Shehu et al. *Fixed Point Theory and Applications* (2015) 2015:123
DOI 10.1186/s13663-015-0375-5

Fixed Point Theory and Applications
a SpringerOpen Journal

RESEARCH

Open Access

Iterative approximation of solutions for proximal split feasibility problems

Yekini Shehu[1], Gang Cai[2*] and Olaniyi S Iyiola[3]

*Correspondence:
caigang-aaaa@163.com
[2]School of Mathematics Science, Chongqing Normal University, Chongqing, 400047, China
Full list of author information is available at the end of the article

Abstract

In this paper, our aim is to introduce a viscosity type algorithm for solving proximal split feasibility problems and prove the strong convergence of the sequences generated by our iterative schemes in Hilbert spaces. First, we prove strong convergence result for a 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. Secondly, we prove another strong convergence result for the case where one of the two involved functions is prox-regular. In all our results in this work, our iterative schemes are proposed by way of selecting the step sizes such that their 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. Finally, we give a numerical example to study the efficiency and implementation of our iterative schemes. Our results complement the recent results of Moudafi and Thakur (Optim. Lett. 8:2099-2110, 2014, doi:10.1007/s11590-013-0708-4) and other recent important results in this direction.

MSC: 49J53; 65K10; 49M37; 90C25

Keywords: proximal split feasibility problems; Moreau-Yosida approximate; prox-regularity; strong convergence; Hilbert spaces

1 Introduction

In this paper, we shall assume that H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. Let I denote the identity operator on H. Let C and Q be nonempty, closed, and convex subsets of real Hilbert spaces H_1 and H_2, respectively. The *split feasibility problem* (SFP) is to find a point

$$x \in C \text{ such that } Ax \in Q, \tag{1.1}$$

where $A : H_1 \to H_2$ is a bounded linear operator. The SFP in finite-dimensional Hilbert spaces was first introduced by Censor and Elfving [1] for modeling inverse problems which arise from phase retrievals and in medical image reconstruction [2]. The SFP attracts the attention of many authors due to its application in signal processing. Various algorithms have been invented to solve it (see, for example, [3–9] and references therein). For a more current and up-to-date survey on split feasibility problems, please see [10].

Note that the split feasibility problem (1.1) can be formulated as a fixed point equation by using the fact

$$P_C\bigl(I-\gamma A^*(I-P_Q)A\bigr)x^* = x^*; \tag{1.2}$$

that is, x^* solves the SFP (1.1) if and only if x^* solves the fixed point equation (1.2) (see [11] for the details). This implies that we can use fixed point algorithms (see [12–14]) to solve SFP. A popular algorithm that solves the SFP (1.1) is due to Byrne's CQ algorithm [2] which is found to be a gradient-projection method (GPM) in convex minimization. Subsequently, Byrne [3] applied Krasnoselskii-Mann iteration to the CQ algorithm, and Zhao and Yang [15] applied Krasnoselskii-Mann iteration to the perturbed CQ algorithm to solve the SFP. It is well known that the CQ algorithm and the Krasnoselskii-Mann algorithm for a split feasibility problem do not necessarily converge strongly in the infinite-dimensional Hilbert spaces.

Our goal in this paper is to study the more general case of proximal split minimization problems and to investigate the strong convergence properties of the associated numerical solutions. To begin with, let us consider the following problem: Find a solution $x^* \in H_1$ such that

$$\min_{x\in H_1}\bigl\{f(x)+g_\lambda(Ax)\bigr\}, \tag{1.3}$$

where H_1, H_2 are two real Hilbert spaces, $f: H_1 \to \mathbb{R}\cup\{+\infty\}$, $g: H_2 \to \mathbb{R}\cup\{+\infty\}$ two proper, convex, lower-semicontinuous functions and $A: H_1 \to H_2$ 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.

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, and convex sets C, Q of H_1 and H_2, respectively, problem (1.3) reduces to

$$\min_{x\in H_1}\bigl\{\delta_C(x)+(\delta_Q)_\lambda(Ax)\bigr\} \quad\Leftrightarrow\quad \min_{x\in C}\left\{\frac{1}{2\lambda}\bigl\|(I-P_Q)(Ax)\bigr\|^2\right\} \tag{1.4}$$

which, when $C\cap A^{-1}(Q)\neq\emptyset$, is equivalent to (1.1).

By the differentiability of the Yosida approximate g_λ, see for instance [16], we have the additivity of the subdifferentials and thus we can write

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

This implies that the optimality condition of (1.3) can then be written as

$$0 \in \lambda\partial f(x) + A^*(I-\operatorname{prox}_{\lambda g})(Ax), \tag{1.5}$$

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

$$\partial f(x) := \bigl\{u\in H_1 : f(y) \geq f(x) + \langle u, y-x\rangle, \forall y\in H_1\bigr\}.$$

The inclusion (1.5) in turn yields the following equivalent fixed point formulation:

$$\operatorname{prox}_{\mu\lambda f}\bigl(x^* - \mu A^*(I - \operatorname{prox}_{\lambda g})\bigr)Ax^* = x^*. \tag{1.6}$$

To solve (1.3), (1.6) suggests us to consider the following split proximal algorithm:

$$x_{n+1} = \operatorname{prox}_{\mu_n\lambda f}\bigl(x_n - \mu_n A^*(I - \operatorname{prox}_{\lambda g})\bigr)Ax_n. \tag{1.7}$$

Based on an idea introduced in work of Lopez *et al.* [17], Moudafi and Thakur [18] recently proved weak convergence results for solving (1.3) 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

$$x^* \in \operatorname{argmin} f \text{ such that } Ax^* \in \operatorname{argmin} g, \tag{1.8}$$

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 $\operatorname{argmin} g := \{\bar{y} \in H_2 : g(\bar{y}) \leq g(y), \forall y \in H_2\}$. We will denote the solution set of (1.8) by Γ. Concerning problem (1.8), Moudafi and Thakur [18] introduced a new way of selecting the step sizes: Set $\theta(x_n) := \sqrt{\|\nabla h(x)\|^2 + \|\nabla l(x)\|^2}$ with $h(x) = \frac{1}{2}\|(I - \operatorname{prox}_{\lambda g})Ax\|^2$, $l(x) = \frac{1}{2}\|(I - \operatorname{prox}_{\lambda\mu_n f})x\|^2$ and introduced the following split proximal algorithm.

Split proximal algorithm 1 Given an initial point $x_1 \in H_1$. Assume that x_n has been constructed and $\theta(x_n) \neq 0$, then compute x_{n+1} via the rule

$$x_{n+1} = \operatorname{prox}_{\lambda\mu_n f}\bigl(x_n - \mu_n A^*(I - \operatorname{prox}_{\lambda g})Ax_n\bigr), \quad n \geq 1, \tag{1.9}$$

where the step size $\mu_n := \rho_n \frac{h(x_n)+l(x_n)}{\theta^2(x_n)}$ with $0 < \rho_n < 4$. If $\theta(x_n) = 0$, then $x_{n+1} = x_n$ is a solution of (1.3) and the iterative process stops, otherwise, we set $n := n + 1$ and go to (1.9).

Using the split proximal algorithm (1.9), Moudafi and Thakur [18] proved the following *weak convergence* theorem for approximating a solution of (1.8).

Theorem 1.1 *Assume that f and g are two proper convex lower-semicontinuous functions and that* (1.8) *is consistent* (*i.e.*, $\Gamma \neq \emptyset$). *If the parameters satisfy the conditions $\epsilon \leq \rho_n \leq \frac{4h(x_n)}{h(x_n)+l(x_n)} - \epsilon$* (*for some $\epsilon > 0$ small enough*), *then the sequence $\{x_n\}$ generated by* (1.9) *weakly converges to a solution of* (1.8).

Furthermore, Moudafi and Thakur [18] assumed f to be convex and allowed the function g to be nonconvex. In the case of indicator functions of subsets with $A = I$, such a situation is encountered in a numerical solution to phase retrieval problem in inverse scattering [19] and is therefore of great practical interest. They considered the more general problem of finding a minimizer $\bar{x}$ of f such that $A\bar{x}$ is a critical point of g, namely

$$0 \in \partial f(\bar{x}) \text{ such that } 0 \in \partial_{pg}(A\bar{x}), \tag{1.10}$$

where ∂_{pg} stands for the proximal subdifferential of g (see Definition 4.1 for definition of a proximal subdifferential). In particular, they studied the convergence properties of the following algorithm.

