Journal of Global Optimization (2023) 87:619–639
https://doi.org/10.1007/s10898-022-01246-9

A relaxed inertial and viscosity method for split feasibility problem and applications to image recovery

Haitao Che[1] · Yaru Zhuang[2] · Yiju Wang[2] · Haibin Chen[2]

Received: 27 August 2021 / Accepted: 25 September 2022 / Published online: 14 October 2022

Abstract

In this paper, by combining Polyak's inertial extrapolation technique for minimization problem with the viscosity approximation for fixed point problem, we develop a new type of numerical solution method for split feasibility problem. Under suitable assumptions, we establish the global convergence of the designed method. The given experimental results applied on the sparse reconstruction problem show that the proposed algorithm is not only robust to different levels of sparsity and amplitude of signals and the noise pixels but also insensitive to the diverse values of scalar weight. Further, the proposed algorithm achieves better restoration performance compared with some other algorithms for image recovery.

Keywords Split feasibility problem · Sparse reconstruction · Inertial extrapolation · Convergence · Image recovery

Mathematics Subject Classification 65H10 · 90C33 · 47H09

1 Introduction

Let $C = \{x \in \mathbb{R}^n \mid c(x) \leq 0\}$ be a nonempty closed convex set in $\mathbb{R}^n$ and $Q = \{y \in \mathbb{R}^m \mid q(x) \leq 0\}$ be a nonempty closed convex set in $\mathbb{R}^m$. The split feasibility problem, abbreviated by SFP [4], is to find

✉ Haitao Che
haitaoche@163.com

Yaru Zhuang
1364566799@qq.com

Yiju Wang
wyijumail@163.com

Haibin Chen
chenhaibin508@163.com

1 School of Mathematics and Information Science, Weifang University, Weifang, Shandong, People's Republic of China

2 School of Management Science, Qufu Normal University, Rizhao, Shandong, People's Republic of China

$$x^* \in C \ \text{ such that } Ax^* \in Q, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}(m \ll n)$ is a bounded linear operator. For the SFP, the first numerical method is Byrne's CQ algorithm [1] which is developed based on the following equivalent minimization formulation

$$\min_{x\in C} f(x) = \frac{1}{2}\|(I - P_Q)Ax\|^2, \tag{2}$$

where P_Q is the projection operator from $\mathbb{R}^n \to Q$, I is an identity matrix and $\|\cdot\|$ is denoted by ℓ_2-norm. Since the objective function of the problem is continuously differentiable with gradient

$$\nabla f(x) = A^\top(I - P_Q)Ax,$$

Applying the traditional gradient projection method for constrained optimization [2] to the SFP yields the following iterative scheme

$$x_{k+1} = P_C(x_k - \alpha\nabla f(x_k)) = P_C(x_k - \alpha A^\top(I - P_Q)Ax_k),$$

where $\alpha \in (0, 2/\|A\|^2)$ is the stepsize.

As the stepsize α critically relies on the underlying linear operator norm, it may be very small and hence affects the efficiency of the algorithm. To overcome this drawback, Yang [17] introduced the following stepsize

$$\alpha_k = \frac{\rho_k}{\|\nabla f(x_k)\|},$$

where $\rho_k > 0$ and $\sum_{k=1}^{\infty} \rho_k = \infty, \quad \sum_{k=1}^{\infty} \rho_k^2 < \infty$. However, the global convergence of the algorithm needs some rigorous conditions on set Q and the operator A. To establish the convergence of the algorithm under a milder condition, Lopéz et al. [8] introduced the following stepsize

$$\alpha_k = \frac{\rho_k f(x_k)}{\|\nabla f(x_k)\|}$$

with $\rho_k \in (0, 4)$.

To make the CQ algorithm more easily implemented, Dang [6] makes a relaxation to convex sets C and Q to obtain the following relaxed minimization problem

$$\min_{x\in C_k} f_k(x) = \frac{1}{2}\|(I - P_{Q_k})Ax\|^2$$

where

$$C_k = \{x \in \mathbb{R}^n \mid c(x_k) - \langle \xi_k, x_k - x\rangle \le 0\}, \tag{3}$$

$$Q_k = \{y \in \mathbb{R}^m \mid q(Ax_k) - \langle \omega_k,\ Ax_k - y\rangle \le 0\}, \tag{4}$$

and $\xi_k \in \partial c(x_k)$ and $\omega_k \in \partial q(Ax_k)$. Based on the relaxed formation, Sahu et al. [13] apply the Polyak's inertial extrapolation technique [12] to Byrne's CQ algorithm to obtain the following relaxed inertial CQ algorithm for solving the SFP

$$\begin{cases} y_k = x_k + \theta_k(x_k - x_{k-1}), \\ x_{k+1} = P_{C_k}(y_k - \alpha A^\top(I - P_{Q_k})Ay_k), \end{cases}$$

where $\alpha \in (0, 2/\|A\|^2)$, $\theta_k \in [0, \bar{\theta}_k]$ with $\bar{\theta}_k = \min\{\theta, \frac{1}{\max\{k^2\|x_k - x_{k-1}\|^2, k^2\|x_k - x_{k-1}\|\}}\}$ and $\theta \in [0, 1)$. The numerical experiments given in [13] show that the inertial extrapolation technique can really improve the efficiency of the CQ algorithm.

On the other hand, to accelerate the Halpern fixed point algorithm for a nonexpansive mapping T in real Hilbert space, Sakurai and liduka [22] introduced the following iterative algorithm by using the idea of conjugate gradient methods

$$\begin{cases} y_n = x_n + \alpha d_{n+1} \\ x_{n+1} = \mu\alpha_n x_0 + (1 - \mu\alpha_n) y_n, \end{cases}$$

where $d_{n+1} := \frac{Tx_n - x_n}{\alpha} + \beta_n d_n$.

Motivated by the results of [6, 11, 17, 22], we consider the SFP from another perspective. Turn a constraint problem (2) into an unconstrained problem. With the aid of projection operator, SFP can be formulated as the following minimization problem

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}\|(I - P_C)x\|^2 + \frac{1}{2}\|(I - P_Q)Ax\|^2.$$

It is easy to see that (2) has a solution if and only if the above minimization problem has zero as its optimal objective value. Furthermore, to implement the calculation of projection easily, one can consider its relaxed form

$$\min_{x \in \mathbb{R}^n} f_k(x) = \frac{1}{2}\|(I - P_{C_k})x\|^2 + \frac{1}{2}\|(I - P_{Q_k})Ax\|^2, \tag{5}$$

where C_k and Q_k are respectively defined by (3) and (4).

With this, based on a new reformulation for the SFP, we establish a more efficient method to solve the SFP. The new approach has the following features: (1) The objective function in the new reformulation is convexly quadratic. The corresponding gradient can be computed very easily. (2) A novel conjugate gradient method is devised by replacing $\frac{Tx_n - x_n}{\alpha}$ in [22] with the steepest descent direction. Compared with the classical conjugate gradient methods which need generate the set of conjugate vectors to compute a new vector d_k by using only the previous vector d_{k-1}. Our algorithm does not need to know all elements of the conjugate set, d_k is automatically updated to these vectors. This remarkable property shows that the method requires little storage and computation. (3) A self-adaptive step procedure is introduced to our algorithm, which is implemented by a simple computation and does not need any line search rule. Lipschitz constant can not be the input parameter of our method and this is particularly interesting when the constant is unknown or not easy to approximate. (4) The Polyaks inertial extrapolation technique and the viscosity approximation technique are applied to speed up the convergence of the algorithm. In addition, the validity of the proposed method is theoretically guaranteed under suitable assumptions. Some numerical experiments are made to show the efficiency of the proposed method.

The remainder of this paper is organized as follows. In Sect. 2, we give some basic concepts and related conclusions. In Sect. 3, we present our new designed inertial method for solving the SFP, and establish its global convergence. The Sect. 4 displays some numerical experiments on signal processing and image recovery to illustrate the efficiency of the proposed method.

2 Preliminaries

In this section, we present some basic concepts and related conclusions. For $x \in \mathbb{R}^n$, its projection to closed convex set $\Omega \in \mathbb{R}^n$ is denoted by $P_\Omega(x)$, i.e.,

$$P_\Omega(x) = \arg\min_{y\in\Omega} \|y - x\|.$$

For the projection operator $P_\Omega(\cdot)$, we have the following conclusion [18].

Lemma 1 *Let Ω be non-empty closed convex subset in $\mathbb{R}^n$. Then the followings hold.*

(i) $\langle P_\Omega(x) - x, y - P_\Omega(x)\rangle \geq 0$, *for all* $x \in \mathbb{R}^n$ *and* $y \in \Omega$*;*
(ii) $\|P_\Omega(x) - P_\Omega(y)\|^2 \leq \langle P_\Omega(x) - P_\Omega(y), x - y\rangle$, *for all* $x, y \in \mathbb{R}^n$*;*
(iii) $\|P_\Omega(x) - y\|^2 \leq \|x - y\|^2 - \|P_\Omega(x) - x\|^2$, *for all* $x \in \mathbb{R}^n$ *and* $y \in \Omega$.

For convex but non-differentiable function, its subdifferential is an important analytic tool which is defined as follows [5].

