CARPATHIAN J. MATH. Online version at http://carpathian.ubm.ro
33 (2017), No. 3, 389-397 Print Edition: ISSN 1584 - 2851 Online Edition: ISSN 1843 - 4401

Projection algorithms for composite minimization

XIAONAN YANG and HONG-KUN XU

ABSTRACT. Parallel and cyclic projection algorithms are proposed for minimizing the sum of a finite family of convex functions over the intersection of a finite family of closed convex subsets of a Hilbert space. These algorithms consist of two steps. Once the kth iterate is constructed, an inner circle of gradient descent process is executed through each local function, and then a parallel or cyclic projection process is applied to produce the $(k+1)$ iterate. These algorithms are proved to converge to an optimal solution of the composite minimization problem under investigation upon assuming boundedness of the gradients at the iterates of the local functions and the stepsizes being chosen appropriately.

1. INTRODUCTION

We are concerned with a composite minimization problem, that is, we consider the case where the objective function is decomposed into the sum of a finite family of convex functions and the set of constraints is the intersection of finitely many closed convex subsets of a Hilbert space H. Precisely, the minimization problem under investigation in this paper can be written as

$$\min_{x\in C:=\bigcap_{i=1}^{M} C_i} f(x) := \sum_{j=1}^{N} f_j(x), \tag{1.1}$$

where M, N are positive integers, each set C_i is a nonempty closed convex subset of a Hilbert space H, and each local function $f_i : H \to \mathbb{R}$ is a Fréchet differentiable and convex function. We always assume the feasible set $C \neq \emptyset$.

Notice that optimization problems of form (1.1) arise in many applied areas, in particular, in machine learning and statistics (see [2, 7, 13] for examples and more details).

The convex feasibility problem (CFP) [1, 3] is formulated as

$$\text{finding a point } x^* \text{ with the property: } x^* \in \bigcap_{i=1}^{N} C_i. \tag{1.2}$$

Thus, the composite minimization problem (1.1) can alternatively be rephrased as finding a solution to the convex feasibility problem (1.2) which also minimizes the composite function f as defined in (1.1). Consequently, two points should be taken into consideration of algorithmic approaches to (1.1): (a) the descent property of the values of the objective function, and (b) the (approximate) feasibility of the iterates generated by the algorithm. To illustrate these points we consider the special case where $M = N = 1$. In this case, (1.1) is reduced to the constrained convex minimization:

$$\min_{x\in C_1} f_1(x). \tag{1.3}$$

Received: 28.11.2016. In revised form: 02.03.2017. Accepted: 22.03.2017
2010 *Mathematics Subject Classification.* 90C25, 90C52, 65K10, 47J25.
Key words and phrases. *convex feasibility, composite minimization, projection algorithm.*
Corresponding author: Hong-Kun Xu; xuhk@hdu.edu.cn

The gradient-projection algorithm (GPA) can solve (1.3): GPA generates a sequence $\{x_k\}$ by the recursion process:

$$x_{k+1} = P_{C_1}(x_k - \lambda_k \nabla f_1(x_k)), \tag{1.4}$$

where the initial guess $x_0 \in H$ is chosen arbitrarily, and $\lambda_k > 0$ is the stepsize. Assume:

(A1) The gradient of f_1, ∇f_1, is α-Lipschitz (for some $\alpha \geq 0$):

$$\|\nabla f_1(x) - \nabla f_1(z)\| \leq \alpha \|x - z\|, \quad x, z \in H;$$

(A2) The sequence of stepsizes, $\{\lambda_k\}$, satisfies the condition:

$$0 < \liminf_{k\to\infty} \lambda_k \leq \limsup_{k\to\infty} \lambda_k < \frac{2}{\alpha}.$$

It is then easy to find that both points (a) and (b) hold (actually, (b) is trivial); moreover, the sequence $\{x_k\}$ generated by GPA (1.4) converges [12, 15] weakly to a solution of (1.3) (if any).

