

Solving proximal split feasibility problems without prior knowledge of operator norms

A. Moudafi · B. S. Thakur

Received: 30 April 2013 / Accepted: 4 November 2013 / Published online: 19 November 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract In this paper our interest is in investigating properties and numerical solutions of Proximal Split feasibility Problems. First, we consider the problem of finding a point which minimizes a convex function f such that its image under a bounded linear operator A minimizes another convex function g . Based on an idea introduced in Lopez (Inverse Probl 28:085004, 2012), we propose a split proximal algorithm with a way of selecting the step-sizes such that its implementation does not need any prior information about the operator norm. Because the calculation or at least an estimate of the operator norm $\|A\|$ is not an easy task. Secondly, we investigate the case where one of the two involved functions is prox-regular, the novelty of this approach is that the associated proximal mapping is not nonexpansive any longer. Such situation is encountered, for instance, in numerical solution to phase retrieval problem in crystallography, astronomy and inverse scattering Luke (SIAM Rev 44:169–224, 2002) and is therefore of great practical interest.

Keywords Proximal split feasibility problems · Moreau–Yosida approximate · Prox-regularity

Mathematics Subject Classification (2010) 49J53 · 65K10 · 49M37 · 90C25

A. Moudafi (✉)
Ceregmia-Département scientifique, Université des Antilles et de la Guyane,
Schœlcher 97275, Martinique
e-mail: Abdellatif.Moudafi@martinique.univ-ag.fr

B. S. Thakur
School of Studies in Mathematics, Pt. Ravishankar Shukla University,
Raipur 492010, India
e-mail: balwantst@gmail.com

1 Introduction and preliminaries

The split feasibility problem has received much attention due to its applications in signal processing and image reconstruction [9], with particular progress in intensity-modulated therapy [6]. For a complete and exhaustive study on algorithms for solving convex feasibility problem, including comments about their applications and an excellent bibliography see, for example [1]. Our purpose here is to study the more general case of proximal split minimization problems and to investigate the convergence properties of the associated numerical solutions. To begin with, we are interested in finding a solution $x^* \in H_1$ of the following problem

$$\min_{x \in H_1} \{f(x) + g_\lambda(Ax)\}, \tag{1.1}$$

where H_1, H_2 are two real Hilbert spaces, $f : H_1 \rightarrow \mathbb{R} \cup \{+\infty\}$, $g : H_2 \rightarrow \mathbb{R} \cup \{+\infty\}$ two proper convex lower semicontinuous functions and $A : H_1 \rightarrow H_2$ a bounded linear operator, $g_\lambda(y) = \min_{u \in H_2} \{g(u) + \frac{1}{2\lambda} \|u - y\|^2\}$ stands for the Moreau–Yosida approximate of the function g of parameter λ .

Note that the differentiability of the Yosida-approximate g_λ , see for instance [19], secures the additivity of the subdifferentials and thus we can write

$$\partial(f(x) + g_\lambda(Ax)) = \partial f(x) + A^* \nabla g_\lambda(Ax) = \partial f(x) + A^* \left(\frac{I - \text{prox}_{\lambda g}}{\lambda} \right) (Ax).$$

The optimality condition of (1.1) can be then written as

$$0 \in \lambda \partial f(x^*) + A^*(I - \text{prox}_{\lambda g})(Ax^*), \tag{1.2}$$

where $\text{prox}_{\lambda g}(x) = \text{argmin}_{u \in H_2} \{g(u) + \frac{1}{2\lambda} \|u - x\|^2\}$ stands for the proximal mapping of g and the subdifferential of f at x is the set

$$\partial f(x) := \{u \in H_1 : f(y) \geq f(x) + \langle u, y - x \rangle \text{ for all } y \in H_1\}.$$

Inclusion (1.2) in turn yields to the following equivalent fixed point formulation

$$x^* = \text{prox}_{\mu \lambda f}(x^* - \mu A^*(I - \text{prox}_{\lambda g}))(Ax^*). \tag{1.3}$$

To solve (1.1), relation (1.3) suggests to consider the following split proximal algorithm

$$x_{k+1} = \text{prox}_{\mu_k \lambda f}(x_k - \mu_k A^*(I - \text{prox}_{\lambda g}))(Ax_k); \tag{1.4}$$

Observe that by taking $f = \delta_C$ [defined as $\delta_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise], $g = \delta_Q$ the indicator functions of two nonempty closed convex sets C, Q of H_1 and H_2 respectively, Problem (1.1) reduces to

$$\min_{x \in H_1} \{\delta_C(x) + (\delta_Q)_\lambda(Ax)\} \Leftrightarrow \min_{x \in C} \left\{ \frac{1}{2\lambda} \|(I - P_Q)(Ax)\|^2 \right\}, \tag{1.5}$$

which, when $C \cap A^{-1}(Q) \neq \emptyset$, is equivalent to the so called *Split Feasibility Problem*, namely

$$x \in C \quad \text{such that } Ax \in Q. \tag{1.6}$$

This problem was used for solving an inverse problem in radiation therapy treatment planning [6] and has been well studied both theoretically and practically, see for example [1,5] and the references therein. In this context, (1.4) reduces to the so-called *CQ-algorithm* introduced by Byrne [4]

$$x_{k+1} = P_C(x_k - \mu_k A^*(I - P_Q)(Ax_k)); \tag{1.7}$$

where the step-size μ_k is chosen in $(0, 2/\|A\|^2)$ and P_C, P_Q stand for the orthogonal projections on the closed convex sets C and Q , respectively.

