J. Appl. Numer. Optim. 1 (2019), No. 2, pp. 145-156
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.1.2019.2.05

POLYAK'S GRADIENT METHOD FOR SOLVING THE SPLIT CONVEX FEASIBILITY PROBLEM AND ITS APPLICATIONS

AVIV GIBALI[1,2,*], NGUYEN H. HA[3], NGO T. THUONG[4], TRINH H. TRANG[5], NGUYEN T. VINH[3]

[1]*Department of Mathematics, ORT Braude College, Karmiel 2161002, Israel*
[2]*The Center for Mathematics and Scientific Computation, University of Haifa, Mt. Carmel, Haifa 3498838, Israel*
[3]*Department of Mathematics, University of Transport and Communications, 3 Cau Giay Street, Hanoi, Vietnam*
[4]*College of Science, Vietnam National University, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam*
[5]*Nam Dinh Teachers' Training College, 813 Truong Chinh Street, Nam Dinh city, Vietnam*

Abstract. In this paper, we are concerned with the problem of finding minimum-norm solutions of a split convex feasibility problem in real Hilbert spaces. We study and analyze the convergence of a new self-adaptive CQ algorithm. The main advantage of the algorithm is that there is no need to calculate the norm of the involved operator.

Keywords. Split feasibility problem; Variational inequalities; Strong convergence; Tikhonov regularization method; Gradient projection method.

2010 Mathematics Subject Classification. 47H05, 90C25.

1. INTRODUCTION

In 2012, Censor, Gibali and Reich [9] introduced the *Split Inverse Problem* (SIP) which concerns the following model. Given two vector spaces X and Y and a bounded linear operator $A : X \to Y$. In addition, two inverse problems are involved. The first one, denoted by IP_1, is formulated in the space X and the second one, denoted by IP_2, is formulated in the space Y. Given these data, the split inverse problem (SIP) is formulated as follows:

$$\text{find a point } x^* \in X \text{ that solves } \mathrm{IP}_1$$
$$\text{and such that}$$
$$\text{the point } y^* = Ax^* \in Y \text{ solves } \mathrm{IP}_2.$$

The first important example of the SIP is the *Split Convex Feasibility Problem* (SCFP), which is due to Censor and Elfving [6]. In this problem, IP_1 and IP_2 are nonempty, closed and convex sets $C \subseteq H_1$ and $Q \subseteq H_2$ (H_1, H_2 be two real Hilbert spaces). With this data and given a bounded linear operator $A : H_1 \to H_2$, the SCFP is defined as follows

$$\text{find } x^* \in C \text{ such that } Ax^* \in Q. \tag{1.1}$$

*Corresponding author.
E-mail addresses: avivg@braude.ac.il (A. Gibali), nguyenhaiha89ad@gmail.com (N.H. Ha), ngothithuong.k56@hus.edu.vn (N.T. Thuong), huutrangnd@gmail.com (T.H. Trang), thevinhbn@utc.edu.vn (N.T. Vinh).

Received July 8, 2019; Accepted August 10, 2019.

This model has been applied to solve inverse problems in various fields, such as, signal processing, image reconstruction and intensity-modulated radiation therapy (IMRT) treatment planning; see, e.g., [2, 3, 5, 6]. Minimum-norm solutions of the SCFP, which is the core of this study, play an important role and are more desirable in some scenarios, for example, in the field of IMRT, such solutions are called the least-intensity feasible (LIF) solutions; see Xiao *et al.* [20]. For simplicity, we denote the solution set of (1.1) by $\Omega := \{x^* \in C \mid Ax^* \in Q\}$. The SCFP was introduced by Censor and Elfving [6], however, their algorithm requires the computation of the inverse of A. Byrne in [2, 3] proposed to reformulate (1.1) as the following constrained minimization problem

$$\min_{x\in C} g(x) := \frac{1}{2}\|Ax - P_Q Ax\|^2, \tag{1.2}$$

where P_Q denoted the orthogonal projection onto set Q (will be explained in the sequel). Applying Polyak's gradient method (since g is continuously differentiable and convex with Lipschitz gradient given by $\nabla g(x) = A^*(I-P_Q)Ax$) to (1.2), Byrne's CQ algorithm is as follows

$$x_{k+1} = P_C(x_k - \gamma A^T(I-P_Q)Ax_k) \tag{1.3}$$

where $\gamma \in \left(0, \frac{2}{\|A\|^2}\right)$ and $\|A\|^2$ denotes the operator norm. Observe that another reformulation of (1.1) which is based on the first optimality condition of (1.2) yields the following Variational Inequality (VI) problem of finding $x \in C$ such that

$$\langle \nabla g(x), y - x\rangle \geq 0 \ \ \forall y \in C. \tag{1.4}$$

For more details on the SCFP, VIs and the CQ algorithm, we refer to [1, 2, 3, 4, 11, 13, 19, 22, 23] and the references therein. It is known that (1.3) converges weakly in real Hilbert spaces only; see, e.g., Xu [22]. In order to establish strong convergence, Xu proposed the following iterative step which is based on the Tikhonov regularization.

$$x^{k+1} = P_C((1-\alpha_k\gamma_k)x^k - \gamma_k \nabla g(x^k)), \ \ k \geq 0. \tag{1.5}$$

Assume that the parameters $\{\alpha_k\}$ and $\{\gamma_k\}$ satisfy the following conditions:

(1) $0 < \gamma_k \leq \frac{\alpha_k}{\|A\|^2 + \alpha_k}$;
(2) $\alpha_k \to 0$ and $\gamma_k \to 0$;
(3) $\sum_{k=0}^{\infty} \alpha_k \gamma_k = \infty$;
(4) $(|\gamma_{k+1} - \gamma_k| + \gamma_k|\alpha_{k+1} - \alpha_k|)/(\alpha_{k+1}\gamma_{k+1})^2 \to 0$.