Split proximal algorithm 2 Given an initial point $x_1 \in H_1$. Assume that x_n has been constructed and $\theta(x_n) \neq 0$, then compute x_{n+1} via the rule

$$x_{n+1} = \operatorname{prox}_{\lambda_n \mu_n f}\bigl(x_n - \mu_n A^*(I - \operatorname{prox}_{\lambda_n g})Ax_n\bigr), \quad n \geq 1, \tag{1.11}$$

where the step size $\mu_n := \rho_n \frac{h(x_n)+l(x_n)}{\theta^2(x_n)}$ with $0 < \rho_n < 4$. If $\theta(x_n) = 0$, then $x_{n+1} = x_n$ is a solution of (1.10) and the iterative process stops, otherwise, we set $n := n + 1$ and go to (1.11).

Using (1.11), Moudafi and Thakur [18] proved the following *weak convergence* theorem for the approximation of the solution of (1.10).

Theorem 1.2 *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* (1.10) *and A a bounded linear operator which is surjective with a dense domain. If the parameters satisfy the following conditions: $\sum_{n=1}^{\infty} \lambda_n < \infty$ and $\inf_n \rho_n(\frac{4h(x_n)}{h(x_n)+l(x_n)} - \rho_n) > 0$ and if $\|x_1 - \bar{x}\|$ is small enough, then the sequence $\{x_n\}$ generated by* (1.11) *weakly converges to a solution of* (1.10).

Remark 1.3 We comment here that the split proximal algorithm (1.9) introduced Moudafi and Thakur [18] for approximating a solution of (1.8) has, in general, weak convergence only, unless the underlying Hilbert space is finite-dimensional. Indeed, based on the results of Hundal [20], we can construct a counterexample as follows.

Example 1.4 In the real Hilbert space $H = \ell_2$, Hundal [20] constructed two closed and convex subsets C and Q such that (see also [21–23])

(i) $C \cap Q \neq \emptyset$;

(ii) the sequence $\{x_n\}_{n=1}^{\infty}$ generated by alternating projections,

$$x_n = (P_C \circ P_Q)^n x_1, \quad n \geq 1 \tag{1.12}$$

with $x_1 \in C$, converges weakly, but not strongly.

(Hundal's counterexample settles in the negative the question whether alternating projections onto closed convex subsets of a Hilbert space can have strong convergence, which remained open for nearly 40 years.) Now in problem (1.8), let us take $f = \delta_C$ and $g = \delta_Q$ the indicator functions of two nonempty, closed, and convex sets C, Q of H_1 and H_2, respectively, where $H_1 = \ell_2 = H_2$. Then $\operatorname{prox}_{\lambda \mu_n f}(x) = P_C(x)$ and $\operatorname{prox}_{\lambda g}(x) = P_Q(x)$. Furthermore, take $A = I = A^*$, where I is an identity mapping on ℓ_2. Taking an initial guess $x_1 \in C$ and $\mu_n = 1$, $\forall n \geq 1$, we see that the split proximal algorithm (1.9) generates a sequence $\{x_n\}_{n=1}^{\infty}$ which coincides with the sequence $\{x_n\}_{n=1}^{\infty}$ given in (1.12). Therefore, the split proximal algorithm (1.9) generates *weakly* (*not strongly*) *convergent sequences* to a solution of problem (1.8), in general, in infinite-dimensional real Hilbert spaces.

Example 1.4 naturally gives rise to this question.

Question Can we appropriately modify the split proximal algorithm (1.9) so as to have strong convergence?

It is our aim in this paper to answer the above question in the affirmative. Thus, motivated by the results of Lopez *et al.* [17] and Moudafi and Thakur [18], our aim in this paper is to introduce new iterative schemes for solving problems (1.8) and (1.10) and *prove strong convergence* of the sequences generated by our schemes in real Hilbert spaces. Our results complement the results of Moudafi and Thakur [18] and Shehu [24–26].

2 Preliminaries

We state the following well-known lemmas which will be used in the sequel.

Lemma 2.1 *Let H be a real Hilbert space. Then we have the following well-known results*:

$$\text{(i)}\quad \|x+y\|^2 = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2, \quad \forall x,y \in H,$$
$$\text{(ii)}\quad \|x+y\|^2 \leq \|x\|^2 + 2\langle y, x+y\rangle, \quad \forall x,y \in H.$$

Lemma 2.2 (Xu [27]) *Let $\{a_n\}$ be a sequence of nonnegative real numbers satisfying the following relation*:

$$a_{n+1} \leq (1-\alpha_n)a_n + \alpha_n\sigma_n + \gamma_n, \quad n \geq 0,$$

where

(i) $\{a_n\} \subset [0,1]$, $\sum \alpha_n = \infty$;

(ii) $\limsup \sigma_n \leq 0$;

(iii) $\gamma_n \geq 0$ $(n \geq 1)$, $\sum \gamma_n < \infty$.

Then $a_n \to 0$ as $n \to \infty$.

3 Strong convergence for convex minimization feasibility problem

In this section, we modify algorithm (1.9) above so as to have strong convergence. Below we include such modification. Let $r : H_1 \to H_1$ be a contraction mapping with constant $\alpha \in (0,1)$. Set $\theta(x) := \sqrt{\|\nabla h(x)\|^2 + \|\nabla l(x)\|^2}$ with $h(x) = \frac{1}{2}\|(I - \operatorname{prox}_{\lambda g})Ax\|^2$, $l(x) = \frac{1}{2}\|(I - \operatorname{prox}_{\lambda\mu_n f})x\|^2$ and introduce the following modified split proximal algorithm.

Modified split proximal algorithm 1 Given an initial point $x_1 \in H_1$. Assume that x_n has been constructed and $\theta(x_n) \neq 0$, then compute x_{n+1} via the rule

$$\begin{cases} y_n = x_n - \mu_n A^*(I - \operatorname{prox}_{\lambda g})Ax_n, \\ x_{n+1} = \alpha_n r(x_n) + (1-\alpha_n)\operatorname{prox}_{\lambda\mu_n f} y_n, \quad n \geq 1, \end{cases} \tag{3.1}$$

where the step size $\mu_n := \rho_n \frac{h(x_n)+l(x_n)}{\theta^2(x_n)}$ with $0 < \rho_n < 4$. If $\theta(x_n) = 0$, then $x_{n+1} = x_n$ is a solution of (1.8) and the iterative process stops, otherwise, we set $n := n+1$ and go to (3.1).

Using (3.1), we prove the following strong convergence theorem for approximation of solutions of problem (1.8).

Theorem 3.1 *Assume that f and g are two proper convex lower-semicontinuous functions and that* (1.8) *is consistent* (*i.e.*, $\Gamma \neq \emptyset$). *If the parameters satisfy the following conditions*:

(a) $\lim_{n\to\infty} \alpha_n = 0$;

(b) $\sum_{n=1}^{\infty} \alpha_n = \infty$;
(c) $\epsilon \leq \rho_n \leq \frac{4h(x_n)}{h(x_n)+l(x_n)} - \epsilon$ *for some* $\epsilon > 0$,

the sequence $\{x_n\}$ *generated by* (3.1) *strongly converges to a solution of* (1.8) *which is also the unique solution of the variational inequality* (*VI*),

$$x^* \in \Gamma, \quad \langle (I-r)x^*, x - x^* \rangle \geq 0, \quad x \in \Gamma. \tag{3.2}$$

In other words, x^* *is the unique fixed point of the contraction* $\mathrm{Proj}_\Gamma r$, $x^* = (\mathrm{Proj}_\Gamma r)x^*$.

