Strong convergence of modified gradient-projection algorithm for constrained convex minimization problems

MING TIAN
Civil Aviation University of China
College of Science
Tianjin 300300
CHINA
tianming1963@126.com

LIHUA HUANG
Civil Aviation University of China
College of Science
Tianjin 300300
CHINA
huanglihua666@yeah.net

Abstract: In this article, a modified gradient-projection algorithm (GPA) is introduced, which combines Xu's idea of an alternative averaged mapping approach to the GPA and the general iterative method for nonexpansive mappings in Hilbert space introduced by Marino and Xu. Under suitable conditions, it is proved that the strong convergence of the sequences generated by implicit and explicit schemes to a solution of a constrained convex minimization problem which also solves a certain variational inequality. Obtained results extend and improve some existed results.

Key–Words: Gradient-projection algorithm, Constrained convex minimization, General iterative method, averaged mapping, nonexpansive mapping, fixed point, variational inequality

1 Introduction

The gradient-projection algorithm is a powerful tool for solving constrained convex optimization problems and has been extensively studied (see [1-4] and the references therein). The convergence of the sequence generated by this method depends on the behavior of the gradient of the objective function. If the gradient is only Lipschitz continuous, the strong convergence of the sequence generated by gradient-projection algorithm cannot be guaranteed. Very recently, Xu [5] gave an operator-oriented approach as an alternative to the gradient-projection algorithm and to the relaxed gradient-projection algorithm–namely, an averaged mapping approach. Moreover, he constructed a counterexample which shows that the sequence generated by the gradient-projection algorithm does not converge strongly in the setting of an infinite-dimensional space. He also presented two modifications of gradient-projection algorithm which have strong convergence. One is simply a convex combination of a contraction with the point that is generated by the gradient-projection algorithm, and the other involves additional projections. Both modifications are adaptations of those modifications [13, 18–20] for Rockafellar's proximal point algorithm [21] which has only week convergence in infinite-dimensional Hilbert spaces [15,22]. Moreover, the first modification is of viscosity nature [23,24].

Let H be a real Hilbert space, and C a nonempty closed and convex subset of H. Consider the following constrained convex minimization problem:

$$\min_{x \epsilon C} f(x), \tag{1}$$

where $f : C \to \mathbb{R}$ is a real valued convex and continuously Fréchet differentiable function. Then the gradient-projection algorithm (GPA) generates a sequence $\{x_n\}_{n=0}^{\infty}$ determined by the gradient of f and the metric projection onto C according to the recursive formula:

$$x_{n+1} := Proj_C(I - \lambda_n \nabla f)x_n, \qquad n \geq 0, \tag{2}$$

where the initial guess x_0 is taken from C arbitrarily, the parameters λ_n are positive real numbers, and $Proj_C$ is the metric projection from H onto C. We assume that the minimization problem (1) is consistent, and let S denotes solution set. It is known [1] that if f has a Lipschitz continuous and strongly monotone gradient, the sequence $\{x_n\}_{n=0}^{\infty}$ generated by (2) can be convergent strongly to a minimizer of f in C. If the gradient of f is only assumed to be Lipschitz continuous, it is proved in [5] that the sequence $\{x_n\}$ generated by (2) converges weekly to a minimizer of (1) provided the sequence $\{\lambda_n\}$ satisfies certain conditions that will be made precise in the section 3.

In addition, we know that the Lipschitz continuity of the gradient of f implies that it is inverse strongly monotone, its complement is an averaged mapping. Then the GPA can be rewritten as the composite of a projection and an averaged mapping, which is an averaged mapping, which is also a nonexpansive mapping.

Consequently, the GPA can be also rewritten as an iterative method for nonexpansive mapping.

On the other hand, Giuseppe Marino and Hong-Kun Xu [6] introduced the general iterative method for nonexpansive mappings in Hilbert spaces. Recall that T:$H \to H$ is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \epsilon H$. The set of fixed pointed of T is the set $Fix(T) := \{x \epsilon H : Tx = x\}$, we assume that $Fix(T) \neq \emptyset$. Let h be a contraction such that $\|h(x) - h(y)\| \leq \rho\|x - y\|, \forall x, y \epsilon H$, where $\rho\epsilon[0, 1)$ is a constant. A is strongly positive bounded linear operator; that is, there is a constant $\bar{\gamma} > 0$ such that$\langle Ax, x\rangle \geq \bar{\gamma}\|x\|^2, \forall x \epsilon H$. Starting with an arbitrary initial $x_0 \epsilon H$, define a sequence $\{x_n\}$ recursively by

$$x_{n+1} = \theta_n\gamma h(x_n) + (I - \theta_n A)Tx_n, \qquad n \geq 0. \quad (3)$$

Where γ is a constant such that $0 < \gamma < \bar{\gamma}/\rho$, and $\{\theta_n\}$ is a sequence in (0,1). It is proved in [6] that under certain appropriate conditions imposed on $\{\theta_n\}$, the sequence $\{x_n\}$ generated by (3) converges strongly to the unique solution x^* in Fix(T) of the variational inequality:

$$\langle (A - \gamma h)x^*, x^* - x\rangle \leq 0, \qquad \forall x \epsilon Fix(T). \quad (4)$$

In this paper we will combine the gradient-projection algorithm (2) with the general iterative method (3) and consider the following modified gradient-projection algorithm:

$$\begin{aligned} x_{n+1} &= (I - \theta_n A)Proj_C(I - \lambda_n\nabla f)x_n \\ &\quad + \theta_n\gamma h(x_n), \qquad n \geq 0. \end{aligned} \quad (5)$$

In section 3 we will prove that if the sequence $\{\theta_n\}$ and the sequence$\{\lambda_n\}$ of parameters satisfy appropriate conditions, the sequence $\{x_n\}$ generated by (5) converges strongly to a minimizer of (1), which is also the unique solution of the variational inequality:

$$\langle (A - \gamma h)x^*, x^* - x\rangle \leq 0, \qquad \forall x \epsilon S. \quad (6)$$

Next we introduce some useful properties. Recall that H is a Hilbert space, C a nonempty closed and convex subset of H. $Proj_C$ is the metric projection from H onto C.

Proposition 1 *Given $x\epsilon H$, we have:*

(i) $\langle x - y, Proj_C x - Proj_C y\rangle \geq \|Proj_C x - Proj_C y\|^2, \forall x, y \epsilon H$.

(ii) $\|x - Proj_C x\|^2 \leq \|x - y\|^2 - \|y - Proj_C x\|^2$, $\forall x \epsilon H, \forall y \epsilon C$.

Definition 2 *A mapping $T : H \to H$ is said to be firmly nonexpansive, if and only if $2T - I$ is nonexpansive, or equivalently,*

$$\langle x - y, Tx - Ty\rangle \geq \|Tx - Ty\|^2, \qquad x, y \epsilon H.$$

Alternatively, T is firmly nonexpansive, if and only if T can be expressed as

$$T = \frac{1}{2}(I + S),$$

where $S : H \to H$ is nonexpansive. Projections are firmly nonexpansive.

Definition 3 *A mapping $T : H \to H$ is said to be an averaged mapping, if and only if it can be written as the average of the identity I and a nonexpansive mapping; that is,*

$$T = (1 - \alpha)I + \alpha S,$$

where α is a number in [0,1] and $S : H \to H$ is nonexpansive. Thus firmly nonexpansive mappings (in particular, projections) are $(1/2)$-averaged mappings.

