

RESEARCH

Open Access

Adaptively relaxed algorithms for solving the split feasibility problem with a new step size

Haiyun Zhou^{1,2} and Peiyuan Wang^{2,3*}

Dedicated to Professor Shih-sen Chang on his 80th birthday

*Correspondence:

wangpy629@163.com

²Department of Mathematics,
Shijiazhuang Mechanical
Engineering College, Shijiazhuang,
050003, China

³The Second Training Base, Naval
Aviation Institution, Huludao,
125001, China

Full list of author information is
available at the end of the article

Abstract

In the present paper, we propose several kinds of adaptively relaxed iterative algorithms with a new step size for solving the split feasibility problem in real Hilbert spaces. The proposed algorithms never terminate, while the known algorithms existing in the literature may terminate. Several weak and strong convergence theorems of the proposed algorithms have been established. Some numerical experiments are also included to illustrate the effectiveness of the proposed algorithms.

MSC: Primary 46E20; 47J20; secondary 47J25

Keywords: split feasibility problem; adaptive step size; adaptively relaxed iterative algorithm; convergence

1 Introduction

Since its inception in 1994, the split feasibility problem SFP [1] has been attracting researchers' interest [2, 3] due to its extensive applications in signal processing and image reconstruction [4], with particular progress in intensity-modulated radiation therapy [5, 6].

Let H_1 and H_2 be real Hilbert spaces, C and Q be nonempty closed convex subsets of H_1 and H_2 , respectively, and $A : H_1 \rightarrow H_2$ a bounded linear operator. Then SFP can be formulated as finding a point \hat{x} with the property

$$\hat{x} \in C \quad \text{and} \quad A\hat{x} \in Q. \quad (1.1)$$

The set of solutions for SFP (1.1) is denote by $\Gamma = C \cap A^{-1}(Q)$.

Over the past two decades years or so, the researchers invested and designed various iterative algorithms for solving SFP (1.1); see [6–13]. The most popular algorithm, among them, is Byrne's CQ algorithm, which generates a sequence $\{x_n\}$ by the recursive procedure

$$x_1 \in H_1, \quad x_{n+1} = P_C(x_n - \tau_n A^*(I - P_Q)Ax_n), \quad n \geq 1, \quad (1.2)$$

where the step size τ_n is chosen in the open interval $(0, 2/\|A\|^2)$, while P_C and P_Q are the orthogonal projections onto C and Q , respectively.

We remark in passing that Byrne's CQ algorithm (1.2) is indeed a special case of the classical gradient projection method (GPM). To see this, let us define $f : H_1 \rightarrow \mathbb{R}$ by

$$f(x) = \frac{1}{2} \|(I - P_Q)Ax\|^2, \tag{1.3}$$

then the convex objective f is differentiable and has a Lipschitz gradient given by

$$\nabla f(x) = A^*(I - P_Q)Ax. \tag{1.4}$$

We consider the following convex minimization problem:

$$\min_{x \in C} f(x). \tag{1.5}$$

It is well known that $\hat{x} \in C$ is a solution of problem (1.5) if and only if

$$\langle \nabla f(\hat{x}), x - \hat{x} \rangle \geq 0, \quad \forall x \in C. \tag{1.6}$$

Also, we know that (1.6) holds true if and only if

$$\hat{x} = P_C(I - \tau \nabla f)\hat{x}, \quad \forall \tau > 0. \tag{1.7}$$

Note that if $\Gamma \neq \emptyset$, then $\hat{x} \in \Gamma \Leftrightarrow f(\hat{x}) = \min_{x \in C} f(x) \Leftrightarrow$ (1.6) holds \Leftrightarrow (1.7) holds. Consequently, we can utilize the classical gradient projection method (GPM) below to solve SFP (1.1):

$$x_1 \in C, \quad x_{n+1} = P_C(x_n - \tau_n \nabla f(x_n)), \quad n \geq 1, \tag{1.8}$$

where $\tau_n \in (0, 2/L)$, while L is the Lipschitz constant of ∇f . Noting that $L = \|A\|^2$, we see immediately that (1.8) is exactly CQ algorithm (1.2).

We note that, in algorithms (1.2) and (1.8) mentioned above, the choice of the step size τ_n depends heavily on the operator (matrix) norm $\|A\|$. This means that for actual implementation of CQ algorithm (1.2), one has first to know at least an upper bound of the operator (matrix) norm $\|A\|$, which is in general difficult. To overcome this difficulty, several authors proposed several various of adaptive methods, which permit the step size τ_n to be selected self-adaptively; see [7–9].

Yang [10] considered the following step size:

$$\tau_n := \frac{\rho_n}{\|\nabla f(x_n)\|}, \tag{1.9}$$

where $\{\rho_n\}$ is a sequence of positive real numbers such that

$$\sum_{n=0}^{\infty} \rho_n = \infty, \quad \sum_{n=0}^{\infty} \rho_n^2 < \infty. \tag{1.10}$$

Very recently, López *et al.* [11] introduced another choice of the step size sequence $\{\tau_n\}$ as follows:

$$\tau_n := \frac{\rho_n f_n(x_n)}{\|\nabla f_n(x_n)\|^2}, \tag{1.11}$$

where $\{\rho_n\}$ is chosen in the open interval $(0, 4)$. By virtue of the step size (1.11), López *et al.* [11] introduced four kinds of algorithms for solving SFP (1.1).

We observe that if $\nabla f(x_n) = 0$ for some $n \geq 1$, then the algorithms introduced by López *et al.* [11] have to terminate in the n th step of iterations. In this case x_n is not necessarily a solution of the SFP (1.1), since x_n may be not in C , Algorithm 4.1 in [11] is such a case. To make up the flaw, we introduce a new choice of the step size sequence $\{\tau_n\}$ as follows:

$$\tau_n := \frac{\rho_n f_n(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2}, \quad (1.12)$$

where $\{\rho_n\}$ is chosen in the open interval $(0, 4)$ and $\{\sigma_n\}$ is a sequence of positive numbers in $(0, 1)$, while f_n and ∇f_n are given by, respectively,

$$f_n(x) = \frac{1}{2} \|(I - P_{Q_n})Ax\|^2, \quad (1.13)$$

$$\nabla f_n(x) = A^*(I - P_{Q_n})Ax, \quad (1.14)$$

where $\{Q_n\}$ will be defined in Section 3.

The purpose of this paper is to introduce a new choice of the step size sequence $\{\tau_n\}$ that makes the associated algorithms never terminate. A new stop rule is also given, which ensures that the $(n + 1)$ th iteration x_{n+1} is a solution of SFP (1.1) and the iterative process stops. Several weak and strong convergence results are presented. Numerical experiments are included to illustrate the effectiveness of the proposed algorithms and the applications in signal processing of the CQ algorithm with the step size selected in this paper.

The rest of this paper is organized as follows. In the next section, some necessary concepts and important facts are collected. The weak and strong convergence theorems of the proposed algorithms with step size (1.12) are established in Section 3. Finally in Section 4, we provide some numerical experiments to illustrate the effectiveness and applications of the proposed algorithms with step size (1.12) to inverse problems arising from signal processing.

2 Preliminaries

Throughout this paper, we assume that SFP (1.1) is consistent, *i.e.*, $\Gamma \neq \emptyset$. We denote by \mathbb{R} the set of real numbers. Let H_1 and H_2 real Hilbert spaces and the letter I the identity mapping on H_1 or H_2 . If $f : H_1 \rightarrow \mathbb{R}$ is a differentiable (subdifferentiable) functional, then we denote by ∇f (∂f) the gradient (subdifferential) of f . Given a sequence $\{x_n\}$ in H , $w_w(x_n)$ (resp. $x_n \rightharpoonup x$) denotes the strong (resp. weak) convergence of $\{x_n\}$ to x . The symbols $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote inner product and norm of Hilbert spaces H_1 and H_2 , respectively. Let $T : H_1 \rightarrow H_1$ be a mapping. We use $\text{Fix}(T)$ to denote the set of fixed points of T . We also denote by $\text{dom}(T)$ the domain of T .

Some equalities in Hilbert space H_1 play very important roles for solving linear and non-linear problems arising from real world.

It is well known that in a real Hilbert space H_1 , the following two equalities hold:

$$\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y \rangle + \|y\|^2, \quad (2.1)$$

for all $x, y \in H_1$.

$$\|tx + (1-t)y\|^2 = t\|x\|^2 + (1-t)\|y\|^2 - t(1-t)\|x-y\|^2, \tag{2.2}$$

for all $x, y \in H_1$ and $t \in \mathbb{R}$.

Recall that a mapping $T : \text{dom}(T) \subset H_1 \rightarrow H_1$ is said to be

(i) nonexpansive if

$$\|Tx - Ty\| \leq \|x - y\|, \tag{2.3}$$

for all $x, y \in \text{dom}(T)$;

(ii) firmly nonexpansive if

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(I - T)x - (I - T)y\|^2, \tag{2.4}$$

for all $x, y \in \text{dom}(T)$;

(iii) λ -averaged if there exist some $\lambda \in (0, 1)$ and another nonexpansive mapping

$S : H_1 \rightarrow H_2$ such that

$$T = (1 - \lambda)I + \lambda S. \tag{2.5}$$

The following proposition describes the characterizations of firmly nonexpansive mappings (see [12]).