Definition 1 For convex $f : \mathbb{R}^n \to \mathbb{R}$, its subdifferential at x is defined as

$$\partial f(x) = \{\xi \in \mathbb{R}^n \mid f(y) \geq f(x) + \langle \xi, y - x\rangle, \ \forall\, y \in \mathbb{R}^n\}.$$

Definition 2 Function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be lower semi-continuous at x if $f(x) \leq \liminf_{k\to\infty} f(x_k)$ for any sequence $\{x_k\}$ with $\lim_{k\to\infty} x_k \to x$.

Definition 3 Function $f : \mathbb{R}^n \to \mathbb{R}^n$ is said to be ρ-contracted, if there exists $\rho > 0$ such that

$$\|f(x) - f(y)\| \leq \rho\|x - y\|, \quad \forall\, x, y \in \mathbb{R}^n.$$

Lemma 2 (Lemma 2.5 in [10]) *A mapping $A : \mathbb{R}^n \to \mathbb{R}^n$ is said to be strongly positive, if there exists a constant $r > 0$ such that $\langle Ax, x\rangle \geq r\|x\|^2$ for all $x \in \mathbb{R}^n$. Furthermore, if linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$ is strongly positive bounded with coefficient $r > 0$, then $\|I - \rho A\| \leq 1 - \rho r$ for any $0 < \rho < \|A\|^{-1}$.*

Lemma 3 ([16]) *For nonnegative sequence $\{a_k\}$, if $a_{k+1} \leq (1-\gamma_k)a_k + \gamma_k\delta_k$ with $\gamma_k \in (0, 1)$ and $\delta_k \in \mathbb{R}$ being such that*

$$\sum_{n=1}^{\infty} \gamma_k = \infty, \quad \limsup_{k\to\infty} \delta_k \leq 0 \quad \text{or} \quad \sum_{n=1}^{\infty} |\delta_k\gamma_k| < \infty,$$

then $\lim_{k\to\infty} a_k = 0$.

Lemma 4 ([9]) *If sequence $\{\sigma_k\}$ has a subsequence $\{\sigma_{k_j}\}$ satisfying $\sigma_{k_j} < \sigma_{k_j+1}$, then the sequence defined by $\tau_k = \max\{t \leq k \mid \sigma_t < \sigma_{t+1}\}$ is nondecreasing. Further $\lim_{k\to\infty} \tau_k = \infty$ and $\max\{\sigma_{\tau_k}, \sigma_k\} \leq \sigma_{\tau_k+1}$.*

3 Algorithm and convergence

In this section, we will present a relaxed inertial and viscosity method for solving the SFP and establish its global convergence. Certainly, the objective function of (5) is continuously differentiable with derivative

$$\nabla f_k(x) = (I - P_{C_k})x + A^\top(I - P_{Q_k})Ax.$$

Algorithm 1. A relaxed inertial and viscosity method

Step 1. Choose ρ-contraction function $g : \mathbb{R}^n \to \mathbb{R}^n$ and strongly positive bounded linear operator $D : \mathbb{R}^n \to \mathbb{R}^n$ with coefficient $\mu > 0$. Take initial points $x_0, x_1 \in \mathbb{R}^n$, parameters $\gamma \in (0, \frac{\mu}{\rho})$, $\tau \in (0, 1]$, and tolerance error $\varepsilon > 0$.

Step 2. Take $\theta_1 \in [0, 1)$, $\rho_1 \in (0, 1)$, and set $y_1 = x_1 + \theta_1(x_1 - x_0)$.
If $\|(I - P_{C_1})y_1 + A^\top(I - P_{Q_1})Ay_1\| \le \varepsilon$, terminate;
otherwise, set $d_0 = -t_1\nabla f_1(y_1)$ with $t_1 = \frac{2\rho_1(\|(I-P_{C_1})y_1\|^2+\|(I-P_{Q_1})Ay_1\|^2)}{\|(I-P_{C_1})y_1+A^\top(I-P_{Q_1})Ay_1\|^2}$.
Set $k = 1$ and go to Step 3.

Step 3. Take $\alpha_k \in (0, 1)$, $\beta_k \in [0, \frac{1}{2})$, $\rho_k \in (0, 1)$ and

$$\begin{cases} t_k = \dfrac{2\rho_k(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2)}{\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2}, \\ d_k = - t_k\nabla f_k(y_k) + \tau\beta_k d_{k-1}, \\ z_k = y_k + d_k. \end{cases}$$

Set $x_{k+1} = \alpha_k\gamma g(z_k) + (I - \alpha_k D)P_{C_k} z_k$, $\quad k = k + 1$, go to Step 4.

Step 4. Take $\theta_k \in [0, 1)$ and set $y_k = x_k + \theta_k(x_k - x_{k-1})$. If $\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\| \le \varepsilon$, terminate; otherwise, go to Step 3.

With this, we can present the following algorithm for solving the SFP.

For the algorithm, the iterative scheme $y_k = x_k + \theta_k(x_k - x_{k-1})$ uses the inertia technique introduced by Polyak for solving the minimization problem [12], and the iterative scheme $x_{k+1} = \alpha_k\gamma g(z_k) + (I - \alpha_k D)P_{C_k} z_k$ uses the viscosity approximation technique for solving fixed point problem [10]. In addition, the iterative scheme $d_k = -t_k\nabla f_k(y_k) + \tau\beta_k d_{k-1}$ adopts a conjugate gradient technique. So the algorithm is a new type of solution method for solving the SFP in essence.

To obtain our results, we will reveal the following technical lemmas.

Lemma 5 *Suppose that the following conditions*

(i) $\lim\limits_{k\to\infty} \alpha_k = 0$, $\sum\limits_{i=1}^{\infty} \alpha_k = \infty$,

(ii) $\beta_k \le \alpha_k^2$,

(iii) $\{(I - P_{C_k})y_k\}$ *and* $\{(I - P_{Q_k})Ay_k\}$ *are respectively bounded, hold. Then,* $\{d_k\}$ *is bounded.*

Proof Since $\lim\limits_{k\to\infty} \alpha_k = 0$ by condition (i), we may assume, without loss of generality, that $\alpha_k < \min\{\frac{1}{2\mu-\gamma\rho}, \frac{1}{\gamma}, \frac{1}{\gamma\rho}\}$ for all k.

Then, we prove that $\{d_k\}$ is bounded by induction.

For $k = 0$, this holds trivially. Now, we assume that the assertion holds for some k, and show that it continues to hold for $k + 1$. From conditions (i) and (ii) in the assertion, we conclude that $\lim\limits_{k\to\infty} \beta_k = 0$. Thus, there exists $k_0 > 0$ such that $\beta_k \le \frac{1}{2}$, for all $k \ge k_0$.

Set $M_1 = \max\{\|d_{k_0}\|, 2\sup\limits_{k\ge1} \|t_k[(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k]\|\}$. Then it follows from condition (iii) that $M_1 < \infty$. Assume that $\|d_k\| \le M_1$ for any $k \ge k_0$, then

$$\begin{aligned} \|d_{k+1}\| &= \| - t_{k+1}[(I - P_{C_{k+1}})y_{k+1} + A^\top(I - P_{Q_{k+1}})Ay_{k+1}] + \tau\beta_{k+1}d_k\| \\ &\le \|t_{k+1}[(I - P_{C_{k+1}})y_{k+1} + A^\top(I - P_{Q_{k+1}})Ay_{k+1}]\| + \tau\beta_{k+1}\|d_k\| \le M_1. \end{aligned}$$

This shows that $\|d_k\| \le M_1$ for all k. Hence, $\{d_k\}$ is bounded. □

Lemma 6 *Suppose that the following conditions* *(i)* $\beta_k \le \alpha_k^2$,

(ii) $\lim_{k\to\infty} \frac{\theta_k}{\alpha_k}\|x_k - x_{k-1}\| = 0$,

(iii) $0 < \liminf_{k\to\infty} \rho_k < \limsup_{k\to\infty} \rho_k < 1$, *hold. Then, the sequence* $\{x_k\}$, $\{y_k\}$ *and* $\{z_k\}$ *are all bounded.*

Proof For any solution x^* of the SFP, it holds that $x^* = P_{C_k}x^*$ and $Ax^* = P_{Q_k}(Ax^*)$. Then it follows from $z_k = y_k + d_k$ that

$$\begin{aligned}\|z_k - x^*\| &= \|y_k - x^* - t_k[(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k] + \tau\beta_k d_{k-1}\|\\ &\le \|y_k - x^* - t_k[(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k]\| + \tau\beta_k\|d_{k-1}\|. \end{aligned} \tag{6}$$