Some properties of averaged mappings are gathered in the proposition below.

Proposition 4 *Let the operators S, T, $V : H \to H$ be given.*

(i)If $T = (1 - \alpha)S + \alpha V$, for some $\alpha\epsilon[0, 1]$ and if S is averaged and V is nonexpansive, then T is averaged.

(ii) T is firmly nonexpansive, if and only if the complement $I - T$ is firmly nonexpansive.

(iii) If $T = (1 - \alpha)S + \alpha V$ for some $\alpha\epsilon[0, 1]$, S is firmly nonexpansive and V is nonexpansive, then T is averaged.

(iv) The composition of finitely many averaged mappings is averaged. That is, if each of the mappings $\{T_i\}_{i=1}^N$ is averaged, then so is the composite $T_1 \cdots T_N$. In particular, if T_1 is α_1-averaged and T_2 is α_2-averaged, where $\alpha_1, \alpha_2\epsilon[0, 1]$, then the composite T_1T_2 is α-averaged, where $\alpha = \alpha_1 + \alpha_2 - \alpha_1\alpha_2$.

(v) If the mappings $\{T_i\}_{i=1}^N$ are averaged and have a common fixed point, then

$$\bigcap_{i=1}^N Fix(T_i) = Fix(T_1 \cdots T_N).$$

Here the notations $Fix(T)$ denotes the set of fixed point of the mapping T; that is, $Fix(T) := \{x\epsilon H : Tx = x\}$.

Monotone operators are also very useful in the convergence analysis.

Definition 5 *(See [30] for comprehensive theory of monotone operators.) Let $A : H \to H$ be an operator* *(i) A is monotone, if and only if*

$$\langle x - y, Ax - Ay \rangle \geq 0, \qquad \forall\, x, y \,\epsilon\, H.$$

(ii) Given is a number $\eta > 0$. A is said to be η-strongly monotone, if and only if

$$\langle x - y, Ax - Ay \rangle \geq \eta \|x - y\|^2, \qquad \forall\, x, y \,\epsilon\, H.$$

(iii) Given is a number $\upsilon > 0$. A is said to be υ-inverse strongly monotone (υ-ism), if and only if

$$\langle x - y, Ax - Ay \rangle \geq \upsilon \|Ax - Ay\|^2, \qquad \forall\, x, y \,\epsilon\, H.$$

It is easily seen that, if T is nonexpansive, then $I - T$ is monotone. It is also easily seen that a projection $Proj_C$ is a one-ism.

Inverse strongly (also referred to as co-coercive) monotone operators have widely been applied to solve practical problems in various fields; for instance, in traffic assignment problems (see [10,11]).

The following proposition gathers some results on the relationship between averaged mappings and inverse strongly monotone operators.

Proposition 6 *([7,12]) Let $T : H \to H$ be given. We have*

(i) T is nonexpansive, if and only if the complement $I - T$ is $(1/2)$-ism;

(ii)If T is L-Lipschitz continuous, then T is $1/L$-ism;

(iii) If T is υ-ism, then for $\gamma > 0$, γT is (υ/γ)-ism;

(iv) T is averaged, if and only if the complement $I - T$ is υ-ism for some $\upsilon > 1/2$; indeed, for $\alpha \epsilon [0, 1]$, T is α-averaged, if and only if $I - T$ is $(1/2\alpha)$-ism.

The organization of this paper is as follows. In Sect.2, we introduce some useful lemmas which will be used in the proofs for the main results in Sect.3. In Sect.3, we propose implicit and explicit iterative schemes for solving the constrained convex minimization problem (1), and prove that the sequences generated by our schemes converge strongly to a solution of the constrained convex minimization problem. Such a solution is also a solution of a variational inequality defined over the set of solutions to the constrained convex minimization problem. In Sect.4, we apply this algorithm to the split feasibility problem.

2 Preliminaries

This section collects some lemmas which will be used in the proofs for the main results in the next section. Some of them are known; others are not hard to derive.

Lemma 7 *Assume that $\{a_n\}_{n=0}^{\infty}$ is a sequence of nonnegative real numbers such that*

$$a_{n+1} \leq (1 - \gamma_n) a_n + \gamma_n \delta_n + \beta_n, \qquad n \geq 0,$$

where $\{\gamma_n\}_{n=0}^{\infty}$ and $\{\beta_n\}_{n=0}^{\infty}$ are sequences in [0,1] and $\{\delta_n\}_{n=0}^{\infty}$ is a sequence in $\mathbb{R}$ such that

(i) $\sum\limits_{n=0}^{\infty} \gamma_n = \infty$;

(ii) either $\limsup\limits_{n\to\infty} \delta_n \leq 0$ or $\sum\limits_{n=0}^{\infty} \gamma_n |\delta n| < \infty$;

(iii) $\sum\limits_{n=0}^{\infty} \beta_n < \infty$.

Then $\lim\limits_{n\to\infty} a_n = 0$.

The so-called demiclosed principle for nonexpansive mappings will often be used.

Lemma 8 *(Demiclosed Principle) ([14]) Let C be a closed and convex subset of a Hilbert space H and let $T : C \to C$ be a nonexpansive mapping with $Fix(T) \neq \phi$. If $\{x_n\}_{n=1}^{\infty}$ is a sequence in C weekly converging to x and if $\{(I - T)x_n\}_{n=1}^{\infty}$ converges strongly to y, then $(I - T)x = y$. In particular, if $y = 0$, then $x \epsilon Fix(T)$.*

Lemma 9 *Let H be a Hilbert space, C a closed convex subset of H, $h : H \to H$ a contraction with coefficient $0 < \rho < 1$, and A a strongly positive bounded linear operator with coefficient $\bar{\gamma} > 0$. Then, for $0 < \gamma < \bar{\gamma}/\rho$,*

$$\begin{aligned} &\langle x - y, (A - \gamma h)x - (A - \gamma h)y \rangle \\ \geq\ &(\bar{\gamma} - \gamma\rho)\|x - y\|^2, \forall\, x, y \,\epsilon\, H. \end{aligned}$$

That is, $A - \gamma h$ is strongly monotone with coefficient $\bar{\gamma} - \gamma\rho$.

Recall the metric (nearest point) projection $Proj_C$ from a real Hilbert space H to a closed convex subset C of H is defined as follows: given $x \epsilon H$, $Proj_C x$ is the only point in C with the property

$$\|x - Proj_C x\| = inf\{\|x - y\| : y \,\epsilon\, C\}.$$

$Proj_C$ is characterized as follows.