Proposition 2.1 *Let $T : \text{dom}(T) \subset H_1 \rightarrow H_1$ be a mapping. Then the following statements are equivalent.*

- (i) T is firmly nonexpansive;
- (ii) $I - T$ is firmly nonexpansive;
- (iii) $\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty \rangle$ for all $x, y \in H_1$;
- (iv) T is $\frac{1}{2}$ -averaged;
- (v) $2T - I$ is nonexpansive.

Recall that the metric (nearest point) projection from H_1 onto a nonempty closed convex subset C of H_1 is defined as follows: for each $x \in H_1$, there exists a unique point $P_C x \in C$ with the property:

$$\|x - P_C x\| \leq \|x - y\|, \quad \forall y \in C. \tag{2.6}$$

Now we list some basic properties of P_C below; see [12] for details.

Proposition 2.2

(p₁) *Given $x \in H$ and $z \in C$. Then $z = P_C x$ if and only if we have the inequality*

$$\langle x - z, y - z \rangle \leq 0, \quad \forall y \in C; \tag{2.7}$$

(p₂)

$$\|P_Cx - P_Cy\|^2 \leq \langle x - y, P_Cx - P_Cy \rangle, \quad \text{for all } x, y \in H_1; \quad (2.8)$$

(p₃)

$$\|(I - P_C)x - (I - P_C)y\|^2 \leq \langle (I - P_C)x - (I - P_C)y, x - y \rangle, \quad (2.9)$$

for all $x, y \in H_1$.

(p₄) $2P_C - I$ is nonexpansive;

(p₅)

$$\|P_Cx - P_Cy\|^2 \leq \|x - y\|^2 - \|(I - P_C)x - (I - P_C)y\|^2, \quad x, y \in H_1; \quad (2.10)$$

in particular,

(p₆)

$$\|P_Cx - y\|^2 \leq \|x - y\|^2 - \|(I - P_C)x\|^2, \quad \text{for all } x \in H_1 \text{ and } y \in C. \quad (2.11)$$

From (p₂), (p₃), and (p₄), we see immediately that both P_C and $(I - P_C)$ are firmly non-expansive and $\frac{1}{2}$ -averaged.

Recall that a function $f : H_1 \rightarrow \mathbb{R}$ is called convex if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall \lambda \in (0, 1), \forall x, y \in H_1.$$

It is well known that a differentiable function f is convex if and only if we have the relation

$$f(z) \geq f(x) + \langle \nabla f(x), z - x \rangle, \quad \forall z \in H_1.$$

Recall that an element $\xi \in H_1$ is said to be a subgradient of $f : H_1 \rightarrow \mathbb{R}$ at x if

$$f(z) \geq f(x) + \langle \xi, z - x \rangle, \quad \forall z \in H_1.$$

If the function $f : H_1 \rightarrow \mathbb{R}$ has at least one subgradient at x , it is said to be subdifferentiable at x . The set of subgradients of f at the point x is called the subdifferential of f at x , and is denoted by $\partial f(x)$. A function f is called subdifferentiable if it is subdifferentiable at every $x \in H_1$. If f is convex and differentiable, then $\partial f(x) = \{\nabla f(x)\}$ for every $x \in H_1$. A function f is called subdifferentiable if it is subdifferentiable at every $x \in H_1$. If f is convex and differentiable, then $\partial f(x) = \{\nabla f(x)\}$ for every $x \in H_1$. A function $f : H_1 \rightarrow \mathbb{R}$ is said to be weakly lower semi-continuous (w-lsc) at x if $x_n \rightharpoonup x$ implies

$$f(x) \leq \liminf_n f(x_n).$$

f is said to be w-lsc on H_1 if it is w-lsc at every point $x \in H_1$.

It is well known that for a convex function $f : H_1 \rightarrow \mathbb{R}$, it is w-lsc on H_1 if and only if it is lsc on H_1 .

It is an easy exercise to prove the following conclusions (see [13, 14]).

Proposition 2.3 *Let f be given as in (1.3). Then the following conclusions hold.*

- (i) f is convex and differentiable;
- (ii) $\nabla f(x) = A^*(I - P_Q)Ax, x \in H_1$;
- (iii) f is w -lsc on H_1 ;
- (iv) ∇f is $\|A\|^2$ -Lipschitz:

$$\|\nabla f(x) - \nabla f(y)\| \leq \|A\|^2 \|x - y\|, \quad x, y \in H.$$

The concept of Fejér monotonicity plays a key role in establishing weak convergence theorems. Recall that a sequence $\{x_n\}$ in H_1 is said to be Fejér monotone with respect to (w.r.t.) a nonempty closed convex subset C in H_1 if

$$\|x_{n+1} - z\| \leq \|x_n - z\|, \quad \forall n \geq 1, \forall z \in C.$$

Proposition 2.4 (see [11, 15]) *Let C be a nonempty closed convex in H_1 . If the sequence $\{x_n\}$ is Fejér monotone w.r.t. C , then the following hold:*

- (i) $x_n \rightharpoonup \hat{x}$ if and only if $w_w(x_n) \subset C$;
- (ii) the sequence $\{P_C x_n\}$ converges strongly;
- (iii) if $x_n \rightharpoonup \hat{x} \in C$, then $\hat{x} = \lim_n P_C x_n$.

Proposition 2.5 (see [16]) *Let $\{\alpha_n\}$ be a sequence of nonnegative real numbers such that*

$$\alpha_{n+1} \leq (1 - t_n)\alpha_n + t_n b_n, \quad n \geq 1,$$

where $\{t_n\}$ is a sequence in $(0, 1)$ and b_n is a sequence in \mathbb{R} such that

- (i) $\sum_{n=1}^{\infty} t_n = \infty$;
- (ii) $\overline{\lim} b_n \leq 0$ or $\sum_{n=1}^{\infty} |t_n b_n| < \infty$. Then $\alpha_n \rightarrow 0$ ($n \rightarrow \infty$).

3 Main results

Let $c : H_1 \rightarrow \mathbb{R}$ and $q : H_2 \rightarrow \mathbb{R}$ be convex functions and define level sets of c and q as follows:

$$C = \{x \in H_1 | c(x) \leq 0\} \quad \text{and} \quad Q = \{y \in H_2 | q(y) \leq 0\}. \quad (3.1)$$

Assume that both c and q are subdifferentiable on H_1 and H_2 , respectively, and that ∂c and ∂q are bounded mappings. Given an arbitrary initial data $x_1 \in H_1$. Assume that x_n is the current value for $n \geq 1$. We introduce two sequences of half-spaces as follows:

$$C_n = \{x \in H_1 | c(x_n) \leq \langle \xi_n, x_n - x \rangle\}, \quad (3.2)$$

where $\xi_n \in \partial c(x_n)$, and

$$Q_n = \{y \in H_2 | q(Ax_n) \leq \langle \eta_n, Ax_n - y \rangle\}, \quad (3.3)$$

where $\eta_n \in \partial q(Ax_n)$. Clearly, $C \subseteq C_n$ and $Q \subseteq Q_n$ for all $n \geq 1$.

Construct x_{n+1} via the formula

$$x_{n+1} = P_{C_n}(x_n - \tau_n \nabla f(x_n)), \tag{3.4}$$

where $\{\tau_n\}$ is given as (1.12),

$$f_n(x) = \frac{1}{2} \|(I - P_{Q_n})Ax\|^2 \tag{3.5}$$

and

$$\nabla f_n(x) = A^*(I - P_{Q_n})Ax, \tag{3.6}$$

More precisely, we introduce the following relaxed CQ algorithm in an adaptive way.

Algorithm 3.1 Choose an initial data $x_1 \in H_1$ arbitrarily. Assume that the n th iterate x_n has been constructed then we compute the $(n + 1)$ th iteration x_{n+1} via the formula:

$$x_{n+1} = P_{C_n}(x_n - \tau_n \nabla f(x_n)), \quad n \geq 1, \tag{3.7}$$

where the step size τ_n is chosen in such a way that

$$\tau_n := \frac{\rho_n f_n(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2}, \tag{3.8}$$

with $0 < \rho_n < 4$ and $0 < \sigma_n < 1$. If $x_{n+1} = x_n$ for some $n \geq 1$, then x_n is a solution of the SFP (1.1) and the iterative process stops; otherwise, we set $n := n + 1$ and go on to (3.7) to compute the next iteration x_{n+2} .

We remark in passing that if $x_{n+1} = x_n$ for some $n \geq 1$, then $x_m = x_n$ for all $m \geq n + 1$, consequently, $\lim_{m \rightarrow \infty} x_m = x_n$ is a solution of SFP (1.1). Thus, we may assume that the sequence $\{x_n\}$ generated by Algorithm 3.1 is infinite.

Theorem 3.2 Assume that $\liminf_n \rho_n(4 - \rho_n) \geq \rho > 0$. Then the sequence $\{x_n\}$ generated by Algorithm 3.1 converges weakly to a solution \hat{x} of SFP (1.1), where $\hat{x} = \lim_{n \rightarrow \infty} P_\Gamma x_n$.

