Proximal Point Type Algorithms with Relaxed and Inertial Effects Beyond Convexity

S.-M. Grad* F. Lara†‡ R. T. Marcavillaca §

March 5, 2024

Abstract We show that the recent relaxed-inertial proximal point algorithm due to Attouch and Cabot remains convergent when the function to be minimized is not convex, being only endowed with certain generalized convexity properties. Numerical experiments showcase the improvements brought by the relaxation and inertia features to the standard proximal point method in this setting, too.

Keywords: Proximal point algorithms, Relaxed iterative methods, Inertial iterative methods, Generalized convexity, Prox-convexity.

1 Introduction

sec:1

In this work we show that the relaxed-inertial proximal point algorithm, originally introduced by Attouch and Cabot in [5] for minimizing a proper, lower semicontinuous and convex function with a better performance than the standard proximal point algorithm remains convergent when the convexity of the involved function is replaced with *prox-convexity*. This generalized convexity notion has been recently introduced by two of the authors in [11], motivated by the fact that for virtually all other nonconvex functions, except for the weakly convex ones, the proximity operator is a set-valued mapping, and nonconvex mixed variational inequalities involving functions enjoying it were considered in [10, 15, 16, 23]. One of the key properties of the prox-convex functions is that, when they are also proper and lower semicontinuous, their proximity operators are single-valued firmly nonexpansive mappings. The definition of this class of generalized convex functions is satisfied by some quasiconvex, weakly convex and D.C. (difference of convex) functions, an application in oligopolistic market equilibrium being also presented in [10, 15, 16, 23].

*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: felipelaraobreque@gmail.com; flarao@academicos.uta.cl. Web: felipelara.cl, ORCID-ID: 0000-0002-9965-0921

§Universidad de Tarapacá, Arica, Chile. E-mail: raultm.rt@gmail.com, ORCID-ID: 0000-0003-3748-0768

In [11] it was shown that the proximal point algorithm can be successfully employed for minimizing a prox-convex function over a closed convex set, and a natural follow-up question is whether improvements of this method could stay convergent beyond the convex framework, too. An additional motivation comes from the fact that the relaxed-inertial proximal point algorithm proposed by Attouch and Cabot in [5] in the convex setting has been recently extended by the authors in [12] for minimizing strongly quasiconvex functions over linear subspaces of finitely dimensional vector spaces. In this work we managed to show that this refined version of the proximal point algorithm can be successfully employed for minimizing a proper, lower semicontinuous and prox-convex function over a finitely dimensional linear subspace, too.

The content of the paper is arranged as follows. In Section 2 we gathered the necessary preliminary notions and results in order to make this work virtually self-contained. The considered optimization problem and the relaxed-inertial proximal point algorithm proposed for solving it are presented and investigated in Section 3. Further, Section 4 is dedicated to presenting some computational results that show that in the considered framework the relaxed-inertial version of the proximal point method is indeed faster and cheaper to use than the standard one. A brief fifth section dedicated to conclusions closes the paper.

2 Preliminaries and Basic Definitions

sec:2

The *inner product* of $\mathbb{R}^n$ and the *Euclidean norm* are denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. Given a convex and closed set $K \subseteq \mathbb{R}^n$, the *projection* of $x \in \mathbb{R}^n$ on K is denoted by $P_K(x)$, and the *indicator function* on K by $\delta_K : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

Given any $x, y, z \in \mathbb{R}^n$ and any $\beta \in \mathbb{R}$, the following relations hold:

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

3:points

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

iden:1

Given any extended real-valued function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$, the *effective domain* of h is defined by $\text{dom}\, h := \{x \in \mathbb{R}^n : h(x) < +\infty\}$. It is said that h is *proper* if $\text{dom}\, h$ is nonempty and $h(x) > -\infty$ for all $x \in \mathbb{R}^n$. We denote by $\text{epi}\, h := \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : h(x) \leq t\}$ the *epigraph* of h, by $S_\lambda(h) := \{x \in \mathbb{R}^n : h(x) \leq \lambda\}$ the *sublevel set of h at the height* $\lambda \in \mathbb{R}$ and by $\arg\min_{\mathbb{R}^n} h$ the *set of minimal points of h*. A function h is *lower semicontinuous* (lsc henceforth) at $\overline{x} \in \mathbb{R}^n$ if for any sequence $\{x_k\}_k \in \mathbb{R}^n$ with $x_k \to \overline{x}$, $h(\overline{x}) \leq \liminf_{k \to +\infty} h(x_k)$. Furthermore, the usual convention $\sup_\emptyset h := -\infty$ and $\inf_\emptyset h := +\infty$ is adopted.

A function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ with convex domain is said to be

(a) *convex* if, given any $x, y \in \text{dom}\, h$, then

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

def:convex

(*b*) *quasiconvex* if, given any $x, y \in \operatorname{dom} h$, then

$$h(\lambda x + (1-\lambda)y) \leq \max\{h(x), h(y)\}, \ \forall\, \lambda \in [0,1]. \tag{2.4}$$

def:qcx

It is said that h is *strictly convex* (resp. *strictly quasiconvex*) if the inequality in (2.3) (resp. (2.4)) is strict whenever $x \neq y$ and $\lambda \in \,]0,1[$.

Every convex function is quasiconvex. The function $h : \mathbb{R} \to \mathbb{R}$, with $h(x) := x^3$ is quasiconvex without being convex. Furthermore, we recall that h is convex if and only if epi h is a convex set and that h is quasiconvex if and only if $S_\lambda(h)$ is a convex set for all $\lambda \in \mathbb{R}$.

For algorithmic purposes, the following notions from (see [6, 17, 22]) are useful.

A function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ with a convex domain is said to be:

(*a*) *strongly convex* if there exists $\gamma \in \,]0, +\infty[$ such that for all $x, y \in \operatorname{dom} h$ and all $\lambda \in [0,1]$, we have

$$h(\lambda y + (1-\lambda)x) \leq \lambda h(y) + (1-\lambda)h(x) - \lambda(1-\lambda)\frac{\gamma}{2}\|x-y\|^2; \tag{2.5}$$

strong:convex

(*b*) *strongly quasiconvex* if there exists $\gamma \in \,]0, +\infty[$ such that for all $x, y \in \operatorname{dom} h$ and all $\lambda \in [0,1]$, we have

$$h(\lambda y + (1-\lambda)x) \leq \max\{h(y), h(x)\} - \lambda(1-\lambda)\frac{\gamma}{2}\|x-y\|^2. \tag{2.6}$$

strong:quasiconvex

In these cases, it is said that h is *strongly convex* (resp. *strongly quasiconvex*) *with modulus* $\gamma > 0$. When relations (2.5) or (2.6) hold only for all $x, y \in K \subset \operatorname{dom} h$ it is said that h is *strongly convex (quasiconvex) on* K.