Observe that the splitting of the objective function f into the sum of N (simpler) local functions, and the set C of constraints into the intersection of M (simpler) convex subsets aims at providing more efficient algorithmic approaches to (1.1) by utilizing the simpler structures of the local functions $\{f_j\}$ (for instance, the proximal mappings of f_j are computable [4]) and of the sets $\{C_i\}$ (for instance, the projections P_{C_i} possess closed formulae). This means that when we study algorithms for the composite optimization problem (1.1), we should use individual local functions and individual subsets at each iteration, not the full sum of the local functions $\{f_j\}$, nor the full intersection of the sets $\{C_i\}$.

The purpose of this paper is exactly to provide two such algorithms, which we call parallel and cyclic projection algorithms (see (3.5) and (3.18) in Section 3) for the reason that parallel and cyclic projections play a key role in defining these algorithms.

2. Preliminaries

The fundamental tool of our argument in this paper is the concept of projections. Let H be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, respectively, and let C be a nonempty closed convex subset of H. The (nearest point) projection from H onto C, dented by P_C, is defined by

$$P_C x := \arg\min_{y\in C} \|x - y\|, \quad x \in H.$$

The following properties are pertinent to our argument in Section 3.

Proposition 2.1. *Projections are of the following properties:*

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

(ii) $\langle P_C x - P_C y, x - y\rangle \geq \|P_C x - P_C y\|^2$ *for all* $x, y \in H$*; in particular,* P_C *is nonexpansive, namely,*

$$\|P_C x - P_C y\| \leq \|x - y\|, \quad x, y \in H.$$

(iii) $\|P_C x - y\|^2 \leq \|x - y\|^2 - \|P_C x - x\|^2$ *for all* $x \in H$ *and* $y \in C$.

The following two lemmas are also useful for proving the convergence of our algorithms in this paper.

Lemma 2.1. [12] *Assume* $\{a_k\}$ *is a sequence of nonnegative real numbers with the property:*

$$a_{k+1} \leq a_k + b_k, \quad k \geq 0,$$

where $\{b_k\}$ *is a sequence of nonnegative real numbers such that* $\sum_{k=0}^{\infty} b_k < \infty$*. Then* $\{a_k\}$ *is bounded and* $\lim_{k\to\infty} a_k$ *exists.*

Lemma 2.2. [8, Lemma 2.5] *Let K be a nonempty subset of a Hilbert space H. Assume $\{x_k\}$ is a bounded sequence in H with the properties:*

(a) $\lim_{k\to\infty} \|x_k - z\|$ *exists for each $z \in K$;*
(b) if x' is a weak cluster point of $\{x_k\}$, then $x' \in K$.

Then the full sequence $\{x_k\}$ converges weakly to a point in K.

We need the demiclosedness principle of nonexpansive mappings as follows.

Lemma 2.3. [11, 6] *Let K be a closed convex subset of a Hilbert space H and $T : C \to C$ a nonexpansive mapping (i.e., $\|Tx - Ty\| \le \|x - y\|$ for all $x, y \in K$). Suppose $\{v_k\}$ is a sequence in K such that $v_k \to v$ weakly and $v_k - Tv_k \to 0$ in norm, then $v = Tv$.*

3. Projection Algorithms and Their Convergence Analysis

3.1. **Projection Algorithms for the Convex Feasibility Problem.** We mention three projection algorithms [1, 3, 10, 14] for solving CFP (1.2) which will be used in defining our projection algorithms.

Theorem 3.1. *Beginning with an arbitrarily chosen initial guess $x_0 \in H$, we iterate $\{x_k\}$ in either one of the following three projection algorithms:*

(i) *Sequential projections: $x_{k+1} = P_{C_M} \cdots P_{C_1} x_k$;*
(ii) *Parallel projections: $x_{k+1} = \sum_{j=1}^{M} \beta_j P_{C_j} x_k$, with $\beta_j > 0$ for all j and $\sum_{j=1}^{M} \beta_j = 1$;*
(iii) *Cyclic projections: $x_{k+1} = P_{C_{[k+1]}} x_k$. [Here $[k+1] = i \,(\mathrm{mod}\, M)$, $0 \le i < M$.]*

Then $\{x_k\}$ converges weakly to a solution of CFP (1.2).