Proof Let $z \in \Gamma$ be fixed, and set $y_n = x_n - \tau_n \nabla f_n(x_n)$. By virtue of (2.1), (3.7), and Proposition 2.2(p₆), we have

$$\begin{aligned} \|x_{n+1} - z\|^2 &= \|P_{C_n}y_n - z\|^2 \\ &\leq \|y_n - z\|^2 - \|y_n - P_{C_n}y_n\|^2 \\ &= \|x_n - z - \tau_n \nabla f_n(x_n)\|^2 - \|y_n - P_{C_n}y_n\|^2 \\ &= \|x_n - z\|^2 - 2\tau_n \langle \nabla f_n(x_n), x_n - z \rangle - \|y_n - P_{C_n}y_n\|^2 \\ &\quad + \tau_n^2 \|\nabla f_n(x_n)\|^2. \end{aligned} \tag{3.9}$$

In view of Proposition 2.2, we know that $I - P_{Q_n}$ are firmly nonexpansive for all $n \geq 1$, and from this one derives

$$\begin{aligned} \langle \nabla f_n(x_n), x_n - z \rangle &= \langle (I - P_{Q_n})Ax_n, Ax_n - Az \rangle \\ &= \langle (I - P_{Q_n})Ax_n - (I - P_{Q_n})Az, Ax_n - Az \rangle \\ &\geq \|(I - P_{Q_n})Ax_n\|^2 = 2f_n(x_n), \end{aligned} \tag{3.10}$$

from which it turns out that

$$\begin{aligned} \|x_{n+1} - z\|^2 &\leq \|x_n - z\|^2 - 4\tau_n f_n(x_n) + \tau_n^2 \|\nabla f_n(x_n)\|^2 - \|y_n - P_{C_n}y_n\|^2 \\ &= \|x_n - z\|^2 - \frac{4\rho_n f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} + \frac{\rho_n^2 f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^4} \|\nabla f_n(x_n)\|^2 \\ &\quad - \|y_n - P_{C_n}y_n\|^2 \\ &\leq \|x_n - z\|^2 - \rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} - \|y_n - P_{C_n}y_n\|^2, \end{aligned} \tag{3.11}$$

which in turn allows us to deduce the following conclusions:

- (i) $\{x_n\}$ is Fejér monotone w.r.t. Γ ; in particular,
- (ii) $\{x_n\}$ is a bounded sequence;
- (iii) $\sum_{n=1}^{\infty} \rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} < \infty$; and
- (iv)

$$\sum_{n=1}^{\infty} \|y_n - P_{C_n}y_n\|^2 < \infty. \tag{3.12}$$

By Proposition 2.4(i), to show that $x_n \rightarrow x$, it suffices to show that $w_w(x_n) \subset \Gamma$. To see this, take $x^* \in w_w(x_n)$ and let $\{x_{n_k}\}$ be a sequence of $\{x_n\}$ weakly converging to x^* . By our assumption that $\liminf_n \rho_n(4 - \rho_n) \geq \rho > 0$, without loss of generality, we can assume that $\rho_n(4 - \rho_n) \geq \frac{\rho}{2}$ for all $n \geq 1$. It follows from (3.12) (iii) that

$$\frac{f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \rightarrow 0 \quad (n \rightarrow \infty). \tag{3.13}$$

Note that $\|\nabla f_n(x_n)\| + \sigma_n \leq \|A\|^2 \|x_n - z\| + 1$ for $z \in \Gamma$. This, together with (3.13), implies that $f_n(x_n) \rightarrow 0$, that is, $\|(I - P_{Q_n})Ax_n\| \rightarrow 0$. By our assumption that ∂q is a bounded mapping, we see that there exists a constant $M > 0$ such that $\|\eta_n\| \leq M, \forall \eta_n \in \partial q(Ax_n)$.

Since $P_{Q_n}(Ax_n) \in Q_n$, by the definition of Q_n , we have

$$q(Ax_n) \leq \langle \eta_n, Ax_n - P_{Q_n}(Ax_n) \rangle \leq M \|(I - P_{Q_n})Ax_n\| \rightarrow 0. \tag{3.14}$$

Noting that $Ax_{n_k} \rightarrow \hat{x}$ and using the w-lsc of q , we have $q(Ax^*) \leq \liminf_k q(Ax_{n_k}) \leq 0$, which implies that $Ax^* \in Q$. We next prove $x^* \in C$. Firstly, from (3.12) (iv), we know that $\|y_n - P_{C_n}y_n\| \rightarrow 0$. Notice that

$$\|y_n - x_n\| = \|\tau_n \nabla f_n(x_n)\| \leq \frac{4f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \cdot \frac{\|\nabla f_n(x_n)\|}{\|\nabla f_n(x_n)\| + \sigma_n} \leq \frac{4f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \rightarrow 0,$$

we have

$$\begin{aligned} \|x_n - P_{C_n}x_n\| &\leq \|x_n - y_n\| + \|y_n - P_{C_n}y_n\| + \|P_{C_n}y_n - P_{C_n}x_n\| \\ &\leq 2\|x_n - y_n\| + \|y_n - P_{C_n}y_n\| \rightarrow 0. \end{aligned}$$

Since ∂c is a bounded mapping, we have $M_1 > 0$ such that

$$\|\xi_1\| \leq M_1, \quad \forall \xi_n \in \partial c(x_n).$$

Since $P_{C_n}(x_n) \in C_n$, by the definition of C_n , we have

$$c(x_n) \leq \langle \xi_n, x_n - P_{C_n}x_n \rangle \leq M_1 \|(I - P_{C_n})x_n\| \rightarrow 0.$$

Then w-lsc of C implies that $c(x^*) \leq \underline{\lim}_k c(x_{n_k}) \leq 0$, thus $x^* \in C$ and $w_w(x_n) \subset \Gamma$, completing the proof. \square

We introduce a little more general algorithm as follows.

Algorithm 3.3 Choose an initial data $x_1 \in H_1$ arbitrarily. Assume that the n th iteration x_n has been constructed; then we compute the $(n + 1)$ th iteration x_{n+1} via the formula:

$$x_{n+1} = \beta_n x_n + (1 - \beta_n)P_{C_n}(x_n - \tau_n \nabla f(x_n)), \quad n \geq 1, \tag{3.15}$$

where the step size $\{\tau_n\}$ is as before and $\{\beta_n\}$ is a sequence in $(0, 1)$ satisfying $\overline{\lim}_n \beta_n < 1$. If $x_{n+1} = x_n$ for some $n \geq 1$, then x_n is a solution of the SFP (1.1) and the iterative process stops; otherwise, we set $n := n + 1$ and go on to (3.15) to compute the next iteration x_{n+2} .

We have the following weak convergence theorem.

Theorem 3.4 Assume that $\underline{\lim}_n \rho_n(4 - \rho_n) \geq \rho > 0$. Then the sequence $\{x_n\}$ generated by Algorithm 3.3 converges weakly to a solution \hat{x} of the SFP (1.1) where $\hat{x} = \lim_{n \rightarrow \infty} P_\Gamma x_n$.

Proof Let $z \in \Gamma$ be fixed and set $y_n = x_n - \tau_n \nabla f_n(x_n)$. By virtue of (2.1), (2.2), (3.15), (3.10), and Proposition 2.2(p₆), we have

$$\begin{aligned} \|x_{n+1} - z\|^2 &= \|\beta_n(x_n - z) + (1 - \beta_n)(P_{C_n}y_n - z)\|^2 \\ &= \beta_n \|x_n - z\|^2 + (1 - \beta_n) \|P_{C_n}y_n - z\|^2 - \beta_n(1 - \beta_n) \|x_n - P_{C_n}y_n\|^2 \\ &\leq \beta_n \|x_n - z\|^2 + (1 - \beta_n) \|y_n - z\|^2 - (1 - \beta_n) \|y_n - P_{C_n}y_n\|^2 \\ &= \beta_n \|x_n - z\|^2 + (1 - \beta_n) \|x_n - z - \tau_n \nabla f_n(x_n)\|^2 \\ &\quad - (1 - \beta_n) \|y_n - P_{C_n}y_n\|^2 \\ &= \beta_n \|x_n - z\|^2 + (1 - \beta_n) \|x_n - z\|^2 - 2(1 - \beta_n) \tau_n \langle \nabla f_n(x_n), x_n - z \rangle \\ &\quad + (1 - \beta_n) \tau_n^2 \|\nabla f_n(x_n)\|^2 - (1 - \beta_n) \|y_n - P_{C_n}y_n\|^2 \\ &\leq \|x_n - z\|^2 - 4(1 - \beta_n) \tau_n f_n(x_n) + (1 - \beta_n) \tau_n^2 \|\tau_n \nabla f_n(x_n)\|^2 \end{aligned}$$

$$\begin{aligned}
 & -\beta_n(1-\beta_n)\|y_n - P_{C_n}y_n\|^2 \\
 \leq & \|x_n - z\|^2 - (1-\beta_n)\rho_n(4-\rho_n)\frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \\
 & - (1-\beta_n)\|y_n - P_{C_n}y_n\|^2,
 \end{aligned} \tag{3.16}$$

which implies that

- (i) $\{x_n\}$ is Fejér monotone w.r.t. Γ ; in particular,
- (ii) $\{x_n\}$ is a bounded sequence;
- (iii) $\sum_{n=1}^\infty (1-\beta_n)\rho_n(4-\rho_n)f_n^2(x_n)/(\|\nabla f_n(x_n)\| + \sigma_n)^2 < \infty$; and
- (iv) $\sum_{n=1}^\infty (1-\beta_n)\|y_n - P_{C_n}y_n\|^2 < \infty$.