From condition (iii), one has

$$\begin{aligned}&\|y_k - x^* - t_k[(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k]\|^2\\
&= \|y_k - x^*\|^2 + t_k^2\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2 - 2t_k\langle(I - P_{C_k})y_k\\
&\quad + A^\top(I - P_{Q_k})Ay_k, y_k - x^*\rangle\\
&= \|y_k - x^*\|^2 - t_k[2\langle(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k, y_k - x^*\rangle - t_k\|(I - P_{C_k})y_k\\
&\quad + A^\top(I - P_{Q_k})Ay_k\|^2]\\
&= \|y_k - x^*\|^2 - t_k[2\langle(I - P_{C_k})y_k, y_k - x^*\rangle\\
&\quad + \langle(I - P_{Q_k})Ay_k, Ay_k - Ax^*\rangle - t_k\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2]\\
&= \|y_k - x^*\|^2 - t_k[2\langle(I - P_{C_k})y_k, y_k - P_{C_k}y_k + P_{C_k}y_k - x^*\rangle\\
&\quad + \langle(I - P_{Q_k})Ay_k, Ay_k - P_{Q_k}Ay_k + P_{Q_k}Ay_k - Ax^*\rangle - t_k\|(I - P_{C_k})y_k\\
&\quad + A^\top(I - P_{Q_k})Ay_k\|^2]\\
&\le \|y_k - x^*\|^2 - t_k[2(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2) - t_k\|(I - P_{C_k})y_k\\
&\quad + A^\top(I - P_{Q_k})Ay_k\|^2]\\
&= \|y_k - x^*\|^2 - \frac{4\rho_k(1-\rho_k)(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2)^2}{\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2}\\
&\le \|y_k - x^*\|^2,\end{aligned}$$

where the first inequality uses Lemma 1. Since $\|d_k\| \le M_1$ for $k \ge k_0$, it follows from (6) in this section that

$$\|z_k - x^*\| \le \|y_k - x^*\| + \tau M_1\beta_k. \tag{7}$$

Since

$$\begin{aligned}\|y_k - x^*\| &= \|x_k - x^* + \theta_k(x_k - x_{k-1})\|\\ &\le \|x_k - x^*\| + \theta_k\|x_k - x_{k-1}\|,\end{aligned}$$

one has from Algorithm 1, Lemma 2, (7) in this section and condition (i) that

$$\begin{aligned}\|x_{k+1} - x^*\| &= \|\alpha_k\gamma g(z_k) + (I - \alpha_k D)P_{C_k}z_k - x^*\|\\
&= \|\alpha_k(\gamma g(z_k) - Dx^*) + (I - \alpha_k D)(P_{C_k}z_k - P_{C_k}x^*)\|\\
&\le \alpha_k\|\gamma g(z_k) - Dx^*\| + (1 - \mu\alpha_k)\|z_k - x^*\|\\
&= \alpha_k\|\gamma g(z_k) - \gamma g(x^*) + \gamma g(x^*) - Dx^*\| + (1 - \mu\alpha_k)\|z_k - x^*\|\\
&\le \alpha_k\rho\gamma\|z_k - x^*\| + \alpha_k\|\gamma g(x^*) - Dx^*\| + (1 - \mu\alpha_k)\|z_k - x^*\|\end{aligned}$$

$$
\begin{aligned}
&=(1-\alpha_k(\mu-\gamma\rho))\|z_k-x^*\|+\alpha_k\|\gamma g(x^*)-Dx^*\| \\
&\le(1-\alpha_k(\mu-\gamma\rho))\|x_k-x^*\|+\alpha_k\|\gamma g(x^*)-Dx^*\| \\
&\quad+(1-\alpha_k(\mu-\gamma\rho))\theta_k\|x_k-x_{k-1}\|+(1-\alpha_k(\mu-\gamma\rho))\tau M_1\beta_k \\
&\le(1-\alpha_k(\mu-\gamma\rho))\|x_k-x^*\|+\alpha_k\|\gamma g(x^*)-Dx^*\|+\theta_k\|x_k-x_{k-1}\|+\tau M_1\beta_k \\
&\le(1-\alpha_k(\mu-\gamma\rho))\|x_k-x^*\| \\
&\quad+\alpha_k(\mu-\gamma\rho)\left[\frac{\|\gamma g(x^*)-Dx^*\|}{\mu-\gamma\rho}+\frac{1}{\mu-\gamma\rho}\left(\tau M_1\alpha_k+\frac{\theta_k}{\alpha_k}\|x_k-x_{k-1}\|\right)\right] \\
&\le\max\left\{\|x_k-x^*\|,\frac{\|\gamma g(x^*)-Dx^*\|}{\mu-\gamma\rho}+\frac{1}{\mu-\gamma\rho}\left(\tau M_1\alpha_k+\frac{\theta_k}{\alpha_k}\|x_k-x_{k-1}\|\right)\right\},
\end{aligned}
\tag{8}
$$

where the third inequality follows from (8), the fourth inequality follows from $\beta_k \le \alpha_k^2$, the fifth inequality follows from the fact $0<\alpha_k(\mu-\gamma\rho)<1$ deduced from $\alpha_k < \frac{1}{2\mu-\gamma\rho}$ and $\gamma \in (0, \frac{\mu}{\rho})$. Since sequence $\{\frac{\theta_k}{\alpha_k}\|x_k-x_{k-1}\|\}$ is bounded by condition (ii), there exists $M_2>0$, such that $\frac{\theta_k}{\alpha_k}\|x_k-x_{k-1}\| \le M_2$. By induction, one has

$$
\|x_{k+1}-x^*\|\le\max\left\{\|x_1-x^*\|,\frac{\|\gamma g(x^*)-Dx^*\|}{\mu-\gamma\rho}+\frac{1}{\mu-\gamma\rho}\left(\tau M_1\frac{1}{2\mu-\gamma\rho}+M_2\right)\right\},
$$

which implies that $\{x_k\}$ is bounded. Consequently, $\{y_k\}$ and $\{z_k\}$ are all bounded. □

Now, we show the convergence of the proposed method.

Theorem 1 *Assume the solution set of the SFP, denoted by Γ, is consistent. Then under the assumptions that*

(i) $\lim_{k\to\infty}\alpha_k=0$, $\sum_{i=1}^{\infty}\alpha_k=\infty$;
(ii) $\beta_k\le\alpha_k^2$;
(iii) $\lim_{k\to\infty}\frac{\theta_k}{\alpha_k}\|x_k-x_{k-1}\|=0$;
(iv) $0<\liminf_{k\to\infty}\rho_k<\limsup_{k\to\infty}\rho_k<1$;
(v) $\{(I-P_{C_k})y_k\}$ *and* $\{(I-P_{Q_k})Ay_k\}$ *are respectively bounded, the sequence* $\{x_k\}$ *generated by Algorithm 1 converges to a solution of the SFP.*

Proof From Lemma 5 of the boundedness of $\{d_k\}$, we conclude that there exists $M_3>0$ such that $2\tau|\langle z_k-x^*,d_{k-1}\rangle|\le M_3$ for any solution x^* of the SFP.

It follows from Algorithm 1 that

$$
\begin{aligned}
\|z_k-x^*\|^2&=\|y_k-x^*+d_k\|^2 \\
&=\|y_k-x^*-t_k[(I-P_{C_k})y_k+A^\top(I-P_{Q_k})Ay_k]+\tau\beta_k d_{k-1}\|^2 \\
&\le\|y_k-x^*-t_k[(I-P_{C_k})y_k+A^\top(I-P_{Q_k})Ay_k]\|^2+2\tau\beta_k\langle z_k-x^*,d_{k-1}\rangle \\
&\le\|y_k-x^*\|^2-t_k[2(\|(I-P_{C_k})y_k\|^2+\|(I-P_{Q_k})Ay_k\|^2) \\
&\quad-t_k\|(I-P_{C_k})y_k+A^\top(I-P_{Q_k})Ay_k\|^2]+2\tau\beta_k\langle z_k-x^*,d_{k-1}\rangle \\
&\le\|y_k-x^*\|^2-\frac{4\rho_k(1-\rho_k)(\|(I-P_{C_k})y_k\|^2+\|(I-P_{Q_k})Ay_k\|^2)^2}{\|(I-P_{C_k})y_k+A^\top(I-P_{Q_k})Ay_k\|^2}+M_3\alpha_k^2
\end{aligned}
\tag{9}
$$

$$
\le\|y_k-x^*\|^2+M_3\alpha_k^2, \tag{10}
$$

where the first inequality follows from the fact $\|a+b\|^2 \leq \|a\|^2 + 2\langle a+b, b\rangle$, the second inequality follows from Lemma 1, and the third inequality follows from $\beta_k \leq \alpha_k^2$. Combining this with the fact that

$$\begin{aligned}\|y_k - x^*\|^2 &= \|x_k - x^* + \theta_k(x_k - x_{k-1})\|^2 \\ &= \|x_k - x^*\|^2 + \theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k\|x_k - x^*\|\|x_k - x_{k-1}\|, \end{aligned} \tag{11}$$

yields

$$\begin{aligned}\|x_{k+1} - x^*\|^2 =& \|\alpha_k\gamma g(z_k) + (I - \alpha_k D)P_{C_k}z_k - x^*\|^2 \\ =& \|\alpha_k(\gamma g(z_k) - Dx^*) + (I - \alpha_k D)(P_{C_k}z_k - P_{C_k}x^*)\|^2 \\ \leq& (1-\mu\alpha_k)^2\|z_k - x^*\|^2 + 2\alpha_k\langle\gamma g(z_k) - Dx^*, x_{k+1} - x^*\rangle \\ =& (1-\mu\alpha_k)^2\|z_k - x^*\|^2 + 2\alpha_k\langle\gamma g(z_k) - \gamma g(x^*), x_{k+1} - x^*\rangle + 2\alpha_k\langle\gamma g(x^*) \\ &- Dx^*, x_{k+1} - x^*\rangle \\ \leq& (1-\mu\alpha_k)^2\|z_k - x^*\|^2 + 2\alpha_k\gamma\rho\|z_k - x^*\|\|x_{k+1} - x^*\| + 2\alpha_k\langle\gamma g(x^*) \\ &- Dx^*, x_{k+1} - x^*\rangle \\ \leq& (1-\mu\alpha_k)^2\|z_k - x^*\|^2 + \alpha_k\gamma\rho(\|z_k - x^*\|^2 + \|x_{k+1} - x^*\|^2) + 2\alpha_k\langle\gamma g(x^*) \\ &- Dx^*, x_{k+1} - x^*\rangle \\ =& ((1-\mu\alpha_k)^2 + \alpha_k\gamma\rho)\|z_k - x^*\|^2 + \alpha_k\gamma\rho\|x_{k+1} - x^*\|^2 + 2\alpha_k\langle\gamma g(x^*) \\ &- Dx^*, x_{k+1} - x^*\rangle \\ \leq& ((1-\mu\alpha_k)^2 + \alpha_k\gamma\rho)\|x_k - x^*\|^2 + \alpha_k\gamma\rho\|x_{k+1} - x^*\|^2 + 2\alpha_k\langle\gamma g(x^*) \\ &- Dx^*, x_{k+1} - x^*\rangle \\ &+ [(1-\mu\alpha_k)^2 + \alpha_k\gamma\rho](\theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k\|x_k - x^*\|\|x_k - x_{k-1}\| + M_3\alpha_k^2),\end{aligned}$$