3.2. **The Parallel Projection Algorithm.** We introduce the following parallel projection algorithm (PPA) for solving the composite minimization problem (1.1). Take an initial guess $x_0 \in H$ arbitrarily and then iterate x_{k+1} ($k \ge 0$) by the iteration process:

$$\begin{cases} x_{k,0} = x_k, & \text{(3.5a)} \\ x_{k,j} = x_{k,j-1} - \lambda_k \nabla f_j(x_{k,j-1}), \quad j = 1, 2, \cdots, N, & \text{(3.5b)} \\ x_{k+1} = \displaystyle\sum_{i=1}^{M} \beta_i P_{C_i} x_{k,N}, & \text{(3.5c)} \end{cases}$$

where, for each $1 \le i \le M$, P_{C_i} is the projection from H onto C_i, and $\beta_i > 0$ is such that $\sum_{i=1}^{M} \beta_i = 1$.

3.3. **Convergence of PPA.**

Lemma 3.4. *Let $\{x_k\}_{k=0}^{\infty}$ be generated by Algorithm (3.5). Suppose*

$$\|\nabla f_j(x_{k,j-1})\| \le L_j, \quad j = 1, 2, \cdots, N, \quad k \ge 0, \tag{3.6}$$

where, for each $1 \le j \le N$, $L_j \ge 0$ is a constant. Set $L = \sum_{i=1}^{N} L_i$. Then, for each $x \in C$, we have

$$\|x_{k+1} - x\|^2 \le \|x_k - x\|^2 - 2\lambda_k[f(x_k) - f(x)] + \lambda_k^2 L^2. \tag{3.7}$$

Proof. The proof given below is some minor modifications of the proof of [9, Lemma 2.1]. However we include it here for the sake of completeness.

Using the convexity of the norm and nonexpansivity of projections, we immediately get, for $x \in C$,

$$\begin{aligned} \|x_{k+1} - x\|^2 &= \left\| \sum_{i=1}^{M} \beta_i (P_{C_i} x_{k,N} - P_{C_i} x) \right\|^2 \\ &\leq \sum_{i=1}^{M} \beta_i \|P_{C_i} x_{k,N} - P_{C_i} x\|^2 \\ &\leq \sum_{i=1}^{M} \beta_i \|x_{k,N} - x\|^2 \\ &= \|x_{k,N} - x\|^2. \end{aligned} \tag{3.8}$$

On the other hand, for each $1 \leq j \leq N$, we have

$$\begin{aligned} \|x_{k,j} - x\|^2 &= \|(x_{k,j-1} - x) - \lambda_k \nabla f_j(x_{k,j-1})\|^2 \\ &= \|x_{k,j-1} - x\|^2 - 2\lambda_k \langle \nabla f_j(x_{k,j-1}), x_{k,j-1} - x\rangle + \lambda_k^2 \|\nabla f_j(x_{k,j-1})\|^2. \end{aligned}$$

Using (3.6) and the inequality

$$f_j(x) \geq f_j(x_{k,j-1}) + \langle \nabla f_j(x_{k,j-1}), x - x_{k,j-1} \rangle$$

we obtain

$$\|x_{k,j} - x\|^2 \leq \|x_{k,j-1} - x\|^2 - 2\lambda_k [f_j(x_{k,j-1}) - f_j(x)] + \lambda_k^2 L_j^2.$$

Adding up the above inequalities over $j = 1, 2, \cdots, N$ yields

$$\begin{aligned} \|x_{k,N} - x\|^2 &\leq \|x_k - x\|^2 - 2\lambda_k \sum_{j=1}^{N} [f_j(x_{k,j-1}) - f_j(x)] + \lambda_k^2 \sum_{j=1}^{N} L_j^2 \\ &= \|x_k - x\|^2 - 2\lambda_k [f(x_k) - f(x)] \\ &\quad + \lambda_k^2 \sum_{j=1}^{N} L_j^2 - 2\lambda_k \sum_{j=1}^{N} [f_j(x_{k,j-1}) - f_j(x_k)]. \end{aligned} \tag{3.9}$$