By our assumptions on $\{\beta_n\}$ and $\{\rho_n\}$, we have $\frac{f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \rightarrow 0$ and $y_n - P_{C_n}y_n \rightarrow 0$, the rest of the arguments follow exactly from the corresponding parts of Theorem 3.2, we omit its details. This completes the proof. \square

We remark that Theorem 3.4 generalizes Theorem 3.2, that is, if we take $\beta_n \equiv 0$ in Theorem 3.4, then we can obtain Theorem 3.2. It is really interesting work to compare convergence rate of Algorithms 3.1 and 3.3.

Generally speaking, Algorithms 3.1 and 3.3 have only the weak convergence in the frame work of infinite-dimensional spaces, and therefore the modifications of Algorithms 3.1 and 3.3 are needed in order to realize the strong convergence. Considerable efforts have been made and several interesting results have been reported recently; see [17–20]. Below is our modification of Algorithms 3.1 and 3.3.

Algorithm 3.5 Choose an arbitrary initial data $x_1 \in H_1$. Assume that the n th iteration $x_n \in H_1$ has been constructed. Set

$$y_n = P_{C_n}(x_n - \tau_n \nabla f(x_n)), \quad n \geq 1, \tag{3.17}$$

with the step size τ_n given by (1.12), and define two half-spaces Y_n and Z_n by

$$Y_n = \left\{ z \in H_1 : \|y_n - z\|^2 \leq \|x_n - z\|^2 - \rho_n(4-\rho_n)\frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \right\}, \tag{3.18}$$

$$Z_n = \{z \in H_1 : \langle x_1 - x_n, z - x_n \rangle \leq 0\}. \tag{3.19}$$

The $(n + 1)$ th iterate x_{n+1} is then constructed in the formula:

$$x_{n+1} = P_{Y_n \cap Z_n}(x_1). \tag{3.20}$$

If $x_{n+1} = x_n$ for some $n \geq 1$, then x_n is a solution of SFP (1.1) and the iterative process stops; otherwise, we set $n := n + 1$ and go on to (3.17)-(3.20) to compute the next iteration x_{n+2} .

Theorem 3.6 Assume that $\liminf_n \rho_n(4-\rho_n) \geq \rho > 0$. Then the sequence $\{x_n\}$ generated by Algorithm 3.5 converges strongly to a solution x^* of SFP (1.1), where $x^* = P_\Gamma(x_1)$.

Proof Firstly, we show that

$$Y_n \cap Z_n \supseteq \Gamma \neq \emptyset, \tag{3.21}$$

for all $n \geq 1$. Indeed, in view of (3.11), we have $\Gamma \subset Y_n$ for all $n \geq 1$. To show (3.21) holds, it suffices to show that $\Gamma \subset Z_n$ for all $n \geq 1$. We complete the proof by induction. Since $Z_1 = H_1$, $\Gamma \subset Z_1$. Assume that $\Gamma \subset Z_k$ for some $k \geq 1$; we plan to show $\Gamma \subset Z_{k+1}$. Since $\Gamma \subset Y_k \cap Z_k$, and $Y_k \cap Z_k \neq \emptyset$ closed convex, $x_{n+1} = P_{Y_n \cap Z_n}(x_1)$ is well defined. It follows from Proposition 2.2(p₁) that

$$\langle x_{k+1} - z, x_{k+1} - x_1 \rangle \leq 0, \quad \forall z \in \Gamma. \quad (3.22)$$

This implies that $z \in Z_{k+1}$ and hence $\Gamma \subset Z_{k+1}$. Consequently, $\Gamma \subset Z_n$ for all $n \geq 1$, and thus (3.21) holds true.

From the definition of Z_n and Proposition 2.2(p₁), we see that $x_n = P_{Z_n}x_1$. It then follows from (3.20) that

$$\|x_n - x_1\| = \|P_{Z_n}x_1 - x_1\| \leq \|x_{n+1} - x_n\| = \|P_{Z_{n+1}}x_1 - x_1\| \leq \|P_{\Gamma}x_1 - x_1\|. \quad (3.23)$$

This derives that $\lim_n \|x_n - x_1\|$ exists, denoted by d .

Noting that $x_{n+1} \in Z_n$, we have

$$\langle x_{n+1} - x_n, x_n - x_1 \rangle \geq 0. \quad (3.24)$$

By virtue of (2.1) and (3.24), we obtain

$$\|x_{n+1} - x_1\|^2 - \|x_n - x_1\|^2 = \|x_{n+1} - x_n\|^2 + 2\langle x_{n+1} - x_n, x_n - x_1 \rangle \geq \|x_{n+1} - x_n\|^2. \quad (3.25)$$

From this one derives that $x_{n+1} - x_n \rightarrow 0$ ($n \rightarrow \infty$).

Since $x_{n+1} \in Y_n$, we have

$$\|y_n - x_{n+1}\|^2 \leq \|x_n - x_{n+1}\|^2 - \rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2}, \quad (3.26)$$

from which it turns out that

$$\|y_n - x_{n+1}\| \leq \|x_n - x_{n+1}\| \rightarrow 0 \quad (3.27)$$

and

$$\rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \leq \|x_n - x_{n+1}\|^2 \rightarrow 0. \quad (3.28)$$

At this point, we show $w_w(x_n) \subset \Gamma$. To end this, take $\hat{x} \in w_w(x_n)$. Then there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $x_{n_j} \xrightarrow{w} \hat{x}$. By our assumption that $\underline{\lim}_n \rho_n(4 - \rho_n) \geq \rho > 0$, from (3.28) we conclude that

$$\frac{f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \rightarrow 0, \quad (3.29)$$

which implies that $f_n(x_n) \rightarrow 0$, since $\{\|\nabla f_n(x_n)\| + \sigma_n\}$ is bounded. Notice that $P_{Q_n}(Ax_n) \in Q_n$, and ∂q is a bounded mapping, we have

$$q(Ax_n) \leq \langle \eta_n, Ax_n - P_{Q_n}(Ax_n) \rangle \leq M \|(I - P_{Q_n})Ax_n\| \rightarrow 0. \quad (3.30)$$

Since $Ax_{n_j} \rightarrow A\hat{x}$ and q is w-lsc on H_2 , we derive

$$q(A\hat{x}) \leq \liminf_j q(Ax_{n_j}) \leq 0,$$

which implies that $A\hat{x} \in Q$.

We next show $\hat{x} \in C$. Indeed, from (3.27) we have

$$\|y_n - x_n\| \leq \|y_n - x_{n+1}\| + \|x_{n+1} - x_n\| \rightarrow 0. \tag{3.31}$$

From (3.29), we have also

$$\tau_n \|\nabla f_n(x_n)\| = \frac{\rho_n f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \cdot \frac{\|\nabla f_n(x_n)\|}{\|\nabla f_n(x_n)\| + \sigma_n} \leq \frac{4f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \rightarrow 0. \tag{3.32}$$

Consequently, it follows from (3.31) and (3.32) that

$$\begin{aligned} \|x_n - P_{C_n}x_n\| &\leq \|x_n - y_n\| + \|y_n - P_{C_n}x_n\| \\ &\leq \|x_n - y_n\| + \tau_n \|\nabla f_n(x_n)\| \rightarrow 0. \end{aligned} \tag{3.33}$$

Since $P_{C_n}(x_n) \in C_n$, noting ∂c is a bounded mapping, we immediately obtain

$$c(x_n) \leq \langle \xi_n, x_n - P_{C_n}x_n \rangle \leq M_2 \|(I - P_{C_n})x_n\| \rightarrow 0.$$

Then the w-lsc of c ensures that

$$c(\hat{x}) \leq \liminf_j c(x_{n_j}) \leq 0,$$

from which it turns out that $\hat{x} \in C$, and thus $\hat{x} \in \Gamma$. It follows from (3.21) that $\hat{x} \in Z_n$, which implies that

$$\langle x_1 - x_n, \hat{x} - x_n \rangle \leq 0, \tag{3.34}$$

for all $n \geq 1$. Thus, from (3.34) we obtain

$$\|x_n - \hat{x}\|^2 \leq \langle \hat{x} - x_1, \hat{x} - x_n \rangle, \tag{3.35}$$

in particular, we have

$$\|x_{n_j} - \hat{x}\|^2 \leq \langle \hat{x} - x_1, \hat{x} - x_{n_j} \rangle,$$

consequently, $x_{n_j} \rightarrow \hat{x}$, since $x_{n_j} \xrightarrow{w} \hat{x}$.