where the first inequality follows from the fact $\|a+b\|^2 \leq \|a\|^2 + 2\langle a+b, b\rangle$ and $\|I - \alpha_k D\| \leq 1 - \mu\alpha_k$, the second inequality uses the Cauchy-Schwarz inequality and ρ-contraction of function g, the third inequality follows from $2\langle a, b\rangle \leq \|a\|^2 + \|b\|^2$, the last inequality follows from (10) and (11). Thus

$$\begin{aligned}(1-\alpha_k\gamma\rho)\|x_{k+1} - x^*\|^2 \leq& ((1-\mu\alpha_k)^2 + \alpha_k\gamma\rho)\|x_k - x^*\|^2 + 2\alpha_k\langle\gamma g(x^*) - Dx^*, x_{k+1} - x^*\rangle \\ &+ [(1-\mu\alpha_k)^2 + \alpha_k\gamma\rho](\theta_k^2\|x_k - x_{k-1}\|^2 \\ &+ 2\theta_k\|x_k - x^*\|\|x_k - x_{k-1}\| + M_3\alpha_k^2),\end{aligned}$$

and hence

$$\begin{aligned}\|x_{k+1} - x^*\|^2 \leq& \frac{(1-\mu\alpha_k)^2 + \alpha_k\gamma\rho}{1-\alpha_k\gamma\rho}\|x_k - x^*\|^2 + \frac{2\alpha_k}{1-\alpha_k\gamma\rho}\langle\gamma g(x^*) - Dx^*, x_{k+1} - x^*\rangle \\ &+ \frac{(1-\mu\alpha_k)^2 + \alpha_k\gamma\rho}{1-\alpha_k\gamma\rho}(\theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k\|x_k - x^*\|\|x_k - x_{k-1}\| + M_3\alpha_k^2) \\ =& \frac{1 - 2\mu\alpha_k + (\mu\alpha_k)^2 + \alpha_k\gamma\rho}{1-\alpha_k\gamma\rho}\|x_k - x^*\|^2 + \frac{2\alpha_k}{1-\alpha_k\gamma\rho}\langle\gamma g(x^*) - Dx^*, x_{k+1} - x^*\rangle \\ &+ \frac{(1-\mu\alpha_k)^2 + \alpha_k\gamma\rho}{1-\alpha_k\gamma\rho}(\theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k\|x_k - x^*\|\|x_k - x_{k-1}\| + M_3\alpha_k^2) \\ =& \left(1 - \frac{2\alpha_k(\mu - \gamma\rho)}{1-\alpha_k\gamma\rho}\right)\|x_k - x^*\|^2 + \frac{(\mu\alpha_k)^2}{1-\alpha_k\gamma\rho}\|x_k - x^*\|^2 \\ &+ \frac{2\alpha_k}{1-\alpha_k\gamma\rho}\langle\gamma g(x^*) - Dx^*, x_{k+1} - x^*\rangle\end{aligned}$$

$$
\begin{aligned}
&+\frac{(1-\mu\alpha_k)^2+\alpha_k\gamma\rho}{1-\alpha_k\gamma\rho}(\theta_k^2\|x_k-x_{k-1}\|^2+2\theta_k\|x_k-x^*\|\|x_k-x_{k-1}\|+M_3\alpha_k^2)\\
\le&\left(1-\frac{2\alpha_k(\mu-\gamma\rho)}{1-\alpha_k\gamma\rho}\right)\|x_k-x^*\|^2+\frac{2\alpha_k(\mu-\gamma\rho)}{1-\alpha_k\gamma\rho}\left\{\frac{\langle\gamma g(x^*)-Dx^*,x_{k+1}-x^*\rangle}{\mu-\gamma\rho}\right.\\
&+\frac{\mu^2\alpha_k}{2(\mu-\gamma\rho)}M_4^2+\frac{(1-\mu\alpha_k)^2+\alpha_k\gamma\rho}{2\alpha_k(\mu-\gamma\rho)}(\theta_k^2\|x_k-x_{k-1}\|^2\\
&\left.+2\theta_kM_4\|x_k-x_{k-1}\|+M_3\alpha_k^2)\right\},
\end{aligned}
$$

where the second inequality follows from $\alpha_k\in(0,1)$ and uses the fact that there exists $M_4>0$ such that $\|x_k-x^*\|\le M_4$ for all k.

Set $\xi_k=\frac{2\alpha_k(\mu-\gamma\rho)}{1-\alpha_k\gamma\rho}$ and

$$
\begin{aligned}
\delta_k=&\frac{\langle\gamma g(x^*)-Dx^*,x_{k+1}-x^*\rangle}{\mu-\gamma\rho}+\frac{\mu^2\alpha_k}{2(\mu-\gamma\rho)}M_4^2\\
&+\frac{(1-\mu\alpha_k)^2+\alpha_k\gamma\rho}{2\alpha_k(\mu-\gamma\rho)}(\theta_k^2\|x_k-x_{k-1}\|^2+2\theta_kM_4\|x_k-x_{k-1}\|+M_3\alpha_k^2).
\end{aligned}
$$

Then

$$
\|x_{k+1}-x^*\|^2\le(1-\xi_k)\|x_k-x^*\|^2+\xi_k\delta_k. \tag{12}
$$

Now, we arrive at the state of proving the convergence of sequence $\{x_k\}$. For this, we break the argument into two cases.

Case 1 There exists $k_0\in\mathbb{N}$ such that sequence $\{\|x_k-x^*\|^2\}_{k\ge k_0}$ is nonincreasing. In this case, sequence $\{\|x_k-x^*\|^2\}$ converges and

$$
\lim_{k\to\infty}(\|x_k-x^*\|^2-\|x_{k+1}-x^*\|^2)=0. \tag{13}
$$

It follows from (8) that

$$
\begin{aligned}
\|x_{k+1}-x^*\|^2&\le((1-\alpha_k(\mu-\gamma\rho))\|z_k-x^*\|+\alpha_k\|\gamma g(x^*)-Dx^*\|)^2\\
&\le(\|z_k-x^*\|+\alpha_k\|\gamma g(x^*)-Dx^*\|)^2\\
&=\|z_k-x^*\|^2+\alpha_k(2\|\gamma g(x^*)-Dx^*\|\|z_k-x^*\|+\alpha_k\|\gamma g(x^*)-Dx^*\|^2)\\
&\le\|z_k-x^*\|^2+2\alpha_kM_5,
\end{aligned}\tag{14}
$$

where the second inequality uses the fact that there exists $M_5>0$ such that $\|\gamma g(x^*)-Dx^*\|\|z_k-x^*\|+\frac{1}{2}\alpha_k\|\gamma g(x^*)-Dx^*\|^2\le M_5$ for all k.

On the other way, by $y_k=x_k+\theta_k(x_k-x_{k-1})$ and the fact $\|x_k-x^*\|\le M_4$ that

$$
\|y_k-x^*\|^2\le\|x_k-x^*\|^2+\theta_k^2\|x_k-x_{k-1}\|^2+2\theta_kM_4\|x_k-x_{k-1}\|. \tag{15}
$$

By condition (ii), it follows from (9), (14) and (15) that

$$
\begin{aligned}
&\frac{4\rho_k(1-\rho_k)(\|(I-P_{C_k})y_k\|^2+\|(I-P_{Q_k})Ay_k\|^2)^2}{\|(I-P_{C_k})y_k+A^\top(I-P_{Q_k})Ay_k\|^2}\\
&\le\|y_k-x^*\|^2-\|z_k-x^*\|^2+2\tau\beta_k\langle z_k-x^*,d_{k-1}\rangle\\
&\le\|x_k-x^*\|^2+\theta_k^2\|x_k-x_{k-1}\|^2+2\theta_kM_4\|x_k-x_{k-1}\|-\|x_{k+1}-x^*\|^2 \qquad (16)\\
&+2\alpha_kM_5+2\tau\beta_k\langle z_k-x^*,d_{k-1}\rangle\\
&\le\|x_k-x^*\|^2-\|x_{k+1}-x^*\|^2+2\theta_kM_4\|x_k-x_{k-1}\|+\theta_k^2\|x_k-x_{k-1}\|^2+2\alpha_kM_5+M_3\alpha_k^2. \qquad (17)
\end{aligned}
$$

Since $\lim_{k\to\infty} \theta_k \|x_k - x_{k-1}\| = 0$ by condition (iii), it follows from (13), and conditions (i)(ii) that

$$\lim_{k\to\infty} \frac{(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2)^2}{\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2} = 0. \tag{18}$$

