity in Optimization and Economics”. Academic Press, New York, (1981).
- [STA] [46] I.M. STANCU-MINASIAN. “Fractional Programming: Theory, Methods and Applications”. Kluwer Academic Publishers, (1997).
- [STU] [47] J. ŠTULLER, Ordered modified Gram-Schmidt orthogonalization revised, *J. Comput. Appl. Math.*, **63**, 221–227, (1995).
- [VTV] [48] L. VAN VINH, V.N. TRAN, P.T. VUONG, A second-order dynamical system for equilibrium problems, *Numer. Algor.*, **91**, 327–351, (2022).
- [VER] [49] B.S. VERKHOVSKY, Information protection based on extraction of square roots of gaussian integers, *Int. J. Commun. Netw. Syst. Sci.*, **4**, 133–138, (2011).
- [V1] [50] J.P. VIAL, Strong convexity of sets and functions, *J. Math. Economics*, **9**, 187–205, (1982).
- [WAN] [51] M. WANG, The existence results and Tikhonov regularization method for generalized mixed variational inequalities in Banach spaces, *Ann. Math. Phys.*, **7**, 151–163, (2017).
- [YEM] [52] L.H. YEN, L.D. MUU, A parallel subgradient projection algorithm for quasiconvex equilibrium problems under the intersection of convex sets, *Optimization*, **71**, 4447–4462, (2022).
- [YMU] [53] L.H. YEN, L.D. MUU, An extragradient algorithm for quasiconvex equilibrium problems without monotonicity, *J. Global Optim.*, DOI: 10.1007/s10898-023-01291-y, (2023).
- [ZHA] [54] Y.-B. ZHAO, The iterative methods for monotone generalized variational inequalities, *Optimization*, **42**, 285–307, (1997).