The determination of the step-size in (1.7) (idem for its Krasnoselskii–Mann version, see for instance [21–23], and also for (1.4), see for example [9] and references therein) depends on the operator norm which computation (or at least estimate) is not an easy task. To overcome this difficulty, Lopez et al. [10] introduce a new choice of the step-size sequence (μ_k) and propose the following algorithm:

Algorithm 3.1 Given an initial arbitrarily point $x_0 \in H_1$. Assume that $x_k \in C$ has been constructed and $\nabla \tilde{h}(x_k) \neq 0$; then compute x_{k+1} via the rule

$$x_{k+1} = P_C(x_k - \mu_k A^*(I - P_Q))(Ax_k), \tag{1.8}$$

where $\mu_k := \rho_k \frac{\tilde{h}(x_k)}{\|\nabla \tilde{h}(x_k)\|^2}$ with $0 < \rho_k < 4$ and $\tilde{h}(x) = \frac{1}{2}\|(I - P_Q)Ax\|^2$. If $\nabla \tilde{h}(x_k) = 0$, then $x_{k+1} = x_k$ is a solution of (2.2) and the iterative process stops; otherwise, we set $k := k + 1$ and go to (1.8).

They proved the weak convergence of the sequence generated by (1.8) if (2.2) is consistent and $\inf_k \rho_k(4 - \rho_k) > 0$.

At this stage we would like to emphasize that our interest, in the first part of the present paper, is in solving (1.1) in the case $\operatorname{argmin} f \cap A^{-1}(\operatorname{argmin} g) \neq \emptyset$, or in other words: in finding a minimizer x^* of f such that Ax^* minimizes g , namely

$$x^* \in \operatorname{argmin} f \quad \text{such that } Ax^* \in \operatorname{argmin} g, \tag{1.9}$$

f, g being two proper, lower semicontinuous convex functions, $\operatorname{argmin} f := \{\bar{x} \in H_1 : f(\bar{x}) \leq f(x) \forall x \in H_1\}$ and $\operatorname{argmin} g := \{\bar{y} \in H_2 : g(\bar{y}) \leq g(y) \forall y \in H_2\}$. Γ will denote the solution set.

This problem was considered, for instance in [5] and [14], however, the iterative methods proposed to solve it need to know a priori the norm (or at least an estimate of the norm) of the bounded linear operator A . To avoid this difficulty, inspired by the idea introduced in [10], we develop in Sect. 2 an algorithm which is designed to address a way of selecting the step-sizes such that its implementation does not need any prior information about the operator norm and prove its related convergence result. This result is an extension of Theorem 3.5-[10].

In the second part of this paper, we will still assume f to be convex, but we allow the function g to be nonconvex. In the case of indicator functions of subsets with $A = I$, such situation is encountered in numerical solution to phase retrieval problem in inverse scattering [11] and is therefore of great practical interest. Here, we consider the more general problem of finding a minimizer \bar{x} of f such that $A\bar{x}$ is a critical point of g , namely

$$0 \in \partial f(\bar{x}) \quad \text{such that } 0 \in \partial_P g(A\bar{x}), \tag{1.10}$$

where ∂_P stands for the Proximal subdifferential of g which will be define in Sect. 3.

At this time the nonconvex theory is much less developed than the convex one. A notable exception, in the fixed-point context, is the work by Luke [12], who studies the convergence of a projection/reflection algorithm in a prox-regular setting. Nevertheless, the *fixed point operator is assumed to be locally firmly nonexpansive*. In [15] a proximal approach was also developed for finding critical points of uniformly prox-regular functions. Here, in the case of variable regularization parameters (λ_k) , we will prove in Sect. 3 the convergence of our Split Proximal Algorithm if the bounded linear operator is surjective. The latter assumption is always satisfied in inverse problems in which a priori information is available about the representation of the target solution in a frame, see for instance [7] and the references therein.