Note that every strongly convex function is strongly quasiconvex, and every strongly quasiconvex function is strictly quasiconvex. The Euclidean norm $\|\cdot\|$ is strongly quasiconvex without being strongly convex on any bounded convex set $K \subseteq \mathbb{R}^n$ (see [17, Theorem 2]) and the function $x \mapsto x^3$ is strictly quasiconvex without being strongly quasiconvex on $\mathbb{R}$.

Given a proper function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$, the convex subdifferential of h at $\overline{x} \in \operatorname{dom} h$ is defined by

$$\partial h(\overline{x}) := \{\xi \in \mathbb{R}^n : \ h(y) \geq h(\overline{x}) + \langle \xi, y - \overline{x}\rangle, \ \forall\, y \in \mathbb{R}^n\}, \tag{2.7}$$

subd:usual

and by $\partial h(x) = \emptyset$ if $x \notin \operatorname{dom} h$.

For $\gamma > 0$ we define the *Moreau envelope of parameter* γ of h by

$${}^{\gamma}h(z) = \inf_{x \in \mathbb{R}^n} \left(h(x) + \frac{1}{2\gamma}\|z - x\|^2 \right). \tag{2.8}$$

The *proximity operator of parameter* $\gamma > 0$ of a function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ at $x \in \mathbb{R}^n$ is defined as

$$\operatorname{Prox}_{\gamma h} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n : x \mapsto \operatorname{arg\,min}_{y \in \mathbb{R}^n} \left\{ h(y) + \frac{1}{2\gamma}\|y - x\|^2 \right\}. \tag{2.9}$$

gammah-def

When h is proper, convex and lower semicontinuous, $\mathrm{Prox}_{\gamma h}$ turns out to be a single-valued operator, and the following the well know identity holds

$$\overline{x} = \mathrm{Prox}_{\gamma h}(z) \iff \langle \overline{x} - z, x - \overline{x} \rangle \geq \gamma \left(h(\overline{x}) - h(x) \right), \ \forall\, x \in \mathbb{R}^n. \tag{2.10}$$

Moreover, when $\gamma = 1$ we write Prox_h instead of Prox_{1h}.

We denote

$$\mathrm{Prox}_h(K, z) := \mathrm{Prox}_{(h+\delta_K)}(z) = \arg\min_{y \in K} \left\{ h(y) + \frac{1}{2\gamma} \|y - x\|^2 \right\}. \tag{2.11}$$ prx

We recall the following generalized convexity notion [10, 11].

def:pconvex **Definition 2.1.** *Let K be a closed set in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper function such that $K \cap \mathrm{dom}\, h \neq \emptyset$. It is said that h is prox-convex on K if there exists $\alpha > 0$ such that for every $z \in K$, $\mathrm{Prox}_h(K, z) \neq \emptyset$, and*

$$\overline{x} \in \mathrm{Prox}_h(K, z) \implies h(\overline{x}) - h(x) \leq \alpha \langle \overline{x} - z, x - \overline{x} \rangle, \ \forall\, x \in K. \tag{2.12}$$ prox:all

The scalar $\alpha > 0$ for which (2.12) holds is said to be the *prox-convex value* of the function h on K. When $K = \mathbb{R}^n$ we say that h is *prox-convex*.

Clearly, every convex function is prox-convex with prox-convex value $\alpha = 1$ (see [11, Proposition 3.4]). The reverse statement does not hold as the function $h : [0, 1] \to \mathbb{R}$ given by $h(x) = -x^2 - x$ shows (see [11, Example 3.1]).

Recall that for a prox-convex function h with prox-convex value $\alpha > 0$, we have

$$\overline{x} \in \mathrm{Prox}_h(K, z) \implies z - \overline{x} \in \partial \left(\frac{1}{\alpha}(h + \delta_K) \right)(\overline{x}).$$

Some interesting properties for prox-convex functions are listed below.

single-val **Lemma 2.1.** ([11, Proposition 3.3]) *Let K be a closed set in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ a proper prox-convex function on K such that $K \cap \mathrm{dom}\ h \neq \emptyset$. Then the map $z \to \mathrm{Prox}_h(K, z)$ is single-valued and firmly nonexpansive.*

If h is prox-convex with prox-convex value α, then (see [11, Page 322]) we know that $\mathrm{Prox}_{(1/\alpha)h} = \mathrm{Prox}_h$ is a singleton, hence

$$^{\frac{1}{\alpha}}h(z) = \min_{x \in K} \left(h(x) + \frac{\alpha}{2} \|z - x\|^2 \right) = h(\mathrm{Prox}_h(z)) + \frac{\alpha}{2} \|z - \mathrm{Prox}_h(z)\|^2. \tag{2.13}$$

Consequently, $^{1/\alpha}h(z) \in \mathbb{R}$ for all $z \in \mathbb{R}^n$. Furthermore, we have the following:

Lemma 2.2. ([11, Proposition 3.6]) *Let K be a closed set in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lower semicontinuous and prox-convex with prox-convex value $\alpha > 0$ on K such that $K \cap \mathrm{dom}\ h \neq \emptyset$. Then $^{1/\alpha}h : \mathbb{R}^n \to \mathbb{R}$ is Fréchet differentiable everywhere and*

$$\nabla(^{\frac{1}{\alpha}}h) = \alpha \left(I - \mathrm{Prox}_{\frac{1}{\alpha}h} \right), \tag{2.14}$$ gradient:formula

is α-Lipschitz continuous.

not:equals **Remark 2.1.** *We note that strongly quasiconvex functions and prox-convex functions are not related each other. Indeed, the function* $x \in \mathbb{R}_+ \mapsto \sqrt{x}$ *is strongly quasiconvex on convex and bounded intervals (see [18, Proposition 16]) without being prox-convex, while the constant function* $x \in \mathbb{R}^n \mapsto \alpha \in \mathbb{R}$ *is convex (hence also prox-convex) without being strongly quasiconvex.*

The following result is due to Opial [21].

lm:opial **Lemma 2.3.** *Let* $C \subseteq \mathbb{R}^n$ *be a nonempty set and* $\{x^k\}_k \subseteq \mathbb{R}^n$ *a sequence in* $\mathbb{R}^{\ltimes}$. *Assume that*

(*a*) *for every* $z \in C$, $\lim_{k\to\infty} ||x^k - z||$ *exists;*

(*b*) *every cluster point of* $\{x^k\}_k$, *as* $k \to +\infty$, *belongs to* C.

Then $\{x^k\}$ *converges, as* $k \to +\infty$, *to a point in* C.

The following lemma essentially proved in [2, Theorem 2.1] (see also [4, Lemma A.4] for a short and direct proof) will be used to analyze the convergence of the proposed algorithm.

lm:alv.att **Lemma 2.4.** *Let the sequences* $\{\varphi_k\}_k$, $\{s_k\}_k$, $\{\theta_k\}_k$ *and* $\{\delta_k\}_k$ *in* $[0, +\infty[$ *and let* $\theta \in \mathbb{R}$ *be such that* $\varphi_0 = \varphi_{-1}$, $0 \leq \theta_k \leq \theta < 1$ *and*

$$\varphi_{k+1} - \varphi_k + s_{k+1} \leq \theta_k(\varphi_k - \varphi_{k-1}) + \delta_k, \ \forall\ k \geq 0. \tag{2.15}$$ eq:alv.att02

Then the following assertions hold.

