INCREMENTAL PROXIMAL METHOD FOR NONSMOOTH CONVEX OPTIMIZATION WITH FIXED POINT CONSTRAINTS OF QUASI-NONEXPANSIVE MAPPINGS

HARUHI OISHI, YU KOBAYASHI, AND HIDEAKI IIDUKA

ABSTRACT. This paper considers the convex optimization problem of minimizing the sum of convex functions over the intersection of fixed point sets of quasi-nonexpansive mappings in a real Hilbert space, and presents an incremental proximal method for solving it. A convergence analysis of the method for a sufficiently small constant step size indicates that it approximates a solution of the problem. For a diminishing step size, the method is guaranteed to solve the problem under certain assumptions. Moreover, a convergence-rate analysis for a diminishing step size is also provided to assess its efficiency. Finally, we numerically compare the performance of the proposed method with those of the existing incremental and parallel methods for a concrete convex optimization problem. These numerical comparisons show that the proposed method with a diminishing step size optimizes the problem faster than the existing ones.

1. INTRODUCTION

In this paper, we deal with the problem of optimizing the sum of convex functions over the intersection of fixed points of *quasi-nonexpansive mappings* [8, Problem 2.1] (see [4, 10, 11] for applications of the problem). Algorithms for solving this problem have been proposed in [8, 18]. Reference [8] proposed parallel and incremental subgradient methods for solving the problem and provided convergence as well as convergence-rate analyses, while reference [18] presented a parallel proximal method together with convergence and convergence-rate analyses. Subgradient methods [5, 6, 7, 12] and proximal methods [9, 17] were presented for optimizing the sum of convex functions over the intersection of fixed points of *nonexpansive mappings*.

The current paper presents an incremental proximal method for solving the problem. The method is obtained by combining the incremental method in [8] with the proximal one in [9]. We also present a convergence analysis for a constant step size and a diminishing step size. The analysis shows that the proposed method with a small constant step size may approximate a solution to the problem (Theorem 3.1) and that with a diminishing step size it converges to a solution under certain assumptions (Theorem 3.2). We also provide a convergence-rate analysis with a diminishing step size (Theorem 3.3). Finally, we numerically compare the performance of the proposed method with those of the existing methods [8, 18] and show its efficiency.

This paper is organized as follows. Section 2 gives the mathematical preliminaries and states the main problem. Section 3 presents the incremental proximal method for solving

2010 *Mathematics Subject Classification.* 65K05, 90C25, 90C90.

Key words and phrases. Fixed point; Incremental computing; Nonsmooth convex optimization; Proximal method; Quasi-nonexpansive mapping.

the main problem and analyzes its convergence. Section 4 numerically compares the behaviors of the proposed method and the existing ones. Section 5 concludes the paper with a brief summary.

2. MATHEMATICAL PRELIMINARIES

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. We use the standard notation $\mathbb{N}$ for the natural numbers including zero and $\mathbb{R}^N$ for an N-dimensional Euclidean space.

2.1. **Convexity, proximal point, and subdifferentiability.** A function $f\colon H \to \mathbb{R}$ is said to be convex if, for all $x, y \in H$ and for all $\alpha \in [0,1]$, $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$. A function f is said to be *strictly convex* [3, Definition 8.6] if, for all $x, y \in H$ and for all $\alpha \in (0,1)$, $x \neq y$ implies $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$. f is strongly convex with constant β [3, Definition 10.5] if there exists $\beta > 0$ such that, for all $x, y \in H$ and for all $\alpha \in (0,1)$, $f(\alpha x + (1-\alpha)y) + (\beta\alpha(1-\alpha)/2)\|x-y\|^2 \leq \alpha f(x) + (1-\alpha) f(y)$.

Let $f\colon H \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex. Then, the *proximity operator* of f [3, Definition 12.23], [15], denoted by Prox_f, maps every $x \in H$ to the unique minimizer of $f(\cdot) + (1/2)\|x - \cdot\|^2$; i.e.,

$$\left\{\mathrm{Prox}_f(x)\right\} = \operatorname*{argmin}_{y \in H} \left[f(y) + \frac{1}{2}\|x-y\|^2 \right].$$

The uniqueness and existence of $\mathrm{Prox}_f(x)$ are guaranteed for all $x \in H$ [3, Definition 12.23], [14]. We call $\mathrm{Prox}_f(x)$ the *proximal point* of f at x. Let $\mathrm{dom}(f) := \{x \in H \colon f(x) < +\infty\}$ be the domain of a function $f\colon H \to (-\infty, +\infty]$.

The *subdifferential* [3, Definition 16.1] of f is defined for all $x \in H$ by

$$\partial f(x) := \{u \in H \colon f(y) \geq f(x) + \langle y - x, u \rangle \ (y \in H)\}.$$

We call $u \in \partial f(x)$ the *subgradient* of f at x.

2.2. **Quasi-nonexpansivity and demiclosedness.** The fixed point set of a mapping $Q\colon H \to H$ is denoted by

$$\mathrm{Fix}(Q) := \{x \in H \colon Q(x) = x\}.$$

Q is said to be *quasi-nonexpansive* [3, Definition 4.1(iii)] if $\|Q(x) - y\| \leq \|x - y\|$ for all $x \in H$ and for all $y \in \mathrm{Fix}(Q)$. When a quasi-nonexpansive mapping has one fixed point, its fixed point set is closed and convex [2, Proposition 2.6]. Q is said to be *quasi-firmly nonexpansive* [1, Section 3] if, for all $x \in H$ and for all $y \in \mathrm{Fix}(Q)$,

$$\|Q(x) - y\|^2 + \|(\mathrm{Id} - Q)(x)\|^2 \leq \|x - y\|^2,$$

where $\mathrm{Id}(x) := x$ $(x \in H)$. Any quasi-firmly nonexpansive mapping satisfies the quasi-nonexpansivity condition. Moreover, Q is quasi-firmly nonexpansive if and only if $R := 2Q - \mathrm{Id}$ is quasi-nonexpansive [3, Proposition 4.2], which implies that $(1/2)(\mathrm{Id} + R)$ is quasi-firmly nonexpansive when R is quasi-nonexpansive.