2 Convex minimization feasibility problem

Now, we are in a position to introduce a new way of selecting the step-sizes. To that end, we set $\theta(x_k) := \sqrt{\|\nabla h(x)\|^2 + \|\nabla l(x)\|^2}$ with $h(x) = \frac{1}{2}\|(I - \text{prox}_{\lambda g})Ax\|^2$, $l(x) = \frac{1}{2}\|(I - \text{prox}_{\mu_k \lambda f})x\|^2$ and introduce the following split proximal algorithm:

Split proximal algorithm Given an initial point $x_0 \in H_1$. Assume that x_k has been constructed and $\theta(x_k) \neq 0$; then compute x_{k+1} via the rule

$$x_{k+1} = \text{prox}_{\lambda \mu_k f}(x_k - \mu_k A^*(I - \text{prox}_{\lambda g})(Ax_k)), \tag{2.1}$$

where the stepsize $\mu_k := \rho_k \frac{h(x_k)+l(x_k)}{\theta^2(x_k)}$ with $0 < \rho_k < 4$. If $\theta(x_k) = 0$, then $x_{k+1} = x_k$ is a solution of (1.9) and the iterative process stops; otherwise, we set $k := k + 1$ and go to (2.1).

Observe that by taking $f = \delta_C$, $g = \delta_Q$ the indicator functions of two nonempty closed convex sets C, Q of H_1 and H_2 respectively, we recover [10, Algorithm 3.1]. Indeed, since $\text{prox}_{\lambda \mu_k f} = P_C$, the iterates x_k belong to C and thus $\nabla l(x_k) = 0$. So $\theta(x_k)$ reduces to $\nabla h(x_k)$.

Recall that the proximal mapping of g is firmly nonexpansive, namely

$$\langle \text{prox}_{\mu g}y - \text{prox}_{\mu g}y', y - y' \rangle \geq \|\text{prox}_{\mu g}y - \text{prox}_{\mu g}y'\|^2 \quad \forall y, y' \in H_2,$$

and it is also the case for complement $I - \text{prox}_{\mu g}$, see for example [7].

In order to prove the main result of this section, the idea is to apply the following well-known result on weak convergence in Hilbert spaces (the Opial’s lemma) with $S := \Gamma$. The original proof of this Lemma in [16] requires S to be closed and convex. An argument given in [20] shows that we do not need convexity.

Lemma 2.1 see [16,20].

Let H be a Hilbert space and (x_k) a sequence in H such that there exists a nonempty closed set $S \subset H$ satisfying:

- (i) For every $z \in S$, $\lim_k \|x_k - z\|$ exists.
- (ii) Any weak-cluster point of the sequence (x_k) belongs in S .

Then, there exists $\bar{x} \in S$ such that (x_k) weakly converges to \bar{x} .

The following Theorem contains the main convergence result of this section.

Theorem 2.2 Assume that f and g are two proper convex lower-semicontinuous functions and that (1.9) is consistent (i.e., $\Gamma \neq \emptyset$). If the parameters satisfy the following conditions $\varepsilon \leq \rho_k \leq \frac{4h(x_k)}{h(x_k)+l(x_k)} - \varepsilon$ (for some $\varepsilon > 0$ small enough), then the sequence (x_k) generated by (2.1) weakly converges to a solution of (1.9).

Proof Let $z \in \Gamma$ and note that $\nabla h(x) = A^*(I - \text{prox}_{\mu_k g})Ax$, $\nabla l(x) = (I - \text{prox}_{\mu_k \lambda f})x$. Using the fact that $\text{prox}_{\lambda \mu_k f}$ is nonexpansive, z verifies (1.9) (since minimizers of any function are exactly fixed-points of its proximal mapping) and having in hand

$$\langle \nabla h(x_k), x_k - z \rangle = \langle (I - \text{prox}_{\mu_k g})Ax_k, Ax_k - Az \rangle \geq \|(I - \text{prox}_{\lambda g})Ax_k\|^2 = 2h(x_k),$$

thank to the fact that $I - \text{prox}_{\mu_k g}$ is firmly nonexpansive, we can write

$$\begin{aligned} \|x_{k+1} - z\|^2 &\leq \|x_k - z\|^2 + \mu_k^2 \|\nabla h(x_k)\|^2 - 2\mu_k \langle \nabla h(x_k), x_k - z \rangle \\ &\leq \|x_k - z\|^2 + \mu_k^2 \|\nabla h(x_k)\|^2 - 4\mu_k h(x_k) \\ &= \|x_k - z\|^2 + \rho_k^2 \frac{(h(x_k) + l(x_k))^2}{(\theta^2(x_k))^2} \|\nabla h(x_k)\|^2 - 4\rho_k \frac{h(x_k) + l(x_k)}{\theta^2(x_k)} h(x_k) \\ &\leq \|x_k - z\|^2 + \rho_k^2 \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} - 4\rho_k \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} \frac{h(x_k)}{h(x_k) + l(x_k)} \\ &= \|x_k - z\|^2 - \rho_k \left(\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k \right) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)}. \end{aligned}$$

Thus

$$\|x_{k+1} - z\|^2 \leq \|x_k - z\|^2 - \rho_k \left(\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k \right) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)}. \tag{2.2}$$

The sequence (x_k) is thus Fejer monotone with respect to Γ which assures the existence of the limit

$$l(z) := \lim_{k \rightarrow +\infty} \|x_k - z\|^2 < +\infty, \tag{2.3}$$

and hence the first condition of Lemma 2.1 is satisfied. The latter in turn implies that (x_k) is bounded.