Observing

$$\begin{aligned} f_j(x_{k,j-1}) - f_j(x_k) &\geqslant \langle \nabla f_j(x_k), x_{k,j-1} - x_k \rangle \\ &\geqslant -L_j \|x_{k,j-1} - x_k\| \end{aligned}$$

and

$$\begin{aligned} \|x_{k,j-1} - x_k\| &= \|x_{k,j-1} - x_{k,0}\| \\ &= \left\| \sum_{l=1}^{j-1} (x_{k,l} - x_{k,l-1}) \right\| \\ &= \left\| \sum_{l=1}^{j-1} \lambda_k \nabla f_l(x_{k,l-1}) \right\| \\ &\leq \lambda_k \sum_{l=1}^{j-1} L_l, \end{aligned}$$

we derive from (3.9) that

$$\begin{aligned}\|x_{k,N}-x\|^2 &\le \|x_k-x\|^2-2\lambda_k[f(x_k)-f(x)]+\lambda_k^2\sum_{j=1}^N L_j^2+2\lambda_k^2\sum_{j=1}^N L_j\left(\sum_{l=1}^{j-1}L_l\right)\\ &=\|x_k-x\|^2-2\lambda_k[f(x_k)-f(x)]+\lambda_k^2\left(\sum_{j=1}^N L_j\right)^2\\ &=\|x_k-x\|^2-2\lambda_k[f(x_k)-f(x)]+\lambda_k^2L^2.\end{aligned}$$

This together with (3.8) proves (3.7). □

Assume that the sequence of stepsizes $\{\lambda_k\}$ satisfies the condition

$$0<\lambda_k,\quad \lim_{k\to\infty}\lambda_k=0,\quad \sum_{k=0}^\infty\lambda_k=\infty. \tag{3.10}$$

Let

$$S^*:=\left\{x^*\in C: f(x^*)=\inf_{x\in C}f(x)\right\}\quad\text{and}\quad f^*:=\inf_{x\in C}f(x)$$

be the set of optimal solutions and the optimal value of the composite minimization problem (1.1), respectively. We shall always assume from now and onwards that $S^*\neq\emptyset$.

Lemma 3.5. *Let $\{x_k\}$ be generated by PPA (3.5) and assume (3.6).*

(i) *If $\{x_k\}$ is bounded and $\{\lambda_k\}$ satisfies (3.10), then* $\liminf_{k\to\infty}f(x_k)=\inf_{x\in C}f(x)$.
(ii) *If $\lambda_k>0$ (for all k) and $\sum_{k=1}^\infty\lambda_k^2<\infty$, then* $\lim_{k\to\infty}\|x_k-x\|$ *exists for $x\in S^*$; hence, $\{x_k\}$ is bounded.*

Proof. (i) If $\liminf_{k\to\infty}f(x_k)>f^*$, then there exist some $\varepsilon_0>0$ and $k_0\ge 0$ such that $f(x_k)>f^*+\varepsilon_0$ for all $k\ge k_0$. Since $\lambda_k\to 0$, we may also assume $\lambda_kL^2<\varepsilon_0$ for all $k\ge k_0$. It then turns out from (3.7) that, for $x\in S^*$ and $k\ge k_0$,

$$\varepsilon_0\lambda_k\le\|x_k-x\|^2-\|x_{k+1}-x\|^2.$$

This implies that the series $\sum_{k=k_0}^\infty\lambda_k$ is convergent, which contradicts (3.10). Therefore, we must have $\liminf_{k\to\infty}f(x_k)=f^*$.

We next turn to (ii). Again using (3.7) we obtain, for $x\in S^*$ and $k\ge 0$,

$$\|x_{k+1}-x\|^2\le\|x_k-x\|^2+\lambda_k^2L^2.$$

The conclusion of (ii) now follows immediately from Lemma 2.1. □

Consider the distance function from $x\in H$ to the solution set S^*:

$$d_{S^*}(x):=\inf\{\|x-z\|: z\in S^*\},\quad x\in H.$$

Corollary 3.1. *Let $\{x_k\}$ be generated by PPA (3.5) and assume (3.6). Assume also $\lambda_k>0$ (for all k) and $\sum_{k=0}^\infty\lambda_k^2<\infty$. Then $\lim_{k\to\infty}d_{S^*}(x_k)$ exists. Moreover, if dim $H<\infty$, then $\{x_k\}$ converges to an optimal solution of (1.1) if and only if $\lim_{k\to\infty}d_{S^*}(x_k)=0$.*