Proof Let $x^* \in \Gamma$. Observe that $\nabla h(x) = A^*(I - \mathrm{prox}_{\mu_n g})Ax$, $\nabla l(x) = (I - \mathrm{prox}_{\mu_n \lambda f})x$. Using the fact that $\mathrm{prox}_{\mu_n \lambda f}$ is nonexpansive, x^* verifies (1.8) (since minimizers of any function are exactly fixed points of its proximal mapping) and having in hand

$$\langle \nabla h(x_n), x_n - x^* \rangle = \langle (I - \mathrm{prox}_{\mu_n g})Ax_n, Ax_n - Ax^* \rangle \geq \| (I - \mathrm{prox}_{\mu_n g})Ax_n \|^2 = 2h(x_n),$$

thanks to the fact that $I - \mathrm{prox}_{\mu_n g}$ is firmly nonexpansive, we can write

$$\begin{aligned} \|y_n - x^*\|^2 &= \|x_n - x^*\|^2 + \mu_n^2 \|\nabla h(x_n)\|^2 - 2\mu_n \langle \nabla h(x_n), x_n - x^* \rangle \\ &\leq \|x_n - x^*\|^2 + \mu_n^2 \|\nabla h(x_n)\|^2 - 4\mu_n h(x_n) \\ &= \|x_n - x^*\|^2 + \rho_n^2 \frac{(h(x_n) + l(x_n))^2}{(\theta^2(x_n))^2} \|\nabla h(x_n)\|^2 - 4\rho_n \frac{h(x_n) + l(x_n)}{\theta^2(x_n)} h(x_n) \\ &\leq \|x_n - x^*\|^2 + \rho_n^2 \frac{(h(x_n) + l(x_n))^2}{\theta^2(x_n)} - 4\rho_n \frac{(h(x_n) + l(x_n))^2}{\theta^2(x_n)} \frac{h(x_n)}{h(x_n) + l(x_n)} \\ &= \|x_n - x^*\|^2 - \rho_n \left(\frac{4h(x_n)}{h(x_n) + l(x_n)} - \rho_n \right) \frac{(h(x_n) + l(x_n))^2}{\theta^2(x_n)}. \end{aligned} \tag{3.3}$$

From (3.1) and (3.3), we obtain

$$\begin{aligned} \|x_{n+1} - x^*\| &= \|\alpha_n r(x_n) + (1-\alpha_n) \mathrm{prox}_{\lambda \mu_n f} y_n - x^*\| \\ &\leq \alpha_n \|r(x_n) - x^*\| + (1-\alpha_n) \|y_n - x^*\| \\ &\leq \alpha_n \big[\|r(x_n) - r(x^*)\| + \|r(x^*) - x^*\| \big] + (1-\alpha_n)\|y_n - x^*\| \\ &\leq \alpha_n \alpha \|x_n - x^*\| + \alpha_n \|r(x^*) - x^*\| + (1-\alpha_n)\|x_n - x^*\| \\ &= \big[1 - \alpha_n(1-\alpha)\big] \|x_n - x^*\| + \alpha_n \|r(x^*) - x^*\| \\ &\leq \max\left\{ \|x_n - x^*\|, \frac{\|r(x^*) - x^*\|}{1-\alpha} \right\} \\ &\quad \vdots \\ &\leq \max\left\{ \|x_1 - x^*\|, \frac{\|r(x^*) - x^*\|}{1-\alpha} \right\}. \end{aligned} \tag{3.4}$$

Therefore, $\{x_n\}$ and $\{y_n\}$ are bounded.

The rest of the proof will be divided into two parts.

Case 1. Suppose that there exists $n_0 \in \mathbb{N}$ such that $\{\|y_n - x^*\|\}_{n=n_0}^{\infty}$ is nonincreasing. Then $\{\|y_n - x^*\|\}_{n=1}^{\infty}$ converges and $\|y_n - x^*\|^2 - \|y_{n+1} - x^*\|^2 \to 0$, $n \to \infty$. From (3.3) and (3.4), we have

$$\begin{aligned}
&\rho_n\left(\frac{4h(x_n)}{h(x_n)+l(x_n)}-\rho_n\right)\frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)} \\
&\quad\le \|x_n-x^*\|^2-\|y_n-x^*\|^2 \\
&\quad\le \left(\alpha_{n-1}\|r(x_{n-1})-x^*\|+(1-\alpha_{n-1})\|y_{n-1}-x^*\|\right)^2-\|y_n-x^*\|^2 \\
&\quad\le \|y_{n-1}-x^*\|^2-\|y_n-x^*\|^2+2\alpha_{n-1}\|r(x_{n-1})-x^*\|\|y_{n-1}-x^*\|+\alpha_{n-1}^2\|r(x_{n-1})-x^*\|^2.
\end{aligned}$$

Condition (a) above implies that

$$\rho_n\left(\frac{4h(x_n)}{h(x_n)+l(x_n)}-\rho_n\right)\frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)}\to 0,\quad n\to\infty.$$

Hence, we obtain

$$\frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)}\to 0,\quad n\to\infty. \tag{3.5}$$

Consequently, we have

$$\lim_{n\to\infty}\left(h(x_n)+l(x_n)\right)=0\quad\Leftrightarrow\quad\lim_{n\to\infty}h(x_n)=0\quad\text{and}\quad\lim_{n\to\infty}l(x_n)=0,$$

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

$$\begin{aligned}
&\|\nabla h(x_n)\|=\|\nabla h(x_n)-\nabla x^*\|\le\|A\|^2\|x_n-x^*\|\quad\text{and} \\
&\|\nabla l(x_n)\|=\|\nabla l(x_n)-\nabla x^*\|\le\|x_n-x^*\|.
\end{aligned}$$

We observe that

$$0<\mu_n<4\frac{h(x_n)+l(x_n)}{\theta^2(x_n)}\to 0,\quad n\to\infty,$$

implies that $\mu_n\to 0$, $n\to\infty$. Hence, we have from (3.1) that

$$\|y_n-x_n\|=\mu_n\|A^*(I-\operatorname{prox}_{\lambda g})Ax_n\|\le\mu_n M_1\to 0,\quad n\to\infty,$$

for some $M_1>0$.

From $\lim_{n\to\infty}\frac{1}{2}\|(I-\operatorname{prox}_{\lambda\mu_n f})x_n\|^2=\lim_{n\to\infty}l(x_n)=0$ and $\lim_{n\to\infty}\|y_n-x_n\|=0$, we have

$$\|y_n-\operatorname{prox}_{\lambda\mu_n f}x_n\|\le\|y_n-x_n\|+\|(I-\operatorname{prox}_{\lambda\mu_n f})x_n\|\to 0,\quad n\to\infty.$$

So,

$$\begin{aligned}
\|y_n-\operatorname{prox}_{\lambda\mu_n f}y_n\|&\le\|y_n-\operatorname{prox}_{\lambda\mu_n f}x_n\|+\|\operatorname{prox}_{\lambda\mu_n f}x_n-\operatorname{prox}_{\lambda\mu_n f}y_n\| \\
&\le\|y_n-\operatorname{prox}_{\lambda\mu_n f}x_n\|+\|x_n-y_n\|\to 0,\quad n\to\infty
\end{aligned}$$

and

$$\|x_n - \operatorname{prox}_{\lambda\mu_n f} y_n\| \leq \|x_n - y_n\| + \|y_n - \operatorname{prox}_{\lambda\mu_n f} y_n\| \to 0, \quad n \to \infty.$$

Also, observe that from (3.1), we obtain $\|x_{n+1} - \operatorname{prox}_{\lambda\mu_n f} y_n\| \to 0$, $n \to \infty$. We then have

$$\|x_{n+1} - x_n\| \leq \|x_{n+1} - \operatorname{prox}_{\lambda\mu_n f} y_n\| + \|x_n - \operatorname{prox}_{\lambda\mu_n f} y_n\| \to 0, \quad n \to \infty.$$

Now, let z be a weak cluster point of $\{x_n\}$, there exists a subsequence $\{x_{n_j}\}$ which weakly converges to z. The lower-semicontinuity of h then implies that

$$0 \leq h(z) \leq \liminf_{j\to\infty} h(x_{n_j}) = \lim_{n\to\infty} h(x_n) = 0.$$

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

Likewise, the lower-semicontinuity of l implies that

$$0 \leq l(z) \leq \liminf_{j\to\infty} l(x_{n_j}) = \lim_{n\to\infty} l(x_n) = 0.$$

That is, $l(z) = \frac{1}{2}\|(I - \operatorname{prox}_{\mu_n\lambda f})z\| = 0$, *i.e.*, z is a fixed point of the proximal mapping of f or equivalently $0 \in \partial f(z)$. In other words, z is a minimizer of f. Hence, $z \in \Gamma$.