Now, Summing up inequality (2.2) with respect to $k = 0, 1, \dots$, we obtain

$$\sum_{k=0}^{\infty} \rho_k \left(\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k \right) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} \leq \|x_0 - z\|^2 < +\infty. \tag{2.4}$$

On the other hand, the assumption $\varepsilon \leq \rho_k \leq \frac{4h(x_k)}{h(x_k) + l(x_k)} - \varepsilon$ together with (2.4) ensure that

$$\sum_{k=0}^{\infty} \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} < +\infty. \tag{2.5}$$

Consequently,

$$\lim_{k \rightarrow +\infty} (h(x_k) + l(x_k)) = 0 \Leftrightarrow \lim_{k \rightarrow +\infty} h(x_k) = 0 \quad \text{and} \quad \lim_{k \rightarrow +\infty} l(x_k) = 0, \tag{2.6}$$

because $\theta^2(x_k) := \|\nabla h(x_k)\|^2 + \|\nabla l(x_k)\|^2$ is bounded. This follows from the fact that ∇h is Lipschitz continuous with constant $\|A\|^2$, ∇l is nonexpansive and (x_k) is bounded. More precisely, for any z which solves (1.9), we have

$$\begin{aligned} \|\nabla h(x_k)\| &= \|\nabla h(x_k) - \nabla h(z)\| \leq \|A\|^2 \|x_k - z\| \quad \text{and} \\ \|\nabla l(x_k)\| &= \|\nabla l(x_k) - \nabla l(z)\| \leq \|x_k - z\|. \end{aligned}$$

Now, let \tilde{x} be a weak cluster point of (x_k) , there exists a subsequence (x_{k_v}) which weakly converges to \tilde{x} . The lower-semicontinuity of h then implies that

$$0 \leq h(\tilde{x}) \leq \liminf_{v \rightarrow +\infty} h(x_{k_v}) = \lim_{k \rightarrow +\infty} h(x_k) = 0.$$

That is $h(\tilde{x}) = \frac{1}{2} \|(I - \text{prox}_{\lambda g})A\tilde{x}\|^2 = 0$, i.e. $A\tilde{x}$ is a fixed point of the proximal mapping of g or equivalently $0 \in \partial g(A\tilde{x})$. In other words $A\tilde{x}$ is a minimizer of g .

Likewise, the lower-semicontinuity of l implies that

$$0 \leq l(\tilde{x}) \leq \liminf_{v \rightarrow +\infty} l(x_{k_v}) = \lim_{k \rightarrow +\infty} l(x_k) = 0.$$

That is $l(\tilde{x}) = \frac{1}{2} \|(I - \text{prox}_{\mu_k \lambda f})\tilde{x}\|^2 = 0$, i.e. \tilde{x} is a fixed point of the proximal mapping of f or equivalently $0 \in \partial f(\tilde{x})$. In other words \tilde{x} is also a minimizer of f and thus a solution of problem (1.9). The second condition of Lemma 2.1 is also verified and consequently the whole sequence (x_k) converges weakly to a solution of problem (1.9). This completes the proof. □

Remark 2.3 i) Where the bounded linear operator A is the identity operator, (1.9) is nothing else than the problem of finding a common minimizer of f and g and (2.1) reduces to the following relaxed split proximal algorithm

$$x_{k+1} = \text{prox}_{\lambda\mu_k f}((1 - \mu_k)x_k + \mu_k \text{prox}_{\lambda g}(x_k)). \quad (2.7)$$

- ii) We would like also to emphasize that by taking $f = \delta_C$, $g = \delta_Q$ the indicator functions of two nonempty closed convex sets C, Q of H_1 and H_2 respectively, (2.1) reduces to [10, Algorithm 3.1] and we recover the corresponding convergence result, namely [10, Theorem 3.5].
- iii) It is worth mentioning that our approach works for split equilibrium and split null point problems considered in [5] and [14], respectively. To that end, just replace the proximal mappings of the convex functions f and g by the resolvent operators associated to two monotone equilibrium bifunctions and two maximal monotone operators, respectively.

3 A step towards the nonconvex case

Throughout this section g is assumed to be prox-regular. The following problem

$$0 \in \partial f(\bar{x}) \quad \text{such that} \quad 0 \in \partial_P g(A\bar{x}), \quad (3.1)$$

is very general in the sense that it includes, as special cases, g convex and g lower- \mathcal{C}^2 function which is of great importance in optimization and can be locally expressed as a difference $g - \frac{r}{2} \|\cdot\|^2$, where g is a finite convex function, hence a large core of problems of interest in variational analysis and optimization. It should be noticed that examples abound of practitioners needing algorithms for solving nonconvex problems, for instance, in crystallography, astronomy and more recently in inverse scattering, see for example [12].