(*a*) *For all* $k \geq 1$, *we have*

$$\varphi_k + \sum_{j=1}^{k} s_j \leq \varphi_0 + \frac{1}{1-\theta} \sum_{j=0}^{k-1} \delta_j. \tag{2.16}$$ eq:alv.att01

(*b*) *If* $\sum_{k=0}^{\infty} \delta_k < +\infty$, *then* $\lim_{k\to\infty} \varphi_k$ *exists, i.e., the sequence* $\{\varphi_k\}_k$ *converges to some element in* $[0, +\infty[$.

Finally, we also consider the following lemma for our convergence analysis.

lm:inverse1 **Lemma 2.5.** ([3, Lemma A.2]) *Consider* $\phi :]0, 2[\to]0, 1[$ *a fuction given by*

$$\phi(\rho) = \frac{2(2-\rho)}{4 - \rho + \sqrt{16\rho - 7\rho^2}}.$$

Then its inverse $\phi^{-1} :]0, 1[\to]0, 2[$ *is given by*

$$\phi^{-1}(\beta) = \frac{2(\beta-1)^2}{2(\beta-1)^2 + 3\beta - 1}.$$

lm:quadratic **Lemma 2.6.** ([3, Lemma A.3]) *Let* $q : \mathbb{R} \to \mathbb{R}$ *be given by* $q(x) := ax^2 - bx + c$. *Assume that* $b, c > 0$, $b^2 - 4ac > 0$ *and define*

$$\beta := \frac{2c}{b + \sqrt{b^2 - 4ac}} > 0. \tag{2.17}$$ eq:root_b

(*a*) *If* $a = 0$*, then* $q(\cdot)$ *is a decreasing affine function and* $\beta > 0$ *is its unique root.*

(*b*) *If* $a > 0$ (*resp.* $a < 0$)*, then* $q(\cdot)$ *is a convex* (*resp. concave*) *quadratic function and* $\beta > 0$ *is its smallest* (*resp. largest*) *root.*

In both cases (*a*) *and* (*b*)*,* $\beta > 0$ *is a root of* $q(\cdot)$*, and* $q(\cdot)$ *is decreasing on the interval* $[0, \beta]$.

3 A Relaxed-Inertial Proximal Point Algorithm

sec:3

Let K be a linear space in $\mathbb{R}^n$, and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper function such that $K \subseteq \text{dom } h$. We consider the constrained optimization problem

$$\min_{x \in K} h(x). \tag{COP}$$ COP

The precise statement of Algorithm 1 that is proposed for solving (COP), originally introduced in [5] for minimizing convex functions, is given below.

rippa-pc

Algorithm 1 RIPPA for Prox-convex Functions (RIPPA-pc)

Step 0. (Initialization). Let $x^0 = x^{-1} \in K$, $\theta \in [0, 1[$, $0 < \rho' \leq \rho'' < 2$ and set $k = 0$.

Step 1. Choose $\theta_k \in [0, \theta]$ and set

$$y^k = x^k + \theta_k(x^k - x^{k-1}), \qquad \text{[extrapolation step]} \tag{3.1}$$ step:extr01

and define

$$z^k = \text{Prox}_h(K, y^k) \qquad \text{[proximal step]}. \tag{3.2}$$ step:sqcx1

Step 2. If $z^k = y^k$, then Stop, and $y^k \in \arg\min_K h$. Otherwise, choose $\rho_k \in [\rho', \rho'']$ and update

$$x^{k+1} = (1 - \rho_k)y^k + \rho_k z^k \qquad \text{[relaxation step]}. \tag{3.3}$$ eq:relax.step

Step 3. Let $k = k + 1$ and go to Step 1.

Remark 3.1. (*i*) *In virtue of relations* (3.1) *and* (3.3)*,* K *should be assumed to be a linear subspace and not only as a closed and convex set. We note that the authors in [1, 2, 9, 19] considered their algorithms on the whole space, while in [13, 24]* 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.*

(*ii*) *The proximal step* (3.2) *is well defined by the prox-convexity of* h, *that is,* $\mathrm{Prox}_h(K, y^k) \neq \emptyset$ *and it is uniquely defined by Lemma 2.1. Moreover, Algorithm 1 generalizes the classical proximal point algorithm for minimizing prox-convex functions [11, Theorem 4.1]. Indeed, the classical proximal point algorithm is obtained by taking* $\theta = 0$ *and* $\rho' = \rho'' = 1$, *in which case* $y^k = x^k$, *see Corollary 3.1 below.*

(*iii*) *In view of* (2.14), *the update of the Algorithm 1 can be written in an equivalent way as*

$$\begin{cases} y^k &= x^k + \theta_k (x^k - x^{k-1}), \\ x^{k+1} &= y^k - \rho_k \nabla (^1h)(y^k). \end{cases}$$

Hence Algorithm 1 can be seen as an inertial gradient method applied to the Moreau envelope of h.

3.1 Convergence analysis

sec:3-1

We start the convergence analysis of Algorithm 1 by proving the following result.

lmm:auxR1

Proposition 3.1. *Let* K *be a linear space in* $\mathbb{R}^n$ *and* $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be a proper, lower semicontinuous and prox-convex on* K *function such that* $K \subseteq \mathrm{dom}\, h$ *and* $\arg\min_K h \neq \emptyset$. *Let* $\{x^k\}_k$, $\{y^k\}_k$ *and* $\{z^k\}_k$ *be the sequences generated by Algorithm 1. Then, for every* $x^* \in \mathbb{R}^n$, *we have*

$$||x^{k+1} - x^*||^2 + \rho_k (2 - \rho_k) ||y^k - z^k||^2 + \frac{2\rho_k}{\alpha} \left(h(z^k) - h(x^*) \right) \leq ||y^k - x^*||^2, \ \forall\, k \geq 0. \tag{3.4}$$

eq:aux.03

Proof. From the relaxation step (3.3) of Algorithm 1 and using identity (2.2), we obtain

$$\begin{aligned} \|x^{k+1} - x^*\|^2 - \|y^k - x^*\|^2 = {} & \rho_k (\|z^k - x^*\|^2 - \|y^k - x^*\|^2) \\ & - \rho_k (1 - \rho_k) \|y^k - z^k\|^2. \end{aligned} \tag{3.5}$$

eq:1

Furthermore, it follows from relation (2.1) that

$$\|z^k - x^*\|^2 - \|y^k - x^*\|^2 = \|z^k - y^k\|^2 + 2\langle z^k - y^k, y^k - x^* \rangle. \tag{3.6}$$

eq:2

On the other hand, from the proximal step (3.2), we have for every $k \geq 0$ that,

$$\begin{aligned} & z^k = \mathrm{Prox}_h(K, y^k) \implies y^k - z^k \in \partial \left(\frac{1}{\alpha} h + \delta_K \right) (z^k) \\ & \Longleftrightarrow \alpha \langle y^k - z^k, x - z^k \rangle \leq h(x) - h(z^k), \ \forall\, x \in K, \end{aligned}$$

