Computational and Applied Mathematics (2022) 41:5
https://doi.org/10.1007/s40314-021-01703-3

The iterative method for solving the proximal split feasibility problem with an application to LASSO problem

Xiaojun Ma[1] · Hongwei Liu[1] · Xiaoyin Li[1]

Received: 28 June 2021 / Revised: 30 September 2021 / Accepted: 13 November 2021 /
Published online: 9 December 2021

Abstract
In this paper, we investigate strong convergence of the iterative algorithm for solving the proximal split feasibility problem in real Hilbert spaces. The algorithm is motivated by the inertial method, the viscosity-type method and the split proximal algorithm with a self-adaptive stepsize. A strong convergence theorem for the proposed algorithm is established without requiring firm-nonexpasiveness of the involved operators. An application of our obtained results is offered. Finally, some numerical experiments are provided for illustration and comparison.

Keywords Proximal split feasibility problem · Viscosity-type method · Non-expansiveness · Strong convergence

Mathematics Subject Classification 47J20 · 49J40 · 35A15 · 47J25

1 Introduction

It is well known that the split feasibility problem (SFP) plays a key role in signal processing Byrne (2004) and medical image reconstruction Byrne (2002). Therefore, many numerical algorithms have been developed to solve the SFP; see Byrne (2004, 2002); Censor et al. (2005); Dong et al. (2020); López et al. (2012); Reich et al. (2020); Sahu et al. (2020) and the references therein.

The original model of the SFP was considered by Censor and Elfving Censor and Elfving (1994) for modeling inverse problems, and a classical method for solving the SFP is

Communicated by Joerg Fliege.

✉ Xiaojun Ma
fzhuangmaxj@163.com

Hongwei Liu
hwliuxidian@163.com

Xiaoyin Li
lixiaoyin2020@163.com

1 Xidian University School of Mathematics and Statistics, Xi'an 710126, Shaanxi, China

Byrne's CQ algorithm Byrne (2004, 2002). It is easily shown that the SFP can be got from the proximal split feasibility problem (SFP) which is a generalization of proximal split minimization problems in Moudafi and Thakur (2014). For the proximal SFP, there are numerous iteratively algorithms for the study of its convergence properties; see Abbas et al. (2018); Moudafi and Thakur (2014); Shehu and Iyiola (2017, ?, 2018); Wang and Xu (2014); Shehu and Iyiola (2018) and references therein. To be specific, Moudafi and Thakur Moudafi and Thakur (2014) discussed its weak convergence by introducing a split proximal algorithm in which the self-adaptive stepsize was not determined by an Armijo-like rule Dong et al. (2018); Gibali et al. (2018); Qu and Xiu (2005); Shehu and Gibali (2020). This rule often results in additional computation costs. Based on the inertial idea (which is viewed as a procedure of speeding up the convergence properties; see Kesornprom and Cholamjiak (2019); Sahu et al. (2020); Suantai et al. (2018); Shehu et al. (2020); Iyiola et al. (2018)), Shehu et al. Shehu and Iyiola (2017) used the inertial technique to modify Moudafi et al.'s algotithm and obtained the weak convergence in real Hilbert spaces. To obtain its strong convergence, various related algorithms have been proposed in recent years. For instance, Abbas et al. Abbas et al. (2018) presented two divergent one-step methods; Shehu et al. Shehu and Iyiola (2017, 2018) combined Mann-type, accelerated hybrid viscosity, and steepest-descent methods to ensure it; Wang and Xu (2014) proposed the proximal gradient method. However, the study of strong convergence of the algorithm Shehu and Iyiola (2017) with new inertial effects has yet to be founded. We also observe that Shehu et al.'s algorithm Shehu and Iyiola (2017) does not require the estimation of the operator norm, but the convergence of their algorithm requires firm-nonexpasiveness of the involved operators. These observations bring us the following concern:

Question: Can we prove a strong convergence result for the proximal SFP employing a new modification of the inertial split proximal algorithm Shehu and Iyiola (2017) *under weaker conditions than firm-nonexpasiveness of the involved operators?*

Inspired and motivated by the works in Moudafi (2000); Shehu and Iyiola (2017), in this paper, we propose an iterative algorithm for solving the proximal SFP. The algorithm consists of the inertial method, the viscosity-type algorithm, and the split proximal algorithm with a self-adaptive stepsize. The strong convergence of the offered algorithm is established but without firm-nonexpasiveness of the mappings involved. We also provide an application of our main results for solving the split feasibility problems in Hilbert spaces. Finally, three numerical examples are listed for illustrating the effectiveness of the proposed algorithm.

The paper is arranged as follows. In Sect. 2, some basic concepts and lemmas used in subsequent sections are proposed. The main results are presented in Sect. 3. Numerical experiments are provided in Sect. 4. We give some conclusions in the final section.

2 Preliminaries

The symbol $\rightharpoonup$ stands for the weak convergence, the symbol $\to$ represents the strong convergence. Let H_1 and H_2 be real Hilbert spaces. For a proper lower semi-continuous convex (lsc) function $F : H_1 \to]-\infty, \infty]$, its domain is denoted by $\mathrm{dom} F$, i.e., $\mathrm{dom} F := \{x \in H_1 : F(x) < \infty\}$. The proximal operator $\mathrm{prox}_{\tau G} : H_2 \to H_2$ is defined by

$$\mathrm{prox}_{\tau G}(x) := \arg\min_{y \in H_2} \left\{ G(y) + \frac{1}{2\tau} \|x - y\|^2 \right\},$$

where $\tau > 0$ and $G : H_2 \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex, and lower semi-continuous (lsc) function.

In view of Combettes and Hirstoaga (2005), the proximal mapping $\text{prox}_{\tau G}$ is firmly nonexpansive, that is

$$\begin{aligned}&\|\text{prox}_{\tau G}(x) - \text{prox}_{\tau G}(y)\|^2 \leq \|x - y\|^2 \\ &- \|\left(x - \text{prox}_{\tau G}(x)\right) - \left(y - \text{prox}_{\tau G}(y)\right)\|^2, \ \forall x, \ y \in H_2,\end{aligned}$$

and its fixed point set is the set of minimizers of G. Let C be a nonempty closed and convex subset of H_1, and then, the orthogonal projection of x onto C is defined by

$$P_C x := \arg\min \{\|x - y\| \mid y \in C\}, \ \forall x \in H_1.$$

Definition 2.1 Let $h : H_1 \to H_1$ be a mapping, and then

(i) h is called *nonexpansive* if

$$\|hx - hy\| \leq \|x - y\|, \ \forall x, \ y \in H_1.$$

(ii) h is said to be *firmly nonexpansive* if

$$\langle hx - hy, x - y\rangle \geq \|hx - hy\|^2, \ \forall x, \ y \in H_1.$$

(iii) Let $D \subset H_1$ be a set and let $h : D \to \mathbb{R} \cup \{+\infty\}$ be named *weak lower semi-continuity* if $x_n \rightharpoonup x$, the following statement holds:

$$\liminf_{n\to\infty} h(x_n) \geq h(x).$$

Lemma 2.1 *Let $\nu \in]0, 1[$, for all $x, y, z \in H_1$, and then*

(i) $\|\nu x + (1 - \nu)y\|^2 = \nu\|x\|^2 + (1 - \nu)\|y\|^2 - \nu(1 - \nu)\|x - y\|^2$;
(ii) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y\rangle$;
(iii) $\langle x - y, x - z\rangle = \frac{1}{2}\|x - y\|^2 + \frac{1}{2}\|x - z\|^2 - \frac{1}{2}\|y - z\|^2$.

Lemma 2.2 *Goebel and Reich (1984) Let $C \subset H_1$ be a nonempty closed convex set and let P_C be the metric projection from H_1 to C. Then, the following statements hold:*

(i) $\langle x - P_C x, y - P_C x\rangle \leq 0$ *for all* $x \in H_1$ *and* $y \in C$;
(ii) $\|P_C x - P_C y\| \leq \|x - y\|$ *for all* $x, \ y \in H_1$.