We start with some elementary facts of prox-regularity. We denote by $B(x, \varepsilon)$ (respectively $B[x, \varepsilon]$) the open (respectively closed) ball around x with radius ε .

Let $g : H_2 \rightarrow \mathbb{R} \cup \{+\infty\}$ be a function and let $\bar{x} \in \text{dom}g$, i.e., $g(\bar{x}) < +\infty$. Poliquin–Rockafellar [17] introduced the concept of a proximal subdifferential and then they investigated the limiting proximal subdifferential (see also Bernard and Thibault [2] for the Hilbert setting). More precisely, the proximal subdifferential $\partial_P g(\bar{x})$ is defined as follows

Definition 3.1 A vector v is in $\partial_P g(\bar{x})$ if there exist some $r > 0$ and $\varepsilon > 0$ such that for all $x \in B(\bar{x}, \varepsilon)$,

$$\langle v, x - \bar{x} \rangle \leq g(x) - g(\bar{x}) + \frac{r}{2} \|x - \bar{x}\|^2. \quad (3.2)$$

When $g(\bar{x}) = +\infty$, one puts $\partial_P g(\bar{x}) = \emptyset$.

Before stating the definition of prox-regularity of g and properties of its proximal mapping, we recall that g is locally l.s.c at \bar{x} if its epigraph is closed relative to a neighborhood of $(\bar{x}, g(\bar{x}))$, prox-bounded if g is minorized by a quadratic function, and recall that for $\varepsilon > 0$, the g -attentive ε -localisation of $\partial_P g$ around (\bar{x}, \bar{v}) , is the mapping $T_\varepsilon : H_2 \rightarrow 2^{H_2}$ defined by

$$T_\varepsilon(x) = \begin{cases} \{v \in \partial_P g(x), \|v - \bar{v}\| < \varepsilon\} & \text{if } \|x - \bar{x}\| < \varepsilon \text{ and } |g(x) - g(\bar{x})| < \varepsilon, \\ \emptyset & \text{otherwise.} \end{cases}$$

Definition 3.2 A function g is said to be prox-regular at \bar{x} for $\bar{v} \in \partial_P g(\bar{x})$ if there exist some $r > 0$ and $\varepsilon > 0$ such that for all $x, x' \in B(\bar{x}, \varepsilon)$ with $|g(x) - g(x')| < \varepsilon$ and all $v \in B(\bar{v}; \varepsilon)$ with $v \in \partial_P g(x)$ one has

$$g(x') \geq g(x) + \langle v, x' - x \rangle - \frac{r}{2} \|x' - x\|^2. \tag{3.3}$$

If the property holds for all vectors $\bar{v} \in \partial_P g(\bar{x})$, the function is said to be prox-regular at \bar{x} .

Fundamental insights into the properties of a function g come from the study of its Moreau–Yosida regularization g_λ and the associated proximal mapping $\text{prox}_{\lambda g}$ defined for $\lambda > 0$ respectively by

$$g_\lambda(x) = \inf_{u \in H_2} \{g(u) + \frac{1}{2\lambda} \|u - x\|^2\} \quad \text{and} \quad \text{prox}_{\lambda g}(x) := \arg \min_{u \in H_2} \{g(u) + \frac{1}{2\lambda} \|u - x\|^2\}.$$

The latter is a fundamental tool in optimization and it was shown that a fixed point iteration on the proximal mapping could be used to develop a simple optimization algorithm, namely, the proximal point algorithm.

Note also, see for example Sect. 1 in [8], that *local minima are zeroes of the Proximal subdifferential and that the Proximal subdifferential and the convex one coincide in the convex case.*

Now, let us state the following proposition which summarizes some important consequences of the prox-regularity, see [17, Theorem 4.4] and [2, Lemma 3.1], Theorem 3.4.

Proposition 3.1 *Suppose that g is locally lower semicontinuous at \bar{x} and prox-regular at \bar{x} for $\bar{v} = 0$ with respect to r and ε . Let T_ε be the g -attentive ε -localisation of $\partial_P g$ around (\bar{x}, \bar{v}) . Then for each $\lambda \in]0, 1/r[$ there is a neighborhood U_λ of \bar{x} such that, on U_λ*

- i) the mapping $\text{prox}_{\lambda g}$ is single-valued and Lipschitz continuous with constant $\frac{1}{1-\lambda r}$ and $\text{prox}_{\lambda g}(x) = (I + \lambda T_\varepsilon)^{-1}(x) = [\text{singleton}]$.*
- ii) g_λ is differentiable (more precisely g is of class \mathcal{C}^{1+}) with*

$$\nabla g_\lambda(x) = \frac{x - \text{prox}_{\lambda g}(x)}{\lambda} = (\lambda I + T_\varepsilon^{-1})^{-1}(x).$$

Now, let us prove the following key property of the proximal mapping complement.