Xu [22] proved that (1.5) converges strongly to the minimum-norm solution of the SCFP (1.1). Other related results with different assumptions can be found in [8, 24]. Instead of reformulating the SCFP as a constrained minimization problem (1.2), the following unconstrained minimization is also applied

$$f(x) := \frac{1}{2}\|x - P_C x\|^2 + \frac{1}{2}\|Ax - P_Q Ax\|^2. \tag{1.6}$$

Since function f is continuously differentiable and convex with gradient $\nabla f(x) = x - P_C x + A^*(I-P_Q)Ax$, Qu and Liu [16] introduced the following gradient descent method for finding a minimum-norm solution of the SCFP (1.1), that is,

$$x^{k+1} = x^k - \gamma_k \nabla f_{\alpha_k}(x^k), \ \ k \geq 0, \tag{1.7}$$

where

$$f_{\alpha_k}(x^k) = \nabla f(x^k) + \alpha_k x^k$$

and

$$0 < \gamma_k < \frac{\alpha_k}{(1 + \|A\|^2 + \alpha_k)^2}.$$

Since the above two methods require the knowledge or approximation of the operator norm $\|A\|$, a computational effort which might be expensive or even not practical, a natural question arises:

Can we introduce a new CQ type method which does not require the computation of the operator norm and still converges strongly?

Motivated and inspired by the works of Lopéz et al. [13], Qu and Liu [16], Tian and Zhang [17], Wang and Xu [18], Xu [22] and Yao, Jigang and Liou [24], we give an answer to the above question by introducing a new self-adaptive CQ type algorithm for finding minimum-solutions of the SCFP which converges strongly in real Hilbert spaces. The outline of the paper as follows. We recall some useful definitions and lemmas in Section 2. In Section 3, our new algorithm is introduced and analyzed. Finally in Section 4, we present a numerical example to illustrate and compare our algorithm's efficient.

2. Preliminaries

Throughout this paper H (might be with subscripts) denotes a real Hilbert space with inner product $\langle , \rangle$ and induced norm $\| \|$. We denote the strong (weak) convergence of a sequence $\{x^k\}$ to a point x by $x^k \to x$ ($x^k \rightharpoonup x$). Recall the following simple norm inequality

$$\|tx + (1-t)y\|^2 \le t\|x\|^2 + (1-t)\|y\|^2, \tag{2.1}$$

for all $x, y \in H$ and for all $t \in [0,1]$. Given a nonempty, closed and convex set $C \subset H$, for every element $x \in H$, there exists a unique nearest point in C, denoted by $P_C x$ such that

$$||x - P_C x|| = \inf\{||x - y|| \mid \ y \in C\},$$

where P_C is called the *metric projection* of H onto C.

Lemma 2.1. *The metric projection P_C has the following basic properties:*

(1) $\langle x - P_C x, y - P_C x\rangle \le 0$ *for all $x \in H$ and $y \in C$;*
(2) $\|P_C x - P_C y\| \le \|x - y\|$ *for all $x, y \in H$;*
(3) $\|P_C x - P_C y\|^2 \le \langle x - y, P_C x - P_C y\rangle$ *for all $x, y \in H$;*

Next we present three technical lemmas which are useful to our convergence analysis.

Lemma 2.2. (Maingé [14]) *Let $\{\Gamma_n\}$ be a sequence of real numbers that does not decrease at infinity, in the sense that there exists a subsequence $\{\Gamma_{n_j}\}$ of $\{\Gamma_n\}$ such that $\Gamma_{n_j} < \Gamma_{n_j+1}$ for all $j \ge 0$. Also consider the sequence of integers $\{\tau(n)\}_{n \ge n_0}$ defined by*

$$\tau(n) = \max\{k \le n \mid \Gamma_k < \Gamma_{k+1}\}.$$

Then $\{\tau(n)\}_{n \ge n_0}$ is a nondecreasing sequence verifying $\lim\limits_{n\to\infty} \tau(n) = \infty$ and, for all $n \ge n_0$,

$$\max\{\Gamma_{\tau(n)}, \Gamma_n\} \le \Gamma_{\tau(n)+1}.$$

Lemma 2.3. (Xu [21]) *Assume that $\{a_k\}$ is a sequence of nonnegative real numbers such that*

$$a_{k+1} \le (1-\alpha_k)a_k + \alpha_k \gamma_k + b_k, \quad k \in \mathbb{N},$$

where $\{\alpha_k\}$ is a sequence in $(0,1)$, $\{b_k\}$ is a sequence of nonnegative real numbers and $\{\gamma_k\}$ is a sequence of real numbers such that

(1) $\sum\limits_{k=0}^{\infty} \alpha_k = \infty$,
(2) $\sum\limits_{k=0}^{\infty} b_k < \infty$,
(3) $\limsup_{k\to\infty} \gamma_k \le 0$.

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

We end this section by recalling a new fundamental tool which will be helpful for proving strong convergence of our algorithm.