Lemma 10 *Let C be a closed and convex subset of a real Hilbert space H. Given $x \epsilon H$ and $y \epsilon C$. Then $y = Proj_C x$ if and only if there holds the inequality*

$$\langle x - y, y - z \rangle \geq 0, \qquad \forall\, z \,\epsilon\, C.$$

Lemma 11 *Assume A is a strongly positive bounded linear operator on a Hilbert space H with coefficient $\bar{\gamma} > 0$ and $0 < s \leq \|A\|^{-1}$. Then $\|I - sA\| \leq 1 - s\bar{\gamma}$.*

Proof: Recall that a standard result in functional analysis is that if V is bounded linear self-adjoint operator on H, then

$$\|V\| = sup\{|\langle Vx, x\rangle| : x\epsilon H, \|x\| = 1\}.$$

Now for $x\epsilon H$ with $\|x\| = 1$, we see that

$$\langle (I - sA)x, x\rangle = 1 - s\langle Ax, x\rangle \geq 1 - s\|A\| \geq 0$$

(i.e., $I - sA$ is positive). It follows that

$$\begin{aligned} \|I - sA\| &= sup\{\langle (I - sA)x, x\rangle : x\epsilon H, \|x\| = 1\} \\ &= sup\{1 - s\langle Ax, x\rangle : x\epsilon H, \|x\| = 1\} \\ &\leq 1 - s\bar{\gamma} \quad by(2). \end{aligned}$$

We adopt the following notation:

- ◦ $x_n \to x$ means that $x_n \to x$ strongly.
- ◦ $x_n \rightharpoonup x$ means that $x_n \to x$ weakly.
- ◦ $\omega_w(x_n) := \{x : \exists x_{nj} \rightharpoonup x\}$ is the week ω-limit set of the sequence $\{x_n\}_{n=1}^{\infty}$.

3 Main results

Let H be a real Hilbert space and C a closed and convex subset of Hilbert space H. Let $f : C \to \mathbb{R}$ be a real valued convex and continuously Fréchet differentiable function. If the gradient of f is Lipschitz continuous, namely, there is a constant $L > 0$ such that

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \qquad \forall\, x, y \,\epsilon\, C. \quad (7)$$

Throughout the rest of this paper, we always assume that the closed and convex subset C satisfies $C \pm C \subset C$, and A is strongly positive bounded linear operator; that is, there is a constant $\bar{\gamma} > 0$ such that

$$\langle Ax, x\rangle \geq \bar{\gamma}\|x\|^2, \qquad \forall\, x \,\epsilon\, C.$$

Recall also that a contraction on C is a self-mapping h such that

$$\|h(x) - h(y)\| \leq \rho\|x - y\|, \qquad \forall\, x, y \,\epsilon\, C.$$

where $\rho\epsilon[0, 1)$ is a constant.

Denote by Π the collection of all contraction on C; namely,

$$\Pi = \{h : h \quad is \quad a \quad contraction \quad on \quad C\}.$$

Now given $h \,\epsilon\, \Pi$ with contraction coefficient $0 \leq \rho < 1$, $s \,\epsilon\, (0, 1)$ such that $s < \|A\|^{-1}$ and $0 < \gamma < \bar{\gamma}/\rho$. Consider a mapping Xs on C defined by

$$Xs(x) = s\gamma h(x) + (I - sA)Proj_C(I - \lambda_s\nabla f)(x). \quad (8)$$

where λ_s is positive real parameters, and λ_s is continuous with the aspect s.

It is easy to see that Xs is a contraction. Indeed, by Lemma 11, and noticing that $Proj_C(I - \lambda_s\nabla f)$ is nonexpansive, we have:

$$\begin{aligned} & \|Xs(x) - Xs(y)\| \\ \leq\; & \|Proj_C(I - \lambda_s\nabla f)x - Proj_C(I - \lambda_s\nabla f)y\| \\ & \cdot\|I - sA\| + s\gamma\|h(x) - h(y)\| \\ \leq\; & (1 - s\bar{\gamma})\|x - y\| + s\gamma\rho\|x - y\|) \\ =\; & (1 - s(\bar{\gamma} - \gamma\rho))\|x - y\|. \end{aligned} \quad (9)$$

Hence Xs has a unique fixed point, denoted x_s, which solves uniquely the fixed point equation

$$x_s = s\gamma h(x_s) + (I - sA)Proj_C(I - \lambda_s\nabla f)(x_s). \quad (10)$$

Note that x_s indeed depends on h for simplicity of notation throughout the rest of this paper. We will also always use γ to mean a number in $(0, \bar{\gamma}/\rho)$.

Then the following proposition summarizes the basic properties of $\{x_s\}$.

Proposition 12 *Let x_s be defined via (10).*
(i) $\{x_s\}$ is bounded for $s \,\epsilon\, (0, \|A\|^{-1})$.
(ii) $\lim_{s\to 0}\|x_s - Proj_C(I - \lambda_s\nabla f)(x_s)\| = 0$.
(iii) x_s defines a continuous curve from $(0, \|A\|^{-1})$ into C.

Proof: First observe that for $s\epsilon(0, \|A\|^{-1})$, we have $\|I - sA\| \leq 1 - s\bar{\gamma}$ by Lemma 5. To show (i), we take a $p \,\epsilon\, S$, noticing that $p = Proj_C(I - \lambda_s\nabla f)(p)$. Then we have

$$\begin{aligned} & \|x_s - p\| \\ =\; & \|s(\gamma h(x_s) - Ap) \\ & +(I - sA)(Proj_C(I - \lambda_s\nabla f)(x_s) - p)\| \\ \leq\; & s\|\gamma h(x_s) - Ap\| + (1 - s\bar{\gamma})\|x_s - p\| \\ =\; & s\|\gamma h(x_s) - \gamma h(p) + \gamma h(p) - Ap\| \\ & +(1 - s\bar{\gamma})\|x_s - p\| \\ \leq\; & s\gamma\rho\|x_s - p\| + s\|\gamma h(p) - Ap\| \\ & +(1 - s\bar{\gamma})\|x_s - p\| \\ =\; & (1 - s\bar{\gamma} + s\gamma\rho)\|x_s - p\| + s\|\gamma h(p) - Ap\|. \end{aligned}$$

It follows that

$$\|x_s - p\| \leq \frac{\|\gamma h(p) - Ap\|}{\bar{\gamma} - \gamma\rho}$$

Hence $\{x_s\}$ is bounded.

(ii) We have

$$\begin{aligned} & \|x_s - Proj_C(I - \lambda_s\nabla f)(x_s)\| \\ =\; & \|s\gamma h(x_s) - sAProj_C(I - \lambda_s\nabla f)(x_s)\| \\ =\; & s\|\gamma h(x_s) - AProj_C(I - \lambda_s\nabla f)(x_s)\| \to 0. \end{aligned}$$

since the boundedness of $\{x_s\}$ implies that $\{h(x_s)\}$ and $\{AProj_C(I-\lambda_s\nabla f)(x_s)\}$.