Lemma 2.3 *Saejung and Yotkaew (2012) Suppose that $\{s_n\}_{n=1}^{\infty}$ is a sequence of nonnegative real numbers, such that*

$$s_{n+1} \leq (1 - \alpha_n)s_n + \alpha_n\delta_n, \ \forall n \geq 1,$$

where

(i) $\{\alpha_n\}_{n=1}^{\infty} \subset]0, 1[$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,
(ii) *if* $\limsup_{k\to\infty} \delta_{n_k} \leq 0$ *for every subsequence* $\{s_{n_k}\}$ *of* $\{s_n\}$ *fulfilling* $\liminf_{k\to\infty} \|s_{n_k+1} - s_{n_k}\| \geq 0$.

Then, $\lim_{n\to\infty} s_n = 0$.

3 Strong convergence

In this section, we first let H_1 and H_2 be real Hilbert spaces and $A : H_1 \to H_2$ be a bounded linear operator with its adjoint A^*, $F : H_1 \to \mathbb{R} \cup \{+\infty\}$ and $G : H_2 \to \mathbb{R} \cup \{+\infty\}$ be proper, convex, and lower semi-continuous (lsc) functions. Now, we consider the following proximal SFP:

$$\text{Find a solution } z^* \in H_1 \text{ such that } \min\{F(x) + G_\tau(Ax) : x \in H_1\},$$

where $\tau > 0$ and $G_\tau(x) := \min_{y \in H_2} \left\{G(y) + \frac{1}{2\tau}\|y - x\|^2\right\}$ can be regarded as the Moreau–Yosida approximate of the function G of parameter τ. If such point exists, then its solutions set is denoted by Γ. Similar as in Abbas et al. (2018); Shehu and Iyiola (2017), we also offer the following definitions used for the rest of the paper.

Given any $\tau > 0$ and $x \in H_1$, we define

$$\begin{aligned} E(x) &= \tfrac{1}{2}\|(I - \text{prox}_{\tau G})Ax\|^2; \\ L(x) &= \tfrac{1}{2}\|(I - \text{prox}_{\tau F})x\|^2 \text{ and} \\ \theta(x) &= \sqrt{\|\nabla E(x) + \nabla L(x)\|^2}. \end{aligned}$$

Then, the Lipschitz gradients ∇E and ∇L of E and L, respectively, are

$$\begin{aligned} \nabla E(x) &= A^*(I - \text{prox}_{\tau G})Ax \text{ and} \\ \nabla L(x) &= (I - \text{prox}_{\tau F})x, \end{aligned}$$

whose Lipschitz constants are $\|A\|^2$ and 1, respectively. Before describing our algorithm, the following conditions are required in convergence analysis.

(A1) The solution set of the proximal SFP is nonempty, that is, $\Gamma \neq \emptyset$.
(A2) Let $\{\tilde{\tau}_n\} \subset [0, \tilde{\theta}[$ with $\tilde{\theta} > 0$ be a positive sequence, such that $\tilde{\tau}_n = o(\gamma_n)$, i.e., $\lim_{n\to\infty} \frac{\tilde{\tau}_n}{\gamma_n} = 0$ where the sequence $\{\gamma_n\} \subset]0, 1[$ fulfills $\sum_{n=1}^{\infty} \gamma_n = \infty$ and $\lim_{n\to\infty} \gamma_n = 0$.
(A3) The mapping $f : H_1 \to H_1$ is $\tilde{\rho}$- contractive with constant $\tilde{\rho} \in [0, 1[$.
(A4) $\inf \kappa_n(2 - \kappa_n) > 0$.

Below, our iterative scheme is stated in Algorithm 1.

Algorithm 1

Step 0. Take $x_0, x_1 \in H_1$ and $0 < \kappa_n < 2$. Choose a sequence $\{\sigma_n\} \subset [0, \sigma[\subset [0, 1[$ and $0 < \gamma_n < 1$.
Step 1. Given x_{n-1}, x_n $(n \geq 1)$ and compute

$$\begin{aligned} &w_n = x_n + \sigma_n\left(x_n - x_{n-1}\right), \; y_n = w_n - \lambda_n(\nabla E(w_n) + \nabla L(w_n)), \\ &x_{n+1} = \gamma_n f(x_n) + (1 - \gamma_n)y_n, \end{aligned} \tag{3.1}$$

where the stepsize λ_n is updated via

$$\lambda_n = \kappa_n \frac{E(w_n) + L(w_n)}{\theta(w_n)^2}, \tag{3.2}$$

where $\theta(w_n) = \sqrt{\|\nabla E(w_n) + \nabla L(w_n)\|^2}$.

Step 2. If $\nabla E(w_n) = \nabla L(w_n) = 0$ and $w_n = x_n$, then stop, x_n is a solution of the proximal SFP. Otherwise, set $n := n + 1$ and return to Step 1.

Remark 3.1 In Algorithm 1, if $\Gamma \neq \emptyset$ and $\nabla E(w_n) = \nabla L(w_n) = 0$ and $w_n = x_n$, then $x_n \in \Gamma$.

Proof Since if $A^*(I - \text{prox}_{\tau G})Aw_n = (I - \text{prox}_{\tau F})w_n = 0$, then this shows that $w_n \in \Gamma$. Additionally, $A^*(I - \text{prox}_{\tau G})Aw_n = (I - \text{prox}_{\tau F})w_n = 0$ yields from (3.1) of Algorithm 1 that $y_n = w_n$. This, together with $w_n = x_n$, implies that $y_n = x_n \in \Gamma$. Thus, $x_n \in \Gamma$. □

Remark 3.2 In Algorithm 1, the inertial parameter σ_n is chosen as

$$\sigma_n = \begin{cases} \min\left\{\frac{\tilde{\tau}_n}{\|x_n - x_{n-1}\|}, \sigma\right\} & \text{if } x_n \neq x_{n-1}, \\ \sigma, & \text{otherwise.} \end{cases} \tag{3.3}$$

In what follows, the proof of the following lemma and main theorem does not involve firm-nonexpasiveness of the operator $I - \text{prox}_{\tau(\cdot)}$.

Theorem 3.1 *Suppose that Conditions* $(A1) - (A4)$ *hold. The sequence* $\{x_n\}$ *generated by Algorithm 1 converges strongly to a point* $z \in \Gamma$, *where* $z = P_\Gamma o f(z)$.

Proof Let $z \in \Gamma$. Since $\text{prox}_{\tau(\cdot)}$ is nonexpansive, z solves the proximal SFP due to minimizers of any function are exactly fixed points of its proximal mapping, and furthermore, by Lemma 2.1 (iii), we derive

$$\begin{aligned}
&\langle w_n - z, -\nabla E(w_n)\rangle \\
&= \langle w_n - z, A^*(\text{prox}_{\tau G} - I)Aw_n\rangle \\
&= \langle Aw_n - Az, (\text{prox}_{\tau G} - I)Aw_n\rangle \\
&= \langle Aw_n - \text{prox}_{\tau G}Aw_n + \text{prox}_{\tau G}Aw_n - Az, (\text{prox}_{\tau G} - I)Aw_n\rangle \\
&= \langle \text{prox}_{\tau G}Aw_n - Az, \text{prox}_{\tau G}Aw_n - Aw_n\rangle - \|\text{prox}_{\tau G}Aw_n - Aw_n\|^2 \\
&= \frac{1}{2}\left(\|\text{prox}_{\tau G}Aw_n - Az\|^2 + \|\text{prox}_{\tau G}Aw_n - Aw_n\|^2 - \|Aw_n - Az\|^2\right) \\
&\quad -\|\text{prox}_{\tau G}Aw_n - Aw_n\|^2 \\
&\le -\frac{1}{2}\|\text{prox}_{\tau G}Aw_n - Aw_n\|^2 \\
&= -E(w_n),
\end{aligned}$$