Taking $x = x^*$, we have $\langle z^k - y^k, z^k - x^* \rangle \leq \frac{1}{\alpha}(h(x^*) - h(z^k))$. This imples

$$\langle z^k - y^k, y^k - x^* \rangle \leq -\|z^k - y^k\|^2 + \frac{1}{\alpha} \left(h(x^*) - h(z^k) \right). \tag{3.7}$$

eq:3

It follows from (3.6) and (3.7) that

$$\|z^k - x^*\|^2 - \|y^k - x^*\|^2 \leq -\|z^k - y^k\|^2 + \frac{2}{\alpha}(h(x^*) - h(z^k)).$$

The assertion in (3.4) follows by replacing the last relation in (3.5). □

In what follows, for any $x^* \in \mathbb{R}^n$ we consider the anchor sequence $\{\varphi_k\}_k$ defined as

$$\varphi_k := ||x^k - x^*||^2. \tag{3.8}$$ eq:varphi

As a consequence of the previous result, we have the following statement.

prop:01 **Proposition 3.2.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \operatorname{dom} h$ and $\arg\min_K h \neq \emptyset$. Let $\{x^k\}_k$, $\{y^k\}_k$ and $\{z^k\}_k$ be the sequences generated by Algorithm 1. Then, for every $k \geq 0$, we have*

$$\varphi_{k+1} - \varphi_k - \theta_k(\varphi_k - \varphi_{k-1}) + \frac{2-\rho_k}{\rho_k}||x^{k+1} - y^k||^2 + \frac{2\rho_k}{\alpha}\left(h(z^k) - h(x^*)\right) \leq (\theta_k^2 + \theta_k)||x^k - x^{k-1}||^2. \tag{3.9}$$ eq:5

Proof. The proof follows in the lines of the one of [12, Proposition 3.2] by using Proposition 3.1 instead of [12, Proposition 3.1], so we omit it. □

Our first main result, which shows that the sequences $\{x^k\}_k$, $\{y^k\}_k$ and $\{z^k\}_k$, generated by Algorithm 1 converge to an optimal solution to problem (COP) for prox-convex functions, is given below.

th:main **Theorem 3.1.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \operatorname{dom} h$ and $\arg\min_K h \neq \emptyset$, $0 < \rho' \leq \rho'' < 2$, $\{\rho_k\}_k \subseteq [\rho', \rho'']$, $\theta \in [0, 1[$, $\{\theta_k\}_k \subseteq [0, \theta]$ and $\{x^k\}_k$, $\{y^k\}_k$ and $\{z^k\}_k$, be the sequences generated by Algorithm 1. If*

$$\sum_{k=0}^{\infty} \theta_k ||x^k - x^{k-1}||^2 < +\infty, \tag{3.10}$$ eq:th:main.01

then the following assertions hold:

(*a*) *For every $\overline{x} \in \arg\min_K h$, the limit $\lim_{k\to\infty} \|x^k - \overline{x}\|$ exists and*

$$\lim_{k\to+\infty} \|x^{k+1} - y^k\| = \lim_{k\to+\infty} \|z^k - y^k\| = \lim_{k\to+\infty} h(z^k) - h(\overline{x}) = 0. \tag{3.11}$$ eq:6

(*b*) *$\{x^k\}_k$ converges to a point $\overline{x} \in \arg\min_K h$ and $\liminf_{k\to+\infty} h(x^k) = \min_K h$. Moreover, the sequences $\{y^k\}_k$ and $\{z^k\}_k$ converge both to $\overline{x}$, too.*

Proof. (a): Let us fix $\overline{x} \in \arg\min_K h$. As $x^{-1} = x^0$, it holds from definition of φ_k that $\varphi_{-1} = \varphi_0$. Taking in Proposition 3.2, $\delta_k := (\theta_k^2 + \theta_k)||x^k - x^{k-1}||^2$ and

$$s_{k+1} := \frac{2-\rho_k}{\rho_k}||x^{k+1} - y^k||^2 + \frac{2\rho_k}{\alpha}\left(h(z^k) - h(\overline{x})\right),$$

we observe that condition (2.15) in Lemma 2.4 is fulfilled. Hence, we conclude from assumption (3.10) and Lemma 2.4(b) that $\lim_{k\to\infty}||x^k - \overline{x}||$ exists and, in particular, $\{x^k\}_k$ is bounded. Also, from assumption (3.10) and Lemma 2.4(a), we conclude that $\sum_{k=0}^{\infty} s_{k+1} < +\infty$, and so, $s_{k+1} \to 0$ as $k \to +\infty$. Thus due to the boundedness of $\{\rho_k\}_k$ it hold that $\lim_{k\to+\infty}\|x^{k+1} - y^k\| = 0$ and $\lim_{k\to+\infty}(h(z^k) - h(\overline{x})) = 0$.

Finally, since $x^{k+1} - y^k = \rho_k(z^k - y^k)$ by relation (3.3) and $\{\rho_k\}_k$ is bounded, it follows that $\lim_{k\to+\infty}\|z^k - y^k\| = 0$.

(b): Let us verify the hypotheses of Opial's lemma (see Lemma 2.3). Its item (a) is fulfilled due to the first part of (a). It remains to verify (b). For that purpose let $\widehat{x} \in K$ be a cluster point of $\{x^k\}_k$ (because $\{x^k\}_k$ is bounded by part (a)). Then, there exists a subsequence $\{x^{k_l}\}_l$ of $\{x^k\}_k$ such that $x^{k_l} \to \widehat{x}$ as $l \to +\infty$. Applying (3.11) to the subsequence $\{x^{k_l}\}_l$ and since $\lim_{l\to+\infty} x^{k_l} = \widehat{x}$, we conclude that

$$\lim_{l\to+\infty} y^{k_l} = \lim_{l\to+\infty} z^{k_l} = \widehat{x}. \qquad (3.12)$$ useful:1

Now, since h is prox-convex function and $z^k = \mathrm{Prox}_h(K, y^k)$, we have

$$h(z^k) - h(x) \leq \alpha\langle z^k - y^k, x - z^k\rangle, \ \forall\, x \in K. \qquad (3.13)$$

Replace k by k_l in the previous inequality. Then, taking $\liminf_{l\to+\infty}$ and using (3.12), $x^{k_l} \to \widehat{x}$, as $l \to +\infty$, and the lower semicontinuity of h, we obtain

$$h(\widehat{x}) \leq \liminf_{l\to+\infty} h(z^{k_l}) \leq h(x), \ \forall\, x \in K. \qquad (3.14)$$