Lemma 2.4. (He and Yang [12]) *Assume that $\{s_k\}$ is a sequence of nonnegative real numbers such that for all $k \in \mathbb{N}$*

$$s_{k+1} \le (1-\alpha_k)s_k + \alpha_k \delta_k,$$
$$s_{k+1} \le s_k - \eta_k + \gamma_k,$$

where $\{\alpha_k\}$ is a sequence in $(0,1)$, $\{\eta_k\}$ is a sequence of nonnegative real numbers, and $\{\delta_k\}$ and $\{\gamma_k\}$ are two sequences in $\mathbb{R}$ such that

(1) $\sum_{k=0}^{\infty} \alpha_k = \infty$,
(2) $\lim_{k\to\infty} \gamma_k = 0$,
(3) $\lim_{k\to\infty} \eta_{n_k} = 0$ *implies that* $\limsup_{k\to\infty} \delta_{n_k} \le 0$ *for any subsequence* $\{n_k\}$ *of* $\{n\}$.

Then $\lim_{s\to\infty} s_k = 0$.

3. The algorithm and its analysis

In this section, we present our new CQ scheme which is motivated by (1.7) and use a self-adaptive step size to solve the SCFP (1.1). For the convergence proof of the algorithm we assume the following conditions.

Condition 3.1. The solution set of (1.1) is nonempty, that is, $\Omega \neq \emptyset$;

Condition 3.2. The positive sequences $\{\beta_k\}$ and $\{\rho_k\}$ satisfy:

$$\{\beta_k\} \subset (0,1), \ \lim_{k\to\infty} \beta_k = 0, \ \sum_{k=0}^{\infty} \beta_k = \infty, \tag{3.1}$$

$$\inf_k \rho_k \left(4 - \frac{\rho_k}{1-\beta_k}\right) > 0. \tag{3.2}$$

Algorithm 3.1. (Gradient type algorithm for solving (1.1))

Initialization: Choose two sequences $\{\beta_k\}$ and $\{\rho_k\}$ that fulfils (3.1) and (3.2), arbitrary starting point $x^0 \in H_1$ and set $k = 0$.

Iterative Step: Given the current iterate x^k, if

$$\nabla f(x^k) = x^k - P_C x^k + A^*(I - P_Q)Ax^k = 0,$$

then stop. x^k is a solution of (1.1) (where f is defined in (1.6)). Otherwise, compute

$$\lambda_k = \rho_k \frac{\|x - P_C x\|^2 + \|Ax - P_Q Ax\|^2}{2\|x - P_C x + A^*(I - P_Q)Ax\|^2}$$

and update the next iterate as

$$x^{k+1} = (1-\beta_k)x^k - \lambda_k \nabla f(x^k). \tag{3.3}$$

Set $k \leftarrow k+1$ and return to **Iterative Step**.

We start the analysis of the algorithm by showing the validity of its stopping rule.

Lemma 3.1. *Assume that Conditions 3.1 and 3.2 hold. If $\nabla f(x^k) = 0$ in Algorithm 3.1, then $x^k \in \Omega$.*

Proof. We have

$$
\begin{aligned}
0 &= \langle \nabla f(x^k), x^k - z\rangle \\
&= \langle x^k - P_C x^k, x^k - z\rangle + \langle (I-P_Q)Ax^k, Ax^k - Az\rangle \\
&= \langle x^k - P_C x^k, x^k - z\rangle + \langle (I-P_Q)Ax^k - (I-P_Q)Az, Ax^k - Az\rangle \\
&\geq \|x^k - P_C x^k\|^2 + \|(I-P_Q)Ax^k\|^2
\end{aligned}
$$

and this implies that $x^k \in C$ and $Ax^k \in Q$, which completes the proof. □

Next we establish a property of monotonicity which will be used in the sequel.

Lemma 3.2. *Assume that Conditions 3.1 and 3.2 hold and let $\{x^k\}$ be any sequence generated by Algorithm 3.1. Then, for each $z \in \Omega$, the following inequality holds:*

$$
\|x^{k+1} - z\|^2 \leq \beta_k \|z\|^2 + (1-\beta_k)\|x^k - z\|^2 - \rho_k\left(4 - \frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2}.
$$

Proof. By Lemma 2.1(2) and (3.3), we have

$$
\begin{aligned}
\|x^{k+1} - z\|^2 &= \|(1-\beta_k)x^k - \lambda_k \nabla f(x^k) - z\|^2 \\
&= \left\|\beta_k(-z) + (1-\beta_k)\left(x^k - \frac{\lambda_k}{1-\beta_k}\nabla f(x^k) - z\right)\right\|^2 \qquad (3.4)\\
&\leq \beta_k\|z\|^2 + (1-\beta_k)\left\|x^k - \frac{\lambda_k}{1-\beta_k}\nabla f(x^k) - z\right\|^2. \qquad (3.5)
\end{aligned}
$$

Note that

$$
\begin{aligned}
\langle \nabla f(x^k), x^k - z\rangle &= \langle x^k - P_C x^k, x^k - z\rangle + \langle (I-P_Q)Ax^k, Ax^k - Az\rangle \\
&= \langle x^k - P_C x^k, x^k - z\rangle + \langle (I-P_Q)Ax^k - (I-P_Q)Az, Ax^k - Az\rangle \\
&\geq \|x^k - P_C x^k\|^2 + \|(I-P_Q)Ax^k\|^2 \\
&= 2f(x^k).
\end{aligned}
$$