and

$$\begin{aligned}
\langle w_n - z, -\nabla L(w_n)\rangle &= \langle w_n - z, (\text{prox}_{\tau F} - I)w_n\rangle \\
&= \langle w_n - \text{prox}_{\tau F}w_n + \text{prox}_{\tau F}w_n - z, (\text{prox}_{\tau F} - I)w_n\rangle \\
&= \langle \text{prox}_{\tau F}w_n - z, \text{prox}_{\tau F}w_n - w_n\rangle - \|\text{prox}_{\tau F}w_n - w_n\|^2 \\
&= \frac{1}{2}\left(\|\text{prox}_{\tau F}w_n - z\|^2 + \|\text{prox}_{\tau F}w_n - w_n\|^2 - \|w_n - z\|^2\right) \\
&\quad -\|\text{prox}_{\tau F}w_n - w_n\|^2 \\
&\le -\frac{1}{2}\|\text{prox}_{\tau F}w_n - w_n\|^2 \\
&= -L(w_n);
\end{aligned}$$

combining with (3.1) and (3.2) yields that

$$\begin{aligned}\|y_n - z\|^2 &= \|w_n - \lambda_n(\nabla E(w_n) + \nabla L(w_n)) - z\|^2\\ &= \|w_n - z\|^2 + \lambda_n^2\|\nabla E(w_n) + \nabla L(w_n)\|^2\\ &\quad +2\lambda_n\langle w_n - z, -(\nabla E(w_n) + \nabla L(w_n))\rangle\\ &= \|w_n - z\|^2 + \lambda_n^2\|\nabla E(w_n) + \nabla L(w_n)\|^2\\ &\quad +2\lambda_n\langle w_n - z, -\nabla E(w_n)\rangle + 2\lambda_n\langle w_n - z, -\nabla L(w_n)\rangle\\ &\le \|w_n - z\|^2 + \lambda_n^2\|\nabla E(w_n) + \nabla L(w_n)\|^2 - 2\lambda_n(E(w_n) + L(w_n))\\ &= \|w_n - z\|^2 + \kappa_n^2\frac{(E(w_n) + L(w_n))^2}{\|\nabla E(w_n) + \nabla L(w_n)\|^4}(\|\nabla E(w_n) + \nabla L(w_n)\|^2)\\ &\quad -2\kappa_n\frac{(E(w_n) + L(w_n))^2}{\|\nabla E(w_n) + \nabla L(w_n)\|^2}\\ &= \|w_n - z\|^2 + \kappa_n(\kappa_n - 2)\frac{(E(w_n) + L(w_n))^2}{\|\nabla E(w_n) + \nabla L(w_n)\|^2}.\end{aligned} \tag{3.4}$$

After arrangement, we have

$$\|y_n - z\|^2 \le \|w_n - z\|^2 + \kappa_n(\kappa_n - 2)\frac{(E(w_n) + L(w_n))^2}{\|\nabla E(w_n) + \nabla L(w_n)\|^2}. \tag{3.5}$$

From (A4) and (3.5), we have

$$\|y_n - z\| \le \|w_n - z\|. \tag{3.6}$$

By the definition of w_n, we get

$$\begin{aligned}\|w_n - z\| &= \|x_n + \sigma_n(x_n - x_{n-1}) - z\|\\ &\le \|x_n - z\| + \sigma_n\|x_n - x_{n-1}\|\\ &= \|x_n - z\| + \gamma_n \cdot \frac{\sigma_n}{\gamma_n}\|x_n - x_{n-1}\|.\end{aligned} \tag{3.7}$$

According to (3.3), we have $\sigma_n\|x_n - x_{n-1}\| \le \tilde{\tau}_n\ \forall n \ge 1$, which, together with $\lim\limits_{n\to\infty}\frac{\tilde{\tau}_n}{\gamma_n} = 0$, yields that

$$\lim_{n\to\infty}\frac{\sigma_n}{\gamma_n}\|x_n - x_{n-1}\| \le \lim_{n\to\infty}\frac{\tilde{\tau}_n}{\gamma_n} = 0.$$

Therefore, there is a constant $M_1 > 0$, such that

$$\frac{\sigma_n}{\gamma_n}\|x_n - x_{n-1}\| \le M_1,\ \ \forall n \ge 1,$$

which, along with (3.6) and (3.7), yields that

$$\|y_n - z\| \le \|w_n - z\| \le \|x_n - z\| + \gamma_n M_1. \tag{3.8}$$

From (3.1) and (3.8), it follows that:

$$\begin{aligned}
\|x_{n+1}-z\| &= \|\gamma_n f(x_n)+(1-\gamma_n)y_n - z\| \\
&= \|\gamma_n (f(x_n)-z)+(1-\gamma_n)(y_n - z)\| \\
&\le \gamma_n \|f(x_n)-z\|+(1-\gamma_n)\|y_n - z\| \\
&= \gamma_n \|f(x_n)-f(z)+f(z)-z\|+(1-\gamma_n)\|y_n - z\| \\
&\le \gamma_n \tilde{\rho}\|x_n-z\|+\gamma_n\|f(z)-z\|+(1-\gamma_n)\|y_n - z\| \\
&\le \gamma_n \tilde{\rho}\|x_n-z\|+(1-\gamma_n)(\|x_n-z\|+\gamma_n M_1\|)+\gamma_n\|f(z)-z\| \\
&\le (1-\gamma_n(1-\tilde{\rho}))\|x_n-z\|+\gamma_n(1-\tilde{\rho})\frac{\|f(z)-z\|+M_1}{1-\tilde{\rho}} \\
&\le \max\left\{\|x_n-z\|, \frac{M_1+\|f(z)-z\|}{1-\tilde{\rho}}\right\} \\
&\le \cdots \le \max\left\{\|x_1-z\|, \frac{M_1+\|f(z)-z\|}{1-\tilde{\rho}}\right\}.
\end{aligned}$$

This means that the sequence $\{x_n\}$ is bounded. Hence, the sequences $\{y_n\}$, $\{f(x_n)\}$ and $\{w_n\}$ are also bounded.

By (3.1) and the convexity of $\|\cdot\|^2$, we get that

$$\begin{aligned}
\|x_{n+1}-z\|^2 &= \|\gamma_n f(x_n)+(1-\gamma_n)y_n - z\|^2 \\
&= \|\gamma_n (f(x_n)-z)+(1-\gamma_n)(y_n - z)\|^2 \\
&\le \gamma_n \|f(x_n)-z\|^2+(1-\gamma_n)\|y_n - z\|^2 \\
&\le \gamma_n (\|f(x_n)-f(z)\|+\|f(z)-z\|)^2+(1-\gamma_n)\|y_n - z\|^2 \\
&\le \gamma_n (\tilde{\rho}\|x_n-z\|+\|f(z)-z\|)^2+(1-\gamma_n)\|y_n - z\|^2 \\
&\le \gamma_n (\|x_n-z\|+\|f(z)-z\|)^2+(1-\gamma_n)\|y_n - z\|^2 \\
&= \gamma_n \|x_n-z\|^2+\gamma_n\left(\|f(z)-z\|^2+2\|x_n-z\|\|f(z)-z\|\right) \\
&\quad +(1-\gamma_n)\|y_n - z\|^2 \\
&\le \gamma_n \|x_n-z\|^2+(1-\gamma_n)\|y_n - z\|^2+\gamma_n M_2;
\end{aligned}$$

for some $M_2 > 0$. Combining with (3.5), we derive that

$$\begin{aligned}
\|x_{n+1}-z\|^2 \le \gamma_n \|x_n-z\|^2 &+ (1-\gamma_n)\|w_n - z\|^2 \\
&+(1-\gamma_n)\kappa_n(\kappa_n-2)\frac{(E(w_n)+L(w_n))^2}{\|\nabla E(w_n)+\nabla L(w_n)\|^2}+\gamma_n M_2. \qquad (3.9)
\end{aligned}$$