At this point, by virtue of (3.19), we have

$$\langle x_1 - x_n, z - x_n \rangle \leq 0, \quad \forall z \in \Gamma, \tag{3.36}$$

in particular, we have

$$\langle x_1 - x_{n_j}, z - x_{n_j} \rangle \leq 0, \quad \forall z \in \Gamma. \tag{3.37}$$

Thus, upon taking the limit as $j \rightarrow \infty$ in (3.37), we obtain

$$\langle x_1 - \hat{x}, z - \hat{x} \rangle \leq 0, \quad \forall z \in \Gamma. \tag{3.38}$$

This implies that $\hat{x} = P_\Gamma x_1$ by Proposition 2.2(p1). Therefore $\{x_n\}$ converges strongly to $\hat{x} = P_\Gamma x_1$ because of the uniqueness of $P_\Gamma x_1$. This completes the proof. \square

Algorithm 3.7 Choose an arbitrary initial data $x_1 \in H_1$. Assume that the n th iteration $x_n \in H_1$ has been constructed. Set

$$y_n = \beta_n x_n + (1 - \beta_n) P_{C_n}(x_n - \tau_n \nabla f(x_n)), \tag{3.39}$$

with the step size τ_n given by (1.12) and the relaxed factor β_n in $[0, 1)$ satisfying $\overline{\lim}_n \beta_n < 1$. Define two half-spaces Y_n and Z_n by

$$Y_n = \left\{ z \in H_1 : \|y_n - z\|^2 \leq \|x_n - z\|^2 - \rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} - (1 - \beta_n) \|x_n - \tau_n \nabla f_n(x_n) - P_{C_n}(x_n - \tau_n \nabla f_n(x_n))\|^2 \right\}, \tag{3.40}$$

$$Z_n = \{z \in H_1 : \langle x_1 - x_n, z - x_n \rangle \leq 0\}. \tag{3.41}$$

The $(n + 1)$ th iteration x_{n+1} is then constructed by the formula:

$$x_{n+1} = P_{Y_n \cap Z_n}(x_1). \tag{3.42}$$

If $x_{n+1} = x_n$ for some $n \geq 1$, then x_n is a solution of SFP (1.1) and the iterative process stops; otherwise, we set $n := n + 1$ and go on to (3.39)-(3.41) to compute the next iteration x_{n+2} .

Along the proof lines of Theorem 3.6 we can prove the following.

Theorem 3.8 Assume that $\underline{\lim}_n \rho_n(4 - \rho_n) \geq \rho > 0$; then the sequence $\{x_n\}$ generated by Algorithm 3.7 converges strongly to a solution x^* of SFP (1.1), where $x^* = P_\Gamma(x_1)$.

The proof of Theorem 3.8 is similar to that of Theorem 3.6, and therefore we omit its details.

We next turn our attention to another kind of algorithm.

Algorithm 3.9 Choose an arbitrary initial data $x_1 \in H_1$. Assume that the n th iteration $x_n \in H_1$ has been constructed; then we compute the $(n + 1)$ th iteration x_{n+1} via the recursion:

$$x_{n+1} = P_{C_n}(\alpha_n g(x_n) + (1 - \alpha_n)(x_n - \tau_n \nabla f(x_n))), \quad n \geq 1, \tag{3.43}$$

where the step size τ_n is given by (1.12), $g : H_1 \rightarrow H_1$ is a contraction with contractive coefficient $\delta \in (0, 1)$, and $\{\alpha_n\}$ is a real sequence in $(0, 1)$. If $x_{n+1} = x_n$ for some $n \geq 1$, then

x_n is an approximate solution of SFP (1.1) (the approximate rule will be given below) and the iterative process stops; otherwise, we set $n := n + 1$ and go on to (3.43) to compute the next iteration x_{n+2} .

We point out that if $x_{n+1} = x_n$ for some $n \geq 1$, then (3.43) reduces to

$$x_n = P_{C_n}(\alpha_n g(x_n) + (1 - \alpha_n)(x_n - \tau_n \nabla f(x_n))), \quad n \geq 1. \tag{3.44}$$

This implies that $x_n \in C_n$ and hence $x_n \in C$. Write

$$e(x_n, \tau_n) = \|x_n - P_{C_n}(x_n - \tau_n \nabla f_n(x_n))\|.$$

Then it follows from (3.44) that

$$e(x_n, \tau_n) \leq \alpha_n \|g(x_n) - x_n + \tau_n \nabla f_n(x_n)\|, \quad n \geq 1. \tag{3.45}$$

Such an x_n is called an approximate solution of SFP (1.1). If $e(x_n, \tau_n) = 0$, then x_n is a solution of SFP (1.1).

Theorem 3.10 *Assume that $\{\alpha_n\}$ and $\{\rho_n\}$ satisfy conditions (C_1) $\alpha_n \rightarrow 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$ and (C_2) $\liminf_n \rho_n(4 - \rho_n) > 0$, respectively. Then the sequence $\{x_n\}$ generated by Algorithm 3.9 converges strongly to a solution x^* of SFP (1.1), where $x^* = P_{\Gamma}g(x^*)$, equivalently, x^* solves the following variational inequality:*

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

Proof First of all, we show there exists a unique $x^* \in \Gamma$ such that $x^* = P_{\Gamma}g(x^*)$. Indeed, since $P_{\Gamma}g : H_1 \rightarrow H_1$ is a contraction with the contractive coefficient $\delta \in (0, 1)$, by the Banach contractive mapping principle, we conclude that there exists a unique $x^* \in H_1$ such that $x^* = P_{\Gamma}g(x^*) \in \Gamma$, equivalently, x^* solves the following variational inequality:

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

Write $y_n = x_n - \tau_n \nabla f_n(x_n)$ and $z_n = \alpha_n g x_n + (1 - \alpha_n)y_n$. Then (3.43) can be rewritten as

$$x_{n+1} = P_{C_n} z_n. \tag{3.46}$$

Noting that $x^* \in \Gamma$ and $Q \subseteq Q_n$ for all $n \geq 1$, we have $Ax^* \in Q_n$ for all $n \geq 1$, and hence

$$(I - P_{Q_n})Ax^* = 0.$$

Since $I - P_{Q_n}$ is firmly nonexpansive, we have

$$\begin{aligned} \langle \nabla f(x_n), x_n - x^* \rangle &= \langle (I - P_{Q_n})Ax_n, Ax_n - Ax^* \rangle \\ &= \langle (I - P_{Q_n})Ax_n - (I - P_{Q_n})Ax^*, Ax_n - Ax^* \rangle \\ &\geq \|(I - P_{Q_n})Ax_n\|^2 = 2f_n(x_n). \end{aligned} \tag{3.47}$$

By virtue of (2.1) and (3.47), we obtain

$$\begin{aligned}
 \|y_n - x^*\|^2 &= \|x_n - x^* - \tau_n \nabla f_n(x_n)\|^2 \\
 &= \|x_n - x^*\|^2 - 2\tau_n \langle \nabla f_n(x_n), x_n - x^* \rangle + \tau_n^2 \|\nabla f_n(x_n)\|^2 \\
 &\leq \|x_n - x^*\|^2 - 4\tau_n f_n(x_n) + \tau_n^2 \|\nabla f_n(x_n)\|^2 \\
 &= \|x_n - x^*\|^2 - 4\rho_n \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \\
 &\quad + \frac{\rho_n^2 f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \cdot \frac{\|\nabla f_n(x_n)\|^2}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \\
 &\leq \|x_n - x^*\|^2 - \rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2}, \tag{3.48}
 \end{aligned}$$

in particular, we have

$$\|y_n - x^*\| \leq \|x_n - x^*\|, \tag{3.49}$$

for all $n \geq 1$.

We now estimate $\|z_n - x^*\|^2$. By virtue of definition of the norm $\|\cdot\|$ and Schwarz's inequality, we obtain

$$\begin{aligned}
 \|z_n - x^*\|^2 &= \langle z_n - x^*, z_n - x^* \rangle \\
 &= \alpha_n \langle g(x_n) - x^*, z_n - x^* \rangle + (1 - \alpha_n) \langle y_n - x^*, z_n - x^* \rangle \\
 &= \alpha_n \langle g(x_n) - g(x^*), z_n - x^* \rangle + \alpha_n \langle g(x^*) - x^*, z_n - x^* \rangle \\
 &\quad + (1 - \alpha_n) \langle y_n - x^*, z_n - x^* \rangle \\
 &\leq \frac{\delta^2 \alpha_n}{2} \|x_n - x^*\|^2 + \frac{\alpha_n}{2} \|z_n - x^*\|^2 + \alpha_n \langle g(x^*) - x^*, z_n - x^* \rangle \\
 &\quad + \frac{1 - \alpha_n}{2} \|y_n - x^*\|^2 + \frac{1 - \alpha_n}{2} \|z_n - x^*\|^2,
 \end{aligned}$$

from which it turns out that

$$\|z_n - x^*\|^2 \leq \delta^2 \alpha_n \|x_n - x^*\|^2 + 2\alpha_n \langle g(x^*) - x^*, z_n - x^* \rangle + (1 - \alpha_n) \|y_n - x^*\|^2. \tag{3.50}$$