To prove (iii), take $s, s_0 \,\epsilon\, (0, \|A\|^{-1})$ and calculate,

$$\begin{aligned}
& \|x_s - x_{s_0}\| \\
= \;& \|s\gamma h(x_s) + (I-sA)Proj_C(I-\lambda_s\nabla f)x_s \\
& -s_0\gamma h(x_{s_0}) \\
& -(I-s_0A)Proj_C(I-\lambda_{s_0}\nabla f)x_{s_0}\| \\
= \;& \|s\gamma h(x_s) - s\gamma h(x_{s_0}) + s\gamma h(x_{s_0}) \\
& -s_0\gamma h(x_{s_0}) + (I-sA)Proj_C(I-\lambda_s\nabla f)x_s \\
& -(I-sA)Proj_C(I-\lambda_s\nabla f)x_{s_0} \\
& +(I-sA)Proj_C(I-\lambda_s\nabla f)x_{s_0} \\
& -(I-s_0A)Proj_C(I-\lambda_{s_0}\nabla f)x_{s_0}\| \\
\le \;& s\gamma\rho\|x_s - x_{s_0}\| + \gamma|s-s_0|\|h(x_{s_0})\| \\
& +(1-s\bar{\gamma})\|x_s - x_{s_0}\| \\
& +\|(I-sA)Proj_C(I-\lambda_s\nabla f)x_{s_0} \\
& -(I-s_0A)Proj_C(I-\lambda_{s_0}\nabla f)x_{s_0}\| \\
= \;& (s\gamma\rho + 1 - s\bar{\gamma})\|x_s - x_{s_0}\| \\
& +\gamma|s-s_0|\|h(x_{s_0})\| \\
& +\|s_0AProj_C(I-\lambda_{s_0}\nabla f)x_{s_0} \\
& -sAProj_C(I-\lambda_s\nabla f)x_{s_0} \\
& +Proj_C(I-\lambda_s\nabla f)x_{s_0} \\
& -Proj_C(I-\lambda_{s_0}\nabla f)x_{s_0}\| \\
\le \;& (s\gamma\rho + 1 - s\bar{\gamma})\|x_s - x_{s_0}\| \\
& +\gamma|s-s_0|\|h(x_{s_0})\| \\
& +\|(I-\lambda_s\nabla f)x_{s_0} - (I-\lambda_{s_0}\nabla f)x_{s_0}\| \\
& +\|s_0AProj_C(I-\lambda_{s_0}\nabla f)x_{s_0} \\
& -sAProj_C(I-\lambda_s\nabla f)x_{s_0}\| \\
= \;& (s\gamma\rho + 1 - s\bar{\gamma})\|x_s - x_{s_0}\| \\
& +\gamma|s-s_0|\|h(x_{s_0})\| + |\lambda_s - \lambda_{s_0}|\|\nabla f(x_{s_0})\| \\
& +|s_0 - s|\|AProj_C(I-\lambda_s\nabla f)(x_{s_0})\| \\
& +\|s_0AProj_C(I-\lambda_{s_0}\nabla f)(x_{s_0}) \\
& -s_0AProj_C(I-\lambda_s\nabla f)(x_{s_0)}\| \\
\le \;& (s\gamma\rho + 1 - s\bar{\gamma})\|x_s - x_{s_0}\| \\
& +\gamma|s-s_0|\|h(x_{s_0})\| \\
& +|\lambda_s - \lambda_{s_0}|\|\nabla f(x_{s_0})\| \\
& +|s-s_0|\|AProj_C(I-\lambda_s\nabla f)(x_{s_0})\| \\
& +s_0|\lambda_s\lambda_{s_0}|\|A\|\|\nabla f(x_{s_0})\| \\
= \;& (s\gamma\rho + 1 - s\bar{\gamma})\|x_s - x_{s_0}\| \\
& +|s-s_0|(\gamma\|h(x_{s_0})\| \\
& +\|AProj_C(I-\lambda_s\nabla f)(x_{s_0})\|) \\
& +|\lambda_s - \lambda_{s_0}|(\|\nabla f(x_{s_0})\| + s_0\|A\|\|\nabla f(x_{s_0})\|).
\end{aligned}$$

It follows that

$$\begin{aligned}
& (s\bar{\gamma} - s\gamma\rho)\|x_s - x_{s_0}\| \\
\le \;& |s-s_0|(\gamma\|h(x_{s_0})\| \\
& +\|AProj_C(I-\lambda_s\nabla f)(x_{s_0})\| \\
& +|\lambda_s - \lambda_{s_0}|(\|\nabla f(x_{s_0})\| + s_0\|A\|\|\nabla f(x_{s_0})\|).
\end{aligned}$$

Consequently,

$$\begin{aligned}
& \|x_s - x_{s_0}\| \\
\le \;& (\frac{\gamma\|h(x_{s_0})\| + \|AProj_C(I-\lambda_s\nabla f)(x_{s_0})\|}{s\bar{\gamma} - s\gamma\rho} \\
& +\frac{|\lambda_s - \lambda_{s_0}|\|\nabla f(x_{s_0})\|}{s\bar{\gamma} - s\gamma\rho} \\
& +\frac{|\lambda_s - \lambda_{s_0}|s_0\|A\|\|\nabla f(x_{s_0})\|}{s\bar{\gamma} - s\gamma\rho} \\
& \to 0,
\end{aligned}$$

as $s \to s_0$.

This shows that x_s is continuous because of the continuity of λ_s.

Our first result below shows that $\{x_s\}$ converges in norm to a minimizer of the minimization problem (1) which solves some variational inequality.

Theorem 13 *Assume that the minimization problem (1) is consistent and let S denote its solution set. Assume that the gradient ∇f satisfies the Lipschitz condition (7). Let $\{\lambda_s\}$ satisfies $0 < \lambda_s < 2/L$. Then we have that $\{x_s\}$ converges in norm as $s \to 0$ to a minimizer of (1) which is also the unique solution of the variational inequality*

$$\langle (A-\gamma h)x^*, x^* - x\rangle \le 0, \qquad \forall\, x \,\epsilon\, S. \tag{11}$$

Equivalently, we have $Proj_S(I - A + \gamma h)x^* = x^*$.

Proof: We first show the uniqueness of a solution of the variational inequality (11), which is indeed a consequence of the strong monotonicity of $A - \gamma h$. Suppose both $x^* \,\epsilon\, S$ and $\hat{x} \,\epsilon\, S$ are solutions to (11); then

$$\langle (A-\gamma h)x^*, x^* - \hat{x}\rangle \le 0 \tag{12}$$

and

$$\langle (A-\gamma h)\hat{x}, \hat{x} - x^*\rangle \le 0 \tag{13}$$

Adding up (12) and (13) gets

$$\langle (A-\gamma h)x^* - (A-\gamma h)\hat{x}, x^* - \hat{x}\rangle \le 0. \tag{14}$$

The strong monotonicity of $A - \gamma h$ (Lemma 9) implies that $x^* = \hat{x}$ and the uniqueness is proved. Below we use $x^* \,\epsilon\, S$ to denote the unique solution of (11).