Substituting (3.8) into (3.9), then there exists $M_3 > 0$, such that

$$\begin{aligned}
\|x_{n+1}-z\|^2 &\le \gamma_n \|x_n-z\|^2+(1-\gamma_n)(\|x_n-z\|+\gamma_n M_1)^2 \\
&\quad +(1-\gamma_n)\kappa_n(\kappa_n-2)\frac{(E(w_n)+L(w_n))^2}{\|\nabla E(w_n)+\nabla L(w_n)\|^2}+\gamma_n M_2 \\
&= \gamma_n \|x_n-z\|^2+(1-\gamma_n)\|x_n-z\|^2+(1-\gamma_n)(\gamma_n M_1)^2 \\
&\quad +2(1-\gamma_n)\gamma_n M_1\|x_n-z\|+(1-\gamma_n)\kappa_n(\kappa_n-2) \\
&\quad \times\frac{(E(w_n)+L(w_n))^2}{\|\nabla E(w_n)+\nabla L(w_n)\|^2}+\gamma_n M_2 \\
&\le \|x_n-z\|^2+(1-\gamma_n)\kappa_n(\kappa_n-2)\frac{(E(w_n)+L(w_n))^2}{\|\nabla E(w_n)+\nabla L(w_n)\|^2}+\gamma_n M_3.
\end{aligned}$$

That is

$$(1-\gamma_n)\kappa_n(2-\kappa_n)\frac{(E(w_n)+L(w_n))^2}{\|\nabla E(w_n)+\nabla L(w_n)\|^2} \le \|x_n-z\|^2-\|x_{n+1}-z\|^2+\gamma_n M_3. \tag{3.10}$$

By Lemma 2.1 (i, ii) and (3.8), we derive that

$$\begin{aligned}
\|x_{n+1}-z\|^2 &= \|\gamma_n f(x_n)+(1-\gamma_n)y_n-z\|^2 \\
&= \|\gamma_n(f(x_n)-f(z))+(1-\gamma_n)(y_n-z)+\gamma_n(f(z)-z)\|^2 \\
&\le \|\gamma_n(f(x_n)-f(z))+(1-\gamma_n)(y_n-z)\|^2+2\gamma_n\langle f(z)-z, x_{n+1}-z\rangle \\
&\le \gamma_n\|f(x_n)-f(z)\|^2+(1-\gamma_n)\|y_n-z\|^2+2\gamma_n\langle f(z)-z, x_{n+1}-z\rangle \\
&\le \gamma_n\tilde{\rho}^2\|x_n-z\|^2+(1-\gamma_n)\|y_n-z\|^2+2\gamma_n\langle f(z)-z, x_{n+1}-z\rangle \\
&\le \gamma_n\tilde{\rho}\|x_n-z\|^2+(1-\gamma_n)\|w_n-z\|^2+2\gamma_n\langle f(z)-z, x_{n+1}-z\rangle.
\end{aligned} \tag{3.11}$$

From the definition of w_n, it follows that:

$$\begin{aligned}
\|w_n-z\|^2 &= \|x_n+\sigma_n(x_n-x_{n-1})-z\|^2 \\
&= \|x_n-z\|^2+\sigma_n^2\|x_n-x_{n-1}\|^2+2\sigma_n\langle x_n-z, x_n-x_{n-1}\rangle \\
&\le \|x_n-z\|^2+\sigma_n^2\|x_n-x_{n-1}\|^2+2\sigma_n\|x_n-z\|\|x_n-x_{n-1}\|.
\end{aligned} \tag{3.12}$$

Let $M=\sup\limits_{n\ge 1}\{\sigma\|x_n-x_{n-1}\|, 2\|x_n-z\|\}$. Combining (3.11) and (3.12), we obtain that

$$\begin{aligned}
\|x_{n+1}-z\|^2 &\le (1-\gamma_n\,(1-\tilde{\rho}))\,\|x_n-z\|^2+\sigma_n^2\|x_n-x_{n-1}\|^2 \\
&\quad +2\gamma_n\langle f(z)-z, x_{n+1}-z\rangle+2\sigma_n\|x_n-z\|\|x_n-x_{n-1}\| \\
&= (1-\gamma_n\,(1-\tilde{\rho}))\,\|x_n-z\|^2+(1-\tilde{\rho})\tfrac{2}{1-\tilde{\rho}}\gamma_n\langle f(z)-z, x_{n+1}-z\rangle \\
&\quad +\sigma_n\|x_n-x_{n-1}\|(\sigma_n\|x_n-x_{n-1}\|+2\|x_n-z\|) \\
&\le (1-\gamma_n(1-\tilde{\rho}))\|x_n-z\|^2+(1-\tilde{\rho})\tfrac{2}{1-\tilde{\rho}}\gamma_n\langle f(z)-z, x_{n+1}-z\rangle \\
&\quad +2M\sigma_n\|x_n-x_{n-1}\| \\
&= (1-\gamma_n(1-\tilde{\rho}))\|x_n-z\|^2 \\
&\quad +(1-\tilde{\rho})\gamma_n\left(\tfrac{2}{1-\tilde{\rho}}\langle f(z)-z, x_{n+1}-z\rangle+\tfrac{2M\sigma_n}{(1-\tilde{\rho})\gamma_n}\|x_n-x_{n-1}\|\right).
\end{aligned}$$

After arrangement, there exists $M>0$, such that

$$\begin{aligned}
\|x_{n+1}-z\|^2 &\le (1-(1-\tilde{\rho})\gamma_n)\|x_n-z\|^2 \\
&\quad +(1-\tilde{\rho})\gamma_n\left(\tfrac{2}{1-\tilde{\rho}}\langle f(z)-z, x_{n+1}-z\rangle+\tfrac{2M\sigma_n}{(1-\tilde{\rho})\gamma_n}\|x_n-x_{n-1}\|\right).
\end{aligned} \tag{3.13}$$

Next, we let

$$\begin{aligned}
s_n &= \|x_n-z\|^2, \\
\alpha_n &= (1-\tilde{\rho})\gamma_n, \\
\delta_n &= \tfrac{2}{1-\tilde{\rho}}\langle f(z)-z, x_{n+1}-z\rangle+\tfrac{2M\sigma_n}{(1-\tilde{\rho})\gamma_n}\|x_n-x_{n-1}\|.
\end{aligned}$$

Then, (3.13) reduces to the following inequality:

$$s_{n+1}\le(1-\alpha_n)s_n+\alpha_n\delta_n,\ \forall\, n\ge 1.$$

Clearly, Lemma 2.3 (i) is satisfied. Now, it needs to verify that Lemma 2.3 (ii) is also satisfied. Suppose that $\{\|x_{n_k}-z\|\}$ is the subsequence of $\{\|x_n-z\|\}$ and satisfies $\liminf_{k\to\infty}(\|x_{n_k+1}-$

$z\| - \|x_{n_k} - z\|) \geq 0$. Then

$$\begin{aligned}&\liminf_{k\to\infty} (\|x_{n_k+1} - z\|^2 - \|x_{n_k} - z\|^2)\\&\quad = \liminf_{k\to\infty} \big(\big(\|x_{n_k+1} - z\| - \|x_{n_k} - z\|\big)\,(\|x_{n_k+1} - z\| + \|x_{n_k} - z\|)\big) \geq 0.\end{aligned} \tag{3.14}$$

By $\lim\limits_{k\to\infty} \gamma_{n_k} = 0$, (3.10) and (3.14), one has

$$\begin{aligned}&\limsup_{k\to\infty} (1-\gamma_{n_k})\kappa_{n_k}(2-\kappa_{n_k})\tfrac{(E(w_{n_k})+L(w_{n_k}))^2}{\|\nabla E(w_{n_k})+\nabla L(w_{n_k})\|^2}\\&\qquad \leq \limsup_{k\to\infty} (\|x_{n_k+1} - z\|^2 - \|x_{n_k} - z\|^2 + \gamma_{n_k} M_3)\\&\qquad \leq \limsup_{k\to\infty} (\|x_{n_k+1} - z\|^2 - \|x_{n_k} - z\|^2) + \limsup_{k\to\infty} \gamma_{n_k} M_3\\&\qquad = -\liminf_{k\to\infty} (\|x_{n_k+1} - z\|^2 - \|x_{n_k} - z\|^2) \leq 0.\end{aligned}$$

Now, we have

$$\lim_{k\to\infty}\left(\kappa_{n_k}(2-\kappa_{n_k})\frac{(E(w_{n_k})+L(w_{n_k}))^2}{\|\nabla E(w_{n_k})+\nabla L(w_{n_k})\|^2}\right) = 0.$$