Next, we prove that $\{x_n\}$ converges strongly to x^*, where x^* is the unique solution of the VI (3.2). First observe that there is some $z \in \omega_w(x_n) \subset \Gamma$ (where $\omega_w(x_n) := \{x : \exists x_{n_j} \rightharpoonup x\}$ is the weak w-limit set of the sequence $\{x_n\}_{n=1}^{\infty}$) such that

$$\limsup_{n\to\infty} \langle r(x^*) - x^*, x_n - x^*\rangle = \langle r(x^*) - x^*, z - x^*\rangle \leq 0. \tag{3.6}$$

Applying Lemma 2.1(ii) to (3.1), we have

$$\begin{aligned}
\|y_{n+1} - x^*\|^2 \leq \|x_{n+1} - x^*\|^2 &= \|\alpha_n(r(x_n) - x^*) + (1-\alpha_n)(\operatorname{prox}_{\lambda\mu_n f} y_n - x^*)\|^2 \\
&= \|\alpha_n(r(x_n) - r(x^*)) + (1-\alpha_n)(\operatorname{prox}_{\lambda\mu_n f} y_n - x^*) + \alpha_n(r(x^*) - x^*)\|^2 \\
&\leq \|\alpha_n(r(x_n) - r(x^*)) + (1-\alpha_n)(\operatorname{prox}_{\lambda\mu_n f} y_n - x^*)\|^2 \\
&\quad + 2\alpha_n\langle r(x^*) - x^*, x_{n+1} - x^*\rangle \\
&\leq \alpha_n\|r(x_n) - r(x^*)\|^2 + (1-\alpha_n)\|y_n - x^*\|^2 + 2\alpha_n\langle r(x^*) - x^*, x_{n+1} - x^*\rangle \\
&\leq \alpha_n\|r(x_n) - r(x^*)\|^2 + (1-\alpha_n)\|x_n - x^*\|^2 + 2\alpha_n\langle r(x^*) - x^*, x_{n+1} - x^*\rangle \\
&\leq \alpha_n\alpha^2\|x_n - x^*\|^2 + (1-\alpha_n)\|x_n - x^*\|^2 + 2\alpha_n\langle r(x^*) - x^*, x_{n+1} - x^*\rangle \\
&= [1 - (1-\alpha^2)\alpha_n]\|x_n - x^*\|^2 + 2\alpha_n\langle r(x^*) - x^*, x_{n+1} - x^*\rangle.
\end{aligned} \tag{3.7}$$

Now, using Lemma 2.2 in (3.7), we have $\|x_n - x^*\| \to 0$. That is, $x_n \to x^*$, $n \to \infty$.

Case 2. Assume that $\{\|y_n - x^*\|\}$ is not a monotonically decreasing sequence. Set $\Gamma_n = \|y_n - x^*\|^2$ and let $\tau : \mathbb{N} \to \mathbb{N}$ be a mapping for all $n \geq n_0$ (for some n_0 large enough) by

$$\tau(n) := \max\{k \in \mathbb{N} : k \leq n, \Gamma_k \leq \Gamma_{k+1}\}.$$

Clearly, τ is a nondecreasing sequence such that $\tau(n) \to \infty$ as $n \to \infty$ and

$$0 \le \Gamma_{\tau(n)} \le \Gamma_{\tau(n)+1}, \quad \forall n \ge n_0.$$

This implies that $\|y_{\tau(n)} - x^*\| \le \|y_{\tau(n)+1} - x^*\|$, $\forall n \ge n_0$. Thus $\lim_{n\to\infty} \|y_{\tau(n)} - x^*\|$ exists.

By (3.3) and (3.4), we obtain

$$\begin{aligned}
&\rho_{\tau(n)}\left(\frac{4h(x_{\tau(n)})}{h(x_{\tau(n)}) + l(x_{\tau(n)})} - \rho_{\tau(n)}\right)\frac{(h(x_{\tau(n)}) + l(x_{\tau(n)}))^2}{\theta^2(x_{\tau(n)})} \\
&\quad \le \|x_{\tau(n)} - x^*\|^2 - \|y_{\tau(n)} - x^*\|^2 \\
&\quad \le \big(\alpha_{\tau(n)-1}\|r(x_{\tau(n)-1}) - x^*\| + (1 - \alpha_{\tau(n)-1})\|y_{\tau(n)-1} - x^*\|\big)^2 - \|y_{\tau(n)} - x^*\|^2 \\
&\quad \le \|y_{\tau(n)-1} - x^*\|^2 - \|y_{\tau(n)} - x^*\|^2 + 2\alpha_{\tau(n)-1}\|r(x_{\tau(n)-1}) - x^*\|\|y_{\tau(n)-1} - x^*\| \\
&\qquad + \alpha^2_{\tau(n)-1}\|r(x_{\tau(n)-1}) - x^*\|^2.
\end{aligned}$$

Using condition (a) in the last inequality above, we have

$$\rho_{\tau(n)}\left(\frac{4h(x_{\tau(n)})}{h(x_{\tau(n)}) + l(x_{\tau(n)})} - \rho_{\tau(n)}\right)\frac{(h(x_{\tau(n)}) + l(x_{\tau(n)}))^2}{\theta^2(x_{\tau(n)})} \to 0, \quad n \to \infty.$$

Hence, we obtain

$$\frac{(h(x_{\tau(n)}) + l(x_{\tau(n)}))^2}{\theta^2(x_{\tau(n)})} \to 0, \quad n \to \infty. \tag{3.8}$$

Consequently, we have

$$\lim_{n\to\infty}\big(h(x_{\tau(n)}) + l(x_{\tau(n)})\big) = 0 \quad \Leftrightarrow \quad \lim_{n\to\infty} h(x_{\tau(n)}) = 0 \quad \text{and} \quad \lim_{n\to\infty} l(x_{\tau(n)}) = 0.$$

Furthermore, we observe that

$$0 < \mu_{\tau(n)} < 4\frac{h(x_{\tau(n)}) + l(x_{\tau(n)})}{\theta^2(x_{\tau(n)})} \to 0, \quad n \to \infty,$$

implies that $\mu_{\tau(n)} \to 0$, $n \to \infty$. This implies from (3.1) that

$$\|y_{\tau(n)} - x_{\tau(n)}\| = \mu_{\tau(n)}\|A^*(I - \mathrm{prox}_{\lambda g})Ax_{\tau(n)}\| \le \mu_{\tau(n)}M^* \to 0, \quad n \to \infty,$$

for some $M^* > 0$. Since $\{x_{\tau(n)}\}$ is bounded, there exists a subsequence of $\{x_{\tau(n)}\}$, still denoted by $\{x_{\tau(n)}\}$, which converges weakly to z. Observe that since $\lim_{n\to\infty} \|x_{\tau(n)} - y_{\tau(n)}\| = 0$, we also have $y_{\tau(n)} \rightharpoonup z$. By similar argument as above in Case 1, we can show that $z \in \Gamma$ and $\lim_{n\to\infty} \|x_{\tau(n)+1} - x_{\tau(n)}\| = 0$. Using (3.3) and (3.1), we obtain

$$\begin{aligned}
\|y_{\tau(n)+1} - x^*\|^2 &\le \|x_{\tau(n)+1} - x^*\|^2 \\
&= \|\alpha_{\tau(n)}\big(r(x_{\tau(n)}) - x^*\big) + (1 - \alpha_{\tau(n)})\big(\mathrm{prox}_{\lambda\mu_{\tau(n)}f}\, y_{\tau(n)} - x^*\big)\|^2 \\
&= \|\alpha_{\tau(n)}\big(r(x_{\tau(n)}) - r(x^*)\big) + (1 - \alpha_{\tau(n)})\big(\mathrm{prox}_{\lambda\mu_{\tau(n)}f}\, y_{\tau(n)} - x^*\big)
\end{aligned}$$