This implies that $\widehat{x} \in \arg\min_K h$, and hence, every cluster point of the sequence $\{x^k\}_k$ belongs to $\arg\min_K h$, which proves the part (b) of Opial's lemma. Therefore, we conclude from Lemma 2.3 that the whole sequence $\{x^k\}_k$ converges to some $\overline{x} \in \arg\min_K h$. Finally, since $x^k \to \overline{x} \in \arg\min_K h$, as $k \to +\infty$, it follows from the first two limits of (3.11) that $y^k \to \overline{x}$ and $z^k \to \overline{x}$ as $k \to +\infty$ and the proof is complete. □

Remark 3.2. *In virtue of Remark 2.1, Algorithm 1 deals with a different class of generalized convex functions than the relaxed-inertial proximal point algorithm proposed in [12] which was considered for strongly quasiconvex functions.*

sec:3-2

3.2 Sufficient Conditions

In this section, we provide sufficient conditions for ensuring the fulfillment of assumption (3.10). To that end, we first show the following result.

prop:02 **Proposition 3.3.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \operatorname{dom} h$ and $\arg\min_K h \neq \emptyset$. Let $\{x^k\}_k$, $\{y^k\}_k$ and $\{z^k\}_k$ be the sequences generated by Algorithm 1. Given any $\overline{x} \in \arg\min_K h$, for $k \geq 0$ we set $\varphi_k := \|x^k - \overline{x}\|^2$ as in* (3.8)*. Then, for any $k \geq 0$ one has*

$$\begin{aligned}\varphi_{k+1} - \varphi_k - \theta_k(\varphi_k - \varphi_{k-1}) - \nu_k\|x^k - x^{k-1}\|^2 \\ \leq \frac{2-\rho''}{\rho''}(\theta_k - 1)||x^{k+1} - x^k||^2,\end{aligned} \tag{3.15}$$

eq:8

where $\nu_k := 2\left(1 - \frac{1}{\rho''}\right)\theta_k^2 + \frac{2}{\rho''}\theta_k$.

Proof. Similar to [12, Proposition 4.1]. □

The following result extends, in some sense, [1, Proposition 2.5] which was proved for the relaxed-inertial proximal point algorithm for solving monotone inclusion problems (see also, [4, Theorem 3.5] and [3, Theorem 2] for aplications to splitting methods) from the convex to the prox-convex case.

thM:main2 **Theorem 3.2.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \operatorname{dom} h$ and $\arg\min_K h \neq \emptyset$, $0 < \rho' \leq \rho'' < 2$, $\{\rho_k\}_k \subseteq [\rho', \rho'']$, $\theta \in [0, 1[$, and $\{\theta_k\}_k$ is nondecreasing satisfying the following (for some $\beta > 0$)*

$$0 \leq \theta_k \leq \theta_{k+1} \leq \theta < \beta < 1, \ \forall\, k \geq 0, \tag{3.16}$$

eq:alpha

and

$$\rho'' = \rho''(\beta, \rho') := \frac{2(\beta-1)^2}{2(\beta-1)^2 + 3\beta - 1}. \tag{3.17}$$

eq:rho

If $\{x^k\}_k$ is the sequence generated by Algorithm 1, then

$$\sum_{k=1}^{\infty} ||x^k - x^{k-1}||^2 < +\infty. \tag{3.18}$$

eq:sum

As a consequence, the sequence $\{x^k\}_k$ converges to some point $\overline{x} \in \arg\min_K h$ and $\liminf_{k\to+\infty} h(x^k) = \min_K h$.

Proof. For any $\overline{x} \in \arg\min_K h$, we set φ_k as in (3.8). Define

$$\Gamma_k := \varphi_k - \theta_k\varphi_{k-1} + \nu_k||x^k - x^{k-1}||^2, \ \forall\, k \geq 0. \tag{3.19}$$

eq:def.mu_k

Let $k \geq 0$. Since $\varphi_k \geq 0$ and $\{\theta_k\}_k$ is nondecreasing, it follows from definition of Γ_k and (3.3) that

$$\begin{aligned}\Gamma_{k+1}-\Gamma_k &= \left(\varphi_{k+1}-\varphi_k-\theta_k(\varphi_k-\varphi_{k-1})-\nu_k\|x^k-x^{k-1}\|^2\right)+\nu_{k+1}||x^{k+1}-x^k||^2\\ &\le \left(\frac{2-\rho''}{\rho''}(\theta_k-1)+\nu_{k+1}\right)||x^{k+1}-x^k||^2\\ &\le \left((\frac{2}{\rho''}-1)\theta_{k+1}-\frac{2}{\rho''}+1+\nu_{k+1}\right)||x^{k+1}-x^k||^2\\ &= -q(\theta_{k+1})||x^{k+1}-x^k||^2, \end{aligned} \tag{3.20}$$

eq:10

where $q:\mathbb{R}\to\mathbb{R}$ is a quadratic function defined by

$$q(x):=2\left(\frac{1}{\rho''}-1\right)x^2-\left(\frac{4}{\rho''}-1\right)x+\left(\frac{2}{\rho''}-1\right). \tag{3.21}$$

eq:def-q1

We show (3.18) following the same lines of [4, Theorem 2]. To this end, we claim that $q(\theta_{k+1})$ admits a uniform positive lower bound. Indeed, we have from (3.17) and Lemma 2.5(a), that

$$\beta:=\frac{2(2-\rho'')}{4-\rho''+\sqrt{16\rho''-7\rho''^2}}.$$

From which and Lemma 2.6 with $q(\cdot)$ as in (3.21), $a=2(1/\rho''-1)$, $b=4/\rho''-1$ and $c=2/\rho''-1$, we conclude that $q(\beta)=0$ and $q(\cdot)$ is decreasing on $[0,\beta]$. Thus in view of (3.16), we deduce

$$q(\theta_{k+1})\ge q(\theta)>q(\beta)=0,$$

that shows the claim. Now, combining the claim with (3.20), we deduce

$$q(\theta)||x^{k+1}-x^k||^2\le \Gamma_k-\Gamma_{k+1}, \tag{3.22}$$

eq:d7

thus $\{\Gamma_k\}_k$ is non increasing.