Thus, we get

$$\lim_{k\to\infty}\frac{(E(w_{n_k})+L(w_{n_k}))^2}{\|\nabla E(w_{n_k})+\nabla L(w_{n_k})\|^2} = 0.$$

As a result, we have

$$\lim_{k\to\infty}(E(w_{n_k})+L(w_{n_k})) = 0 \Leftrightarrow \lim_{k\to\infty} E(w_{n_k}) = 0 \text{ and } \lim_{k\to\infty} L(w_{n_k}) = 0. \tag{3.15}$$

Since $\theta^2_{n_k} = \|\nabla E(w_{n_k}) + \nabla L(w_{n_k})\|^2$ is bounded. This follows from the fact that ∇E is Lipschitz continuous with constant $\|A\|^2$, ∇L is nonexpansive and $\{w_{n_k}\}$ is bounded. More precisely, for any z^* which solves the proximal SFP, we have

$$\|\nabla E(w_{n_k})\| = \|\nabla E(w_{n_k}) - \nabla E(z^*)\| \leq \|A\|^2\|w_{n_k} - z^*\|$$

and

$$\|\nabla L(w_{n_k})\| = \|\nabla L(w_{n_k}) - \nabla L(z^*)\| \leq \|w_{n_k} - z^*\|.$$

By (3.1) and (3.2), we get

$$\begin{aligned}\|y_{n_k} - w_{n_k}\|^2 &= \|\lambda_{n_k}(\nabla E(w_{n_k}) + \nabla L(w_{n_k}))\|^2\\&= \frac{\kappa^2_{n_k}(E(w_{n_k})+L(w_{n_k}))^2}{\|\nabla E(w_{n_k})+\nabla L(w_{n_k})\|^2}\\&\leq \frac{4(E(w_{n_k})+L(w_{n_k}))^2}{\|\nabla E(w_{n_k})+\nabla L(w_{n_k})\|^2}\\&\to 0, \text{ as } k \to \infty.\end{aligned}$$

This shows that

$$\lim_{k\to\infty}\|y_{n_k} - w_{n_k}\| = 0. \tag{3.16}$$

From $\lim_{k\to\infty}\gamma_{n_k}=0$ and (3.1), it follows that:

$$\|x_{n_k}-w_{n_k}\|=\sigma_{n_k}\|x_{n_k}-x_{n_k-1}\|=\gamma_{n_k}\frac{\sigma_{n_k}}{\gamma_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0,\ \text{as } k\to\infty. \tag{3.17}$$

Using (3.16) and (3.17), we have

$$\lim_{k\to\infty}\|y_{n_k}-x_{n_k}\|=0. \tag{3.18}$$

By (3.1) and $\lim_{k\to\infty}\gamma_{n_k}=0$, we have

$$\|x_{n_k+1}-y_{n_k}\|=\gamma_{n_k}\|f(x_{n_k})-y_{n_k}\|\to 0,\ \text{as } k\to\infty.$$

This deduces that

$$\left\|x_{n_k+1}-x_{n_k}\right\|\le\|x_{n_k+1}-y_{n_k}\|+\|y_{n_k}-x_{n_k}\|\to 0,\ \text{as } k\to\infty. \tag{3.19}$$

Since the sequence $\{x_{n_k}\}$ is bounded, then there exists a subsequence $\left\{x_{n_{k_i}}\right\}$ of $\{x_{n_k}\}$ converging weakly to a point $z^*\in H_1$, such that

$$\limsup_{k\to\infty}\left\langle f(z)-z,x_{n_k}-z\right\rangle=\lim_{i\to\infty}\left\langle f(z)-z,x_{n_{k_i}}-z\right\rangle=\langle f(z)-z,z^*-z\rangle.$$

Thanks to (3.17), we have

$$w_{n_{k_i}}\rightharpoonup z^*,\ \text{as } i\to\infty.$$

By the weak lower semi-continuity of E, we arrive at

$$0\le E(z^*)\le\liminf_{i\to\infty}E\left(w_{n_{k_i}}\right)=\lim_{k\to\infty}E\left(w_{n_k}\right)=0.$$

This means that $E(z^*)=\frac{1}{2}\|(I-\text{prox}_{\tau G})Az^*\|^2=0$, that is, 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. Similarly, by the weak lower semi-continuity of L, we have $0\le L(z^*)\le\liminf_{i\to\infty}L\left(w_{n_{k_i}}\right)=\lim_{k\to\infty}L\left(w_{n_k}\right)=0$. This means that $L(z^*)=\frac{1}{2}\|(I-\text{prox}_{\lambda_n\tau F})z^*\|^2=0$, that is, 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. Therefore, $z^*\in\Gamma$. From the definition of $z=P_\Gamma of(z)$, it yields that

$$\limsup_{k\to\infty}\langle f(z)-z,x_{n_k}-z\rangle=\lim_{i\to\infty}\left\langle f(z)-z,x_{n_{k_i}}-z\right\rangle=\langle f(z)-z,z^*-z\rangle\le 0,$$

which together with (3.19) implies that

$$\begin{aligned}&\limsup_{k\to\infty}\langle f(z)-z,x_{n_k+1}-z\rangle\\&\quad\le\limsup_{k\to\infty}\langle f(z)-z,x_{n_{k+1}}-x_{n_k}\rangle+\limsup_{k\to\infty}\langle f(z)-z,x_{n_k}-z\rangle\\&\quad\le 0.\end{aligned}$$

Hence

$$\limsup_{k\to\infty}\delta_{n_k}=\limsup_{k\to\infty}\left\{\frac{2}{1-\tilde{\rho}}\langle f(z)-z,x_{n_k+1}-z\rangle+\frac{2M\sigma_{n_k}}{(1-\tilde{\rho})\gamma_{n_k}}\|x_{n_k}-x_{n_k-1}\|\right\}\le 0.$$

Employing Lemma 2.3, we conclude that $\lim_{n\to\infty}\|x_n-z\|=0$. □

If $F \equiv \delta_C$ [defined as $\delta_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise] and $G \equiv \delta_Q$, the indicator functions of the nonempty, closed, and convex sets $C \subset H_1$ and $Q \subset H_2$, respectively, then the proximal SFP reduces to the following SFP:

$$\text{Find } z^* \in C \text{ such that } Az^* \in Q.$$

Furthermore, we derive the following strongly convergent corollary from Theorem 3.1.

Corollary 3.1 *Let H_1, H_2, C, Q, A, A^*, and Γ be the same as above description. Suppose that $\Gamma \neq \emptyset$, $\{\sigma_n\} \subset [0, \sigma[\subset [0, 1[$ and Conditions (A1)–(A4) hold. Let $x_0,\ x_1 \in H_1$ and $\{x_n\}$ be a sequence generated by*

$$\begin{cases} w_n = x_n + \sigma_n(x_n - x_{n-1}),\ y_n = w_n - \lambda_n(\nabla E(w_n) + \nabla L(w_n)), \\ x_{n+1} = \gamma_n f(x_n) + (1 - \gamma_n)y_n, \end{cases}$$

where σ_n is defined in (3.3) and the stepsize λ_n can be computed via

$$\lambda_n = \kappa_n \frac{E(w_n) + L(w_n)}{\theta(w_n)^2},$$

where $0 < \kappa_n < 2$, $L(w_n) = \frac{1}{2}\|(I - P_C)w_n\|^2$, $E(w_n) = \frac{1}{2}\|(I - P_Q)Aw_n\|^2$ and $\theta(w_n) = \sqrt{\|\nabla E(w_n) + \nabla L(w_n)\|^2}$.

Then, the iterative sequence $\{x_n\}$ produced above strongly converges to $z \in \Gamma$, where $z = P_\Gamma o f(z)$.

Remark 3.3 Theorem 3.1 improves the result of [Shehu and Iyiola (2017), Theorem 3.2], because strong convergence of our method is obtained without assuming firm-nonexpasiveness of the operator $I - \text{prox}_{\tau(\cdot)}$.