We estimate the second term on the right-hand side of (3.5) as follows:

$$
\begin{aligned}
&\left\|x^k - \frac{\lambda_k}{1-\beta_k}\nabla f(x^k) - z\right\|^2 \\
&= \|x^k - z\|^2 + \frac{\lambda_k^2}{(1-\beta_k)^2}\|\nabla f(x^k)\|^2 - \frac{2\lambda_k}{1-\beta_k}\langle \nabla f(x^k), x^k - z\rangle \\
&\leq \|x^k - z\|^2 + \frac{\lambda_k^2}{(1-\beta_k)^2}\|\nabla f(x^k)\|^2 - \frac{4\lambda_k}{1-\beta_k} f(x^k) \\
&\leq \|x^k - z\|^2 + \frac{\rho_k^2 f^2(x^k)}{(1-\beta_k)^2\|\nabla f(x^k)\|^2} - \frac{4\rho_k f^2(x^k)}{(1-\beta_k)\|\nabla f(x^k)\|^2}. \qquad (3.6)
\end{aligned}
$$

From (3.5) and (3.6), we arrive at

$$
\begin{aligned}
\|x^{k+1} - z\|^2 &\leq \beta_k\|z\|^2 + (1-\beta_k)\|x^k - z\|^2 + \frac{\rho_k^2 f^2(x^k)}{(1-\beta_k)\|\nabla f(x^k)\|^2} - \frac{4\rho_k f^2(x^k)}{\|\nabla f(x^k)\|^2} \\
&= \beta_k\|z\|^2 + (1-\beta_k)\|x^k - z\|^2 - \rho_k\left(4 - \frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2}
\end{aligned}
$$

which completes the proof. □

Lemma 3.3. *Assume that Conditions 3.1 and 3.2 hold, then the sequence $\{x^k\}$ generated by Algorithm 3.1 is bounded.*

Proof. By Lemma 3.2 and (3.2), we have

$$\begin{aligned}\|x^{k+1}-z\|^2 &\le \beta_k\|z\|^2+(1-\beta_k)\|x^k-z\|^2-\rho_k\left(4-\frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2}\\ &\le \beta_k\|z\|^2+(1-\beta_k)\|x^k-z\|^2.\end{aligned}$$

It follows that

$$\|x^{k+1}-z\|^2 \le \max\{\|z\|^2,\|x^k-z\|^2\}.$$

By induction,

$$\|x^{k+1}-z\|^2 \le \max\{\|z\|^2,\|x^0-z\|^2\},$$

which implies the boundedness of $\{x^k\}$. □

Lemma 3.4. *Assume that Conditions 3.1 and 3.2 hold. Let $\{x^k\}$ be a sequence generated by Algorithm 3.1. Then the following inequality holds, for all $z\in\Omega$ and $k\in\mathbb{N}$,*

$$\|x^{k+1}-z\|^2 \le (1-\beta_k)\|x^k-z\|^2+\beta_k\left[\beta_k\|z\|^2+2(1-\beta_k)\langle x^k-z,-z\rangle+2\lambda_k\langle\nabla f(x^k),z\rangle\right].$$

Proof. By (3.2) and (3.6), we have

$$\begin{aligned}\left\|x^k-\frac{\lambda_k}{1-\beta_k}\nabla f(x^k)-z\right\|^2 &\le \|x^k-z\|^2+\frac{\rho_k^2 f^2(x^k)}{(1-\beta_k)^2\|\nabla f(x^k)\|^2}-\frac{4\rho_k f^2(x^k)}{(1-\beta_k)\|\nabla f(x^k)\|^2}\\ &= \|x^k-z\|^2-\frac{\rho_k f^2(x^k)}{(1-\beta_k)\|\nabla f(x^k)\|^2}\left(4-\frac{\rho_k}{1-\beta_k}\right)\\ &\le \|x^k-z\|^2.\end{aligned}$$

From (3.4) and Lemma 3.2, we obtain

$$\begin{aligned}\|x^{k+1}-z\|^2 &\le \left\|\beta_k(-z)+(1-\beta_k)\left(x^k-\frac{\lambda_k}{1-\beta_k}\nabla f(x^k)-z\right)\right\|^2\\ &\le \beta_k^2\|z\|^2+(1-\beta_k)^2\left\|x^k-\frac{\lambda_k}{1-\beta_k}\nabla f(x^k)-z\right\|^2\\ &\qquad +2\beta_k(1-\beta_k)\left\langle x^k-\frac{\lambda_k}{1-\beta_k}\nabla f(x^k)-z,-z\right\rangle\\ &\le \beta_k^2\|z\|^2+(1-\beta_k)^2\|x^k-z\|^2+2\beta_k(1-\beta_k)\langle x^k-z,-z\rangle\\ &\qquad +2\beta_k\lambda_k\langle\nabla f(x^k),z\rangle\\ &\le (1-\beta_k)\|x^k-z\|^2+\beta_k\left[\beta_k\|z\|^2+2(1-\beta_k)\langle x^k-z,-z\rangle+2\lambda_k\langle\nabla f(x^k),z\rangle\right].\end{aligned}$$

This completes the proof. □

After proving the above, we can establish the main convergence theorem of Algorithm 3.1.

Theorem 3.1. *Assume that Conditions 3.1 and 3.2 hold. Then the sequence $\{x^k\}$ generated by Algorithm 3.1 converges strongly to the minimum-norm element of Ω.*

We provide two proofs of Theorem 3.1. The first proof is as follows.