$$
\begin{aligned}
&\quad + \alpha_{\tau(n)}\big(r\big(x^*\big) - x^*\big)\big\|^2 \\
&\leq \big\|\alpha_{\tau(n)}\big(r(x_{\tau(n)}) - r\big(x^*\big)\big) + (1-\alpha_{\tau(n)})\big(\operatorname{prox}_{\lambda\mu_{\tau(n)}f}\, y_{\tau(n)} - x^*\big)\big\|^2 \\
&\quad + 2\alpha_{\tau(n)}\big\langle r\big(x^*\big) - x^*, x_{n+1} - x^*\big\rangle \\
&\leq \alpha_{\tau(n)}\big\|r(x_{\tau(n)}) - r\big(x^*\big)\big\|^2 + (1-\alpha_{\tau(n)})\big\|y_{\tau(n)} - x^*\big\|^2 \\
&\quad + 2\alpha_{\tau(n)}\big\langle r\big(x^*\big) - x^*, x_{n+1} - x^*\big\rangle \\
&\leq \alpha_{\tau(n)}\alpha^2\big\|x_{\tau(n)} - x^*\big\|^2 + (1-\alpha_{\tau(n)})\big\|y_{\tau(n)} - x^*\big\|^2 \\
&\quad + 2\alpha_{\tau(n)}\big\langle r\big(x^*\big) - x^*, x_{n+1} - x^*\big\rangle \\
&\leq \alpha_{\tau(n)}\alpha^2\big[\big\|y_{\tau(n)} - x^*\big\| + \|x_{\tau(n)} - y_{\tau(n)}\|\big]^2 + (1-\alpha_{\tau(n)})\big\|y_{\tau(n)} - x^*\big\|^2 \\
&\quad + 2\alpha_{\tau(n)}\big\langle r\big(x^*\big) - x^*, x_{n+1} - x^*\big\rangle \\
&= \alpha_{\tau(n)}\alpha^2\big[\big\|y_{\tau(n)} - x^*\big\|^2 + 2\big\|y_{\tau(n)} - x^*\big\|\|x_{\tau(n)} - y_{\tau(n)}\| + \|x_{\tau(n)} - y_{\tau(n)}\|^2\big] \\
&\quad + (1-\alpha_{\tau(n)})\big\|y_{\tau(n)} - x^*\big\|^2 + 2\alpha_{\tau(n)}\big\langle r\big(x^*\big) - x^*, x_{n+1} - x^*\big\rangle \\
&= \big(1 - \big(1-\alpha^2\big)\alpha_{\tau(n)}\big)\big\|y_{\tau(n)} - x^*\big\|^2 \\
&\quad + \alpha_{\tau(n)}\alpha^2\big[2\big\|y_{\tau(n)} - x^*\big\|\|x_{\tau(n)} - y_{\tau(n)}\| + \|x_{\tau(n)} - y_{\tau(n)}\|^2\big] \\
&\quad + 2\alpha_{\tau(n)}\big\langle r\big(x^*\big) - x^*, x_{n+1} - x^*\big\rangle,
\end{aligned}
$$

which implies that, for all $n \geq n_0$,

$$
\begin{aligned}
0 &\leq \big\|y_{\tau(n)+1} - x^*\big\|^2 - \big\|y_{\tau(n)} - x^*\big\|^2 \\
&\leq \alpha_{\tau(n)}\big[2\big\langle r\big(x^*\big) - x^*, x_{\tau(n)+1} - x^*\big\rangle - \big(1-\alpha^2\big)\big\|y_{\tau(n)} - x^*\big\|^2\big] \\
&\quad + \alpha_{\tau(n)}\alpha^2\big[2\big\|y_{\tau(n)} - x^*\big\|\|x_{\tau(n)} - y_{\tau(n)}\| + \|x_{\tau(n)} - y_{\tau(n)}\|^2\big].
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\big\|y_{\tau(n)} - x^*\big\|^2 &\leq \frac{2}{1-\alpha^2}\big\langle r\big(x^*\big) - x^*, x_{\tau(n)+1} - x^*\big\rangle \\
&\quad + \frac{\alpha^2}{1-\alpha^2}\big[2\big\|y_{\tau(n)} - x^*\big\|\|x_{\tau(n)} - y_{\tau(n)}\| + \|x_{\tau(n)} - y_{\tau(n)}\|^2\big].
\end{aligned}
$$

Hence, we obtain (noting that x^* is the unique solution of the VI (3.2))

$$
\begin{aligned}
\limsup_{n\to\infty}\big\|y_{\tau(n)} - x^*\big\|^2 &\leq \frac{2}{1-\alpha^2}\big\langle r\big(x^*\big) - x^*, z - x^*\big\rangle \\
&\quad + \frac{2\alpha^2}{1-\alpha^2}\limsup_{n\to\infty}\big\|y_{\tau(n)} - x^*\big\|\|x_{\tau(n)} - y_{\tau(n)}\| \\
&\quad + \frac{\alpha^2}{1-\alpha^2}\limsup_{n\to\infty}\|x_{\tau(n)} - y_{\tau(n)}\|^2 \leq 0,
\end{aligned}
$$

which implies that

$$
\lim_{n\to\infty}\big\|y_{\tau(n)} - x^*\big\| = 0.
$$

Therefore,

$$\lim_{n\to\infty} \Gamma_{\tau(n)} = \lim_{n\to\infty} \Gamma_{\tau(n)+1} = 0.$$

Furthermore, for $n \geq n_0$, it is easy to see that $\Gamma_{\tau(n)} \leq \Gamma_{\tau(n)+1}$ if $n \neq \tau(n)$ (that is, $\tau(n) < n$), because $\Gamma_j \geq \Gamma_{j+1}$ for $\tau(n) + 1 \leq j \leq n$. As a consequence, we obtain for all $n \geq n_0$,

$$0 \leq \Gamma_n \leq \max\{\Gamma_{\tau(n)}, \Gamma_{\tau(n)+1}\} = \Gamma_{\tau(n)+1}.$$

Hence, $\lim \Gamma_n = 0$, that is, $\{y_n\}$ converges strongly to x^*. Hence, $\{x_n\}$ converges strongly to x^*. This completes the proof. □

Taking $r(x) = u$ in (3.1), we have the following algorithm.

Given an initial point $x_1 \in H_1$. Assume that x_n has been constructed and $\theta(x_n) \neq 0$, then compute x_{n+1} via the rule

$$\begin{cases} y_n = x_n - \mu_n A^*(I - \operatorname{prox}_{\lambda g})Ax_n, \\ x_{n+1} = \alpha_n u + (1 - \alpha_n) \operatorname{prox}_{\lambda\mu_n f} y_n, \quad n \geq 1, \end{cases} \tag{3.9}$$

where the step size $\mu_n := \rho_n \frac{h(x_n)+l(x_n)}{\theta^2(x_n)}$ with $0 < \rho_n < 4$. If $\theta(x_n) = 0$, then $x_{n+1} = x_n$ is a solution of (1.8) and the iterative process stops, otherwise, we set $n := n + 1$ and go to (3.9).

Corollary 3.2 *Assume that f and g are two proper convex lower-semicontinuous functions and that* (1.8) *is consistent* (*i.e.*, $\Gamma \neq \emptyset$). *If the parameters satisfy the following conditions*:

(a) $\lim_{n\to\infty} \alpha_n = 0$;

(b) $\sum_{n=1}^{\infty} \alpha_n = \infty$;

(c) $\epsilon \leq \rho_n \leq \frac{4h(x_n)}{h(x_n)+l(x_n)} - \epsilon$ *for some* $\epsilon > 0$,

the sequence $\{x_n\}$ *generated by* (3.9) *strongly converges to a solution of* (1.8) *which is closest to u from the solution set* Γ. *In other words*, x^* *is the unique fixed point of the contraction* $\operatorname{Proj}_\Gamma r$, $x^* = (\operatorname{Proj}_\Gamma)u$.

4 Strong convergence for nonconvex minimization feasibility problem

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

$$0 \in \partial f(\bar{x}) \text{ such that } 0 \in \partial_{pg}(A\bar{x}), \tag{4.1}$$

is very general in the sense that it includes, as special cases, g is convex and g is a lower-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, [28]. In what follows, we shall represent the set of solutions of (4.1) by Γ.

Definition 4.1 Let $g : H_2 \to \mathbb{R} \cup \{+\infty\}$ be a function and let $\bar{x} \in \operatorname{dom} g$, *i.e.*, $g(\bar{x}) < +\infty$. A vector v is in proximal subdifferential $\partial_{pg}(\bar{x})$ if there exist some $r > 0$ and $\epsilon > 0$ such that for all $x \in B(\bar{x}, \epsilon)$,

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

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

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 $\epsilon > 0$, the g-attentive ϵ-localization of $\partial_{pg}(\bar{x})$ around $(\bar{x}, \bar{v})$, is the mapping $T_\epsilon : H_2 \to 2^{H_2}$ defined by

$$\begin{cases} \{v \in \partial_{pg}(x), \|v - \bar{v}\| < \epsilon\} & \text{if } \|x - \bar{x}\| < \epsilon \text{ and } |g(x) - g(\bar{x})| < \epsilon, \\ \emptyset & \text{otherwise.} \end{cases}$$

Definition 4.2 A function g is said to be prox-regular at $\bar{x}$ for $\bar{v} \in \partial_{pg}(\bar{x})$ if there exist some $r > 0$ and $\epsilon > 0$ such that for all $x, x' \in B(\bar{x}, \epsilon)$ with $|g(x) - g(x')| < \epsilon$ and all $v \in B(\bar{v}, \epsilon)$ with $v \in \partial_{pg}(\bar{x})$ one has

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

If the property holds for all vectors $\bar{v} \in \partial_{pg}(\bar{x})$, 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 $\operatorname{prox}_{\lambda g}$ defined for $\lambda > 0$, respectively, by

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

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, Section 1 in [29], that local minima are zeros of the proximal subdifferential and that the proximal subdifferential and the convex one coincide in the convex case.