Note that

$$\begin{aligned}\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2 \leq &(\|(I - P_{C_k})y_k\| + \|A\|\|(I - P_{Q_k})Ay_k\|)^2 \\ \leq &2\max\{1, \|A\|^2\}(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2).\end{aligned}$$

Then,

$$\begin{aligned}&\frac{(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2)^2}{\|(I - P_{C_k})y_k + A^\top(I - P_{Q_k})Ay_k\|^2} \\ &\geq \frac{1}{2}\min\left\{1, \frac{1}{\|A\|^2}\right\}(\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2).\end{aligned}$$

Hence, one has from (18) that

$$\lim_{k\to\infty} (\|(I - P_{C_k})y_k\|^2 + \|(I - P_{Q_k})Ay_k\|^2) = 0. \tag{19}$$

By (iii) in Lemma 1, (10) and (15), we have

$$\begin{aligned}\|P_{C_k}z_k - z_k\|^2 \leq &\|z_k - x^*\|^2 - \|P_{C_k}z_k - x^*\|^2 \\ \leq &\|y_k - x^*\|^2 + M_3\alpha_k^2 - \|P_{C_k}z_k - x^*\|^2 \\ \leq &\|x_k - x^*\|^2 + \theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k M_4\|x_k - x_{k-1}\| \\ &+ M_3\alpha_k^2 - \|P_{C_k}z_k - x^*\|^2.\end{aligned} \tag{20}$$

According to Algorithm 1, one has

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &= \|\alpha_k\gamma g(z_k) + (I - D\alpha_k)P_{C_k}z_k - x^*\|^2 \\ &= \|\alpha_k(\gamma g(z_k) - Dx^*) + (I - D\alpha_k)(P_{C_k}z_k - x^*)\|^2 \\ &\leq (\alpha_k\|\gamma g(z_k) - Dx^*\| + \|I - D\alpha_k\|\|P_{C_k}z_k - x^*\|)^2 \\ &\leq \|P_{C_k}z_k - x^*\|^2 + \alpha_k M_6,\end{aligned}$$

where the second inequality uses the fact that there exists $M_6 > 0$ such that $\alpha_k\|\gamma g(z_k) - Dx^*\|^2 + 2\|\gamma g(z_k) - Dx^*\|\|P_{C_k}z_k - x^*\| \leq M_6$ for all k and $\alpha_k < \frac{1}{\gamma}$. Then it follows from (20) and $y_k = x_k + \theta_k(x_k - x_{k-1})$ that

$$\begin{aligned}\|P_{C_k}z_k - z_k\|^2 \leq &\|y_k - x^*\|^2 + M_3\alpha_k^2 - \|P_{C_k}z_k - x^*\|^2 \\ \leq &\|x_k - x^*\|^2 + \theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k M_4\|x_k - x_{k-1}\| + M_3\alpha_k^2 \\ &- \|x_{k+1} - x^*\|^2 + \alpha_k M_6.\end{aligned}$$

This implies that

$$\lim_{k\to\infty} \|P_{C_k}z_k - z_k\| = 0. \tag{21}$$

From conditions (i), (iii) and Lemma 6, we derive the boundedness about $\{z_k\}$ and

$$\begin{aligned}\lim_{k\to\infty} \|x_{k+1} - P_{C_k}z_k\| &= \lim_{k\to\infty} \alpha_k\|\gamma g(z_k) - DP_{C_k}z_k\| = 0, \\ \lim_{k\to\infty} \|y_k - x_k\| &= \lim_{k\to\infty} \theta_k\|x_k - x_{k-1}\| = 0.\end{aligned} \tag{22}$$

Further, it follows from (19) and conditions (i)-(ii) that

$$\begin{aligned}\lim_{k\to\infty}\|z_k - y_k\| &= \lim_{k\to\infty}\| - t_k\nabla f_k(y_k) + \tau\beta_k d_{k-1}\| \\ &\le \lim_{k\to\infty}(\|t_k\|\|\nabla f_k(y_k)\| + {\alpha_k}^2\|\tau d_{k-1}\|) = 0. \end{aligned} \tag{23}$$

By the fact that

$$\|x_{k+1} - x_k\| \le \|x_{k+1} - P_{C_k}z_k\| + \|P_{C_k}z_k - z_k\| + \|z_k - y_k\| + \|y_k - x_k\|,$$

together with (21)–(23), we obtain

$$\lim_{k\to\infty}\|x_{k+1} - x_k\| = 0. \tag{24}$$

By the non-expansiveness of the projection operator, one has

$$\begin{aligned}\|x_k - P_{C_k}x_k\| &\le \|x_k - y_k\| + \|y_k - z_k\| + \|z_k - P_{C_k}z_k\| \\ &\quad + \|P_{C_k}z_k - P_{C_k}y_k\| + \|P_{C_k}y_k - P_{C_k}x_k\| \\ &\le 2\|x_k - y_k\| + 2\|y_k - z_k\| + \|z_k - P_{C_k}z_k\|.\end{aligned}$$

By (21), (22) and (23), one has

$$\lim_{k\to\infty}\|x_k - P_{C_k}x_k\| = 0. \tag{25}$$

Applying the fact

$$\begin{aligned}\|Ax_k - P_{Q_k}(Ax_k)\| &\le \|A(x_k - y_k)\| + \|Ay_k - P_{Q_k}(Ay_k)\| + \|P_{Q_k}(Ay_k) - P_{Q_k}(Ax_k)\| \\ &\le 2\|A\|\|x_k - y_k\| + \|Ay_k - P_{Q_k}(Ay_k)\|,\end{aligned}$$

to (19) and (22), one has

$$\lim_{k\to\infty}\|Ax_k - P_{Q_k}(Ax_k)\| = 0. \tag{26}$$

Since $w \in \partial q(y)$ is bounded on any bounded set, there exists a constant $M_7 > 0$ such as $\|\omega_k\| \le M_7$ for $k \in N$. For $\omega_k \in \partial q(Ax_k)$, from (26), one has

$$\begin{aligned}q(Ax_k) &\le \langle \omega_k, Ax_k - P_{Q_k}(Ax_k)\rangle \\ &\le M_7\|Ax_k - P_{Q_k}(Ax_k)\| \overset{k\to\infty}{\longrightarrow} 0.\end{aligned}$$

Lemma 6 shows that $\{x_k\}$ is bounded. Accordingly, there exists subsequence $\{x_{k_j}\}$ of $\{x_k\}$ such that $x_{k_j} \to \bar{x} \in \mathbb{R}^n$, and hence $Ax_{k_j} \to A\bar{x} \in \mathbb{R}^m$. Then, from the lower-semicontinuity of function q, it holds that

$$q(A\bar{x}) \le \liminf_{j\to\infty} q(Ax_{k_j}) = 0,$$

which shows that $A\bar{x} \in Q$.

Now, we prove that $\bar{x} \in C$. In fact, by definition of C_k, one has

$$\begin{aligned}c(x_k) &\le \langle \xi_k, x_k - P_{C_k}(x_k)\rangle \\ &\le M_8\|x_k - P_{C_k}(x_k)\|,\end{aligned}$$

where $M_8 > 0$ is a constant such that $\|\xi_k\| \le M_8$ for all $k \in \mathbb{N}$. Thus, it follows from (25) that

$$\lim_{k\to\infty} c(x_k) = 0.$$

Then, from the lower-semicontinuity of function c, it holds that

$$c(\bar{x}) \leq \liminf_{j\to\infty} c(x_{k_j}) = 0,$$

which shows $\bar{x} \in C$ and hence $\bar{x}$ is a solution of the SFP. By the arbitrariness of the cluster point of the sequence $\{x_k\}$, we conclude that any cluster point of sequence $\{x_k\}$ is a solution to the SFP.

Next, we show that $\limsup_{k\to\infty}\langle D\hat{x} - \gamma g(\hat{x}), \hat{x} - x_k\rangle \leq 0$ for any solution $\hat{x} \in \Gamma$ of the variational inequality

$$\langle D\hat{x} - \gamma g(\hat{x}), \hat{x} - x\rangle \leq 0,\ \forall x \in \Gamma.$$

By the basic property of the projection operator, we know that $\hat{x}$ is a solution of the variational inequality above iff $\hat{x} = P_\Gamma(I - D + \gamma g)\hat{x}$.

Since $\{x_k\}$ is bounded and its any cluster point is a solution of the SFP, it holds that

$$\limsup_{k\to\infty}\langle D\hat{x} - \gamma g(\hat{x}), \hat{x} - x_k\rangle \leq 0. \tag{27}$$

To show that $\{x_k\}$ globally converges to x^*. We take $x^* = \hat{x}$. For $\xi_k = \frac{2\alpha_k(\mu-\gamma\rho)}{1-\alpha_k\gamma\rho}$ and

$$\begin{aligned}\delta_k =&\frac{\langle\gamma g(x^*) - Dx^*, x_{k+1} - x^*\rangle}{\mu - \gamma\rho} + \frac{\mu^2\alpha_k}{2(\mu-\gamma\rho)}M_4^2\\ &+ \frac{(1-\mu\alpha_k)^2 + \alpha_k\gamma\rho}{2\alpha_k(\mu-\gamma\rho)}(\theta_k^2\|x_k - x_{k-1}\|^2 + 2\theta_k M_4\|x_k - x_{k-1}\| + M_3\alpha_k^2),\end{aligned}$$

It follows from condition (i) that $\xi_k \to 0$, $\sum_{k=1}^{\infty}\xi_k = \infty$. By (27), condition (i) and (iii), it holds that $\limsup_{k\to\infty}\delta_k \leq 0$. From Lemma 3, taking $\gamma_k = \xi_k$ and $a_k = \|x_k - x^*\|$, we have

$$\lim_{k\to\infty}\|x_k - x^*\| = 0.$$

Thus, $\{x_k\}$ converges to x^*.