Proof. Let $z:=P_\Omega 0$. From Lemma 3.2, we have

$$\|x^{k+1}-z\|^2 \le \beta_k\|z\|^2+(1-\beta_k)\|x^k-z\|^2-\rho_k\left(4-\frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2} \tag{3.7}$$

or

$$\rho_k\left(4-\frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2} \le \beta_k\|z\|^2+(1-\beta_k)\|x^k-z\|^2-\|x^{k+1}-z\|^2.$$

So, we obtain

$$\rho_k\left(4-\frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2}\le\|x^k-z\|^2-\|x^{k+1}-z\|^2+\beta_k\|z\|^2. \tag{3.8}$$

Next, we consider two possible cases.

Case 1. Put $\Gamma_k:=\|x^k-z\|^2$ for all $k\in\mathbb{N}$. Assume that there is a $k_0\ge 0$ such that, for each $k\ge n_0$, $\Gamma_{k+1}\le\Gamma_k$. In this case, $\lim_{k\to\infty}\Gamma_k$ exists and $\lim_{k\to\infty}(\Gamma_k-\Gamma_{k+1})=0$. Since $\lim_{k\to\infty}\beta_k=0$, it follows from (3.8) that

$$\lim_{k\to\infty}\rho_k\left(4-\frac{\rho_k}{1-\beta_k}\right)\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2}=0.$$

From the assumption

$$\inf_k\rho_k(4-\frac{\rho_k}{1-\beta_k})>0,$$

we obtain

$$\lim_{k\to\infty}\frac{f^2(x^k)}{\|\nabla f(x^k)\|^2}=0,$$

i.e.,

$$\lim_{k\to\infty}\frac{(\|x^k-P_Cx^k\|^2+\|Ax^k-P_QAx^k\|^2)^2}{\|x-P_Cx^k+A^*(I-P_Q)Ax^k\|^2}=0. \tag{3.9}$$

On the other hand, we have

$$\begin{aligned}\frac{(\|x^k-P_Cx^k\|^2+\|Ax^k-P_QAx^k\|^2)^2}{\|x-P_Cx^k+A^*(I-P_Q)Ax^k\|^2}&\ge\frac{(\|x^k-P_Cx^k\|^2+\|Ax^k-P_QAx^k\|^2)^2}{2(\|x-P_Cx^k\|^2+\|A\|^2\|(I-P_Q)Ax^k\|^2)}\\&\ge\frac{(\|x^k-P_Cx^k\|^2+\|Ax^k-P_QAx^k\|^2)^2}{2\max(1,\|A\|^2)(\|x-P_Cx^k\|^2+\|(I-P_Q)Ax^k\|^2)}\\&=\frac{\|x^k-P_Cx^k\|^2+\|Ax^k-P_QAx^k\|^2}{2\max(1,\|A\|^2)}.\end{aligned}$$

This together with (3.9) implies that

$$\lim_{k\to\infty}\|x^k-P_Cx^k\|=0, \tag{3.10}$$

$$\lim_{k\to\infty}g(x^k)=\lim_{k\to\infty}\frac{1}{2}\|(I-P_Q)Ax^k\|^2=0. \tag{3.11}$$

We now show that $\omega_w(\{x^k\})\subset\Omega$. Let $\bar{x}\in\omega_w(\{x^k\})$ be an arbitrary element. Since $\{x^k\}$ is bounded (by Lemma 3.3), there exists a subsequence $\{x^{k_j}\}$ of $\{x^k\}$ such that $x^{k_j}\rightharpoonup\bar{x}$. We have from (3.10) that $\bar{x}=P_C\bar{x}\in C$. From the weak lower semicontinuity of g, we obtain

$$0\le g(\bar{x})\le\liminf_{j\to\infty}g(x^{k_j})=\lim_{k\to\infty}g(x^k)=0.$$

We immediately deduce that $g(\bar{x})=0$, i.e., $A\bar{x}\in Q$. The choice of $\bar{x}$ in $\omega_w(\{x^k\})$ is arbitrary, and so we conclude that $\omega_w(\{x^k\})\subset\Omega$. Using Lemma 3.4, we have

$$\begin{aligned}\|x^{k+1}-z\|^2&\le(1-\beta_k)\|x^k-z\|^2+\beta_k\left[\beta_k\|z\|^2+2(1-\beta_k)\langle x^k-z,-z\rangle+2\lambda_k\langle\nabla f(x^k),z\rangle\right]\\&\le(1-\beta_k)\|x^k-z\|^2+\beta_k\Big[\beta_k\|z\|^2+2(1-\beta_k)\langle x^k-z,-z\rangle+2\lambda_k\big\|\nabla f(x^k)\big\|\|z\|\Big].\end{aligned}$$

From Lemma 2.3, it remains to show that

$$\limsup_{k\to\infty}\langle x^k-z,-z\rangle\le 0.$$

Indeed, since $z=P_\Omega 0$, by using the property of the projection (Lemma 2.1 (1)), we arrive at

$$\limsup_{k\to\infty}\langle x^k-z,-z\rangle=\max_{\hat{z}\in\omega_w(\{x^k\})}\langle\hat{z}-z,-z\rangle\le 0.$$

By applying Lemma 2.3 to (3.12) with the data:

$$a_k := \|x^k - z\|^2, \quad \alpha_k := \beta_k, \quad b_k := 0,$$
$$\gamma_k := \beta_k\|z\|^2 + 2(1-\beta_k)\langle x^k - z, -z\rangle + 2\lambda_k\big\|\nabla f(x^k)\big\|\|z\|,$$

we immediately deduce that the sequence $\{x^k\}$ converges strongly to $z = P_\Omega 0$. Furthermore, it follows again from Lemma 2.1 (1) that

$$\langle p - z, -z\rangle \le 0 \ \ \forall p \in \Omega.$$

Hence

$$\|z\|^2 \le \langle p, z\rangle \le \|z\|\big\|p\big\| \ \ \forall p \in \Omega,$$

from which we infer that z is the minimum-norm solution of the SCFP (1.1).