Now, let us state the following key property of the proximal mapping complement, which was proved in Remark 3.2 of Moudafi and Thakur [18].

Lemma 4.3 (Moudafi and Thakur [18]) *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-localization of ∂_{pg} around $(\bar{x}, \bar{v})$. Then for each $\lambda \in (0, \frac{1}{r})$ and x_1, x_2 in a neighborhood U_λ of $\bar{x}$, one has*

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

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

Now, the regularization parameters λ are allowed to vary in algorithm (3.1), namely considering possibly variable parameters $\lambda \in (0, \frac{1}{r} - \epsilon)$ (for some $\epsilon > 0$ small enough) and $\mu_n > 0$, our interest is in studying the convergence properties of the following algorithm.

Modified split proximal algorithm 2 Let $r : H_1 \to H_1$ be a contraction mapping with constant $\alpha \in (0,1)$. Given an initial point $x_1 \in H_1$. Assume that x_n has been constructed and $\theta(x_n) \neq 0$, then compute x_{n+1} via the rule

$$\begin{cases} y_n = x_n - \mu_n A^*(I - \operatorname{prox}_{\lambda_n g})Ax_n, \\ x_{n+1} = \alpha_n r(x_n) + (1-\alpha_n)\operatorname{prox}_{\lambda \mu_n f} y_n, \quad n \geq 1, \end{cases} \tag{4.2}$$

where the step size $\mu_n := \rho_n \frac{h(x_n)+l(x_n)}{\theta^2(x_n)}$ with $0 < \rho_n < 4$. If $\theta(x_n) = 0$, then $x_{n+1} = x_n$ is a solution of (4.1) and the iterative process stops, otherwise, we set $n := n+1$ and go to (4.2).

Theorem 4.4 *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* (4.1) *and A a bounded linear operator which is surjective with a dense domain. If the parameters satisfy the following conditions*:

(a) $\lim_{n\to\infty} \alpha_n = 0$;
(b) $\sum_{n=1}^{\infty} \alpha_n = \infty$;
(c) $\epsilon \leq \rho_n \leq \frac{4h(x_n)}{h(x_n)+l(x_n)} - \epsilon$ *for some* $\epsilon > 0$;
(d) $\sum_{n=1}^{\infty} \lambda_n < \infty$;

and if $\|x_1 - \bar{x}\|$ is small enough, then the sequence $\{x_n\}$ generated by (4.2) *strongly converges to a solution of* (4.1) *which is also the unique solution of the variational inequality* (*VI*)

$$\bar{x} \in \Gamma, \quad \langle (I-r)\bar{x}, x - \bar{x} \rangle \geq 0, \quad x \in \Gamma. \tag{4.3}$$

In other words, $\bar{x}$ is the unique fixed point of the contraction $\operatorname{Proj}_\Gamma r$, $\bar{x} = (\operatorname{Proj}_\Gamma r)\bar{x}$.

Proof Using the fact that $\operatorname{prox}_{\lambda_n \mu_n f}$ is nonexpansive, $\bar{x}$ verifies (4.1) (critical points of any function are exactly fixed points of its proximal mapping) and having in mind Lemma 4.3, we can write

$$\begin{aligned} \|y_n - \bar{x}\|^2 &= \|x_n - \bar{x}\|^2 + \mu_n^2 \|\nabla h(x_n)\|^2 - 2\mu_n \langle \nabla h(x_n), x_n - \bar{x} \rangle \\ &\leq \|x_n - \bar{x}\|^2 + \mu_n^2 \|\nabla h(x_n)\|^2 - 2\mu_n \left(2h(x_n) - \frac{\lambda_n r \|A\|^2}{(1-\lambda_n r)^2} \|x_n - \bar{x}\|^2 \right) \\ &= \|x_n - \bar{x}\|^2 + 2\mu_n \frac{\lambda_n r \|A\|^2}{(1-\lambda_n r)^2} \|x_n - \bar{x}\|^2 - 4\mu_n h(x_n) + \mu_n^2 \|\nabla h(x_n)\|^2 \\ &\leq \|x_n - \bar{x}\|^2 + 2\rho_n \frac{h(x_n)+l(x_n)}{\|\nabla h(x_n)\|^2 + \|\nabla l(x_n)\|^2} \frac{\lambda_n r \|A\|^2}{(1-\lambda_n r)^2} \|x_n - \bar{x}\|^2 \\ &\quad - \rho_n \left(\frac{4h(x_n)}{h(x_n)+l(x_n)} - \rho_n \right) \frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)} \\ &\leq \left(1 + \lambda_n \rho_n \left(\frac{2h(x_n)}{\|\nabla h(x_n)\|^2} + \frac{2l(x_n)}{\|\nabla l(x_n)\|^2} \right) \frac{r\|A\|^2}{(1-\lambda_n r)^2} \|x_n - \bar{x}\|^2 \right) \end{aligned}$$

$$
\begin{aligned}
&- \rho_n \left(\frac{4h(x_n)}{h(x_n)+l(x_n)} - \rho_n \right) \frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)} \\
&= \left(1 + \lambda_n \rho_n \left(1 + \frac{2h(x_n)}{\|\nabla h(x_n)\|^2} \right) \frac{r\|A\|^2}{(1-\lambda_n r)^2} \|x_n - \bar{x}\|^2 \right) \\
&\quad - \rho_n \left(\frac{4h(x_n)}{h(x_n)+l(x_n)} - \rho_n \right) \frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)}.
\end{aligned} \tag{4.4}
$$

Recall that A is surjective with a dense domain $\Leftrightarrow \exists \gamma > 0$ such that $\|A^* x\| \geq \gamma \|x\|$ (see, for example, Theorem II.19 of Brézis [30]). This ensures that

$$
\frac{2h(x_n)}{\|\nabla h(x_n)\|^2} = \frac{\|(I - \operatorname{prox}_{\lambda_n g})(Ax_n)\|^2}{\|A^*(I - \operatorname{prox}_{\lambda_n g})(Ax_n)\|^2} \leq \frac{1}{\gamma^2}.
$$

The conditions on the parameters λ_n and ρ_n assure the existence of a positive constant M such that

$$
\|y_n - \bar{x}\|^2 \leq (1 + M\lambda_n)\|x_n - \bar{x}\|^2 - \rho_n \left(\frac{4h(x_n)}{h(x_n)+l(x_n)} - \rho_n \right) \frac{(h(x_n)+l(x_n))^2}{\theta^2(x_n)}. \tag{4.5}
$$

Using (4.5) in (4.2) (taking into account that $1 + x \leq e^x$, $x \geq 0$), we obtain

$$
\begin{aligned}
\|x_{n+1} - \bar{x}\| &\leq \left\| \alpha_n r(x_n) + (1-\alpha_n) \operatorname{prox}_{\lambda \mu_n f} y_n - \bar{x} \right\| \\
&\leq \alpha_n \left\| r(x_n) - \bar{x} \right\| + (1-\alpha_n)\|y_n - \bar{x}\| \\
&\leq \alpha_n \left(\left\| r(x_n) - r(\bar{x}) \right\| + \left\| r(\bar{x}) - \bar{x} \right\| \right) + (1-\alpha_n)\|y_n - \bar{x}\| \\
&\leq \alpha_n \alpha \|x_n - \bar{x}\| + \alpha_n \left\| r(\bar{x}) - \bar{x} \right\| + (1-\alpha_n)(1 + M\lambda_n)^{\frac{1}{2}} \|x_n - \bar{x}\| \\
&\leq \left(e^{M\lambda_n} \right)^{\frac{1}{2}} \left(\left[1 - \alpha_n(1-\alpha) \right] \|x_n - \bar{x}\| + \alpha_n \left\| r(\bar{x}) - \bar{x} \right\| \right) \\
&\leq \left(e^{M\lambda_n} \right)^{\frac{1}{2}} \left(\max \left\{ \|x_n - \bar{x}\|, \frac{\|r(\bar{x}) - \bar{x}\|}{1-\alpha} \right\} \right) \\
&= e^{\frac{M}{2}\lambda_n} \left(\max \left\{ \|x_n - \bar{x}\|, \frac{\|r(\bar{x}) - \bar{x}\|}{1-\alpha} \right\} \right) \\
&\ \ \vdots \\
&\leq e^{\frac{M}{2} \sum_{n=1}^{\infty} \lambda_n} \left(\max \left\{ \|x_1 - \bar{x}\|, \frac{\|r(\bar{x}) - \bar{x}\|}{1-\alpha} \right\} \right).
\end{aligned}
$$

Therefore, $\{x_n\}$ and $\{y_n\}$ are bounded.