Case 2 The sequence $\{\|x_k - x^*\|^2\}$ is not monotonically decreasing. Denote $\sigma_k = \|x_k - x^*\|^2$. Without loss of generality, we may suppose there exists subsequence $\{\sigma_{k_j}\}$ of sequence $\{\sigma_k\}$ such that $\sigma_{k_j} < \sigma_{k_j+1}$ for all $j \in \mathbb{N}$.

Define

$$\tau_k = \max\{t \leq k \mid \sigma_t < \sigma_{t+1}\}.$$

Then by Lemma 4, $\{\tau_k\}$ is a nondecreasing sequence with $\lim_{k\to\infty}\tau_k = \infty$ and $\sigma_{\tau_k} \leq \sigma_{\tau_k+1},\ \forall k \geq k_0$.

Following the argument of (16) and (18), one has

$$\begin{aligned}&\frac{4\rho_{\tau_k}(1-\rho_{\tau_k})(\|(I - P_{C_{\tau_k}})y_{\tau_k}\|^2 + \|(I - P_{Q_{\tau_k}})Ay_{\tau_k}\|^2)^2}{\|(I - P_{C_{\tau_k}})y_{\tau_k} + A^\top(I - P_{Q_{\tau_k}})Ay_{\tau_k}\|^2}\\ &\quad\leq \|x_{\tau_k} - x^*\|^2 - \|x_{\tau_k+1} - x^*\|^2 + 2\theta_{\tau_k}M_4\|x_{\tau_k} - x_{\tau_k-1}\| + \theta_{\tau_k}^2\|x_{\tau_k} - x_{\tau_k-1}\|^2\\ &\quad+ 2\alpha_{\tau_k}M_5 + M_3\alpha_{\tau_k}^2,\end{aligned}$$

and

$$\lim_{k\to\infty} \frac{(\|(I - P_{C_{\tau_k}})y_{\tau_k}\|^2 + \|(I - P_{Q_{\tau_k}})Ay_{\tau_k}\|^2)^2}{\|(I - P_{C_{\tau_k}})y_{\tau_k} + A^\top(I - P_{Q_{\tau_k}})Ay_{\tau_k}\|^2} = 0.$$

Following the argument of (24)-(26) and (27), we have

$$\begin{aligned}
&\lim_{k\to\infty} \|x_{\tau_k+1} - x_{\tau_k}\| = 0,\\
&\lim_{k\to\infty} (\|x_{\tau_k} - P_{C_{\tau_k}} x_{\tau_k}\| + \|Ax_{\tau_k} - P_{Q_{\tau_k}}(Ax_{\tau_k})\|) = 0,\\
&\limsup_{k\to\infty} \langle Dx^* - \gamma g(x^*), x^* - x_{\tau_k}\rangle \le 0.
\end{aligned}$$

Then for ξ_k and δ_k defined in (12), it holds that $\lim_{k\to\infty} \xi_{\tau_k} = 0$ and $\limsup_{k\to\infty} \delta_{\tau_k} \le 0$. From (12), one has

$$\|x_{\tau_k+1} - x^*\|^2 \le (1 - \xi_{\tau_k})\|x_{\tau_k} - x^*\|^2 + \xi_{\tau_k}\delta_{\tau_k}. \tag{28}$$

Furthermore,

$$\|x_{\tau_k} - x^*\|^2 \le \delta_{\tau_k}.$$

Thus,

$$\limsup_{k\to\infty} \|x_{\tau_k} - x^*\| \le 0,$$

and hence,

$$\lim_{k\to\infty} \|x_{\tau_k} - x^*\| = 0.$$

By (28),

$$\lim_{k\to\infty} \|x_{\tau_k+1} - x^*\|^2 \le \limsup_{k\to\infty} \|x_{\tau_k+1} - x^*\|^2 \le \limsup_{k\to\infty} \|x_{\tau_k} - x^*\|^2 = 0.$$

It follows from Lemma 4 that for all $k \ge k_0$,

$$0 \le \sigma_k \le \max\{\sigma_{\tau_k}, \sigma_{\tau_k+1}\} = \sigma_{\tau_k+1}.$$

Thus

$$\lim_{k\to\infty} \sigma_k = 0.$$

This means that $\{x_k\}$ globally converges to x^*, and the proof is completed. □

From [14], due to that $\|x_k - x_{k-1}\|$ is known before choosing θ_k, we can readily deduce condition (iii) in numerical computation. In practice, parameter θ_k can be taken as $0 \le \theta_k \le \overline{\theta}_k$, where

$$\overline{\theta}_k = \begin{cases} \min\left\{\dfrac{\varepsilon_k}{\|x_k - x_{k-1}\|}, \varepsilon_1\right\}, & x_k \ne x_{k-1},\\ \varepsilon_1, & otherwise, \end{cases}$$

where $\{\varepsilon_k\}$ is a positive sequence such that $\varepsilon_k = o(\alpha_k)$ and $\theta_k \in [0, \varepsilon_1] \subset [0, 1)$.

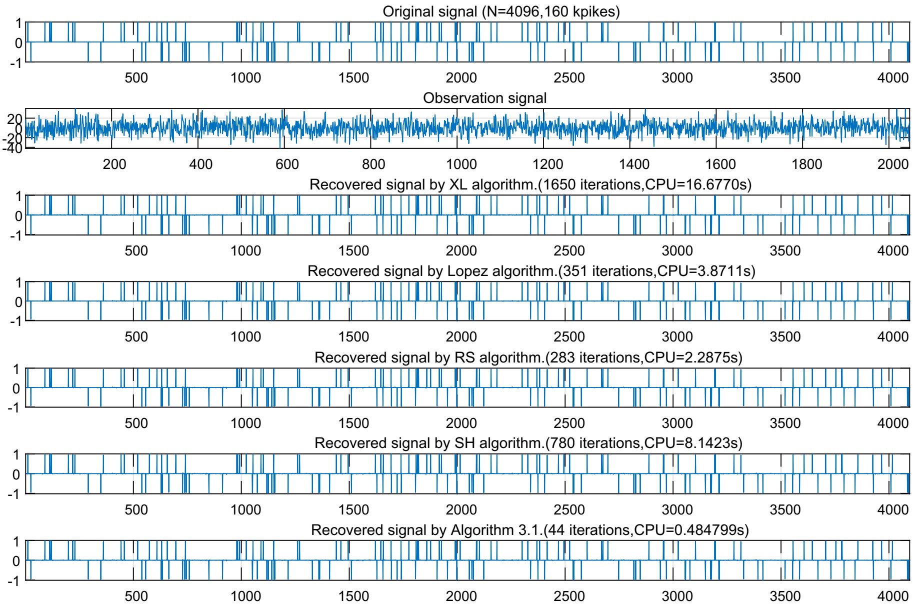


Fig. 1 From top to bottom: original signal, observation data, recovered signal by XL algorithm, Lо́pez algorithm, RS algorithm, SH algorithm, and Algorithm 3.1 in case $n = 4096$, $m = 2048$, $k = 160$

Fig. 2 Original images

Fig. 3 Degraded images

4 Experimental results

In this section, we will make some numerical experiments on signal processing and image recovery to illustrate the superiority of the proposed algorithm to the existing solution methods.

Fig. 4 Comparison of recovered images by using different algorithms, when the number of iterations is 500. From left to right: GK algorithms, ABA algorithms, PNN algorithms, Lṕez algorithms, RS algorithms, SH algorithms, and Algorithm 1

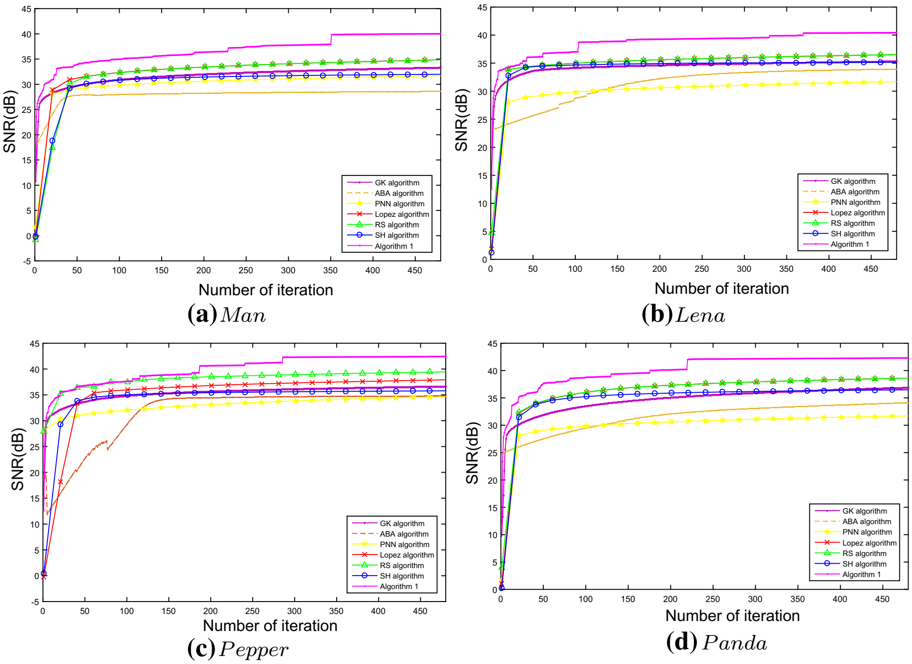


Fig. 5 Comparison of SNR for recovered images by using different algorithms

Table 1 Comparison of SNR by using seven algorithms, when the number of iterations is 500