Remark 3.2 If the assumptions of Proposition (3.1) hold true, then $\forall \lambda \in (0, \frac{1}{r})$ and $\forall x_1, x_2 \in U_\lambda$, one has

$$\begin{aligned} & \langle (I - \text{prox}_{\lambda g})(x_1) - (I - \text{prox}_{\lambda g})(x_2), x_1 - x_2 \rangle \\ & \geq \|(I - \text{prox}_{\lambda g})(x_1) - (I - \text{prox}_{\lambda g})(x_2)\|^2 - \frac{\lambda r}{(1 - \lambda r)^2} \|x_1 - x_2\|^2. \end{aligned}$$

Observe that when $r = 0$ which amount to saying that g is convex, we recover the fact that the mapping $I - \text{prox}_{\lambda g}$ is firmly nonexpansive.

Indeed, let $x_i \in U_\lambda$, $i = 1, 2$. Then $v_i = \frac{x_i - \text{prox}_{\lambda g}(x_i)}{\lambda} \in T_\varepsilon(\text{prox}_{\lambda g}(x_i))$. Invoking the prox-regularity of g , we have the monotonicity of $T_\varepsilon + rI$, see for instance [2, Theorem 3.4]. This implies, for the pairs (x_1, v_1) and (x_2, v_2) , that

$$\langle v_1 - v_2, \text{prox}_{\lambda g}(x_1) - \text{prox}_{\lambda g}(x_2) \rangle \geq -r \|\text{prox}_{\lambda g}(x_1) - \text{prox}_{\lambda g}(x_2)\|^2$$

We also have

$$\begin{aligned} \langle v_1 - v_2, x_1 - x_2 \rangle &= \langle v_1 - v_2, x_1 - \text{prox}_{\lambda g}(x_1) - (x_2 - \text{prox}_{\lambda g}(x_2)) \rangle \\ &\quad + \langle v_1 - v_2, \text{prox}_{\lambda g}(x_1) - \text{prox}_{\lambda g}(x_2) \rangle \\ &\geq \lambda^{-1} \|x_1 - \text{prox}_{\lambda g}(x_1) - (x_2 - \text{prox}_{\lambda g}(x_2))\|^2 \\ &\quad - r \|\text{prox}_{\lambda g}(x_1) - \text{prox}_{\lambda g}(x_2)\|^2, \end{aligned}$$

hence

$$\begin{aligned} & \langle (I - \text{prox}_{\lambda g})(x_1) - (I - \text{prox}_{\lambda g})(x_2), x_1 - x_2 \rangle \\ & \geq \|(I - \text{prox}_{\lambda g})(x_1) - (I - \text{prox}_{\lambda g})(x_2)\|^2 - \lambda r \|\text{prox}_{\lambda g}(x_1) - \text{prox}_{\lambda g}(x_2)\|^2. \end{aligned}$$

Combining the last inequality with Lipschitz continuity of the proximal mapping, i.e.,

$$\|\text{prox}_{\lambda g}(x_1) - \text{prox}_{\lambda g}(x_2)\| \leq \frac{1}{1 - \lambda r} \|x_1 - x_2\|,$$

we obtain the desired result.

We state also a lemma which will be needed in the sequel.

Lemma 3.3 (See Polyak [18, Lemma 2.2.2]).

Let (a_k) , (β_k) and (γ_k) , $k \in \mathbb{N}$ be three sequences of nonnegative numbers satisfying $a_{k+1} \leq (1 + \beta_k)a_k + \gamma_k$. If $\sum_{k=0}^\infty \beta_k < +\infty$ and $\sum_{k=0}^\infty \gamma_k < +\infty$, then (a_k) is convergent.

Now, the regularization parameters λ are allowed to vary in the algorithm (2.1), namely considering possibly variable parameters $\lambda_k \in (0, \frac{1}{r} - \varepsilon)$ (for some $\varepsilon > 0$ small enough) and $\mu_k > 0$, our interest is in studying the convergence properties of the following algorithm:

Split proximal algorithm Given an initial point $x_0 \in H_1$. Assume that x_k has been constructed and $\theta(x_k) \neq 0$; then compute x_{k+1} via the rule

$$x_{k+1} = \text{prox}_{\lambda_k \mu_k f}(x_k - \mu_k A^*(I - \text{prox}_{\lambda_k g})(Ax_k)), \tag{3.4}$$

where the stepsize $\mu_k := \rho_k \frac{h(x_k) + l(x_k)}{\theta^2(x_k)}$ with $0 < \rho_k < 4$. If $\theta(x_k) = 0$, then $x_{k+1} = x_k$ is a solution of (3.1) and the iterative process stops; otherwise, we set $k := k + 1$ and go to (3.4).