4 Numerical experiments

In this section, we provide numerical experiments relative to the proximal SFP. For the first example, we compare Alg. 1 with Abbas et al.'s Algorithms 3.1-3.2 (shortly, AMMOAlg. 3.1-3.2) Abbas et al. (2018), Shehu et al.'s Algorithm 3.1 (SIAlg. 3.1) Shehu and Iyiola (2017), and Shehu et al.'s Algorithm (AHVSDM) Shehu and Iyiola (2018). All the programs are implemented in MATLAB R2017a on a PC Desktop Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHZ computer with RAM 8.00 GB.

In the first example, we study the proximal SFP in the case $\arg\min F \cap A^{-1}(\arg\min G) \neq \emptyset$, or in other words: in finding a minimizer z^* of F, such that Az^* minimizes G, that is

$$\text{Find } z^* \in H_1 \text{ such that } z^* \in \arg\min_{x\in H_1}\ F(x) \text{ and } Az^* \in \arg\min_{y\in H_2}\ G(y), \tag{4.1}$$

where $F : H_1 \to \mathbb{R}$ and $G : H_2 \to \mathbb{R}$ are proper and lower semi-continuous convex (lsc) functions, $\arg\min\ F = \{z^* \in H_1 : F(z^*) \leq F(x)\ \forall x \in H_1\}$ and $\arg\min\ G = \{y^* \in H_2 : G(y^*) \leq G(x)\ \forall x \in H_2\}$, the solution set is denoted by Γ.

Example 4.1 Kesornprom and Cholamjiak (2019) Let $H_1 = H_2 = \mathbb{R}^N$ and $F(x) = \frac{1}{2}d_C^2(x)$, where $C \subset \mathbb{R}^N$ is a unit ball and $G(x) = \frac{1}{2}\|x\|^2$. Set $Ax = x$, $x \in \mathbb{R}^N$. Observe that $0 \in \Gamma$ and $\Gamma \neq \emptyset$. For AMOOAlg. 3.1-3.2 and SIAlg. 3.1, we take $\kappa_n = 1.9$, $\gamma_n = \frac{1}{n+1}$ and $\alpha_n = \frac{1}{10^4(n+1)}$. For AHVSDM, we set $\lambda_n = 10^{-4}$, $\gamma_n = \frac{1.99}{n+1}$, $\mu = 1$, $\tilde{F} = I$ (which

Table 1 Results for Example 4.1

N	ϵ	AMOOAlg. 3.1		AMOOAlg. 3.2		SIAlg. 3.1		AHVSDM		Alg. 1	
		Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time
6000	10^{-5}	167	0.0074	15	9.6778e-04	44	0.0022	111	0.0034	7	6.2313e-04
	10^{-7}	264	0.0123	17	0.0012	2831	0.1568	1113	0.0320	7	7.1777e-04
8000	10^{-5}	169	0.0104	15	0.0015	49	0.0032	119	0.0043	5	6.4476e-04
	10^{-7}	266	0.0189	17	0.0480	3265	0.2216	1192	0.0460	7	9.8156e-04

is a contraction mapping in Shehu and Iyiola (2018) and I is an identity mapping on H_1) and $\beta_n = \frac{0.001}{(n+1)^2}$. For Alg. 1, we adopt $\kappa_n = 1.9$, $\gamma_n = \frac{1}{n^2}$, $\sigma = 0.3$ and $\tilde{\tau}_n = \frac{1}{n^3}$. For all tests, we use the condition $\|x_n - \text{prox}_{\tau,F}(x_n)\| + \|Ax_n - \text{prox}_{\tau,G}(Ax_n)\| < \epsilon$ to terminate all the algorithms and choose $x_0 = [0, 0, 0, \cdots, 0]$ and $x_1 = [1, \cdots, 1] \in \mathbb{R}^N$ and $\tau = 5$. To ensure that all algorithms have a common convergence point in this experiment, we set $f(x) = 0$. The results are summarized in Table 1.

Remark 4.1 The numerical results of Example 4.1 are described in Table 1, the observations we obtain are the following:

(1) The iterative rule proposed in this note implements efficiently and readily. More significantly, it converges fast.
(2) Our proposed algorithm converges faster than some existing algorithms in terms of the number of iterations and execution time under different dimensions of the problem.

For the SFP, we list the following numerical examples and compare Alg. 1 with Gibali et al.'s Algorithm 3.1 (shortly, GMV Alg. 3.1) Gibali et al. (2019) and Suantai et al.'s Algorithm 3.1(SKC Alg. 3.1) Suantai et al. (2018).

Example 4.2 Kesornprom and Cholamjiak (2019) Let $H_1 = H_2 = L^2([0, 1])$ with norm $\|x\|_{L^2} = \left(\int_0^1 x(t)^2 dt\right)^{\frac{1}{2}}$ and inner product $\langle x, y\rangle = \int_0^1 x(t)y(t)dt,\ x, y \in L^2([0, 1])$. Let $C = \{x \in L^2([0, 1]) : \|x\|_{L^2} \leq 1\}$ and $Q = \{x \in L^2([0, 1]) : \langle x, t\rangle = 0\}$. Set $Ax(t) = \frac{x(t)}{2}$. Observe that $0 \in \Gamma$, and so, $\Gamma \neq \emptyset$. For SKC Alg. 3.1 and GMV Alg. 3.1, we fix $\sigma = 0.3$, $\tilde{\tau}_n = \frac{1}{n^2}$, $\gamma_n = \frac{1}{10^4 n}$, $\kappa_n = 1.6$, $f(x) = 0.01x$, $\beta_n = 0.7$ and $\sigma_n = \max(0, \sigma_n - 0.1)$. For Alg. 1, we choose $\kappa_n = 1.6$, $\sigma = 0.3$, $f(x) = 0.01x$, $\gamma_n = \frac{1}{10^4 n}$,$\tilde{\tau}_n = \frac{1}{n^2}$. For López Alg. 5.1, we adopt $\kappa_n = 9 \times 10^{-5}$ and $\gamma_n = \frac{10^{-5}}{n}$. For all algorithms, we regard the condition $\|x_{n+1} - x_n\|_{L^2} < \epsilon$ as a stopping criterion. We choose two types of starting points:

Case 1: $x_0 = t^4,\ x_1 = t + 1$;

Case 2: $x_0 = e^t,\ x_1 = 3e^t$.

Before conducting our numerical experiments, we first recall that the projections on sets C and Q have respective formulas, that is

$$P_C(x) = \begin{cases} \frac{x}{\|x\|_{L^2}}, & \text{if } \|x\|_{L^2} > 1, \\ x, & \text{if } \|x\|_{L^2} \leq 1. \end{cases} \quad \text{and} \quad P_Q(x) = x - \frac{\langle t, x\rangle}{\|t\|_{L^2}} t.$$

The numerical results are shown in Table 2.

Table 2 Results for Example 4.2

(x_0, x_1)	ϵ	GMV Alg.3.1		SKC Alg.3.1		Alg.1		López Alg. 5.1	
		Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time
Case 1	10^{-2}	7	0.2846	11	0.4040	3	0.0745	7	0.1754
	10^{-4}	88	2.8905	51	1.7017	3	0.2806	14	0.3000
Case 2	10^{-2}	13	0.3332	6	0.1545	3	0.0603	9	0.1758
	10^{-4}	17	2.3068	19	1.7090	3	0.0983	16	2.2206

Remark 4.2 According to Table 2, it shows that Alg. 1 behaves better than the compared algorithms with respect to the number of iterations and execution time under various cases of the problem.

Example 4.3 LASSO problem Sahu et al. (2020)

In this subsection, we employ SFP to model a real problem which is the recovery of a sparse signal. We take advantage of the well-known LASSO problem whose form is the following:

$$\min \left\{ \frac{1}{2} \|Ax - b\|^2 : x \in \mathbb{R}^N, \|x\|_1 \leq \kappa \right\}, \tag{4.2}$$

where $A \in \mathbb{R}^{M\times N}$, $M < N$, $b \in \mathbb{R}^M$ and $\kappa > 0$. This problem is devoted to finding a sparse solution of SFP. The system A is generated from a standard normal distribution with mean zero and unit variance. We generate the true sparse signal z^* from uniformly distribution in the interval $[-2, 2]$ with random k position nonzero, while the rest is kept zero. The sample data $b = Az^*$.