Case 2. Assume that there exists a subsequence $\{\Gamma_{k_m}\} \subset \{\Gamma_k\}$ such that $\Gamma_{k_m} \le \Gamma_{k_m+1}$ for all $m \in \mathbb{N}$. In this case, we can define $\tau : \mathbb{N} \to \mathbb{N}$ by

$$\tau(k) = \max\{n \le k \mid \Gamma_n < \Gamma_{n+1}\}.$$

Then we have from Lemma 2.2 that $\tau(k) \to \infty$ as $k \to \infty$ and $\Gamma_{\tau(k)} < \Gamma_{\tau(k)+1}$. So, we have from (3.8) that

$$\begin{aligned}\rho_{\tau(k)}\left(4 - \frac{\rho_{\tau(k)}}{1-\beta_{\tau(k)}}\right)\frac{f^2(x^{\tau(k)})}{\|\nabla f(x^{\tau(k)})\|^2} &\le \|x^{\tau(k)} - z\|^2 - \|x^{\tau(k)+1} - z\|^2 + \beta_{\tau(k)}\|z\|^2 \\ &\le \beta_{\tau(k)}\|z\|^2. \end{aligned} \tag{3.12}$$

Following the same way as the proof of Case 1, we have that

$$\begin{aligned} &\lim_{k\to\infty}\frac{f^2(x^{\tau(k)})}{\|\nabla f(x^{\tau(k)})\|^2} = 0, \\ &\limsup_{k\to\infty}\langle x^{\tau(k)} - z, -z\rangle = \max_{\tilde{z}\in\omega_w(\{x^{\tau(k)}\})}\langle \tilde{z} - z, -z\rangle \le 0 \end{aligned} \tag{3.13}$$

and

$$\begin{aligned}\|x^{\tau(k)+1} - z\|^2 \le (1-\beta_{\tau(k)})\|x^{\tau(k)} - z\|^2 + \beta_{\tau(k)}\big[\beta_{\tau(k)}\|z\|^2 + 2(1-\beta_{\tau(k)})\langle x^{\tau(k)} - z, -z\rangle \\ + 2\lambda_{\tau(k)}\big\|\nabla f(x^{\tau(k)})\big\|\|z\|\big],\end{aligned} \tag{3.14}$$

where $\beta_{\tau(k)} \to 0$. Since $\Gamma_{\tau(k)} < \Gamma_{\tau(k)+1}$, we have from (3.14) that

$$\|x^{\tau(k)} - z\|^2 \le \beta_{\tau(k)}\|z\|^2 + 2(1-\beta_{\tau(k)})\langle x^{\tau(k)} - z, -z\rangle + 2\lambda_{\tau(k)}\big\|\nabla f(x^{\tau(k)})\big\|\|z\|. \tag{3.15}$$

Combining (3.13) with (3.15) yields

$$\limsup_{k\to\infty}\|x^{\tau(k)} - z\|^2 \le 0,$$

and hence

$$\lim_{k\to\infty}\|x^{\tau(k)} - z\|^2 = 0.$$

From (3.14), we have

$$\limsup_{k\to\infty}\|x^{\tau(k)+1} - z\|^2 \le \limsup_{k\to\infty}\|x^{\tau(k)} - z\|^2.$$

Thus

$$\lim_{k\to\infty}\|x^{\tau(k)+1} - z\|^2 = 0.$$

Therefore, by Lemma 2.2, we obtain

$$0 \le \|x^k - z\| \le \max\{\|x^{\tau(k)} - z\|, \|x^k - z\|\} \le \|x^{\tau(k)+1} - z\| \to 0.$$

Consequently, $\{x^k\}$ converges strongly to $z = P_\Omega 0$. The proof is complete. □

Next, we give the second proof.

Proof. Let $z := P_\Omega 0$. Since $\inf_k \rho_k\big(4-\frac{\rho_k}{1-\beta_k}\big) > 0$, we may assume without loss of generality that there exists $\varepsilon > 0$ such that

$$\rho_k\big(4-\frac{\rho_k}{1-\beta_k}\big) \geq \varepsilon.$$

It follows from (3.7) that, for all k,

$$\|x^{k+1}-z\|^2 \leq \beta_k\|z\|^2 + (1-\beta_k)\|x^k-z\|^2 - \frac{\varepsilon f^2(x^k)}{\|\nabla f(x^k)\|^2}. \tag{3.16}$$