Substituting (3.48) into (3.50) yields

$$\begin{aligned}
 \|z_n - x^*\|^2 &\leq [1 - (1 - \delta^2)\alpha_n] \|x_n - x^*\|^2 + 2\alpha_n \langle g(x^*) - x^*, z_n - x^* \rangle \\
 &\quad - (1 - \alpha_n)\rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2}. \tag{3.51}
 \end{aligned}$$

By virtue of Proposition 2.2(p₆), (3.48), and (3.51), noting that $x^* \in C \subset C_n$ for all $n \geq 1$, we have

$$\begin{aligned}
 \|x_{n+1} - x^*\|^2 &= \|P_{C_n} z_n - x^*\|^2 \\
 &\leq \|z_n - x^*\|^2 - \|z_n - P_{C_n} z_n\|^2
 \end{aligned}$$

$$\begin{aligned} &\leq [1 - (1 - \delta^2)\alpha_n] \|x_n - x^*\|^2 + 2\alpha_n \langle (g - I)x^*, z_n - x^* \rangle \\ &\quad - (1 - \alpha_n)\rho_n(4 - \rho_n) \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} - \|(I - P_{C_n})z_n\|^2. \end{aligned} \tag{3.52}$$

We next show that $\{x_n\}$ is bounded. Using (3.43) and (3.49), we have

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \alpha_n \|g(x_n) - x^*\| + (1 - \alpha_n) \|y_n - x^*\| \\ &\leq \delta\alpha_n \|x_n - x^*\| + \alpha_n \|g(x^*) - x^*\| + (1 - \alpha_n) \|x_n - x^*\| \\ &= [1 - (1 - \delta)\alpha_n] \|x_1 - x^*\| + (1 - \delta)\alpha_n \frac{\|g(x^*) - x^*\|}{1 - \delta} \\ &\leq \max \left\{ \|x_1 - x^*\|, \frac{\|g(x^*) - x^*\|}{1 - \delta} \right\} = M, \end{aligned}$$

for all $n \geq 1$, therefore $\{x_n\}$ is bounded; so are $\{y_n\}$ and $\{z_n\}$.

Finally, we show that $x_n \rightarrow x^*$ ($n \rightarrow \infty$).

Set $s_n = \|x_n - x^*\|^2$ and assume that $\rho_n(4 - \rho_n) \geq \rho$ for all $n \geq 1$. Then (3.52) reduces to

$$\begin{aligned} s_{n+1} - s_n + (1 - \delta^2)\alpha_n s_n + (1 - \alpha_n)\rho \frac{f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} + \|(I - P_{C_n})z_n\|^2 \\ \leq 2\alpha_n \langle (g - I)x^*, z_n - x^* \rangle. \end{aligned} \tag{3.53}$$

We consider two possible cases.

Case 1. $\{s_n\}$ is eventually decreasing, *i.e.*, there exists some integer $n_0 \geq 1$ such that

$$s_{n+1} \leq s_n \quad \text{for all } n \geq n_0,$$

which means that $\lim_n s_n$ exists. Note that $\{z_n\}$ is bounded and $\alpha_n \rightarrow 0$. Letting $n \rightarrow \infty$ in (3.53) yields $\frac{f_n(x_n)}{\|\nabla f_n(x_n)\| + \sigma_n} \rightarrow 0$ and $(I - P_{Q_n})z_n \rightarrow 0$. Since $\{\|\nabla f_n(x_n) + \sigma_n\|\}$ is a bounded sequence, we conclude that $f_n(x_n) \rightarrow 0$ and hence

$$(I - P_{Q_n})Ax_n \rightarrow 0. \tag{3.54}$$

Observe that $\|z_n - y_n\| \leq \alpha_n \|g(x_n) - y_n\| \leq \alpha_n M_1 \rightarrow 0$,

$$\begin{aligned} \|y_n - x_n\| &= \tau_n \|\nabla f_n(x_n)\| = \frac{\rho_n f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \|\nabla f_n(x_n)\| \\ &\leq \frac{4f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \rightarrow 0, \end{aligned}$$

and

$$\begin{aligned} \|x_n - P_{C_n}x_n\| &\leq \|x_n - z_n\| + \|z_n - P_{C_n}z_n\| + \|P_{C_n}z_n - P_{C_n}x_n\| \\ &\leq 2\|x_n - z_n\| + \|(I - P_{C_n})z_n\| \rightarrow 0. \end{aligned} \tag{3.55}$$

We may assume that

$$\overline{\lim}_n \langle (g - I)x^*, z_n - x^* \rangle = \overline{\lim}_n \langle (g - I)x^*, x_n - x^* \rangle = \overline{\lim}_{k \rightarrow \infty} \langle (g - I)x^*, x_{n_k} - x^* \rangle. \tag{3.56}$$

Without loss of generality, we assume that $x_{n_k} \rightarrow \hat{x}$ ($k \rightarrow \infty$); then $Ax_{n_k} \rightarrow A\hat{x}$ ($k \rightarrow \infty$). Since $P_{Q_{n_k}}Ax_{n_k} \in Q_{n_k}$, $\{\eta_{n_k}\} \subset \partial q(Ax_{n_k})$ is a bounded sequence and $(I - P_{Q_{n_k}})Ax_{n_k} \rightarrow 0$ ($k \rightarrow \infty$) by (3.54), we deduce that

$$q(Ax_{n_k}) \leq \langle \eta_{n_k}, Ax_{n_k} - P_{Q_{n_k}}(Ax_{n_k}) \rangle \leq \|\eta_{n_k}\| \|(I - P_{Q_{n_k}})Ax_{n_k}\| \rightarrow 0,$$

as $k \rightarrow \infty$, then w-lsc of q implies that

$$q(A\hat{x}) \leq \liminf_k q(Ax_{n_k}) \leq 0,$$

and thus $A\hat{x} \in Q$.

On the other hand, since $P_{C_{n_k}}(x_{n_k}) \in C_{n_k}$, $\{\xi_{n_k}\} \subset \partial c(x_{n_k})$ is a bounded sequence, and $(I - C_{Q_{n_k}})x_{n_k} \rightarrow 0$ by (3.55), we derive

$$c(x_{n_k}) \leq \langle \xi_{n_k}, x_{n_k} - P_{C_{n_k}}x_{n_k} \rangle \leq \|\xi_{n_k}\| \|(I - P_{C_{n_k}})x_{n_k}\| \rightarrow 0$$

as $k \rightarrow \infty$, then w-lsc of c implies that

$$c(\hat{x}) \leq \liminf_k c(x_{n_k}) \leq 0,$$

and thus $\hat{x} \in C$. Consequently, $\hat{x} \in C \cap A^{-1}(Q) = \Gamma$. It follows from (3.56) and Proposition 2.2(p₁) that

$$\overline{\lim}_n \langle (g - I)x^*, z_n - x^* \rangle = \langle (g - I)x^*, \hat{x} - x^* \rangle \leq 0. \tag{3.57}$$

Taking into account of (3.53), we have

$$s_{n+1} \leq [1 - (1 - \delta^2)\alpha_n]s_n + 2\alpha_n \langle (g - I)x^*, z_n - x^* \rangle. \tag{3.58}$$

Applying Proposition 2.5 to (3.58), we derive that $s_n \rightarrow 0$ as $n \rightarrow \infty$, i.e., $x_n \rightarrow x^*$ as $n \rightarrow \infty$.

Case 2. $\{s_n\}$ is not eventually decreasing. In this case, we can find an integer $n_0 \geq 1$ such that $s_{n_0} < s_{n_0+1}$. Define $J(n) := \{n_0 \leq k \leq n : s_k < s_{k+1}\}$, $n > n_0$. Then $J(n) \neq \emptyset$ and $J(n) \subseteq J(n + 1)$. Define $\tau : \mathbb{N} \rightarrow \mathbb{N}$ by

$$\tau(n) := \max J(n), \quad n > n_0.$$

Then $\tau(n) \rightarrow \infty$ as $n \rightarrow \infty$, $s_{\tau(n)} \leq s_{\tau(n)+1}$ for all $n > n_0$ and $s_n \leq s_{\tau(n)+1}$ for all $n > n_0$; see [20] for details.

Since $s_{\tau(n)} \leq s_{\tau(n)+1}$ for all $n > n_0$, it follows from (3.53) that

$$s_{\tau(n+1)} - s_{\tau(n)} \rightarrow 0, \quad \frac{\rho f_n^2(x_n)}{(\|\nabla f_n(x_n)\| + \sigma_n)^2} \leq M_1 \alpha_{\tau(n)} \rightarrow 0$$

and $\|(I - P_{C_{\tau(n)}})z_{\tau(n)}\| \rightarrow 0$ as $n \rightarrow \infty$.

At this point, by virtue of a similar reasoning to the corresponding parts in case 1, we can deduce that $\overline{\lim}_n \langle (g - I)x^*, z_{\tau(n)} - x^* \rangle = \overline{\lim}_n \langle (g - I)x^*, x_{\tau(n)} - x^* \rangle \leq 0$.