Under certain conditions on matrix A, the solution of the minimization problem (4.2) is equivalent to the ℓ_0- norm solution of the underdetermined linear system. For the SFP , we define $C = \{z | \|z\|_1 \leq \kappa\}$, $\kappa = k$, and $Q = \{b\}$, since the projection onto the closed convex set C does not have a closed form solution. Therefore, we employ the subgradient projection. Thus, we define a convex function $c(z) = \|z\|_1 - \kappa$ and denote C_n by

$$C_n = \{z : c(w_n) + \langle \varepsilon_n, z - w_n \rangle \leq 0\},$$

where $\varepsilon_n \in \partial c(w_n)$. Also, the orthogonal projection of a point $z \in \mathbb{R}^N$ onto C_n can be computed via

$$P_{C_n}(z) = \begin{cases} z, & \text{if } c(w_n) + \langle \varepsilon_n, z - w_n \rangle \leq 0, \\ z - \frac{c(w_n) + \langle \varepsilon_n, z - w_n \rangle}{\|\varepsilon_n\|^2} \varepsilon_n, & \text{otherwise.} \end{cases}$$

The subdifferential ∂c at w_n is

$$\partial c(w_n) = \begin{cases} 1, & \text{if } w_n > 0, \\ [-1, 1], & \text{if } w_n = 0, \\ -1, & \text{if } w_n < 0. \end{cases}$$

To implement our method in this example, we initialize the algorithms at the original and define

$$E_n = \frac{\|x_n - z^*\|}{\max\{1, \|x_n\|\}}.$$

Table 3 Results for Example 4.3

(M, N, k)	GMV Alg. 3.1		SKC Alg. 3.1		Alg. 1		López Alg. 5.1	
	E_n	Time	E_n	Time	E_n	Time	E_n	Time
(240,1024,30)	1.7991e-07	1.5225	4.0650e-06	1.0378	7.5263e-10	0.5413	5.3074e-10	1.1068
(480,2048,60)	1.1657e-07	7.1483	2.5370e-06	4.8965	4.8233e-10	2.5933	6.6102e-10	5.9596
(720,3072,90)	1.2757e-07	39.9267	2.8926e-06	26.7934	5.3908e-10	13.5964	9.7823e-10	14.8142
(960,4096,120)	1.3295e-07	75.8669	2.9951e-06	50.8525	5.5842e-10	25.7045	7.2942e-10	54.7798
(1200,5120,150)	1.4871e-07	122.6915	3.3246e-06	82.4403	6.1196e-10	41.4269	7.1270e-10	87.1176

Table 4 Results of Algorithm 1 with different values of κ_n for Example 4.3

(M, N, k)	$\kappa_n = 1$		$\kappa_n = 1.5$		$\kappa_n = 1.99$	
	E_n	Time	E_n	Time	E_n	Time
(240,1024,30)	1.5381e-09	1.1908	1.7594e-09	0.5842	5.5169e-10	0.5813
(480,2048,60)	2.1818e-09	3.4727	1.5719e-09	3.2670	1.4225e-09	3.3791
(720,3072,90)	1.9466e-09	14.0904	1.1367e-09	14.2298	1.0098e-09	14.0690
(960,4096,120)	1.8093e-09	26.6045	1.5169e-09	26.6880	9.9406e-10	26.6288
(1200,5120,150)	2.3998e-09	42.5273	1.2371e-09	42.8870	9.4754e-10	42.5650

We test the numerical behavior of all algorithms with the same iteration error E_n in different M, N and k and limit the number of iterations to 8000 and report E_n in Table 3. The second problem is the recovery of the signal z^* when $M = 1440,\ N = 6144,\ k = 180$, $M \times N$ matrix A is randomly obtained with independent samples of standard Gaussian distribution. More details, the original signal z^* contains 180 randomly placed ± 1 spikes. The iterative process is started with $x_0 = 0$, the following method of mean square error is used for measuring the recovery accuracy:

$$\mathrm{MSE} = \frac{1}{N}\|x_n - z^*\|^2.$$

For all algorithms, we fix $f(x) = 0.0005x$, $\sigma = 0.9$, $\tilde{\tau}_n = \frac{1}{n^5}$ and $\gamma_n = \frac{1}{10^5 n}$. For SKC Alg. 3.1, we take $\kappa_n = 0.02$. For GMV Alg. 3.1, we adopt $\kappa_n = 1.9$ and $\beta_n = 0.7$. For Alg. 1, we set $\kappa_n = 1.9$. For López Alg. 5.1, we choose $\kappa_n = 1.9$ and $\gamma_n = \frac{10^{-7}}{n}$.

Remark 4.3 It can be observed from Tables 3–4 that the proposed algorithm implements efficiently. Moreover, our method requires less CPU time than some strongly convergent algorithms in the literature to obtain more smaller value of error accuracy E_n in different cases. We also find that when the value of the parameter κ_n is 1.99, our proposed algorithm performs better.

The recovery results of all algorithms are shown in Fig. 1, which stands for the original signal, the mean-squared error (MSE) of the restored signal, and the computing time required for the iterative process.

Remark 4.4 As can be observed from Fig. 1, the signal z^* is estimated with fair degree of accuracy by Algorithm 1. Under the same number of iterations, the execution time of Algorithm 1 is less but a little bigger mean-squared error.

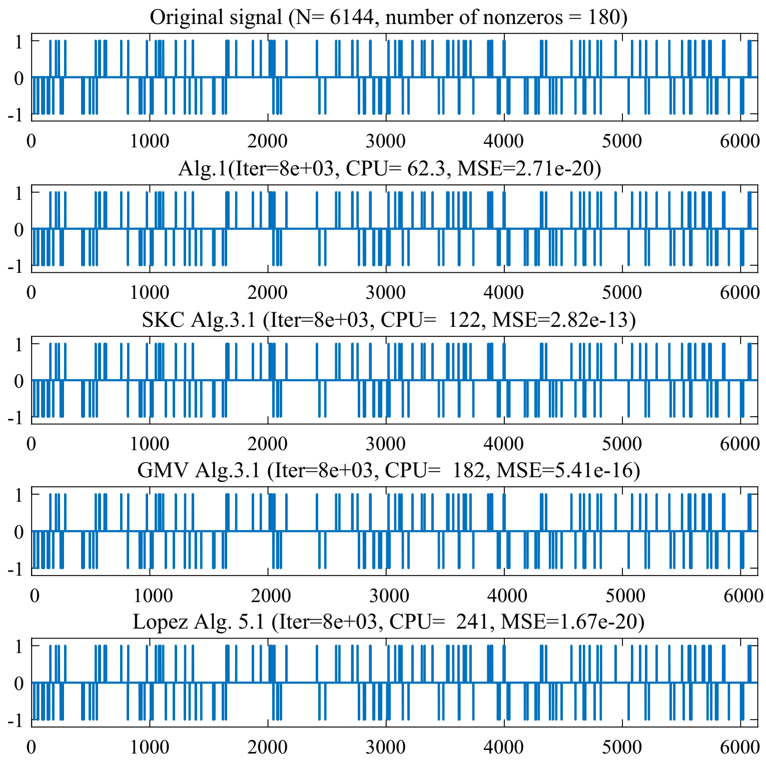

Fig. 1 Comparison of signal processing

Example 4.4 Image deblurring problem Saejung and Yotkaew (2012) We consider the problem here is an image deblurring problem. Fixed a convolution matrix $A \in \mathbb{R}^{m\times n}$ and an unknown image $z \in \mathbb{R}^n$, we derive $b \in \mathbb{R}^m$, which can be viewed as the known degreaded observation. Also, the unknown additive random noise $\nu \in \mathbb{R}^m$ is included, and furthermore, we obtain the image recovery problem as follows:

$$Az = b + \nu. \tag{4.3}$$

This problem obviously is suitable for the setting of SFP with $C = \mathbb{R}^n$; if no noise is added to the observed image b, then $Q = \{b\}$ is a singleton and otherwise $Q = \{x \in \mathbb{R}^m | \|x - (b + \nu)\| \leq \epsilon\}$ for small enough $\epsilon > 0$. In this example, we compare Algorithm 1 with Lopez Alg. 5.1. The test image was corrupted as in He et al. (2016). More precisely, every image was degraded by a 9×9 Gaussian random blur and standard deviation 4, and corrupted by undertaking an additive zero-mean white Gaussian noise with standard deviation 10^{-3}. To measure the quality of the obtained recovered image, we define the following signal-to-noise ratio:

$$\mathrm{SNR} = 20 \log_{10} \frac{\|z\|}{\|\bar{z} - z\|},$$

where z is an original image and $\bar{z}$ is a obtained image. Obviously, when the SNR value is higher, the image is recovered better. For Alg. 1, we take $\kappa_n = 1.6$, $\sigma = 0.3$, $f(x) = 0.01x$, $\gamma_n = \frac{21}{100n}$,$\tilde{\tau}_n = \frac{1}{n^2}$. For Lo̍pez Alg. 5.1, we adopt $\kappa_n = 5 \times 10^{-4}$ and $\gamma_n = \frac{21}{100n}$.