On the other hand, since $x^{-1}=x^0$, it follows from the definition of Γ_k and φ_k that $\Gamma_0=(1-\theta_0)\varphi_0\le\varphi_0=\|x^0-\overline{x}\|^2$. Thus, (3.19), (3.16) and the monotonicity of $\{\Gamma_k\}_k$ yields

$$\varphi_k-\theta_k\varphi_{k-1}\le\varphi_k-\theta_k\varphi_{k-1}+\nu_k||x^k-x^{k-1}||^2=\Gamma_k\le\Gamma_0\le\varphi_0,\ \forall\ k\ge 0.$$

From this last inequality and (3.16), we recursively obtain

$$\varphi_k\le\theta\varphi_{k-1}+\varphi_0\le\cdots\le\theta^k\varphi_0+\varphi_0\sum_{j=0}^{k-1}\theta^j\le\theta^k\varphi_0+\frac{\varphi_0}{1-\theta}. \tag{3.23}$$

eq:X1

Therefore, taking into account $\Gamma_{k+1}\ge-\theta\varphi_k$ (see (3.19)), from (3.22) and (3.23), we deduce

$$\sum_{j=0}^{k}||x^{j+1}-x^j||^2\le\frac{1}{q(\theta)}(\Gamma_0-\Gamma_{k+1})\le\frac{(\varphi_0+\theta\varphi_k)}{q(\theta)}\le\frac{1}{q(\theta)}\left(\frac{\varphi_0}{1-\theta}+\theta^{k+1}\varphi_0\right).$$

Finally, letting $k \to +\infty$, we obtain

$$\sum_{k=0}^{+\infty} ||x^{k+1} - x^k||^2 \leq \frac{\|x^0 - \overline{x}\|^2}{(1-\theta)q(\theta)} < +\infty, \tag{3.24}$$ eq:sum1

which proves (3.18). Finally, it follows from (3.18) and Theorem 3.1 that $\{x^k\}_k$ converges to some point in $\arg\min_K h$, say $\overline{x}$, i.e., $\lim_{k\to+\infty} x^k = \overline{x}$. Moveover, since h is lower semicontinuous, it yields that $\liminf_{k\to+\infty} h(x^k) = \min_K h$. □

When $\theta = 0$ (in which case $\theta_k = 0$ for all $k \geq 0$), Algorithm 1 collapses to a relaxed version of the proximal point algorithm for minimizing prox-convex functions studied in [11, Theorem 4.1], as shown below.

coro:rppa **Corollary 3.1.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \mathrm{dom}\, h$ and $\arg\min_K h \neq \emptyset$ and $\{\rho_k\}_k \subset [\rho', \rho'']$ with $0 < \rho' \leq \rho'' < 2$. Then, for any sequence $\{x^k\}_k$ generated by*

$$x^{k+1} = (1-\rho_k)x^k + \rho_k \operatorname{Prox}_h(K, x^k), \tag{RPPA-pc}$$ rppa

we have

- *(a)* $\displaystyle\sum_{k=0}^{+\infty} \frac{2-\rho_k}{\rho_k} \|x^{k+1} - x^k\|^2 < +\infty$ *and* $\displaystyle\sum_{k=0}^{\infty} \frac{2\rho_k}{\alpha} \left(h(z^k) - h(\overline{x})\right) < +\infty$ *for any* $\overline{x} \in \arg\min_K h$, *where* $z^k := \operatorname{Prox}_h(K, x^k)$.
- *(b)* *For every* $\overline{x} \in \arg\min_K h$, *the limit* $\lim_{k\to+\infty}\|x^k - \overline{x}\|$ *exists, and hence* $\{x^k\}_k$ *is bounded.*
- *(c)* *The sequence* $\{x^k\}_k$, *generated by* (RPPA-pc), *converges to some* $\overline{x} \in \arg\min_K h$ *and* $\liminf_{k\to+\infty} h(x^k) = \min_K h$.

Proof. Let $\overline{x} \in \arg\min_K h$. It follows from (3.4) with $y^k = x^k$ (because $\theta_k = 0$) that

$$\frac{2-\rho_k}{\rho_k}||x^{k+1} - x^k||^2 + \frac{2\rho_k}{\alpha}\left(h(z^k) - h(\overline{x})\right) \leq ||x^k - \overline{x}||^2 - ||x^{k+1} - \overline{x}||^2. \tag{3.25}$$ x1

Adding on both side of the inequality of (3.25), we show (a). Also it holds from (3.25) that $\{\|x^k - \overline{x}\|\}_k$ is non-increasing and bounded, and therefore, convergent. In particular, $\{x^k\}$ is bounded. Thus we have been proved (b).

On the other hand, since $\frac{2-\rho_k}{\rho_k} \geq \frac{2}{\rho''} - 1$, it follows from the first statement of (a) that

$$\sum_{k=0}^{+\infty}\|x^{k+1} - x^k\|^2 < +\infty.$$

Hence, the conclusion in (c) follows immediately from Theorem 3.2. □

cl-cx **Remark 3.3.** *When one takes $\rho'' \leq 1$ in Corollary 3.1, the set K needs not be a linear subspace and can be taken only closed and convex.*

In the particular case of Theorem 3.2 when $\rho_k = \rho' = \rho'' = 1$ for every $k \geq 0$, which corresponds to the absence of relaxation effects in Algorithm 1, we obtain from (3.17) that $\beta = 1/3$, which is the condition on the inertial parameters provided in [2, Proposition 2.1] in our *inertial proximal point method* for solving the prox-convex minimization problem (COP) proposed below.

coro:ippa **Corollary 3.2.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \operatorname{dom} h$ and $\arg\min_K h \neq \emptyset$, $\theta \in [0, 1[$, and $\{\theta_k\}_k$ is nondecreasing satisfying $0 \leq \theta_k \leq \theta_{k+1} \leq \theta < 1/3$. Then for any sequence $\{x^k\}_k$ generated by*

$$\text{(IPPA)} \qquad \begin{cases} k \leftarrow k+1 \\ y^k = x^k + \theta_k(x^k - x^{k-1}) \\ x^{k+1} = \operatorname{Prox}_h(K, y^k), \end{cases} \tag{3.26}$$

we have

(*a*) $\displaystyle\sum_{k=1}^{+\infty} \|x^k - x^{k-1}\|^2 < +\infty$;

(*b*) *for $\{\overline{x}\} \in \arg\min_K h$, the limit $\lim_{k\to+\infty} \|x^k - \overline{x}\|$ exists, and hence $\{x^k\}_k$ is bounded;*

(*c*) *the sequence $\{x^k\}_k$, generated by* (IPPA), *converges to $\overline{x} \in \arg\min_K h$ and $\liminf_{k\to+\infty} h(x^k) = \min_K h$.*

We finish this section by providing a convergence rate result for the proposed algorithm.

rate.conv **Proposition 3.4.** *Let K be a linear space in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and prox-convex on K function such that $K \subseteq \operatorname{dom} h$ and $\arg\min_K h \neq \emptyset$. Let $\{x^k\}_k$, $\{y^k\}_k$ and $\{z^k\}_k$ be the sequences generated by Algorithm 1. Then, we have the following results*

(*a*) *for any $\overline{x} \in \arg\min_K h$ and positive integer k, it holds that*

$$\min_{0\leq j\leq k} \|x^{j+1} - x^j\|^2 \leq \frac{\left(\frac{1+\theta-\theta^2}{q(\theta)(1-\theta)}\right)\|x^0 - \overline{x}\|^2}{k}; \tag{3.27}$$ eq:MT1

(*b*)

$$\min_{0\leq j\leq k} h(z^j) - \min_K h \leq \frac{\alpha\left(1 + \frac{2\theta(1+\theta-\theta^2)}{(1-\theta)^2 q(\theta)}\right)\|x^0 - \overline{x}\|^2}{(2\rho')k}. \tag{3.28}$$ eq:MT2

Proof. (a): The last inequality before (3.24) gives

$$\sum_{j=0}^{k} ||x^{j+1} - x^j||^2 \leq \frac{1}{q(\theta)} \left(\frac{\varphi_0}{1-\theta} + \theta^{k+1}\varphi_0 \right) \leq \frac{1+\theta-\theta^2}{q(\theta)(1-\theta)} \|x^0 - \overline{x}\|^2. \quad (3.29)$$

eq:130

By choosing some $0 \leq j \leq k$ in (3.29), we deduce (3.27).