Noting that $s_{\tau(n)} \leq s_{\tau(n)+1}$, it follows from (3.53) that

$$s_{\tau(n)} \leq \frac{2}{1 - \delta^2} \langle (g - I)x^*, z_{\tau(n)} - x^* \rangle,$$

from which one derives that $\overline{\lim}_n s_{\tau(n)} \leq 0$, and hence $s_{\tau(n)} \rightarrow 0$ as $n \rightarrow \infty$. From this it turns out that $s_{\tau(n)+1} \rightarrow 0$ as $n \rightarrow \infty$, since $s_{\tau(n)+1} - s_{\tau(n)} \rightarrow 0$ as $n \rightarrow \infty$. Consequently, $s_n \rightarrow 0$, as $n \rightarrow \infty$, since $0 \leq s_n \leq s_{\tau(n)+1} \rightarrow 0$ as $n \rightarrow \infty$. This completes the proof. \square

By using an argument like the method in Theorem 3.10, we have the following more general algorithm and convergence theorem.

Algorithm 3.11 Choose an arbitrary initial data $x_1 \in H_1$. Assume that the n th iteration $x_n \in H_1$ has been constructed; then we compute the $(n + 1)$ th iteration x_{n+1} via the recursion

$$x_{n+1} = \beta_n x_n + (1 - \beta_n) P_{C_n} [\alpha_n g(x_n) + (1 - \alpha_n)(x_n - \tau_n \nabla f(x_n))], \quad n \geq 1, \tag{3.59}$$

where $\{\beta_n\}$ is a real sequence in $[0, 1)$ satisfying $\overline{\lim}_n \beta_n < 1$, $\{\alpha_n\}$ is a real sequence in $(0, 1)$ satisfying conditions (C_1) $\alpha_n \rightarrow 0$ and (C_2) $\sum_{n=1}^\infty \alpha_n = \infty$, $g : H_1 \rightarrow H_1$ is a contraction with contractive coefficient $\delta \in (0, 1)$ and τ_n is given by (1.12). If $x_{n+1} = x_n$ for some $n \geq 1$, then x_n is an approximate solution of SFP (1.1) and the iterative process stops; otherwise, we set $n := n + 1$ and go on to (3.59) to compute the next iteration x_{n+2} .

Theorem 3.12 Assume that $\underline{\lim}_n \rho_n(4 - \rho_n) > 0$; the sequence generated by algorithm (3.59) converges strongly to a solution x^* of the SFP (1.1), where $x^* = P_\Gamma g(x^*)$, equivalently, x^* is a solution of the following variational inequality:

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

4 Numerical experiments

In this section, we consider two typical numerical experiments to illustrate the performance of step size (1.12) with CQ-like algorithms. Firstly, we introduce a linear observation model as follows, which covers many problems in signal and image processing:

$$y = Ax + \varepsilon, \quad x \in \mathbb{R}^N, \tag{4.1}$$

where $y \in \mathbb{R}^M$ is the observed or measured data with noisy ε . $A : \mathbb{R}^N \rightarrow \mathbb{R}^M$ denotes the bounded linear observation operator. A is sparse and the range of it is not closed in most inverse problems, thus A is often ill-condition and the problem is also ill-posed. When x is a sparse expansion, finding the solutions of (4.1) can be seen as finding a solution to the least-square problem

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|y - Ax\|^2 \quad \text{subject to} \quad \|x\|_1 < t, \tag{4.2}$$

for any real number $t > 0$.

When we set $C = \{x \in \mathbb{R}^N : \|x\|_1 \leq t\}$ and $Q = \{y\}$, it is a particular case of SFP (1.1); see [11]. Therefore, we continue by applying the CQ algorithm to solve (4.2). We compute the projection onto C through a soft thresholding method; see [11, 21–23].

Next, according to the examples in [11, 22], we also choose two similar particular problems: compressed sensing and image deconvolution, which can be covered by (4.1). The experiments compare the performances of the proposed step size (1.12) with the step size in [11], and analysis some properties of (1.12).

4.1 Compressed sensing

In a general compressed sensing model, we set the hits of a signal $x \in \mathbb{R}^N$ is $N = 2^{12}$. There exist $m=50$ spikes with amplitude ± 1 distributed in the whole domain randomly. The plot can be seen on the top of Figure 1. Then we set the observation dimension $M = 2^{10}$ and a matrix A with $M \times N$ order is also generated arbitrarily. A standard Gaussian distribution noise with variance $\sigma_\varepsilon^2 = 10^{-4}$ is added. Let $t=50$ in (4.2).

For the step sizes (1.11) and (1.12), we always set the constant $\rho = 2$. For Algorithm 3.3, we set $\beta_n = 0.5$. All the processes are started with initial signal $x_0 = 0$ and finished with the stop rule

$$\|x_{n+1} - x_n\| / \|x_n\| < 10^{-3}.$$

We calculated the mean squared error (MSE) for the results

$$MSE = (1/N) \|x^* - x\|,$$

where x^* is an estimated signal of x .

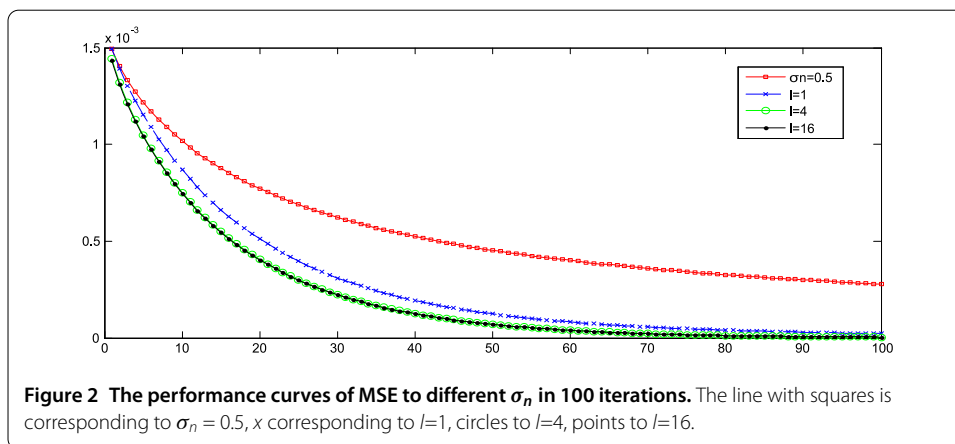
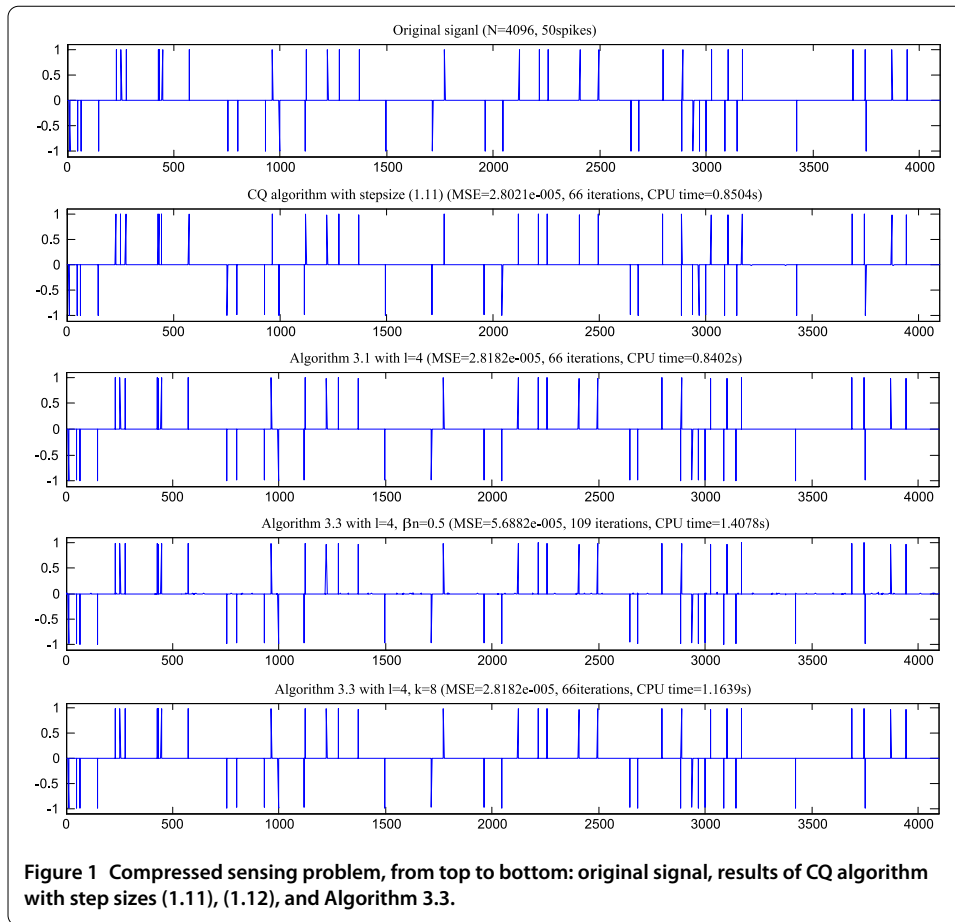
The second and third plots in Figure 1 correspond to the results with step sizes (1.11) and (1.12) to Algorithm 3.1, respectively. The recovered result by Algorithm 3.3 with step size (1.12) is shown in the fourth plot. Especially for the fifth, when we set $\beta_n = (n + 1)^{-k}$, $k = 1, 2, 3, \dots$, when we have $k \geq 3$ the iteration steps of Algorithm 3.3 start to approach the number in the second plot, and the restored precision is a little poorer than the others.