Table 5 Results of all algorithms for Example 4.4

Alg. 1		López Alg. 5.1	
SNR	Time	SNR	Time
25.25	9.08	24.76	8.86

(a) Original image **(b)** Observed image

Fig. 2 Original image (**a**) and observed image (**b**) for Example 4.4

(a) Alg. 1 **(b)** López Alg. 5.1

Fig. 3 Recovered images by Alg. 1 and López Alg. 5.1 in Example 4.4

For all algorithms, we limit the number of iterations to 100 and report numerical results in Table 5 and Figs. 2-3.

Remark 4.5 As shown in Table 5 and Figs. 2-3, we observe that these two methods require the same iterations to recovery images. Concretely, Alg. 1 obtains higher SNR than López Alg. 5.1, but a little longer execution time.

5 Conclusions

In this paper, we obtain a strong convergence result for the proximal split feasibility problems with nonexpansive mappings. We modify the algorithm Shehu and Iyiola (2017) with the viscosity-type algorithm Moudafi (2000), the inertial method and the split proximal algorithm with a self-adaptive stepsize. As practical applications, we consider signal recovery and image deblurring problems. Preliminary numerical experiments confirm the effectiveness of the proposed algorithm in practice.

Funding The second author was supported by National Natural Science Foundation of China (Grant No. 11801430).

References

Abbas M, Alshahrani M, Ansari QH et al (2018) Iterative methods for solving proximal split minimization problems. Numer Algorithms 78:193–215

Anh PK, Vinh NT, Dung VT (2018) A new self-adaptive CQ algorithm with an application to the lasso problem. J Fixed Point Theory Appl 20:142

Byrne C (2002) Iterative oblique projection onto convex sets and the split feasibility problem. Inverse Problems 18:441–453

Byrne C (2004) A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Problems 20:103–120

Censor Y, Elfving T (1994) A multiprojection algorithm using Bregman projection in a product space. Numer Algorithms 8:221–239

Censor Y, Elfving T, Kopf N, Bortfeld T (2005) The multiple-sets split feasibility problem and its applications for inverse problems. Inverse Problems 21:2071–2084

Combettes PL, Hirstoaga SA (2005) Equilibrium programming in Hilbert spaces. J Nonlinear Convex Anal 6:117–136

Combettes PL, Wajs VR (2005) Signal recovery by proximal forward-backward splitting. Multiscale Model Simul 4:1168–1200

Dong QL, Tang YC, Cho YJ, Rassias TM (2018) ’ ’Optimal“ choice of the step length of the projection and contraction methods for solving the split feasibility problem. J Global Optim 71:341–360

Dong QL, He S, Rassias MT (2020) General splitting methods with linearization for the split feasibility problem. J Glob Optim. https://doi.org/10.1007/s10898-020-00963-3

Gibali A, Liu LW, Tang YC (2018) Note on the modified relaxation CQ algorithm for the split feasibility problem. Optim Lett 12:817–830

Gibali A, Mai DT, Vinh NT (2019) A new relaxed CQ algorithm for solving split feasibility problems in Hilbert spaces and its applications. J Ind Manag Optim 15:963–984

Goebel K, Reich S (1984) Uniform convexity, hyperbolic geometry, and nonexpansive mappings. Marcel Dekker, New York

He H, Ling C, Xu HK (2016) An Implementable Splitting Algorithm for the ?1-norm Regularized Split Feasibility Problem. J Sci Comput 67:281–298

Iyiola OS, Ogbuisi FU, Shehu Y (2018) An inertial type iterative method with Armijo linesearch for non-monotone equilibrium problems. Calcolo 55:52

Kesornprom S, Cholamjiak P (2019) Proximal type algorithms involving linesearch and inertial technique for split variational inclusion problem in Hilbert spaces with applications. Optimization. https://doi.org/10.1080/02331934.2019.1638389

Kesornprom S, Pholasa N, Cholamjiak P (2020) On the convergence analysis of the gradient-CQ algorithms for the split feasibility problem. Numer Algorithms 84:997–1017

Lions PL, Mercier B (1979) Splitting algorithms for the sum of two nonlinear operators. SIAM J Numer Anal 16:964–979

López, G, Martin V, Wang F, Xu HK (2012) Solving the split feasibility problem without prior knowledge of matrix norms. Inverse Problems 28

Moudafi A (2000) Viscosity approximation methods for fixed points problems. J Math Anal Appl 241:46–55

Moudafi A, Thakur BS (2014) Solving proximal split feasibility problems without prior knowledge of operator norms. Optim Lett 8:2099–2110

Qu B, Xiu N (2005) A note on the CQ algorithm for the split feasibility problem. Inverse Problems 21:1655–1665

Reich S, Tuyen TM, Ha MTN (2020) An optimization approach to solving the split feasibility problem in Hilbert spaces. J Glob Optim. https://doi.org/10.1007/s10898-020-00964-2

Saejung S, Yotkaew P (2012) Approximation of zeros of inverse strongly monotone operators in Bachna spaces. Nolinear Anal 75:742–750

Sahu DR, Cho YJ, Dong QL, Kashyap MR, Li XH (2020) Inertial relaxed CQ algorithms for solving a split feasibility problem in Hilbert spaces. Numer Algorithms https://doi.org/10.1007/s11075-020-00999-2

Shehu Y, Iyiola OS (2017) Convergence analysis for the proximal split feasibility problem using an inertial extrapolation term method. J Fixed Point Theory Appl 19:2483–2510

Shehu Y, Iyiola OS (2017) Strong convergence result for proximal split feasibility problem in Hilbert spaces. Optimization 12:2275–2290

Shehu Y, Iyiola OS (2018) Accelerated hybrid viscosity and steepest-descent method for proximal split feasibility problems. Optimization 67:475–492

Shehu Y, Iyiola OS (2018) Nonlinear iteration method for proximal split feasibility problems. Math Methods Appl Sci 41:781–802

Shehu Y, Iyiola OS, Ogbuisi FU (2020) Iterative method with inertial terms for nonexpansive mappings: applications to compressed sensing. Numer Algorithms 83:1321–1347

Shehu Y, Gibali A (2020) New inertial relaxed method for solving split feasibilities. Optim Lett https://doi.org/10.1007/s11590-020-01603-1

Suantai S, Pholasa N, Cholamjiak P (2018) The modified inertial relaxed CQ algorithm for solving the split feasibility problems. J Ind Manag Optim 23:1595–1615

Wang Y, Xu H-K (2014) Strong convergence for the the proximal gradient method. J Nonlinear Convex Anal 15:581–593

Yen LH, Muu LD, Huyen NTT (2016) An algorithm for a class of split feasibility problems: application to a model in electricity production. Math Methods Oper Res 84:549–565

Yen LH, Huyen NTT, Muu LD (2019) A subgradient algorithm for a class of nonlinear split feasibility problems: application to jointly constrained Nash equilibrium models. J Global Optim 73:849–868