Theorem 3.4 *Assume that f is a proper convex lower-semicontinuous function, g is locally lower semicontinuous at $A\bar{x}$, prox-bounded and prox-regular at $A\bar{x}$ for $\bar{v} = 0$ with \bar{x} a point which solves (3.1) and A a bounded linear operator which is surjective with a dense domain. If the parameters satisfy the following conditions $\sum_{k=0}^\infty \lambda_k < +\infty$ and $\inf_k \rho_k (\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k) > 0$, and if $\|x_0 - \bar{x}\|$ is small enough, then the sequence (x_k) generated by (3.4) weakly converges to a solution of (3.1).*

Proof Using the fact that $\text{prox}_{\lambda_k \mu_k f}$ is nonexpansive, \bar{x} verifies (3.1) (critical points of any function are exactly fixed-points of its proximal mapping) and having in mind Remark 3.2, we can write

$$\begin{aligned} \|x_{k+1} - \bar{x}\|^2 &\leq \|x_k - \bar{x}\|^2 + \mu_k^2 \|\nabla h(x_k)\|^2 - 2\mu_k \langle \nabla h(x_k), x_k - \bar{x} \rangle \\ &\leq \|x_k - \bar{x}\|^2 + \mu_k^2 \|\nabla h(x_k)\|^2 - 2\mu_k (2h(x_k) - \frac{\lambda_k r \|A\|^2}{(1 - \lambda_k r)^2} \|x_k - \bar{x}\|^2) \\ &= \|x_k - \bar{x}\|^2 + 2\mu_k \frac{\lambda_k r \|A\|^2}{(1 - \lambda_k r)^2} \|x_k - \bar{x}\|^2 - 4\mu_k h(x_k) + \mu_k^2 \|\nabla h(x_k)\|^2 \\ &\leq \|x_k - \bar{x}\|^2 + 2\rho_k \frac{(h(x_k) + l(x_k))}{\|\nabla h(x_k)\|^2 + \|\nabla l(x_k)\|^2} \frac{\lambda_k r \|A\|^2}{(1 - \lambda_k r)^2} \|x_k - \bar{x}\|^2 \\ &\quad - \rho_k (\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} \\ &\leq \left(1 + \lambda_k \rho_k \left(\frac{2h(x_k)}{\|\nabla h(x_k)\|^2} + \frac{2l(x_k)}{\|\nabla l(x_k)\|^2} \right) \frac{r \|A\|^2}{(1 - \lambda_k r)^2} \right) \|x_k - \bar{x}\|^2 \\ &\quad - \rho_k \left(\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k \right) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} \\ &= \left(1 + \lambda_k \rho_k \left(1 + \frac{2h(x_k)}{\|\nabla h(x_k)\|^2} \right) \frac{r \|A\|^2}{(1 - \lambda_k r)^2} \right) \|x_k - \bar{x}\|^2 \\ &\quad - \rho_k \left(\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k \right) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)}. \end{aligned}$$

Recall that

$$(\star) \quad A \text{ is surjective with a dense domain} \Leftrightarrow \exists \gamma > 0 \text{ such that } \|A^*x\| \geq \gamma \|x\|,$$

see for example Brezis [3, Theorem II.19]. This ensures that

$$\frac{2h(x_k)}{\|\nabla h(x_k)\|^2} = \frac{\|(I - \text{prox}_{\lambda_k g})(Ax_k)\|^2}{\|A^*(I - \text{prox}_{\lambda_k g})(Ax_k)\|^2} \leq \frac{1}{\gamma^2}.$$

Conditions on the parameters λ_k and ρ_k assure the existence of a positive constant M such that

$$\|x_{k+1} - \bar{x}\|^2 \leq (1 + M\lambda_k)\|x_k - \bar{x}\|^2 - \rho_k \left(\frac{4h(x_k)}{h(x_k) + l(x_k)} - \rho_k \right) \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)}. \tag{3.5}$$

In view of the inequality (3.5), Lemma 3.3 assures the existence of $\lim_{k \rightarrow +\infty} \|x_k - \bar{x}\|^2$, since $\sum_{k=0}^{\infty} \lambda_k < +\infty$.