(b): Now, combining Proposition 3.2 and Lemma 2.4(a) with $\delta_k = \theta_k(1 + \theta_k)||x^k - x^{k-1}||^2$ and $s_{k+1} = \frac{2-\rho_k}{\rho_k}||x^{k+1} - y^k||^2 + \frac{2\rho_k}{\alpha}\left(h(z^k) - h(\overline{x})\right)$, for every $k \geq 0$, we have

$$\sum_{j=0}^{k} s_{j+1} \leq \varphi_0 - \varphi_k + \frac{1}{1-\theta}\sum_{j=1}^{k-1} \delta_j \leq \varphi_0 + \frac{2\theta}{1-\theta}\sum_{j=1}^{k} ||x^j - x^{j-1}||^2$$
$$\leq \left(1 + \frac{2\theta(1+\theta-\theta^2)}{(1-\theta)^2 q(\theta)}\right) \|x^0 - \overline{x}\|^2. \quad (3.30)$$

eq:015

Since $s_{k+1} \geq \frac{2\rho_k}{\alpha}\left(h(z^k) - h(\overline{x})\right)$ for every k, (3.30) yields

$$\sum_{j=0}^{k} \frac{2\rho_k}{\alpha}\left(h(z^k) - h(\overline{x})\right) \leq \left(1 + \frac{2\theta(1+\theta-\theta^2)}{(1-\theta)^2 q(\theta)}\right) \|x^0 - \overline{x}\|^2.$$

Hence, since $\rho_k \geq \rho'$ for every $k \geq 0$, picking $0 \leq j \leq k$ in last inequality, we deduce (3.28), which completes the proof. □

4 Computational Results

sec:4

While in the convex setting the superiority of a relaxed-inertial proximal point algorithm in comparison with its standard proximal point version has been well-documented (see [5]), in the case of minimizing a prox-convex function no studies about this can be found in the literature. However, the results presented in [15,16,23] for nonconvex mixed variational inequalities involving prox-convex functions suggest that a similar phenomenon should occur. We present in the following some computational results obtained in MATLAB 2019b-Win64 on a Lenovo Yoga 260 Laptop with Windows 10 and an Intel Core i7 6500U CPU with 2.59 GHz and 16GB RAM after implementing the relaxed-inertial proximal point method Algorithm 1 and, for comparison, the "pure" proximal point method [11, (4.1)].

The first example presented below revisits [11, Example 4.1] and exhibits situations where the relaxed-inertial method considered in this work has a superior performance (in terms of both CPU time until reaching a suitable approximate optimal solution to considered problem and the number of necessary iterations for achieving this goal) to its standard proximal point counterpart. The stopping criterion of the algorithm proposed in this paper is activated when the norm (in particular the absolute value) of the difference between y^k and z^k is not larger than an a priori given error $\varepsilon > 0$, while for the "pure" proximal point method this role is played by the difference between two consecutive members of the generated sequence $\{x^k\}_k$.

ex41 **Example 4.1.** *Let $K = \mathbb{R}^2$ and consider the function $h : \mathbb{R}^2 \to \overline{\mathbb{R}}$ given by $h(x_1, x_2) = \delta_{[0,2]}(x_1) + x_2^2 - x_1^2 - x_1$, which is prox-convex on K (see [11, Example 3.1] and [16, Example 2.2]). The global minimum of h is $(2, 0)^\top$, and its proximity operator (on $K = \mathbb{R}^2$)*

$$\operatorname{Prox}_h(z_1, z_2) = \left(\begin{cases} 0, & \text{if } z_1 \leq -2 \\ 2, & \text{if } z_1 > -2 \end{cases}, \ \frac{z_2}{3} \right)^\top, \ z_1, z_2 \in \mathbb{R}.$$

Consider the following parameter constellation (roughly inspired by the convergence analysis results provided in [5, Section 3]), $\theta = 0.3$, $\theta_k = \theta - 1/(5((k+1)^2))$, $\rho' = 0.6$, $\rho'' = 1.4$, $\rho_k = (1/k)\rho' + (1 - 1/k)\rho''$, $k \geq 1$ and $\varepsilon = 10^{-9}$, with the starting points $x^0 = x^{-1} = (0, 999)^\top$. The standard proximal point algorithm required 26 iterations and 0.0755 seconds for delivering an ε-optimal solution to (COP)*, while Algorithm 1 needed only 18 iterations and a CPU time of 0.0262 seconds. Taking $\rho'' = 1.6$, the resources consumed by Algorithm 1 to the same end decreased to only 15 iterations and a CPU time of 0.0203 seconds. When θ was lowered to 0.2, only 14 iterations remained necessary for delivering an ε-optimal solution to* (COP) *after 0.0160 seconds. These values could be improved to 13 iterations and 0.0090 seconds by fine tuning the parameters to $\theta = 0.1$.*

For a coarser admissible error, i.e. $\varepsilon = 10^{-5}$, the "pure" proximal point method needed 18 iterations and 0.0422 seconds for finding an ε-optimal solution to (COP)*, while Algorithm 1 provided it in 10 iterations and 0.0048 seconds in the latest constellation.*

We present below another computational example, inspired by the application in oligopolistic equilibrium proposed in [10, Section 4.2]. More precisely, we are interested in minimizing a separable function in two variables that is prox-convex in one and convex in the other that combines a standard cost function (see [10, 20]) and the classical Huber loss function introduced in [14]. The obtained results (derived by employing the same stopping criteria as in Example 4.1) show that the relaxed-inertial method considered in this work outperforms its standard proximal point counterpart.