Combining (3.12) and (3.16), we obtain the next two inequalities

$$\|x^{k+1}-z\|^2 \leq (1-\beta_k)\|x^k-z\|^2 + \beta_k\delta_k,$$
$$\|x^{k+1}-z\|^2 \leq \|x^k-z\|^2 - \eta_k + \beta_k\|z\|^2,$$

where

$$\eta_k = \frac{\varepsilon f^2(x^k)}{\|\nabla f(x^k)\|^2},$$

$$\delta_k := \beta_k\|z\|^2 + 2(1-\beta_k)\langle x^k - z, -z\rangle + 2\lambda_k\big\|\nabla f(x^k)\big\|\|z\|,$$

$\{\beta_k\} \subset (0,1)$, $\lim_{k\to\infty}\beta_k = 0$ and $\sum_{k=0}^\infty \beta_k = \infty$. In order to use Lemma 2.4 with the data $s_k := \|x^k - z\|^2$, it remains to show that for any subsequence $\{n_k\}$ of $\{n\}$,

$$\eta_{n_k} \to 0 \Longrightarrow \limsup_{k\to\infty} \delta_{n_k} \leq 0.$$

Let $\{n_k\}$ be a subsequence of $\{n\}$ such that $\eta_{n_k} \to 0$. Then, as in the first proof, we can get that every weak limit point of the sequence $\{x^{n_k}\}$ belongs to Ω. We now deduce that

$$\begin{aligned}
\limsup_{k\to\infty}\delta_{n_k} &= \limsup_{k\to\infty}\Big[\beta_{n_k}\|z\|^2 + 2(1-\beta_{n_k})\langle x^{n_k}-z,-z\rangle + 2\lambda_{n_k}\big\|\nabla f(x^{n_k})\big\|\|z\|\Big] \\
&= 2\limsup_{k\to\infty}\langle x^{n_k}-z,-z\rangle \\
&= 2\max_{\bar z\in\omega_w(\{x^{n_k}\})}\langle \bar z - z, -z\rangle \\
&\leq 0.
\end{aligned}$$

Finally, using Lemma 2.4, we have $\|x^k - z\| \to 0$. We thus complete the proof. □

Remark 3.1. (1) Our results focus on the split convex feasibility problem when only one set per each space is involved, but clearly a common product space reformulation for the multiple-sets split convex feasibility problem can be employed, which is similar to [7] (originally due to Pierra [15]) to derive a simultaneous version of Algorithm 3.1.

(2) One main advantage of our algorithm compared to others is that stepsizes are directly computed in each iteration and do not depend on the norm of A. Therefore, Theorem 3.1 improves Chuang [8, Theorem 5.5], Qu and Liu [16, Theorem 3.3], Wang and Xu [18, Theorem 4.3], Xu [22, Theorem 5.5] and Yao, Jigang and Liou [24, Theorem 3.1].

4. Numerical illustration

In this section, we present an illustration and comparison of our new scheme in the infinite dimensional Hilbert space $H_1 = H_2 = L_2[0,1]$ with norm $\|x\| := \sqrt{\int_0^1 |x(t)|^2 dt}$ and inner product given by $\langle x,y\rangle = \int_0^1 x(t)y(t)dt$.

We consider the split convex feasibility problem with the following two nonempty, closed, and convex sets

$$C = \big\{x \in L_2[0,1] \mid \langle x(t), 3t^2\rangle = 0\big\},\;\; Q = \{x \in L_2[0,1] \mid \langle x, t/3\rangle \geq -1\}$$

and the linear bounder operator is given as $(Ax)(t) = x(t)$ (clearly $\|A\| = 1$). Since $A = I$ is the identity, we are actually concern with the so-called convex feasibility problem in $L_2[0,1]$. Since C and Q are hyper-plane and half-space, respectively, the orthogonal projection onto them have an explicit formula, see, for example, [10]

$$P_C(z(t)) = \begin{cases} z(t) - \frac{\langle z(t), 3t^2\rangle}{\|3t^2\|^2_{L^2}} 3t^2, & \text{if} \langle z(t), 3t^2\rangle \neq 0, \\ z(t), & \text{if} \langle z(t), 3t^2\rangle = 0, \end{cases}$$

$$P_Q(z(t)) = \begin{cases} z(t) - \frac{\langle z(t), -t/3\rangle - 1}{\|-t/3\|^2_{L^2}} (-t/3), & \text{if} \langle z(t), t/3\rangle < -1, \\ z(t), & \text{if} \langle z(t), t/3\rangle \geq -1. \end{cases}$$

In Algorithm 3.1, we choose $\rho_k = 2$, $\beta_k = \frac{1}{k}$ and the stopping rule is

$$E_k = \|x^k(t) - P_C x^k(t) + x^k(t) - P_Q x^k(t)\|_{L_2} < 10^{-6}.$$

The results (Sec.=seconds, Iter.=iterations) with different starting points are reported in Table 1 and the error function E_n with respect to each starting point is shown in Figure 1.