To prove that $x_s \to x^*$ ($s \to 0$), we write, for a given $z \,\epsilon\, S$,

$$\begin{aligned}
x_s - z \;=\;& (I-sA)(Proj_C(I-\lambda_s\nabla f)x_s - z) \\
& +s(\gamma h(x_s) - Az)
\end{aligned}$$

to derive that

$$\begin{aligned}
&\|x_s - z\|^2 \\
= \;& \langle (I - sA)(Proj_C(I - \lambda_s \nabla f)x_s - z), x_s - z\rangle \\
& + s\langle \gamma h(x_s) - Az, x_s - z\rangle \\
\le \;& (1 - s\bar{\gamma})\|x_s - z\|^2 + s\langle \gamma h(x_s) - Az, x_s - z\rangle.
\end{aligned}$$

It follows that

$$\begin{aligned}
\|x_s - z\|^2 &\le \frac{1}{\bar{\gamma}}\langle \gamma h(x_s) - Az, x_s - z\rangle \\
&= \frac{1}{\bar{\gamma}}\{\gamma\langle h(x_s) - h(z), x_s - z\rangle \\
&\quad + \langle \gamma h(z) - Az, x_s - z\rangle\} \\
&\le \frac{1}{\bar{\gamma}}\{\gamma\rho\|x_s - z\|^2 \\
&\quad + \langle \gamma h(z) - Az, x_s - z\rangle\}.
\end{aligned}$$

Therefore,

$$\|x_s - z\|^2 \le \frac{1}{\bar{\gamma} - \gamma\rho}\langle \gamma h(z) - Az, x_s - z\rangle. \quad (15)$$

Since $\{x_s\}$ is bounded as $s \to 0$, we see that if $\{s_n\}$ is a sequence in$(0, 1)$ such that $s_n \to 0$ and $x_{s_n} \rightharpoonup \tilde{x}$, then $\tilde{x}\epsilon S$. Indeed, we assume that $\lambda_{s_n} \to \lambda\epsilon(0, 2/L)$, due to $0 < \lambda_s < 2/L$. Set $T_s := Proj_C(I - \lambda_s \nabla f)$ and $T := Proj_C(I - \lambda\nabla f)$. Notice that T is nonexpansive and $Fix(T) = S$. By Proposition 12 (ii), we obtain that

$$\begin{aligned}
&\|x_{s_n} - Tx_{s_n}\| \\
\le \;& \|x_{s_n} - T_{s_n}x_{s_n}\| + \|T_{s_n}x_{s_n} - Tx_{s_n}\| \\
= \;& \|x_{s_n} - T_{s_n}x_{s_n}\| \\
& + \|Proj_C(I - \lambda_{s_n}\nabla f)x_{s_n} - Proj_C(I - \lambda\nabla f)x_{s_n}\| \\
\le \;& \|x_{s_n} - T_{s_n}x_{s_n}\| \\
& + \|(I - \lambda_{s_n}\nabla f)x_{s_n} - (I - \lambda\nabla f)x_{s_n}\| \\
= \;& \|x_{s_n} - T_{s_n}x_{s_n}\| + |\lambda_{s_n} - \lambda|\|\nabla f(x_{s_n})\| \to 0.
\end{aligned}$$

So Lemma 8 guarantees that $\tilde{x}\epsilon Fix(T) = S$. Then by (15), we see $x_{s_n} \to \tilde{x}$.

We next prove that $\tilde{x}$ solves the variational inequality (11). Since

$$x_s = s\gamma h(x_s) + (I - sA)Proj_C(I - \lambda_s\nabla f)x_s. \quad (16)$$

We derive that,

$$(A - \gamma h)x_s = -\frac{1}{s}(I - sA)(I - Proj_C(I - \lambda_s\nabla f))x_s.$$

It follows that, for $x \,\epsilon\, S$,

$$\begin{aligned}
&\langle (A - \gamma h)x_s, x_s - x\rangle \\
= \;& -\frac{1}{s}\langle (I - sA)(I - Proj_C(I - \lambda_s\nabla f))x_s, x_s - x\rangle \\
= \;& -\frac{1}{s}\langle (I - Proj_C(I - \lambda_s\nabla f))x_s \\
& - (I - Proj_C(I - \lambda_s\nabla f))x, x_s - x\rangle \\
& + \langle A(I - Proj_C(I - \lambda_s\nabla f))x_s, x_s - x\rangle \\
\le \;& \langle A(I - Proj_C(I - \lambda_s\nabla f))x_s, x_s - x\rangle. \quad (17)
\end{aligned}$$

Since $I - Proj_C(I - \lambda_s\nabla f)$ is monotone (i.e., $\langle x - y, (I - Proj_C(I - \lambda_s\nabla f))x - (I - Proj_C(I - \lambda_s\nabla f))y\rangle \ge 0$, for $x, y \,\epsilon\, H$. This is due to the non-expansivity of $Proj_C(I - \lambda_s\nabla f)$). Now replacing s in (17) with s_n and letting $n \to \infty$, noticing that $(I - Proj_C(I - \lambda_{s_n}\nabla f))x_{s_n} \to (I - Proj_C(I - \lambda\nabla f))\tilde{x} = 0$ for $\tilde{x} \,\epsilon\, S$, then we obtain

$$\langle (A - \gamma h)\tilde{x}, \tilde{x} - x\rangle \le 0.$$

This is, $\tilde{x} \,\epsilon\, S$ is a solution of (11); hence $\tilde{x} = x^*$ by uniqueness.

In a summary, we have shown that each cluster point of $\{x_s\}$(as $s \to 0$) equals x^*. Therefore, $x_s \to x^*$ as $s \to 0$.

The variational inequality (11) can be rewritten as

$$\langle (I - A + \gamma h)x^* - x^*, x^* - x\rangle \ge 0, \qquad x \,\epsilon\, S.$$

By Lemma 10, this is equivalent to the fixed point equation

$$Proj_S(I - A + \gamma h)x^* = x^*$$

Taking $A = I$ and $\gamma = 1$ in Theorem 7, we get

Corollary 14 *Let $x_s\epsilon C$ be the unique fixed point of the contraction $x \mapsto sh(x) + (1 - s)Proj_C(I - \lambda_s\nabla f)x$. Then $\{x_s\}$ converges in norm as $s \to 0$ to a minimizer of (1) which is also the unique solution of the variational inequality*

$$\langle (I - h)x^*, x^* - x\rangle \le 0, \qquad x \,\epsilon\, S.$$

Next we study a modified gradient-projection algorithm as follows. The initial guess x_0 is taken in C arbitrarily, and the $(n + 1)th$ iterate x_{n+1} is recursively defined by

$$\begin{aligned}
x_{n+1} \;=\;& \theta_n\gamma h(x_n) \\
& + (I - \theta_n A)Proj_C(I - \lambda_n\nabla f)x_n, \\
& \qquad n \ge 0,
\end{aligned}$$

Where $\{\theta_n\}$ is a sequence in $(0, 1)$, and $\{\lambda_n\}$ is a sequence such that

$$0 < \liminf_{n\to\infty}\lambda_n \le \limsup_{n\to\infty}\lambda_n < \frac{2}{L}. \quad (18)$$

Below is the second main result of this paper.

Theorem 15 *Assume that the minimization problem (1) is consistent and let S denote its solution set. Assume that the gradient ∇f satisfies the Lipschitz condition (7). Let a sequence $\{x_n\}_{n=0}^{\infty}$ be generated by the following modified gradient-projection algorithm:*

$$\begin{aligned} x_{n+1} &= \theta_n\gamma h(x_n) \\ &\quad +(I-\theta_n A)Proj_C(I-\lambda_n\nabla f)x_n, \\ &\qquad n=0,1,2,\cdots \end{aligned} \tag{19}$$

Let the sequence of parameters $\{\lambda_n\}_{n=0}^{\infty}$ satisfy the condition (18). In addition, the following conditions are satisfied for $\{\lambda_n\}_{n=0}^{\infty}$ and $\{\theta_n\}_{n=0}^{\infty}\subset(0,1)$:

(C1) $\theta_n\to 0$;

(C2) $\sum_{n=0}^{\infty}\theta_n=\infty$;

(C3) $\sum_{n=0}^{\infty}|\theta_{n+1}-\theta_n|<\infty$;

(C4) $\sum_{n=0}^{\infty}|\lambda_{n+1}-\lambda_n|<\infty$.