Proof. By Lemma 3.4, we get for all $x\in S^*$ and $k\ge 0$

$$\|x_{k+1}-x\|^2\le\|x_k-x\|^2-2\lambda_k(f(x_k)-f^*)+\lambda_k^2L^2.$$

It turns out that

$$d_{S^*}^2(x_{k+1})\le d_{S^*}^2(x_k)-2\lambda_k(f(x_k)-f^*)+\lambda_k^2L^2. \tag{3.11}$$

In particular,

$$d_{S^*}^2(x_{k+1}) \le d_{S^*}^2(x_k) + \lambda_k^2 L^2. \tag{3.12}$$

Since $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$, it follows from Lemma 2.1 that $\lim_{k\to\infty} d_{S^*}^2(x_k)$ exists.

If $\{x_k\}$ converges in norm to an optimal solution x^* of (1.1) (i.e., $x^* \in S^*$), then by continuity of d_{S^*}, we immediately get $\lim_{k\to\infty} d_{S^*}(x_k) = d_{S^*}(x^*) = 0$. Conversely, assume $\dim H < \infty$ and $\lim_{k\to\infty} d_{S^*}(x_k) = 0$. For each $j \ge 1$, we can find $z_j \in S^*$ and some n_j such that

$$\|x_{k_j} - z_j\| < \frac{1}{j}.$$

Since H is finite-dimensional and $\{x_k\}$ is bounded, we may assume $x_{k_j} \to \hat{x}$. It turns out that $z_j \to \hat{x}$ as well. Consequently, $\hat{x} \in S^*$ as S^* is closed. Now we have

$$\lim_{k\to\infty} d_{S^*}(x_k) = \lim_{j\to\infty} d_{S^*}(x_{k_j}) = d_{S^*}(\hat{x}) = 0.$$

□

Lemma 3.6. *Let the sequence $\{x_k\}$ be generated by PPA (3.5) and assume (3.6). Assume in addition that the sequence $\{\lambda_k\}$ of stepsizes satisfies the condition:*

$$\lambda_k > 0 \quad (\text{for all } k \ge 0), \quad \sum_{k=1}^{\infty} \lambda_k = \infty, \quad \sum_{k=1}^{\infty} \lambda_k^2 < \infty. \tag{3.13}$$

(A standard prototype is given by $\lambda_k := \frac{1}{k+1}$ for all $k \ge 0$.) Then every weak cluster point of $\{x_k\}$ is feasible. Moreover, if $\{x_{k_j}\}$ is a subsequence of $\{x_k\}$ weakly convergent to x^* such that $\lim_{j\to\infty} f(x_{k_j}) = f^*$, then $x^* \in S^*$.

Proof. The weak lower-semicontinuity of f implies that $f(x^*) \le \liminf_{j\to\infty} f(x_{k_j}) = f^*$. It turns out that $x^* \in S^*$ provided $x^* \in C$, that is, x^* is feasible. To see this, we proceed as follows by using the convexity of the function $\|\cdot\|^2$ and Proposition 2.1(iii):

$$\begin{aligned}
\|x_{k+1} - x^*\|^2 &= \left\| \sum_{i=1}^{N} \beta_i (P_{C_i} x_{k,N} - x^*) \right\|^2 \\
&\le \sum_{i=1}^{N} \beta_i \|P_{C_i} x_{k,N} - x^*\|^2 \\
&\le \sum_{i=1}^{N} \beta_i (\|x_{k,N} - x^*\|^2 - \|P_{C_i} x_{k,N} - x_{k,N}\|^2) \\
&= \|x_{k,N} - x^*\|^2 - \sum_{i=1}^{N} \beta_i \|P_{C_i} x_{k,N} - x_{k,N}\|^2.
\end{aligned}$$

It follows that

$$\sum_{i=1}^{N} \beta_i \|P_{C_i} x_{k,N} - x_{k,N}\|^2 \le \|x_{k,N} - x^*\|^2 - \|x_{k+1} - x^*\|^2. \tag{3.14}$$