	GK algorithm	ABA algorithm	PNN algorithm	López algorithm	RS algorithm	SH algorithm	Algorithm1
Man	33.37	28.60	28.21	34.83	36.81	31.99	40.02
Lena	35.36	33.96	34.06	36.57	36.60	35.17	40.43
Pepper	36.67	34.77	34.69	37.98	39.51	35.79	42.43
Panda	36.91	34.23	34.88	38.67	38.68	36.50	42.31

Table 2 Comparison of SSIM by using seven algorithms, when the number of iterations is 500

	GK algorithm	ABA algorithm	PNN algorithm	López algorithm	RS algorithm	SH algorithm	Algorithm1
Man	0.9935	0.9839	0.9843	0.9957	0.9973	0.9909	0.9986
Lena	0.9858	0.9749	0.9764	0.9906	0.9927	0.9827	0.9970
Pepper	0.9919	0.9824	0.9822	0.9948	0.9965	0.9896	0.9984
Panda	0.9943	0.9900	0.9910	0.9961	0.9966	0.9929	0.9985

4.1 Signal processing

It is well known that the compressed sensing is a very active research and application area which is based on in the fact that an N-sample signal x with k nonzero elements can available from $k \ll m < n$. More specifically, compressed sensing can be expressed as the following equation

$$y = Ax + \varepsilon, \tag{29}$$

where $x \in \mathbb{R}^n$ is the data to be recovered, $y \in \mathbb{R}^m$ is the vector of observation or measurement that has noised, ε is on behalf of the noise, and $A : \mathbb{R}^n \to \mathbb{R}^m$ is a bounded observation operator. To improve the original image in (29), we need to consider the minimize the value of ε by using the model

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|y - Ax\|^2 + \gamma_1\|x\|_1, \tag{30}$$

where γ_1 is a positive parameter, and $\|\cdot\|_1$ is the ℓ_1-norm, $\|\cdot\|$ is the ℓ_2-norm. The problem (30) involves minimizing an objective function, which has a quadratic error term. Using the theory of convex analysis, we can prove the solution of the constrained least squares problem by using the minimizer

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|y - Ax\|^2 \text{ subject to } \|x\|_1 \le s, \tag{31}$$

where a proper choice $s > 0$.

It is obvious that problem (31) is a special solution to the SFP where $C = \{x \in \mathbb{R}^n \mid \|x\|_1 \le s\}$ and $Q = \{y\}$. In our experiment, we generated a sparse signal that $x \in \mathbb{R}^n$ which $n = 4096$ with $k = 160$ spikes, each spike has amplitude $\{-1, 1\}$. The matrix $A \in \mathbb{R}^{m\times n}$ which $m = 2048$. This signal is plotted at the top of Fig. 1. In this section,we compare our method with similar methods such as XL algorithm [19], López algorithm [8], RS algorithm[15], SH algorithm[13]. According to (29) with white Gaussian noise of variance $\varepsilon^2 = 10^{-4}$. The process shows that the mean squared error is effective in precision, which is started with $s = k$, the initial point $x_1 = (1, 1, \cdots, 1)^T, x_0 = (0, 0, \cdots, 0)^T, \tau = 0.6, \mu = 0.8, \gamma = 0.8$, $\rho_k = 0.7$, $\alpha_k = \frac{1}{10000n}$, $\beta_k = 0.7\alpha_k^2$, $g(x) = 0.9x$ and define $\theta_k = \overline{\theta}_k$ where $\varepsilon_k = \frac{1}{n^3}$ and

Fig. 6 The comparison recovered images of faces from the ORL dataset by using seven algorithms, when the number of iterations is 500. From the top to the bottom: original images, degraded images, recovered images by GK algorithms, ABA algorithms, PNN algorithms, Lόpez algorithms, RS algorithms, SH algorithms, and Algorithm 1

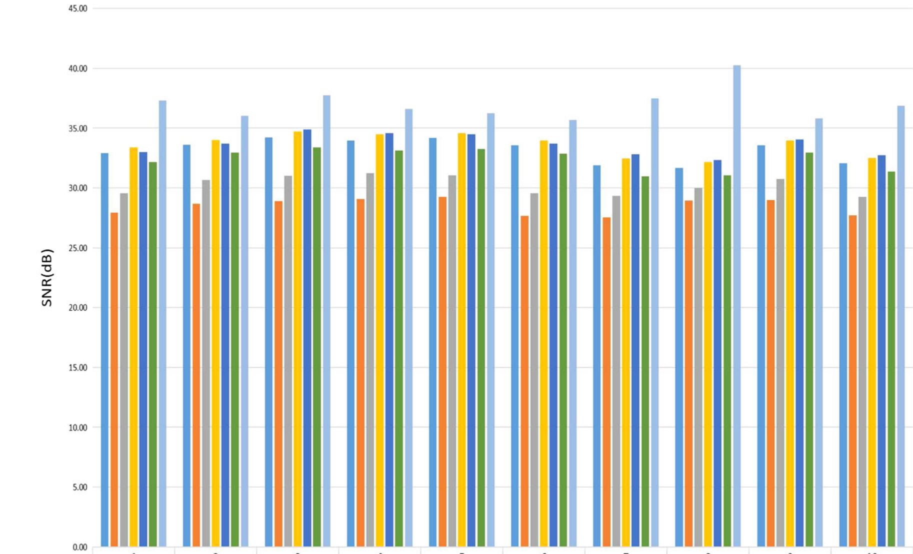
$\varepsilon_1 = 0.5$. The restoration accuracy is measured by the mean squared error as follows

$$MSE = \frac{1}{n}\|\bar{x} - x\|^2 \leq 6 \times 10^{-5},$$

where $\bar{x}$ is the original signal of x. Obviously, Fig. 1 shows that the proposed method can recover sparse signals in fewer iterations and less CPU time.

4.2 Image recovery

In this section, we will make some numerical experiments on some image recovery problems to illustrate the superiority of the proposed algorithm to the existing solution methods. For image recovery problem, it is well known that all greyscale images have $E = M \times N$ pixels, and each pixel value is known to be in the range [0, 255]. We take the hard constraint



	1	2	3	4	5	6	7	8	9	10
GK algorithm	32.91	33.60	34.21	33.95	34.19	33.55	31.91	31.69	33.55	32.06
ABA algorithm	27.95	28.67	28.88	29.08	29.26	27.65	27.53	28.96	28.99	27.71
PNN algorithm	29.58	30.68	31.02	31.25	31.04	29.57	29.34	29.98	30.73	29.26
Lopez algorithm	33.38	34.02	34.70	34.49	34.59	33.98	32.46	32.16	33.94	32.50
RS algorithm	32.98	33.72	34.89	34.57	34.50	33.68	32.82	32.32	34.05	32.73
SH algorithm	32.15	32.94	33.40	33.11	33.25	32.87	30.97	31.07	32.95	31.36
Algorithm 1	37.32	36.04	37.74	36.58	36.23	35.68	37.47	40.24	35.80	36.88

Fig. 7 Comparison of SNR for faces from the ORL dataset by using seven algorithms, when the number of iterations is 500