Then the sequence $\{x_n\}_{n=0}^{\infty}$ converges in norm to a minimizer of (1) which is also the unique solution of the variational inequality

$$\langle(A-\gamma h)x^*,x^*-x\rangle\le 0,\qquad x\,\epsilon\,S. \tag{20}$$

In other words, x^* is the unique fixed point of the contraction $Proj_S(I-A+\gamma h)$; that is, $Proj_S(I-A+\gamma h)x^*=x^*$.

Proof: (1°): The sequence $\{x_n\}_{n=0}^{\infty}$ is bounded. Indeed, we have, for $\bar{x}\,\epsilon\,S$,

$$\begin{aligned} &\|x_{n+1}-\bar{x}\| \\ =\ &\|(I-\theta_n A)[Proj_C(x_n-\lambda_n\nabla f(x_n))-\bar{x}] \\ &+\theta_n(\gamma h(x_n)-A\bar{x})\| \\ \le\ &(1-\theta_n\bar{\gamma})\|x_n-\bar{x}\| \\ &+\|\theta_n(\gamma h(x_n)-\gamma h(\bar{x})+\gamma h(\bar{x})-A\bar{x})\| \\ \le\ &(1-\theta_n\bar{\gamma})\|x_n-\bar{x}\|+\theta_n\gamma\rho\|x_n-\bar{x}\| \\ &+\theta_n\|\gamma h(\bar{x})-A\bar{x}\| \\ =\ &(1-(\bar{\gamma}-\gamma\rho)\theta_n)\|x_n-\bar{x}\|+\theta_n\|\gamma h(\bar{x})-A\bar{x}\| \\ \le\ &\max\{\|x_n-\bar{x}\|,\frac{1}{\bar{\gamma}-\gamma\rho}\|\gamma h(\bar{x})-A\bar{x}\|\}. \end{aligned}$$

So, an induction argument shows that

$$\|x_n-\bar{x}\|\le\max\{\|x_0-\bar{x}\|,\frac{1}{\bar{\gamma}-\gamma\rho}\|\gamma h(\bar{x})-A\bar{x}\|\},\quad n\ge 0.$$

In particular, $\{x_n\}_{n=0}^{\infty}$ is bounded.

(2°): We prove that $\|x_{n+1}-x_n\|\to 0$, as $n\to\infty$. Setting $V_n:=Proj_n(I-\lambda_n\nabla f)$. Let

$$\begin{aligned} M>max\{&\sup_{n\ge 0}\|h(x_n)\|,\ \sup_{k,n\ge 0}\|AV_k(x_n)\|, \\ &\sup_{n\ge 0}\|\nabla f(x_n)\|,\ \sup_{k,n\ge 0}\|V_k(x_n)\|,\|A\|\}. \end{aligned}$$

We compute

$$\begin{aligned} &\|x_{n+1}-x_n\| \\ =\ &\|\theta_n\gamma h(x_n)+(I-\theta_n A)V_n x_n \\ &-[\theta_{n-1}\gamma h(x_{n-1})+(I-\theta_{n-1}A)V_{n-1}x_{n-1}]\| \\ =\ &\|\theta_n\gamma(h(x_n)-h(x_{n-1})) \\ &+\gamma(\theta_n-\theta_{n-1})h(x_{n-1}) \\ &+(I-\theta_n A)(V_n x_n-V_n x_{n-1}) \\ &+(I-\theta_n A)V_n x_{n-1} \\ &-(I-\theta_{n-1}A)V_{n-1}x_{n-1}\| \\ \le\ &\theta_n\gamma\rho\|x_n-x_{n-1}\|+\gamma|\theta_n-\theta_{n-1}|\|h(x_{n-1})\| \\ &+(1-\theta_n\bar{\gamma})\|x_n-x_{n-1}\| \\ &+\|(I-\theta_n A)V_n x_{n-1} \\ &-(I-\theta_{n-1}A)V_{n-1}x_{n-1}\| \\ \le\ &(1-\theta_n(\bar{\gamma}-\gamma\rho))\|x_n-x_{n-1}\| \\ &+|\theta_n-\theta_{n-1}|\gamma M \\ &+\|(I-\theta_n A)V_n x_{n-1} \\ &-(I-\theta_{n-1}A)V_{n-1}x_{n-1}\| \end{aligned} \tag{21}$$

and

$$\begin{aligned} &\|(I-\theta_n A)V_n x_{n-1}-(I-\theta_{n-1}A)V_{n-1}x_{n-1}\| \\ =\ &\|V_n x_{n-1}-\theta_n AV_n x_{n-1}-V_{n-1}x_{n-1} \\ &+\theta_{n-1}AV_{n-1}x_{n-1}\| \\ \le\ &\|V_n x_{n-1}-V_{n-1}x_{n-1}\| \\ &+\|(\theta_n-\theta_{n-1}+\theta_{n-1})AV_n x_{n-1} \\ &-\theta_{n-1}AV_{n-1}x_{n-1}\| \\ \le\ &\|V_n x_{n-1}-V_{n-1}x_{n-1}\| \\ &+|\theta_n-\theta_{n-1}|\|AV_n x_{n-1}\| \\ &+\theta_{n-1}\|A(V_n x_{n-1}-V_{n-1}x_{n-1})\| \\ \le\ &\|V_n x_{n-1}-V_{n-1}x_{n-1}\|(1+M) \\ &+M|\theta_n-\theta_{n-1}| \end{aligned} \tag{22}$$

and

$$\begin{aligned} &\|V_n x_{n-1}-V_{n-1}x_{n-1}\| \\ =\ &\|Proj_C(I-\lambda_n\nabla f)x_{n-1} \\ &-Proj_C(I-\lambda_{n-1}\nabla f)x_{n-1}\| \\ \le\ &\|(I-\lambda_n\nabla f)x_{n-1}-(I-\lambda_{n-1}\nabla f)x_{n-1}\| \\ =\ &|\lambda_n-\lambda_{n-1}|\|\nabla f(x_{n-1})\| \\ \le\ &M|\lambda_n-\lambda_{n-1}|, \end{aligned} \tag{23}$$

combining (21),(22) and (23), we obtain

$$\begin{aligned} \|x_{n+1}-x_n\| \le\ &(1-\theta_n(\bar{\gamma}-\gamma\rho))\|x_n-x_{n-1}\| \\ &+(\gamma M+M)|\theta_n-\theta_{n-1}| \\ &+(M^2+M)|\lambda_n-\lambda_{n-1}|. \end{aligned} \tag{24}$$

Apply Lemma 1 to (24) to conclude that

$$\|x_{n+1}-x_n\|\to 0$$

as $n \to \infty$.