	$x^0(t) = e^t$		$x^0(t) = \cos t$		$x^0 = t^4 + t$		$x^0 = t/5$	
	Sec.	Iter.	Sec.	Iter.	Sec.	Iter.	Sec.	Iter.
Algorithm 3.1	0.0156	9	0.0156	12	0.0312	12	0.0156	11

TABLE 1. Comparison of Algorithm 3.1 with different starting points

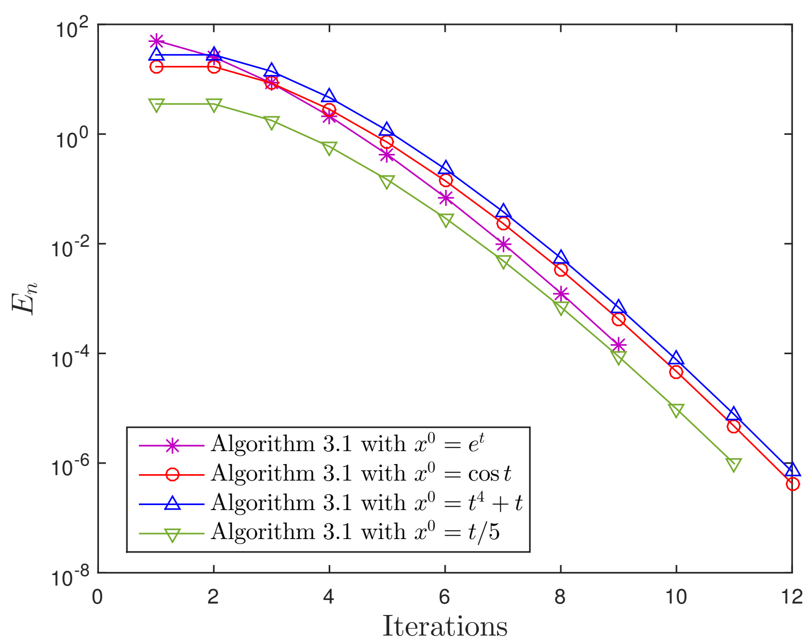

FIGURE 1. Error plotting of E_n for different starting points

Acknowledgments

The fifth author was funded by University of Transport and Communications (UTC) under Grant No. T2019-CB-014.

REFERENCES

[1] Q. H. Ansari, A. Rehan, Split feasibility and fixed point problems, In: Q. H. Ansari (ed.), Nonlinear Analysis, Approximation Theory, Optimization and Applications, pp. 281-322, 2014.

[2] C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, Inverse Probl. 18 (2002), 441-453.

[3] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, Inverse Probl. 20 (2004), 103-120.

[4] L. C. Ceng, Q. H. Ansari, J. C. Yao, An extragradient method for solving split feasibility and fixed point problems, Comput. Math. Appl. 64 (2012), 633-642.

[5] Y. Censor, T. Bortfeld, B. Martin, A. Trofimov, A unified approach for inversion problems in intensity modulated radiation therapy, Phy. Medicine Biol. 51 (2003), 2353-2365.

[6] Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projection in a product space, Numer. Algo. 8 (1994), 221-239.

[7] Y. Censor, A. Segal, On the string averaging method for sparse common fixed point problems, Int. Trans. Oper. Res. 16 (2009), 481-494.

[8] S. S. Chuang, Strong convergence theorems for the split variational inclusion problem in Hilbert spaces, Fixed Point Theory Appl. 2013 (2013), Article ID 350.

[9] Y. Censor, A. Gibali, S. Reich, Algorithms for the split variational inequality problem, Numer. Algo. 59 (2012), 301-323.

[10] A. Cegielski, Iterative Methods for Fixed Point Problems in Hilbert Spaces, Lecture Notes in Mathematics 2057, Springer, Berlin, 2012.

[11] Y. Dang, Y. Gao, The strong convergence of a KM-CQ-like algorithm for a split feasibility problem, Inverse Probl. 27 (2011), Article ID 015007.

[12] S. He, C. Yang, Solving the variational inequality problem defined on intersection of finite level sets, Abstr. Appl. Anal. 2013 (2013), Article ID 942315.

[13] G. López, V. Martín-Márquez, F. Wang, H.K. Xu, Solving the split feasibility problem without prior knowledge of matrix norms, Inverse Probl. 28 (2012), Article ID 085004.

[14] P.E. Maingé: Strong convergence of projected subgradientmethods for nonsmooth and nonstrictly convex minimization, Set-Valued Anal. 16 (2008), 899-912.

[15] G. Pierra, Decomposition through formalization in a product space, Math. Program. 28 (1984), 96-115.

[16] B. Qu, B. Liu, A gradient algorithm for finding minimum-norm solution of the split feasibility problem, Pac. J. Optim. 13 (2017), 355-364.

[17] M. Tian, H.-F. Zhang: Regularized gradient-projection methods for finding the minimum-norm solution of equilibrium and the constrained convex minimization problem, J. Nonlinear Sci. Appl. 9 (2016), 5316-5331.

[18] F. Wang, H.-K. Xu: Approximating curve and strong convergence of the CQ algorithm for the split feasibility problem, J. Inequal. Appl. 2010 (2010), Article ID 102085.

[19] F. Wang, H.-K. Xu, Cyclic algorithms for split feasibility problems in Hilbert spaces, Nonlinear Anal. 74 (2011), 4105-4111.

[20] Y. Xiao, Y. Censor, D. Michalski, J.M. Galvin, The least-intensity feasible solution for aperture-based inverse planning in radiation therapy, Ann. Oper. Res. 119 (2003), 183-203.

[21] H.-K. Xu, Iterative algorithms for nonlinear operators, J. London Math. Soc. 66 (2002), 240-256.

[22] H.-K. Xu, Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces, Inverse Probl. 26 (2010), Article ID 105018.

[23] Q. Yang, The relaxed CQ algorithm for solving the split feasibility problem, Inverse Prob. 20 (2004), 1261–1266.

[24] Y. Yao, W. Jigang, Y.C. Liou: Regularized methods for the split feasibility problem, Abstr. Appl. Anal. 2012 (2012), Article ID 140679.