Example 4.2. *Let $K = \mathbb{R}^2$ and consider the function $h : \mathbb{R}^2 \to \overline{\mathbb{R}}$ given by*

$$h(x_1, x_2) = \delta_{[1,2]}(x_1) + 5x_1 + \ln(1 + 10x_1) + \begin{cases} \frac{x_2^2}{2}, & \text{if } |x_2| \leq 1, \\ |x_2| - \frac{1}{2}, & \text{otherwise}, \end{cases}$$

which is prox-convex on K (see [11, Example 3.3] and [16, Example 2.3(2)] as well as [14]). The global minimum of h is $(1, 0)^\top$, and its proximity operator (on $K = \mathbb{R}^2$) is (cf. [8, 11])

$$\operatorname{Prox}_{\gamma h}(z_1, z_2) = \begin{cases} \left(1, \frac{z_2}{\gamma+1}\right)^\top, & \text{if } |z_2| \leq \gamma + 1, \\ \left(1, z_2 - \gamma \operatorname{sgn}(z_2)\right)^\top, & \text{if } |z_2| > \gamma + 1, \end{cases} \quad z_1, z_2 \in \mathbb{R}, \ \gamma > 0.$$

Considering the same parameter constellation as in Example 4.1 except for $\theta = 0.1$ and $\rho'' = 1.6$, and with the starting points $x^0 = x^{-1} = (2, 999)^\top$, when taking the admissible error $\varepsilon = 10^{-5}$, the standard proximal point algorithm required 744 iterations and 0.0840 seconds for delivering an ε-optimal solution to the considered minimization problem, while Algorithm 1 needed only 629 iterations and a CPU time of 0.0344 seconds. For the allowed error $\varepsilon = 10^{-9}$, the standard proximal point algorithm did the job after 772 iterations in 0.1089 seconds, while Algorithm 1 solved the problem after 726 iterations in 0.0465 seconds. Taking $\theta = 12/(5(1+\sqrt{5}))$ (following [10, Section 4.2]), the performance of the relaxed-inertial method improved significantly, for the admissible error $\varepsilon = 10^{-5}$ Algorithm 1 needed only 149 iterations and a CPU time of 0.0082 seconds, while for $\varepsilon = 10^{-9}$ it provided an ε-optimal solution to the considered minimization problem after 166 iterations in 0.0136 seconds.

Both these examples show that, like in the convex setting, adding inertial and relaxation effects to the standard proximal point algorithm improves its performance when minimizing prox-convex functions, too. Theoretical results similar to the ones provided in [5, Section 3] would be welcome in order to identify constellations where the relaxed-inertial proximal point algorithm surely outperforms its standard counterpart.

5 Conclusions

In order to improve the performances of the standard proximal point algorithm for minimizing a prox-convex function over a finitely dimensional linear subspace we adapted a relaxed-inertial version of it recently proposed in the convex setting. Besides showing the convergence statement, we provided several additional theoretical results and a convergence rate statement, while some numerical experiments confirm the improvements brought by the relaxation and inertia features to the “pure” proximal point method in the prox-convex framework, too.

As subsequent work to the present one we plan to investigate whether the convergence analysis results provided in [5, Section 3] could be adapted to the prox-convex setting, in order to better understand how to fine tune the parameters of the relaxed-inertial proximal point method with the aim of further lowering the costs (in terms of number of iteration and CPU time) of employing it for minimizing prox-convex functions.

Acknowledgements: This research was partially supported by ANID–Chile under project Fondecyt Regular 1220379 (Lara), and by the Hi! PARIS Center and a public grant as part of the Investissement d’avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH (Grad). The authors are grateful to an anonymous referee for comments that contributed to improving the presentation.

References

A-2003 [1] 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).

AA-2001 [2] 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(1-2)**, 3–11, (2001).

MEGM [3] M.M. Alves, J. Eckstein, M. Geremia, J.G. Melo, Relative-error inertial-relaxed inexact versions of Douglas-Rachford and ADMM splitting algorithms, *Comput. Optim. Appl.*, **75**, 389–422, (2020).

MMa [4] M.M. Alves, R.T. Marcavillaca, On inexact relative-error hybrid proximal extragradient, forward-backward and Tsengs modified forward-backward methods with inertial effects, *Set-Valued Var. Anal.*, **28**, 301–325, (2020).

ATC [5] H. Attouch, A. Cabot, Convergence rate of a relaxed inertial proximal algorithm for convex minimization, *O*ptimization, **6**9, 1281–1312, (2020).

BC-2 [6] H.H. Bauschke, P.L. Combettes, "Convex Analysis and Monotone Operators Theory in Hilbert Spaces". *CMS Books in Mathematics*. Springer-Verlag, second edition, (2017).

CMA [7] A. Cambini, L. Martein. "Generalized Convexity and Optimization". Springer-Verlag, Berlin-Heidelberg, (2009).

COP [8] P. L. Combettes, J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim.*, **18**, 1351–1376, (2007).

GOM [9] X. Goudou, J. Munier, The gradient and heavy ball with friction dynamical systems: the quasiconvex case, *Math. Programm.*, **116**, 173–191, (2007).

GL2 [10] S.-M. Grad, F. Lara, Solving mixed variational inequalities beyond convexity, *J. Optim. Theory Appl.*, **190**, 565–580, (2021).

GL1 [11] S.-M. Grad, F. Lara, An extension of the proximal point algorithm beyond convexity, *J. Global Optim.*, **82**, 313–329, (2022).

GLM1 [12] S.-M. Grad, F. Lara, R.T. Marcavillaca, Relaxed-inertial proximal point type algorithms for quasiconvex minimization, *J. Global Optim.*, **85**, 615–635, (2023).

HDT [13] D. Hieu, H.N. Duong, B.H. Thai, Convergence of relaxed inertial methods for equilibrium problems, *J. Appl. Numer. Optim.*, **3**, 215–229, (2021).

HUB [14] P.J. Huber, Robust estimation of a location parameter, *Ann. Math. Statist.*, **35**, 73–101, (1964).

ISD [15] C. Izuchukwu, Y. Shehu, Q.-L. Dong, Two-step inertial forward-reflected-backward splitting based algorithm for nonconvex mixed variational inequalities, *J. Comput. Appl. Math.*, **426**, 115093, (2023).

ISO [16] C. Izuchukwu, Y. Shehu, C.C. Okeke, Extension of forward-reflected-backward method to non-convex mixed variational inequalities, *J. Global Optim.*, **86**, 123–140, (2023).

J-2 [17] M. Jovanović, A note on strongly convex and quasiconvex functions, *Math. Notes*, **60**, 584–585, (1996).

Lara-9 [18] F. Lara, On strongly quasiconvex functions: existence results and proximal point algorithms, *J. Optim. Theory Appl.*, **192**, 891–911, (2022).

MAI [19] P.E. Maingé, Asymptotic convergence of an inertial proximal method for unconstrained quasiconvex minimization, *J. Global Optim.*, **45**, 631–644, (2009).

MUQ [20] L.D. Muu, N.V. Quy, Global optimization from concave minimization to concave mixed variational inequality, *Acta Math. Vietnam.*, **45**, 449–462, (2020).

OPI [21] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, **73**, 591–597, (1967).

POL [22] B.T. Polyak, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Soviet Math.*, **7**, 72–75, (1966).

TSX [23] V.N. Tran, Y. Shehu, R. Xu, P.T. Vuong, Inertial forward-backward-forward method for nonconvex mixed variational inequalities, *Research Square*, DOI: 10.21203/rs.3.rs-2193605/v1, (2022).

VTV [24] L. V. Vinh, V.N. Tran, P.T. Vuong, A second-order dynamical system for equilibrium problems, *Numer. Algor.*, **91**, 327–351, (2022).