For (1.12) we firstly set $\sigma_n = \sigma = 0.5$; then in order to study its effect to the convergence speed of the CQ algorithm, we let it be $\sigma_n = (n + 1)^{-l}$, $l \geq 1$ is an integer. In Figure 2 we can find that when $l \geq 4$ the best MSE curves can be obtained, and it starts to change less. Therefore, σ_n should be as little as possible.

4.2 Image deconvolution

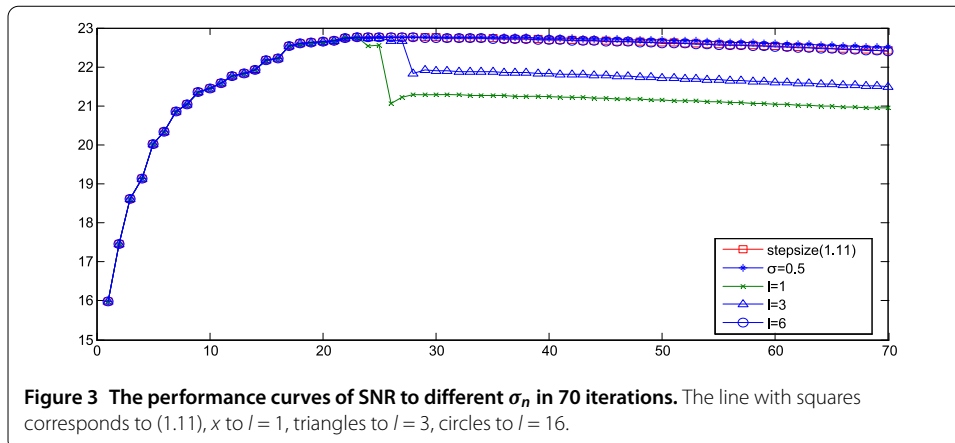
In this subsection, we continue by applying Algorithms 3.1 and 3.3 to recover the blurred Cameraman image. In the experiments, from [22, 24] we employ Haar wavelets and the blur point spread function $h_{ij} = (1 + i^2 + j^2)^{-1}$, for $i, j = -4, \dots, 4$; the noise variance is $\sigma^2 = 2$. The size of the image is $N = M = 256^2$. The threshold value is hand-tuned for the best SNR improvement. t is the sum of all the original pixel values.

We observe the performance of σ_n in (1.12); see Figure 3. We find that at the beginning several steps there are similar SNR curves, however, after 19 iterations, $\sigma = 0.5$ is similar with (1.11). When $1 \leq l < 6$, the curves are worse than the others. While we set $l \geq 6$ the curve starts to be consistent with the curve of (1.11). Therefore, we also know that σ_n should be better as little as possible.



5 Conclusion remarks

In this paper we have proposed several kinds of adaptively relaxed iterative algorithms with a new variable step size τ_n for solving SFP (1.1). The feature is that the new variable step size τ_n contains a sequence of positive numbers in its denominator. Because of this, the proposed algorithms with relaxed iterations will never terminate at any iteration step. On the other hand, unlike the previous known algorithms, our stop rule is that the related iteration process will stop if $x_{n+1} = x_n$ for some $n \geq 1$.



By means of new analysis techniques, we have proved several kinds of weak and strong convergence theorems of the proposed algorithms for solving SFP (1.1), which improved, extended, and complemented those existing in the literature. We remark that all convergence results in this paper still hold true if we use the step size τ_n given by (1.11) to replace the step size given by (1.12). In such a case, the stop rules should be modified. We would like to point out that our Theorems 3.10 and 3.12 are closely related to a sort of variational inequalities.

Finally, numerical experiments have been presented to illustrate the effectiveness of the proposed algorithms and applications in signal processing of the algorithms with the step size selected in this paper. The numerical results tell us that the changes of the choice of the step size σ_n given by (1.12) may affect the convergence rate of the iterative algorithms, and σ_n should be chosen as small as possible; for instance, we can choose σ_n such that $\sigma_n \rightarrow 0$ as $n \rightarrow \infty$.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All the authors contributed equally. All authors read and approved the final manuscript.

Author details

¹Department of Mathematics and Information, Hebei Normal University, Shijiazhuang, 050024, China. ²Department of Mathematics, Shijiazhuang Mechanical Engineering College, Shijiazhuang, 050003, China. ³The Second Training Base, Naval Aviation Institution, Huludao, 125001, China.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (11071053).

Received: 26 February 2014 Accepted: 20 October 2014 Published: 05 Nov 2014

References

1. Censor, Y, Elfving, T: A multiprojection algorithm using Bregman projections in a product space. *Numer. Algorithms* **8**, 221-239 (1994)
2. López, G, Martín-Márquez, V, Xu, HK: Iterative algorithms for the multiple-sets split feasibility problem. In: Censor, Y, Jiang, M, Wang, G (eds.) *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems*, pp. 243-279. Medical Physics Publishing, Madison (2009)
3. Chang, SS, Kim, JK, Cho, YJ, Sim, JY: Weak and strong convergence theorems of solutions to split feasibility problem for nonspreading type mapping in Hilbert spaces. *Fixed Point Theory Appl.* **2014**, Article ID 11 (2014)
4. Stark, H, Yang, Y: *Vector Space Projections: a Numerical Approach to Signal and Image Processing, Neural Nets and Optics*. Wiley, New York (1998)
5. Censor, Y, Bortfeld, T, Martin, B, Bortfeld, T: A unified approach for inversion problems in intensity-modulated radiation therapy. *Phys. Med. Biol.* **51**, 2353-2365 (2003)

6. Censor, Y, Elfving, T, Kopf, N, Bortfeld, T: The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Probl.* **21**, 2071-2084 (2005)
7. Qu, B, Xiu, N: A note on the CQ algorithm for the split feasibility problem. *Inverse Probl.* **21**, 1655-1665 (2005)
8. Li, M: Improved relaxed CQ methods for solving the split feasibility problem. *Adv. Model. Optim.* **13**, 305-318 (2011)
9. Abdellah, B, Muhammad, AN, Mohamed, K, Sheng, ZH: On descent-projection method for solving the split feasibility problems. *J. Glob. Optim.* **54**, 627-639 (2012)
10. Yang, Q: On variable-step relaxed projection algorithm for variational inequalities. *J. Math. Anal. Appl.* **302**, 166-179 (2005)
11. López, G, Martín-Márquez, V, Wang, FH, Xu, HK: Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Probl.* **28**, 085004 (2012)
12. Göebel, K, Kirk, WA: *Topics on Metric Fixed Point Theory*. Cambridge University Press, Cambridge (1990)
13. Aubin, JP: *Optima and Equilibria: An Introduction to Nonlinear Analysis*. Springer, Berlin (1993)
14. Byrne, C: A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl.* **20**, 103-120 (2004)
15. Bauschke, HH, Borwein, JM: On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **38**, 367-426 (1996)
16. Xu, HK: Iterative algorithms for nonlinear operators. *J. Lond. Math. Soc.* **66**, 240-256 (2011)
17. Wang, FH, Xu, HK: Approximating curve and strong convergence of the CQ algorithm for the split feasibility problem. *J. Inequal. Appl.* (2010). doi:10.1155/2010/102085
18. Dang, YZ, Gao, Y: The strong convergence of a KM-CQ-like algorithm for a split feasibility problem. *Inverse Probl.* **27**, 015007 (2011)
19. Yu, X, Shahzad, N, Yao, YH: Implicit and explicit algorithms for solving the split feasibility problem. *Optim. Lett.* **6**, 1447-1462 (2012)
20. Yao, YH, Postolache, M, Liou, YC: Strong convergence of a self-adaptive method for the split feasibility problem. *Fixed Point Theory Appl.* **2013**, Article ID 201 (2013). doi:10.1186/1687-1812-2013-201
21. Daubechies, I, Fornasier, M, Loris, I: Accelerated projected gradient method for linear inverse problems with sparsity constraints. *J. Fourier Anal. Appl.* **14**, 764-792 (2008)
22. Figueiredo, MA, Nowak, RD, Wright, SJ: Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.* **1**, 586-598 (2007)
23. Starck, JL, Murtagh, F, Fadili, JM: *Sparse Image and Signal Processing Wavelets, Curvelets, Morphological Diversity*, p. 166. Cambridge University Press, Cambridge (2010)
24. Mário, ATF: An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.* **12**, 906-917 (2003)

10.1186/1029-242X-2014-448

Cite this article as: Zhou and Wang: **Adaptively relaxed algorithms for solving the split feasibility problem with a new step size.** *Journal of Inequalities and Applications* 2014, **2014**:448

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