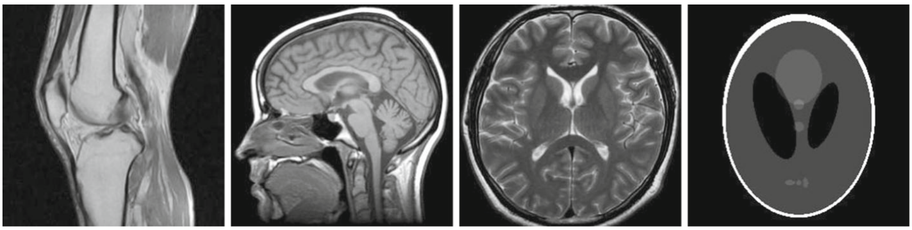


Fig. 8 Original medical images

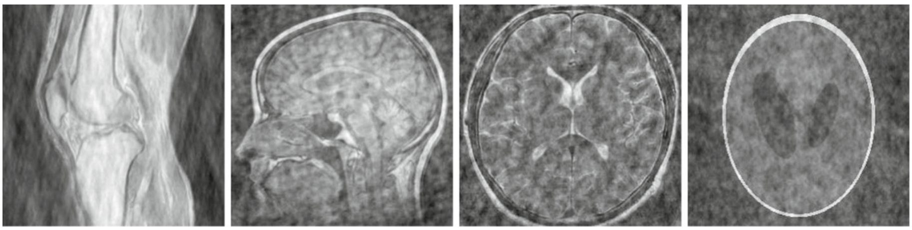


Fig. 9 Sampled images with a sampling rate of 60 percent

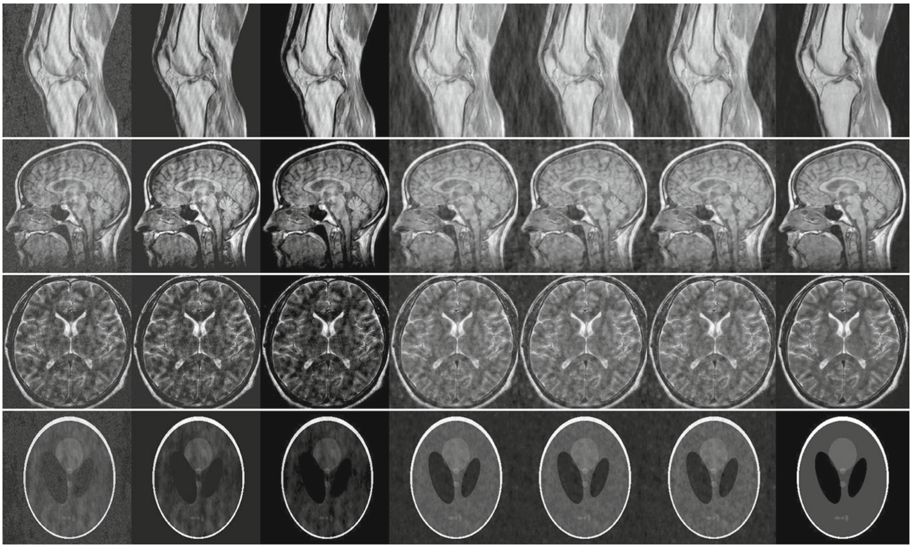


Fig. 10 Comparison of recovered images by using different algorithms, when the number of iterations is 500. From left to right: GK algorithms, ABA algorithms, PNN algorithms, Lóipez algorithms, RS algorithms, SH algorithms, and Algorithm 1

$C := [0, 255]^E$. The image recovery problem can be formulated by the following minimizer

$$\min_{\bar{x} \in C} \|y - A\bar{x}\|,$$

where $\bar{x}$ is an original image, y is the observed image and A is a blurring matrix. In order to measure the quality of recovered images, we choose the signal-to-noise ratio(SNR) in decibel(dB) as

$$SNR := 20 \log_{10} \frac{\|\bar{x}\|}{\|x - \bar{x}\|},$$

where $\bar{x}$ is an original image, and x is a restored image. In other words, the bigger SNR means the better restored image.

Figure 2 shows that the original images of Man, Lena and Panda are 256×256 pixels and the original images of Pepper is 337×337 pixels. In Fig. 3, we display some blurred images which were averaged blur from original images. Measuring the SNR of degraded images of the man, the Lena, the pepper and the panda, the results are 26.28 dB, 28.37 dB, 29.49 dB and 27.26 dB, respectively.

To show the superiority of our designed method, we make a comparison of our method with GK algorithm [20], ABA algorithm [21], PNN algorithm [7], Lóipez algorithm [8], RS algorithm[15], SH algorithm[13]. In Fig. 4, the comparison of recovered images is revealed by seven different algorithms when the number of iterations is 500. Figure 5 demonstrates the trend of the SNR values in dB of seven different algorithms. The comparison of our method with SNR is listed in Table 1, and SSIM of restored images is illustrated in Table 2. It can be seen that the restored images of Algorithm 1 has a good efficiency in dealing with degraded images.

Now, we investigate our method to the ORL dataset of faces. In Fig. 6, original images, degraded images, recovered images given by seven different algorithms are shown from the

Table 3 Comparison of SNR by using seven algorithms, when the number of iterations is 500

	GK algorithms	ABA algorithms	PNN algorithm	Lǿpez algorithm	RS algorithm	SH algorithm	Algorithm1
Leg	18.06	18.68	18.70	24.46	24.53	24.46	30.19
Head	15.85	18.65	18.70	22.01	22.17	22.03	28.69
Brain	13.30	17.29	17.34	21.02	22.19	21.88	28.44
Phantom	12.10	11.07	11.04	26.91	26.52	24.54	70.64

top to the bottom, when the number of iterations is 500. In Fig. 7, the comparison SNR of faces from the ORL dataset are illustrated by using seven algorithms, when the number of iterations is 500. Obviously, the proposed recovery method generates unambiguous face images.

We implement our method to medical images in Fig. 8. In Fig. 9, we can check some images which were randomly sampled by sixty percent from original images. Measuring the SNR of sampled images of Leg, Head, Brain and Phantom, the results are 18.54 dB, 15.97 dB, 15.79 dB and 13.29 dB. In Fig. 10, the comparison of sampled recovered images is displayed by seven different algorithms when the number of iterations is 500. In Table 3, the comparison SNR of recovered sampled images is revealed by using seven algorithms. Clearly, the quality of images recovered by proposed method is better than the others.

Finally, we come to the conclusion that our result provides higher SNR values than those of GK algorithm, ABA algorithm, PNN algorithm, Lǿpez algorithm, RS algorithm and SH algorithm. This implies that the restored images of our algorithm have the better quality and efficiency.

5 Conclusions

In this paper, we developed a relaxed inertial and viscosity method to solve the split feasible problem. Under suitable conditions, the global convergence of the designed algorithm was established. As applications, we implemented our method to solve image recovery problem. Numerical experiments show that our method has a higher efficiency than the well-known methods in the literature.

Acknowledgements H. Che was supported in part by National Natural Science Foundation of China (NSFC) (No. 11401438) and Natural Science Foundation of Shandong Province (No. ZR2019MA022, ZR2020MA027). Y. Wang was supported in part by NSFC (No. 12071250) and Shandong Provincial Natural Science Foundation of Distinguished Young Scholars (No.ZR2021JQ01) . H. Chen was supported in part by National Natural Science Foundation of China (NSFC) (No. 12071249).

Data availability The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

References

1. Byrne, C.: Iterative oblique projection onto convex sets and the split feasibility problem. Inverse Probl. **18**, 441–53 (2002)
2. Calamai, P., Morè, J.: Projected gradient methods for linearly constrained problems. Math. Program. **39**, 93–116 (1987)
3. Candés, E., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**(2), 489–509 (2006)
4. Censor, Y., Elfving, T.: A multiprojection algorithm using Bregman projection in a product space. Numer. Algorithms **8**, 221–239 (1994)
5. Clarke, F.H.: Optimization and nonsmooth analysis (2nd edn), Classics in Applied Mathematics 5. SIAM, Philadelphia (1990)
6. Dang, Y., Sun, J., Xu, H.: Inertial accelerated algorithm for solving a split feasibility problem. J. Ind. Manag. Optim. **13**, 1383–1394 (2017)
7. Liu, Y., Hu, J.: A neural network for $l_1 - l_2$ minimization based on scaled gradient projection: Application to compressed sensing. Neurocomputing **173**(3), 988–993 (2016)
8. López, G., Martn-Márquez, V., Wang, F., Xu, H.: Solving the split feasibility problem without prior knowledge of matrix norm. Inverse Probl. **8**, 374–389 (2012)
9. Maingé, P.: Strong convergence of projected subgradient methods for nonsmooth and non-strictly convex minimization. Set Valued Var. Anal. **16**, 899–912 (2008)
10. Marino, G., Xu, H.: A general iterative method for nonexpansive mappings in Hilbert spaces. J. Math. Anal. Appl. **318**(1), 43–52 (2006)
11. Nocedal, J., Wright, S.: Numerical optimization. Springer-Verlag, New York (1999)
12. Polyak, B.: Some methods of speeding up the convergence of iteration methods. Ussr Comput. Math. Math. Phys. **4**(5), 1–17 (1964)
13. Sahu, D., Cho, Y., Dong, Q., Kashyap, M., Li, X.: Inertial relaxed CQ algorithms for solving a split feasibility problem in Hilbert spaces. Numer. Algorithms **87**, 1075–1095 (2021)
14. Suantai, S., Pholasa, N., Cholamjiak, P.: The modified inertial relaxed CQ algorithm for solving the split feasibility problems. J. Ind. Manag. Optim. **14**(4), 988–993 (2018)
15. Suparatulatorn, R., Charoensawan, P., Poochinapan, K.: Inertial self-adaptive algorithm for solving split feasible problems with applications to image restoration. Math. Methods Appl. Sci. **3**(42), 7268–7284 (2019)
16. Xu, H.: Iterative algorithms for nonlinear operators. J. London Math. Soc. **66**(1), 240–256 (2002)
17. Yang, Q.: On variable-step relaxed projection algorithm for variational inequalities. J. Math. Anal. Appl. **302**, 166–179 (2005)
18. Zarantonello, E.H.: Projections on convex sets in Hilbert space and spectral theory, Contributions to Nonlinear Functional Analysis. Academic Press, New York (1971)
19. Xu B., Liu Q., Huang T.: A discrete-time projection neural network for sparse signal reconstruction with application to face recognition. IEEE Trans. Neural Netw. Learn. Syst. (2018)
20. Garg, K., Baranwal, M.: CAPPA: continuous-time accelerated proximal point algorithm for sparse recovery. IEEE Signal Process. Lett. **27**, 1760–1764 (2020)
21. Abubakar, A., Kumam, P., Mohammad, H., Awwa, A.: A Barzilai-Borwein gradient projection method for sparse signal and blurred image restoration. J. Frankl. Inst. **357**(11), 7266–7285 (2020)
22. Sakurai, K., Iiduka, H.: Acceleration of the halpern algorithm to search for a fixed point of a nonexpansive mapping. Fixed Point Theory Appl. **2014**(1), 202 (2014)