Following the method of proof of Theorem 3.1, we can show that

$$
\lim_{n\to\infty} \left(h(x_n) + l(x_n) \right) = 0 \quad \Leftrightarrow \quad \lim_{n\to\infty} h(x_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} l(x_n) = 0.
$$

If z is a weak cluster point of $\{x_n\}$, then there exists a subsequence $\{x_{n_j}\}$ which weakly converges to z. From the proof of Theorem 3.1, we can show that $0 \in \partial f(z)$ such that $0 \in \partial_{pg}(Az)$.

Finally, from (4.2), we have

$$
\begin{aligned}
\|x_{n+1} - \bar{x}\|^2 &= \left\| \alpha_n \left(r(x_n) - \bar{x} \right) + (1-\alpha_n)(\operatorname{prox}_{\lambda \mu_n f} y_n - \bar{x}) \right\|^2 \\
&= \left\| \alpha_n \left(r(x_n) - r(\bar{x}) \right) + (1-\alpha_n)(\operatorname{prox}_{\lambda \mu_n f} y_n - \bar{x}) + \alpha_n \left(r(\bar{x}) - \bar{x} \right) \right\|^2
\end{aligned}
$$

$$\begin{aligned}
&\le \bigl\|\alpha_n\bigl(r(x_n) - r(\bar{x})\bigr) + (1-\alpha_n)(\operatorname{prox}_{\lambda\mu_n f} y_n - \bar{x})\bigr\|^2 + 2\alpha_n\bigl\langle r(\bar{x}) - \bar{x}, x_{n+1} - \bar{x}\bigr\rangle \\
&\le \alpha_n\bigl\|r(x_n) - r(\bar{x})\bigr\|^2 + (1-\alpha_n)\|y_n - \bar{x}\|^2 + 2\alpha_n\bigl\langle r(\bar{x}) - \bar{x}, x_{n+1} - \bar{x}\bigr\rangle \\
&\le \alpha_n\bigl\|r(x_n) - r(\bar{x})\bigr\|^2 + (1-\alpha_n)(1+M\lambda_n)\|x_n - \bar{x}\|^2 + 2\alpha_n\bigl\langle r(\bar{x}) - \bar{x}, x_{n+1} - \bar{x}\bigr\rangle \\
&\le \alpha_n\alpha^2\|x_n - \bar{x}\|^2 + (1-\alpha_n)\|x_n - \bar{x}\|^2 \\
&\quad + 2\alpha_n\bigl\langle r(\bar{x}) - \bar{x}, x_{n+1} - \bar{x}\bigr\rangle + M\lambda_n\|x_n - \bar{x}\|^2 \\
&= \bigl[1 - \bigl(1-\alpha^2\bigr)\alpha_n\bigr]\|x_n - \bar{x}\|^2 + 2\alpha_n\bigl\langle r(\bar{x}) - \bar{x}, x_{n+1} - \bar{x}\bigr\rangle + \lambda_n M_1, \qquad (4.6)
\end{aligned}$$

for some $M_1 > 0$, from which one concludes that the sequence $\{x_n\}$ strongly converges to a solution of (4.1) using Lemma 2.2. □

Remark 4.5 From Example 1.4, we recall that if $f = \delta_C$ and $g = \delta_Q$ the indicator functions of two nonempty, closed, and convex sets C, Q of H_1 and H_2, respectively, where $H_1 = \ell_2 = H_2$, then $\operatorname{prox}_{\lambda\mu_n f}(x) = P_C(x)$ and $\operatorname{prox}_{\lambda g}(x) = P_Q(x)$. Furthermore, if $A = I = A^*$, where I is an identity mapping on ℓ_2, $\mu_n = 1$, $\forall n \ge 1$, then our modified split proximal algorithm (3.1) for approximation of solutions of problem (1.8) becomes $x_1 \in C$,

$$x_{n+1} = \alpha_n r(x_n) + (1-\alpha_n)(P_C \circ P_Q)x_n, \quad n \ge 1. \qquad (4.7)$$

Noting that $y_n = P_Q x_n$ in (3.1) and following the same method of proof as in Theorem 3.1, we see that our algorithm (4.7) *converges strongly* to a solution of problem (1.8).

We see from Example 1.4 and the above remark that the split proximal algorithm (1.9) generates *weakly* (*not strongly*) *convergent sequences* in general in infinite-dimensional spaces, while our modified split proximal algorithm (3.1) generates *strongly convergent sequences* in infinite-dimensional real Hilbert spaces.

5 Numerical results

In this section, we provide some concrete example including numerical results of the problem considered in Section 3 of this paper. All codes were written in Matlab 2012b and run on HP i5 Dual Core 8.00 GB (7.78 GB usable) RAM laptop.

Example 5.1 Let $C = Q = \{x \in \mathbb{R}^3 : \|x\|_2 \le 1\}$ and take

$$A = 100 \times \begin{pmatrix} 4 & 5 & 7 \\ 6 & 8 & 8 \\ 8 & 7 & 6 \end{pmatrix}.$$

In problem (1.8), let $f = \delta_C$ and $g = \delta_Q$ be the indicator functions of two nonempty, closed, and convex sets C, Q of $\mathbb{R}^3 = H_1 = H_2$. Then

$$\operatorname{prox}_{\lambda\mu_n f}(x) = P_C(x) = \operatorname{prox}_{\lambda g}(x) = P_Q(x) = \begin{cases} x, & \|x\|_2 \le 1, \\ \frac{x}{\|x\|_2}, & \|x\|_2 > 1, x \in \mathbb{R}^3. \end{cases}$$

Hence problem (1.8) becomes: Find some point x in C such that $Ax \in Q$. Now, take $\rho_n = 2$, $\alpha_n = \frac{1}{n+1}$. Also, $h(x_n) = \frac{1}{2}\|(I - P_Q)Ax_n\|_2^2$, $l(x_n) = \frac{1}{2}\|(I - P_C)x_n\|_2^2$ and $\theta(x_n) := \sqrt{\|\nabla h(x_n)\|_2^2 + \|\nabla l(x_n)\|_2^2}$ with $\|\nabla h(x_n)\|_2^2 = \|A^T(I - P_Q)Ax_n\|_2^2$, $\|\nabla l(x_n)\|_2^2 = \|(I - P_C)x_n\|_2^2$. In the implementation, we took $\|A^T(I - P_Q)Ax_n\|_2^2 + \|(I - P_C)x_n\|_2^2 < \frac{1}{10^4}$ as the stopping criterion.

Table 1 Example 5.1, Case I

Time taken	No. of iterations	$\|x_{n+1} - x_n\|_2$
1.8966e–04	1	0.1295
	2	0.0359
	3	0.0064
	4	0.0011
	5	0.0005
	6	0.0004
	7	0.0002

Figure 1 Example 5.1, Case I.