(3°): We prove that $\omega_w(x_n) \subset S$. Let $\hat{x} \epsilon \omega_w(x_n)$ and assume that $x_{nj} \rightharpoonup \bar{x}$ for some subsequence $\{x_{nj}\}_{j=1}^{\infty}$ of $\{x_n\}_{n=0}^{\infty}$. We may further assume that $\lambda_{nj} \to \lambda \,\epsilon\, (0, 2/L)$, due to condition (18). Set $V := Proj_C(I - \lambda\nabla f)$. Notice that V is nonexpansive and $Fix(V) = S$. It turns out that

$$\begin{aligned} & \|x_{nj} - Vx_{nj}\| \\ \le\ & \|x_{nj} - V_{nj}x_{nj}\| + \|V_{nj}x_{nj} - Vx_{nj}\| \\ \le\ & \|x_{nj} - x_{nj+1}\| + \|x_{nj+1} - V_{nj}x_{nj}\| \\ & +\|Proj_C(I - \lambda_{nj}\nabla f)x_{nj} \\ & -Proj_C(I - \lambda\nabla f)x_{nj}\| \\ \le\ & \|x_{nj} - x_{nj+1}\| + \theta_{nj}\|\gamma h(x_{nj}) - AV_{nj}x_{nj}\| \\ & +\|(I - \lambda_{nj}\nabla f)x_{nj} - (I - \lambda\nabla f)x_{nj}\| \\ \le\ & \|x_{nj} - x_{nj+1}\| + |\lambda - \lambda_{nj}|M \\ & +\theta_{nj}(\gamma M + M) \qquad \to 0, \end{aligned}$$

as $j \to \infty$.

So Lemma 8 guarantees that $\omega_w(x_n) \subset Fix(V) = S$.

(4°): We prove that $x_n \to x^*$ as $n \to \infty$, where $x^* \epsilon S$ is the unique solution of the variational inequality (20). First observe that there is some $\hat{x} \epsilon \omega_w(x_n) \subset S$, such that

$$\begin{aligned} & \limsup_{n\to\infty} \langle x_n - x^*, \gamma h(x^*) - Ax^*\rangle \\ =\ & \langle \hat{x} - x^*, \gamma h(x^*) - Ax^*\rangle \le 0. \end{aligned} \tag{25}$$

We now compute

$$\begin{aligned} & \|x_{n+1} - x^*\|^2 \\ =\ & \|\theta_n\gamma h(x_n) + (I - \theta_n A)V_n x_n - x^*\|^2 \\ =\ & \|\theta_n\gamma h(x_n) - \theta_n A(x^*) \\ & +(I - \theta_n A)(V_n x_n - x^*)\|^2 \\ =\ & \|\theta_n\gamma(h(x_n) - h(x^*)) + \theta_n\gamma h(x^*) - \theta_n Ax^* \\ & +(I - \theta_n A)(V_n x_n - x^*)\|^2 \\ \le\ & \|\theta_n\gamma(h(x_n) - h(x^*)) \\ & +(I - \theta_n A)(V_n x_n - x^*)\|^2 \\ & +2\theta_n\langle x_{n+1} - x^*, \gamma h(x^*) - Ax^*\rangle \\ =\ & \|\theta_n\gamma(h(x_n) - h(x^*))\|^2 \\ & +\|(I - \theta_n A)(V_n x_n - x^*)\|^2 \\ & +2\theta_n\gamma\langle h(x_n) - h(x^*), (I - \theta_n A)(V_n x_n - x^*)\rangle \\ & +2\theta_n\langle x_{n+1} - x^*, \gamma h(x^*) - Ax^*\rangle \\ \le\ & \theta_n^2\gamma^2\rho^2\|x_n - x^*\|^2 + (1 - \theta_n\bar{\gamma})^2\|x_n - x^*\|^2 \\ & +2\theta_n\gamma\rho(1 - \theta_n\bar{\gamma})\|x_n - x^*\|^2 \\ & +2\theta_n\langle x_{n+1} - x^*, \gamma h(x^*) - Ax^*\rangle \\ =\ & [1 - (2\theta_n^2\gamma\rho\bar{\gamma} + 2\theta_n\bar{\gamma} - 2\theta_n\gamma\rho \\ & -\theta_n^2\gamma^2\rho^2)]\|x_n - x^*\|^2 \\ & +2\theta_n\langle x_{n+1} - x^*, \gamma h(x^*) - Ax^*\rangle. \end{aligned}$$

It follows that

$$\|x_{n+1} - x^*\|^2 \le (1 - \gamma_n)\|x_n - x^*\|^2 + \delta_n. \tag{26}$$

where

$$\begin{aligned} \gamma_n &= 2\theta_n^2\gamma\rho\bar{\gamma} + 2\theta_n\bar{\gamma} - 2\theta_n\gamma\rho - \theta_n^2\gamma^2\rho^2 \\ \delta_n &= 2\theta_n\langle x_{n+1} - x^*, \gamma h(x^*) - Ax^*\rangle \end{aligned}$$

By (25), we get

$$\limsup_{n\to\infty} \delta_n/\gamma_n \le 0.$$

Now applying Lemma 7 to (26) concludes that $x_n \to x^*$ as $n \to \infty$.

Corollary 16 *Let $\{x_n\}$ be generated by the following hybrid gradient-projection algorithm:*

$$x_{n+1} = \theta_n h(x_n) + (1 - \theta_n)Proj_C(I - \lambda_n\nabla f)x_n, \quad n = 0, 1, 2, \cdots$$

Let the sequence of parameters $\{\lambda_n\}_{n=0}^{\infty}$ satisfy the condition (18). Assume that the sequence $\{\theta_n\}_{n=0}^{\infty}$ and $\{\lambda_n\}_{n=0}^{\infty}$ satisfy conditions (C1)–(C4). Then $\{x_n\}_{n=0}^{\infty}$ converges in norm to a minimizer of (1) which is obtained in Corollary 14.

4 Application of the modified GPA

In this section, we give an application of Theorem 15 to the split feasibility problem (say SFP, for short) which was introduced by Censor and Elfving [26]. Since its inception in 1994, the split feasibility problem (SFP) has received much attention due to its applications in signal processing and image reconstruction, with particular progress in intensity-modulated radiation therapy.

The SFP can mathematically be formulated as the problem of finding a point x with the property

$$x \,\epsilon\, C, \quad and \quad Bx \,\epsilon\, Q, \tag{27}$$

where C and Q are nonempty closed convex subset of Hilbert space H_1 and H_2, respectively. $B : H_1 \to H_2$ is a bounded linear operator.

It is clear that x^* is a solution to the split feasibility problem (27) if and only if $x^* \epsilon C$ and $Bx^* - Proj_Q Bx^* = 0$. We define the proximity function f by

$$f(x) = \frac{1}{2}\|Bx - Proj_Q Bx\|^2,$$

and consider the constrained convex minimization problem

$$\min_{x \epsilon C} f(x) = \min_{x \epsilon C} \frac{1}{2}\|Bx - Proj_Q Bx\|^2. \tag{28}$$

Then, x^* solves the split feasibility problem (27) if and only if x^* solves the minimization problem (28) with the minimize equal to 0. Byrne [28] introduced the so-called CQ algorithm to solve the (SFP).

$$x_{n+1} = Proj_C(I-\lambda B^*(I-Proj_Q)B)x_n, \qquad n \geq 0, \tag{29}$$

where $0 < \lambda < 2/\rho(B^*B)$ and $\rho(B^*B)$ is the spectral radius of the self-adjoint operator B^*B. He obtained that the sequence $\{x_n\}$ generated by (29) converges weekly to a solution of the (SFP).