Note that

$$\begin{aligned}\|x_{k,N}-x_k\| &= \|x_{k,N}-x_{k,0}\| \\ &= \left\|\sum_{i=1}^{N}(x_{k,j}-x_{k,j-1}\right\| \\ &= \left\|\sum_{i=1}^{N}\lambda_k\nabla f_j(x_{k,j-1})\right\| \\ &\le \lambda_k\sum_{i=1}^{N}L_j=\lambda_k L\to 0 \quad (\text{as } k\to\infty).\end{aligned}$$

Now since $\lim_{k\to\infty}\|x_k-x^*\|$ exists, we get $\lim_{k\to\infty}\|x_{k,N}-x_k\|=0$. It turns out from (3.14) that

$$\lim_{k\to\infty}\sum_{i=1}^{N}\beta_i\|P_{C_i}x_{k,N}-x_{k,N}\|^2=0.$$

Equivalently, for each $1\le i\le N$,

$$\lim_{k\to\infty}\|P_{C_i}x_{k,N}-x_{k,N}\|^2=0.$$

Consequently, for each $1\le i\le N$,

$$\lim_{k\to\infty}\|P_{C_i}x_k-x_k\|^2=0.$$

This implies that if z is a weak cluster point of $\{x_k\}$, then $z=P_{C_i}z$ by Lemma 2.3. Hence, $z\in C_i$ for every $1\le i\le N$. Namely, $z\in C$ is feasible. In particular, x^* is feasible. □

We are now in the position to state and prove the main result is this paper.

Theorem 3.2. *Let $\{x_k\}$ be the sequence generated by the parallel projection algorithm (3.5). Assume (3.6) and (3.13). Then we have:*

- (i) *If H is finite-dimensional, then $\{x_k\}$ converges to an optimal solution x^* of the composite minimization problem (1.1) and $\{f(x_k)\}$ converges to the optimal value f^*.*
- (ii) *In a general Hilbert space H, there exists a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ such that*
  - (iia) *$\{x_{k_j}\}$ converges weakly to an optimal solution $x^*\in S^*$, and $\{f(x_{k_j})\}$ converges to the optimal value f^*.*
  - (iib) *If, in addition, the limit of the full sequence $\{f(x_k)\}$ exists as $k\to\infty$, then the full sequence $\{x_k\}$ converges weakly to the optimal solution x^*, and $\{f(x_k)\}$ converges to the optimal value f^*.*

Proof. We first observe the inequality

$$\|x_{k+1}-x\|\le\|x_k-x\|+\lambda_k^2L^2 \tag{3.15}$$

for all $x\in S^*$. This implies that $\{x_k\}$ is bounded and

$$\lim_{k\to\infty}\|x_k-x\| \quad \text{exists for every } x\in S^*. \tag{3.16}$$

Now to see (i), we apply Lemma 3.5 to get a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ such that

$$\lim_{j\to\infty}f(x_{k_j})=\liminf_{k\to\infty}f(x_k)=f^*. \tag{3.17}$$

Since $\dim H<\infty$, we may also assume that $x_{k_j}\to x^*$ (in norm) as $j\to\infty$. Notice that $x^*\in S^*$. Now apply (3.16) with x replaced with x^* to obtain

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

Namely, $x_k \to x^*$ (in norm); hence, $f(x_k) \to f^*$. This proves (i).

To see (iia), we again apply Lemma 3.5 to get a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ such that (3.17) holds. However, in this case we can only assume that $\{x_{k_j}\}$ converges weakly to a point x^*. By Lemma 3.6, we have $x^* \in S^*$.

Finally, to prove (iib) we notice that since the full sequence $\{f(x_k)\}$ converges, we must have

$$\lim_{k\to\infty} f(x_k) = f^*.$$

By Lemma 3.6 again, we have $x' \in S^*$ whenever x' is a weak cluster point of the subsequence $\{x_k\}$. Applying Lemma 2.2, we conclude that the full sequence $\{x_k\}$ converges weakly to a point in S^*. □

3.4. **The Cyclic Projection Algorithm.** Similarly to the parallel projection algorithm (3.5), we can introduce the cyclic projection algorithm (CPA) as follows:

$$\begin{cases} x_{k,0} = x_k, & (3.18a) \\ x_{k,j} = x_{k,j-1} - \lambda_k \nabla f_j(x_{k,j-1}), \quad j = 1, 2, \cdots, N, & (3.18b) \\ x_{k+1} = P_{C_{[k+1]}} x_{k,N}, & (3.18c) \end{cases}$$

where the initial guess $x_0 \in H$ is chosen arbitrarily, and $[k]$ is the mod M function defined in Theorem 3.1(iii).