Now, summing up inequality (3.5) with respect to $k = 0, 1, \dots$ and taking into account the assumption on ρ_k , we obtain

$$\varepsilon^2 \sum_{k=0}^{\infty} \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} \leq \|x_0 - z\|^2 + M \sum_{k=0}^{\infty} \lambda_k \|x_k - z\|^2, \tag{3.6}$$

and by invoking the facts that $(\|x_k - z\|^2)$ is bounded and $\sum_{k=0}^{\infty} \lambda_k < +\infty$, we infer

$$\sum_{k=0}^{\infty} \frac{(h(x_k) + l(x_k))^2}{\theta^2(x_k)} < +\infty. \tag{3.7}$$

Following the proof of Theorem 2.1, we conclude that the sequence (x_k) weakly converges to a solution of (3.1). □

Remark 3.5 In inverse problems, certain physical properties of the target solution are most suitably expressed in terms of the coefficients of its representation with respect to a family of **finite or infinite** vectors $(e_k)_k$ of a Hilbert space which constitutes a frame (see for instance [7] and the references therein), namely there exist two constants such that

$$\forall x \in H \quad \underline{\beta} \|x\|^2 \leq \sum_k |\langle x, e_k \rangle|^2 \leq \overline{\beta} \|x\|^2.$$

The associated frame operator is the injective bounded operator F define on H by $F(x) := (\langle x, e_k \rangle)_k$ and the adjoint of which is the surjective operator define as $F^*((\xi)_k) := \sum_k \xi_k e_k$.

Let $\bar{x} \in H$ be the target solution of the underlying inverse problem in which a priori information (e.g., sparsity, distribution, statistical properties) is available about the coefficients (ξ_k) of the decomposition of \bar{x} in $(e_k)_k$. To recover \bar{x} a natural formulation is a variational problem in the space l^2 of frame coefficients (where a priori information on (ξ_k) can be easily incorporated) that involves two functions f and $g \circ F^*$ with f, g two proper lower semicontinuous convex functions (see for example [7] and the references therein). In this context the assumption on the linear operator $A := F^*$ is satisfied. Furthermore, in the particular case of a Riesz basis, relation (\star) is satisfied with $\gamma = \sqrt{\underline{\beta}}$, see [13].

Acknowledgments The authors would like to thank the anonymous referees for their careful reading, suggestions and pertinent questions which permitted us to improve the first version of this paper.

References

1. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **38**(3), 367–426 (1996)
2. Bernard, F., Thibault, L.: Prox-regular functions in Hilbert spaces. *J. Math. Anal. Appl.* **303**, 1–14 (2005)
3. Brézis, H.: *Analyse fonctionnelle, théorie et applications*. Masson, Paris (1983)
4. Byrne, Ch.: Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Probl.* **18**, 441–453 (2002)
5. Byrne, C.H., Censor, Y., Gibali, A., Reich, S.: Weak and strong convergence of algorithms for the split common null point problem. *J. Nonlinear Convex Anal.* **13**, 759–775 (2012)
6. Censor, Y., Bortfeld, T., Martin, B., Trofimov, A.: A unified approach for inversion problems in intensity-modulated radiation therapy. *Phys. Med. Biol.* **51**, 2353–2365 (2006)
7. Chau, C., Combettes, P.L., Pesquet, J.-C., Wajs, V.R.: A variational formulation for frame-based inverse problems. *Inverse Probl.* **23**, 1495–1518 (2007)
8. Clarke, Z.F.H., Ledyae, Y.S., Stern, R.J., Wolenski, P.R.: *Nonsmooth Analysis and Control Theory*. Springer, New York (1998)
9. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* **4**, 1168–1200 (2005)
10. Lopez, G., Martin-Marquez, V., Wang, F., Xu, H.K.: Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Probl.* **28**, 085004 (2012)
11. Luke, D.R., Burke, J.V., Lyon, R.G.: Optical wavefront reconstruction: theory and numerical methods. *SIAM Rev.* **44**, 169–224 (2002)
12. Luke, D.R.: Finding best approximation pairs relative to a convex and prox-regular set in Hilbert space. *SIAM J. opt.* **19**(2), 714–729 (2008)
13. Mallat, S.G.: *A Wavelet Tour of Signal Processing*, 2nd edn. Academic Press, New York (1999)
14. Moudafi, A.: Split Monotone Variational Inclusions. *J. Optim. Theory Appl.* **150**(2), 275–283 (2011)
15. Moudafi, A.: A proximal iterative approach to a nonconvex optimization problem. *Nonlinear Anal. Theory Methods Appl.* **72**, 704–709 (2010)
16. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **73**, 591–597 (1967)
17. Poliquin, R.A., Rockafellar, R.T.: Prox-regular functions in variational analysis. *Trans. Am. Math. Soc.* **348**, 1805–1838 (1996)
18. Polyak, B.T.: *Introduction to Optimization*. Optimization Software Inc., New York (1987)
19. Rockafellar, R.T., Wets, R.: *Variational Analysis*. Springer, Berlin (1988)
20. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **14**, 877898 (1976)
21. Xu, H.K.: A variable Krasnosel'skii-Mann algorithm and the multiple-set split feasibility problem. *Inverse Probl.* **22**, 2021–2034 (2006)
22. Xu, H.-K.: Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces, *Inverse Probl.* **26**(10), article 105018 (2010)
23. Yao, Y., Chen, R., Marino, G., Liou, Y.C.: Applications of fixed point and optimization methods to the multiple-sets split feasibility problem. *J. Appl. Math.* **2012** Article ID 927530 (2012)