In order to obtain strong convergence iterative sequence to solve the (SFP). We propose the following algorithm:

$$\begin{aligned} & x_{n+1} \\ = \ & (I-\theta_n A)Proj_C(I-\lambda_n B^*(I-Proj_Q)B)x_n, \\ & +\theta_n\gamma h(x_n), \qquad\qquad n \geq 0, \end{aligned} \tag{30}$$

where $h : C \to C$ is a contraction such that $\|h(x) - h(y)\| \leq \rho\|x-y\|, \forall\ x, y\ \epsilon\ C$, and $\rho\ \epsilon\ [0,1)$ is a constant. $A : C \to C$ is strongly positive bounded linear operator; that is, there is a constant $\bar{\gamma} > 0$ such that $\langle Ax, x\rangle \geq \bar{\gamma}\|x\|^2, \forall\ x\ \epsilon\ C$. Where $0 < \gamma < \bar{\gamma}/\rho$. We can show that the sequence $\{x_n\}$ generated by (30) converges strongly to a solution of the (SFP) (27) if the sequence $\{\theta_n\} \subset [0,1]$ and the sequence $\{\lambda_n\}$ of parameters satisfy appropriate conditions.

Applying Theorem 15, we obtain the following result.

Theorem 17 *Assume that the split feasibility problem (27) is consistent. In addition, if $0\epsilon C$ or C is closed convex cone. Let $0 < \liminf_{n\to\infty}\lambda_n \leq \limsup_{n\to\infty}\lambda_n < 2/\|B\|^2$. Let the sequence $\{x_n\}$ be generated by (30). Where the sequence $\{\theta_n\} \subset [0,1]$ and the sequence $\{\lambda_n\}$ satisfy the conditions (C1)–(C4). Then the sequence $\{x_n\}$ converges strongly to a solution of the split feasibility problem (27).*

Proof: By the definition of the proximity function f, we have

$$\nabla f(x) = B^*(I - Proj_Q)Bx,$$

and ∇f is Lipschitz continuous, i.e.,

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x-y\|,$$

where $L = \|B\|^2$. Then the iterative scheme (30) is equivalent to

$$x_{n+1} = \theta_n\gamma h(x_n) + (I-\theta_n A)Proj_C(I-\lambda_n\nabla f)x_n.$$

Due to Theorem 15, we have the conclusion immediately.

Acknowledgements: The research was Supported in part by The Fundamental Research Funds for the Central Universities (the Special Fund of Science in Civil Aviation University of China: No. ZXH2012 K001) and by the Science Research Foundation of Civil Aviation University of China (No.2012KYM03).

References:

[1] E.–S. Levitin and B.–T. Polyak, Constrained minimization methods, *Zh. Vychisl. Mat. Mat. Fiz.* 6, 1996, pp. 787–823.

[2] P.–H. Calamai and J.–J. Moré, Projected gradient methods for linearly constrained problems, *Math. Program.* 39, 1987, pp. 93–116.

[3] B.–T. Polyak, Introduction to optimization, In: Optimization Software, New York, 1987.

[4] M. Su and H.–K. Xu, Remarks on the gradient-projection algorithm, *J. Nonlinear Anal. Optim.*1, 2010, pp. 35–43.

[5] H.–K. Xu, Averaged Mappings and the Gradient-Projection Algorithm, *J. Optim Theory. Appl.* 2011, 150.

[6] G. Marino and H.–K. Xu, A generated iterative method for nonexpansive mappings in Hilbert spaces, *J. Math. Anal. Appl.* 318, 2006, pp. 43–52.

[7] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Probl.* 20, 2004, pp. 103–120.

[8] P.–L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, Optimization 53, 2004, pp. 475–504.

[9] K. Geobel and W.–A. Kirk, Topics in Metric Fixed Point Theory, Cambridge Studies in Advanced Mathematics, vol. 28. Cambridge University Press, Cambridge 1990.

[10] D.–P. Bertsekas and E.–M. Gafni, Projection methods for variational inequalities with applications to the traffic assignment problem, *Math. Program. Stud.* 17, 1982, pp. 139–159.

[11] D. Han and H.–K. Lo, Solving non-additive traffic assignment problems, a descent method for cocoercive variational inequalities, *Eur. J. Oper. Res.* 159, 2004, pp. 529–544.

[12] C. Martinez-Yanes and H.–K. Xu, Strong convergence of the CQ method for fixed-point iteration processes. *Nonlinear Anal.* 64, 2006, pp. 2400–2411.

[13] H.–K. Xu, Iterative algorithms for nonlinear operators, *J. Lond. Math. Soc.* 66, 2002, pp. 240–256.

[14] H. Hundal, An alternating projection that does not converge in norm, *Nonlinear Anal.* 57, 2004, pp. 35–61.

[15] H.–H. Bauschke, J.–V. Burke, F.–R. Deutsch, H.–S. Hundal and J.–D. Vanderwerff, A new proximal point iteration that converges weekly but not in norm. *Proc. Am. Math. Soc.* 133, 2005, pp. 1829–1835.

[16] H.–H. Bauschke, E. Matouskova and S. Reich, Projections and proximal point methods, Convergence results and counterexamples, *Nonlinear Anal.* 56, 2004, pp. 715–738.

[17] E. Matouskova and S. Reich, The Hundal example revisited, *J. Nonlinear Convex Anal.* 4, 2003, pp. 411–427.

[18] M.–V. Solodov and B.–F. Svaiter, Forcing strong convergence of proximal point iterations in a Hilbert space, *Math. Program., Ser.* A 87, 2000, pp. 189–202.

[19] G. Marino and H.–K. Xu, Convergence of generalized proximal point algorithm, *Commun. Pure Appl. Anal.* 3, 2004, pp. 791–808.

[20] H.–K. Xu, A regularization method for the proximal point algorithm, *J. Glob. Optim.* 36, 2006, pp. 115–125.

[21] R.–T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14, 1976, pp. 877–898.

[22] O. Güler, On the convergence of the proximal point algorithm for convex optimization, *SIAM J. Control Optim.* 29, 1991, pp. 403–419.

[23] A. Moudafi, Viscosity approximation methods for fixed-points problems, *J. Math. Anal. Appl.* 241, 2000, pp. 46–55.

[24] H.–K. Xu, Viscosity approximation methods for nonexpansive mappings, *J. Math. Anal. Appl.* 298, 2004, pp. 279–291.

[25] L.–C. Ceng, Q.–H. Ansari and J.–C. Yao, Some iterative methods for finding fixed points and for solving constrained convex minimization problems, *Nonlinear Analysis.* 74, 2011, pp. 5286–5302.

[26] Y. Censor and T. Elfving, A multiprojection algorithm using Bregman projections in a product space, *Numer. Alg.* 8, 1994, pp. 221–239.

[27] M. Su and H.–K. Xu, Remarks on the gradient-projection algorithm, *J. Optim. Theory Appl.* 1(1), 2010, pp. 35–43.

[28] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Probl.* 20, 2004, pp. 103–120.

[29] M. Tian, A general iterative algorithm for nonexpansive mappings in Hilbert spaces, *Nonlinear Analysis.* 73, 2010, pp. 689–694.

[30] H. Brezis, *Operateurs Maximaux Monotones et Semi-Groups de Contractions dans les Espaces de Hilbert*, North–Holland, Amsterdam 1973.