Our proof of Theorem 3.2 shows that the part played by the parallel projection operator $\sum_{i=1}^{M} \beta_i P_{C_i}$ in the proof of Theorem 3.2 is to guarantee that the weak cluster points of the iterates $\{x_k\}$ are feasible (i.e., solutions of CFP (1.2)). This can be done by other projection operators such as the sequential projection operator [5]. Below we consider the cyclic projection operator as shown in Theorem 3.1. Therefore, we have the following result, the proof of which is similar to that of Theorem 3.2 and is thus omitted.

Theorem 3.3. *Let $\{x_k\}$ be generated by the CPA (3.18). Assume (3.6) and (3.13). Then the conclusions of Theorem 3.2 hold.*

Remark 3.1. In the algorithm (3.5) or (3.18), in order to construct the $(k+1)$th iterate x_{k+1}, an inner circle of iteration process is carried out through each local function, namely, the gradient-descent step (3.5b) is performed for each local function f_j. Consequently, the value of f_j may decrease from x_k to x_{k+1} should the stepsize $\lambda_k > 0$ is chosen appropriately. However, it is unclear whether or not the value of the sum function f would decrease from x_k to x_{k+1}.

It is also interesting to know whether or not the conclusion of Theorem 3.2(iib) remains true should the assumption that the limit of the full sequence $\{f(x_k)\}$ exists is removed.

Acknowledgement. The authors were grateful to the reviewers for their careful reading and helpful comments which improved the presentation of this manuscript.

References

[1] Bauschke, H. H. and Borwein, J. M., *On projection algorithms for solving convex feasibility problems*, SIAM Rev., **38** (1996), 1367–1426

[2] Bertsekas, D. P., Incremenal gradients, subgradient, and proximal methods for convex optimization: a survey, in *Optimization for Machine Learning* (S. Sra, S. Nowozin, and S. J. Wright, Eds.), MIT Press, 2011, pp. 85–119

[3] Combettes, P. L., *Hilbertian convex feasibility problem: convergence of projection methods*, Appl. Math. Optim., **35** (1997), 311–330

[4] Combettes, P. L. and Wajs, V. R., *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., **4** (2005), No. 4, 1168–1200

[5] De Pierro, A. R. and Neto, E. S. H., *From convex feasibility to convex constrained optimization using block action projection methods and underrelaxation*, Intl. Trans. Oper. Res., **16** (2009), 495–504

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

[7] Hastie, T., Tibshirani, R. and Friedman, J., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd Ed.), Springer, New York, 2009

[8] Lopez Acedo G. and Xu, H. K., *Iterative methods for strict pseudo-contractions in Hilbert spaces*, Nonlinear Anal., **67** (2007), 2258–2271

[9] Nedic, A. and Bertsekas, D. P., *Incremental subgradient methods for nondifferentiable optimization*, SIAM J. Option, **12** (2001), No. 1, 109–138

[10] O'Hara, J. G., Pillay, P. and Xu, H. K., *Iterative approaches to convex feasibility problems in Banach spaces*, Nonlinear Anal., **64** (2006), 2022–2042

[11] Opial, Z., *Weak convergence of the sequence of successive approximations of nonexpansive mappings*, Bull. Amer. Math. Soc., **73** (1967), 595–597

[12] Polyak, B. T., *Introduction to Optimization*, Optimization Software, New York, 1987

[13] Sra, S., Nowozin, S. and Wright, S. J. (Eds), *Optimization for Machine Learning*, MIT Press, Cambridge, Massachusetts, 2011, 512 pp.

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

[15] Xu, H. K., *Averaged mappings and the gradient-projection algorithm*, J. Optim. Theory Appl., **150** (2011), 360–378

DEPARTMENT OF MATHEMATICS
HANGZHOU DIANZI UNIVERSITY
SCHOOL OF SCIENCE
HANGZHOU 310018, CHINA
E-mail address: 1903059411@qq.com, xuhk@hdu.edu.cn