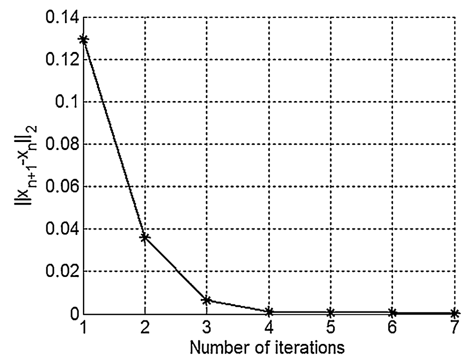


Table 2 Example 5.1, Case II

Time taken	No. of iterations	$\|x_{n+1} - x_n\|_2$
4.2706e–04	1	0.5187
	2	0.1441
	3	0.0256
	4	0.0050
	5	0.0067
	6	0.0045
	7	0.0054
	8	0.0030
	9	0.0014
	10	0.0010
	11	0.0018
	12	0.0008
	13	0.0004
	14	0.0004
	15	0.0002
	16	0.0001

Our iterative scheme (3.1) then becomes

$$\begin{cases} y_n = x_n - \mu_n A^T(I - P_Q)Ax_n, \\ x_{n+1} = \alpha_n r(x_n) + (1 - \alpha_n)P_C y_n, \quad n \geq 1. \end{cases} \tag{5.1}$$

Let $r(x) = \frac{1}{2}(x_1, x_2, x_3)$, $x = (x_1, x_2, x_3)$. We consider initial values for the problem considered in this example.

Case I: Take $x_1 = (0.1, 0.1, 0.1)$. The numerical result of this problem using our algorithm (5.1) with this initial value is listed in Table 1 and the graph is given in Figure 1.

Case II: Take $x_1 = (0.4, 0.4, 0.4)$. The numerical result of this problem using our algorithm (5.1) with this initial value is listed in Table 2 and the graph is given in Figure 2.

Figure 2 Example 5.1, Case II.

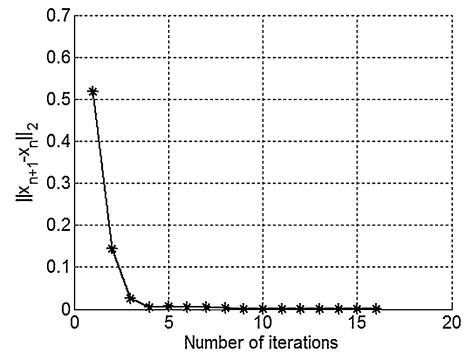
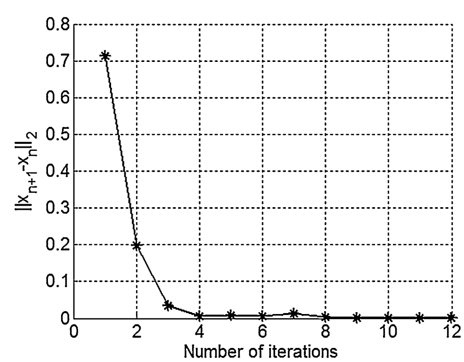


Table 3 Example 5.1, Case III

Time taken	No. of iterations	$\|x_{n+1} - x_n\|_2$
3.2844e–04	1	0.7133
	2	0.1982
	3	0.0352
	4	0.0069
	5	0.0082
	6	0.0060
	7	0.0123
	8	0.0025
	9	0.0007
	10	0.0010
	11	0.0003
	12	0.0002

Figure 3 Example 5.1, Case III.



Case III: Take $x_1 = (0.55, 0.55, 0.55)$. The numerical result of this problem using our algorithm (5.1) with this initial value is listed in Table 3 and the graph is given in Figure 3.

Competing interests

The authors declare that there is no conflict of interests regarding this manuscript.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details

[1]Department of Mathematics, University of Nigeria, Nsukka, Nigeria. [2]School of Mathematics Science, Chongqing Normal University, Chongqing, 400047, China. [3]Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA.

Acknowledgements
This work was supported by the NSF of China (No. 11401063) and Natural Science Foundation of Chongqing (cstc2014jcyjA00016).

Received: 18 December 2014 Accepted: 2 July 2015 Published online: 22 July 2015

References

1. Censor, Y, Elfving, T: A multiprojection algorithm using Bregman projections in a product space. Numer. Algorithms **8**(2-4), 221-239 (1994)
2. Byrne, C: Iterative oblique projection onto convex sets and the split feasibility problem. Inverse Probl. **18**(2), 441-453 (2002)
3. Byrne, C: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. **20**(1), 103-120 (2004)
4. Maingé, PE: The viscosity approximation process for quasi-nonexpansive mappings in Hilbert spaces. Comput. Math. Appl. **59**(1), 74-79 (2010)
5. Qu, B, Xiu, N: A note on the CQ algorithm for the split feasibility problem. Inverse Probl. **21**(5), 1655-1665 (2005)
6. Xu, H-K: A variable Krasnosel'skii-Mann algorithm and the multiple-set split feasibility problem. Inverse Probl. **22**(6), 2021-2034 (2006)
7. Yang, Q: The relaxed CQ algorithm solving the split feasibility problem. Inverse Probl. **20**(4), 1261-1266 (2004)
8. Yang, Q, Zhao, J: Generalized KM theorems and their applications. Inverse Probl. **22**(3), 833-844 (2006)
9. Yao, Y, Jigang, W, Liou, Y-C: Regularized methods for the split feasibility problem. Abstr. Appl. Anal. **2012**, Article ID 140679 (2012)
10. Ansari, QH, Rehan, A: Split feasibility and fixed point problems. In: Ansari, QH (ed.) Nonlinear Analysis: Approximation Theory, Optimization and Applications, pp. 281-322. Springer, New York (2014)
11. Xu, H-K: Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces. Inverse Probl. **26**(10), Article ID 105018 (2010)
12. Yao, Y, Chen, R, Liou, Y-C: A unified implicit algorithm for solving the triple-hierarchical constrained optimization problem. Math. Comput. Model. **55**(3-4), 1506-1515 (2012)
13. Yao, Y, Cho, Y-J, Liou, Y-C: Hierarchical convergence of an implicit double-net algorithm for nonexpansive semigroups and variational inequalities. Fixed Point Theory Appl. **2011**, Article ID 101 (2011)
14. Yao, Y, Liou, Y-C, Kang, SM: Two-step projection methods for a system of variational inequality problems in Banach spaces. J. Glob. Optim. **55**(4), 801-811 (2013)
15. Zhao, J, Yang, Q: Several solution methods for the split feasibility problem. Inverse Probl. **21**(5), 1791-1799 (2005)
16. Rockafellar, RT, Wets, R: Variational Analysis. Springer, Berlin (1988)
17. Lopez, G, Martin-Marquez, V, Wang, F, Xu, H-K: Solving the split feasibility problem without prior knowledge of matrix norms. Inverse Probl. **28**, Article ID 085004 (2012)
18. Moudafi, A, Thakur, BS: Solving proximal split feasibility problems without prior knowledge of operator norms. Optim. Lett. **8**, 2099-2110 (2014). doi:10.1007/s11590-013-0708-4
19. Luke, DR, Burke, JV, Lyon, RG: Optical wavefront reconstruction: theory and numerical methods. SIAM Rev. **44**, 169-224 (2002)
20. Hundal, H: An alternating projection that does not converge in norm. Nonlinear Anal. **57**, 35-61 (2004)
21. Bauschke, HH, Burke, JV, Deutsch, FR, Hundal, HS, Vanderwerff, JD: A new proximal point iteration that converges weakly but not in norm. Proc. Am. Math. Soc. **133**, 1829-1835 (2005)
22. Bauschke, HH, Matouskova, E, Reich, S: Projections and proximal point methods: convergence results and counterexamples. Nonlinear Anal. **56**, 715-738 (2004)
23. Matouskova, E, Reich, S: The Hundal example revisited. J. Nonlinear Convex Anal. **4**, 411-427 (2003)
24. Shehu, Y: Approximation of common solutions to proximal split feasibility problems and fixed point problems. Fixed Point Theory (accepted)
25. Shehu, Y: Iterative methods for convex proximal split feasibility problems and fixed point problems. Afr. Math. (2015). doi:10.1007/s13370-015-0344-5
26. Shehu, Y, Ogbuisi, FU: Convergence analysis for proximal split feasibility problems and fixed point problems. J. Appl. Math. Comput. **48**, 221-239 (2015)
27. Xu, H-K: Iterative algorithm for nonlinear operators. J. Lond. Math. Soc. **66**(2), 1-17 (2002)
28. Luke, DR: Finding best approximation pairs relative to a convex and prox-regular set in Hilbert space. SIAM J. Optim. **19**(2), 714-729 (2008)
29. Clarke, ZFH, Ledyaev, YS, Stern, RJ, Wolenski, PR: Nonsmooth Analysis and Control Theory. Springer, New York (1998)
30. Brézis, H: Analyse Fonctionnelle: Théorie et Applications. Masson, Paris (1983)