Let $x, u \in H$ and $(x_n)_{n \in \mathbb{N}} \subset H$. $\mathrm{Id} - Q$ is said to be *demiclosed* if weak convergence of (x_n) to x and $\lim_{n \to +\infty} \|x_n - Q(x_n) - u\| = 0$ imply $x - Q(x) = u$. $\mathrm{Id} - Q$ is demiclosed when Q is nonexpansive, i.e., $\|Q(x) - Q(y)\| \leq \|x - y\|$ $(x, y \in H)$ [3, Theorem 4.17]. Section 4 provides an example in which Q is quasi-firmly nonexpansive and $\mathrm{Id} - Q$ is demiclosed.

The *metric projection* P_C onto a nonempty, closed convex subset C of H is firmly nonexpansive, i.e., $\|P_C(x)-P_C(y)\|^2+\|(\mathrm{Id}-P_C)(x)-(\mathrm{Id}-P_C)(y)\|^2 \leq \|x-y\|^2$ $(x,y\in H)$. Moreover, $\mathrm{Fix}(P_C)=C$ [3, Proposition 4.8, (4.8)].

2.3. **Main problem and propositions.** This paper deals with the following problem:

Problem 2.1. [8, Problem 2.1], [18, Problem 1.1] Assume that

(A1) $Q_i\colon H\to H$ $(i\in\mathscr{I}:=\{1,2,\ldots,I\})$ is quasi-firmly nonexpansive;
(A2) $f_i\colon H\to\mathbb{R}$ $(i\in\mathscr{I})$ is convex and continuous with $\mathrm{dom}(f_i):=\{x\in H\colon f_i(x)<+\infty\}=H$.

Then,

$$\text{minimize } f(x):=\sum_{i\in\mathscr{I}} f_i(x) \text{ subject to } x\in X:=\bigcap_{i\in\mathscr{I}}\mathrm{Fix}(Q_i),$$

where one assumes that there exists a solution of Problem 2.1.

The following propositions will be used to perform the convergence analysis of the proposed method for solving Problem 2.1.

Proposition 2.1. [3, Propositions 12.26, 12.27, 12.28, and 16.14] *Let $f\colon H\to(-\infty,+\infty]$ be proper, lower semicontinuous, and convex. Then, the following hold:*

(i) *Let $x,p\in H$. $p=\mathrm{Prox}_f(x)$ if and only if $x-p\in\partial f(p)$ (i.e., $\langle y-p,x-p\rangle+f(p)\leq f(y)$ for all $y\in H$).*
(ii) Prox_f *is firmly nonexpansive with* $\mathrm{Fix}(\mathrm{Prox}_f)=\mathrm{argmin}_{x\in H} f(x)$.
(iii) *If f is continuous at $x\in\mathrm{dom}(f)$, $\partial f(x)$ is nonempty. Moreover, $\delta>0$ exists such that $\partial f(B(x;\delta))$ is bounded, where $B(x;\delta)$ stands for a closed ball with center x and radius δ.*

Proposition 2.2. [16, Lemma 3.1] *Suppose that $(x_n)_{n\in\mathbb{N}}\subset H$ weakly converges to $\hat{x}\in H$ and $\bar{x}\neq\hat{x}$. Then, $\liminf_{n\to+\infty}\|x_n-\hat{x}\|<\liminf_{n\to+\infty}\|x_n-\bar{x}\|$.*

Proposition 2.3. [3, Theorem 9.1] *When $f\colon H\to\mathbb{R}$ is convex, f is weakly lower semicontinuous if and only if f is lower semicontinuous.*

Proposition 2.4. [13, Lemma 2.1] *Let $(\Gamma_n)_{n\in\mathbb{N}}\subset\mathbb{R}$ and suppose that $(\Gamma_{n_j})_{j\in\mathbb{N}}$ $(\subset(\Gamma_n)_{n\in\mathbb{N}})$ exists such that $\Gamma_{n_j}<\Gamma_{n_j+1}$ for all $j\in\mathbb{N}$. Define $(\tau(n))_{n\geq n_0}\subset\mathbb{N}$ by $\tau(n):=\max\{k\leq n\colon\Gamma_k<\Gamma_{k+1}\}$ for some $n_0\in\mathbb{N}$. Then, $(\tau(n))_{n\geq n_0}$ is increasing and $\lim_{n\to\infty}\tau(n)=+\infty$. Moreover, $\Gamma_{\tau(n)}\leq\Gamma_{\tau(n)+1}$ and $\Gamma_n\leq\Gamma_{\tau(n)+1}$ for all $n\geq n_0$.*

3. Proposed Incremental Proximal Method

Algorithm 1 is the proposed algorithm for solving Problem 2.1.

Let us consider a network system with I users and assume that user i has its own private objective function f_i and mapping Q_i and tries to minimize f_i over $\mathrm{Fix}(Q_i)$. Moreover, let us assume that user i can communicate with users $(i-1)$ and $(i+1)$, where user $(I+1)$ is user 1 and user 0 is user I. This implies that user $(i-1)$ can transmit $x_{n,i-1}$ to user i at iteration n. Then, at iteration n, since user i has its own objective function f_i, it computes $y_{n,i}:=\mathrm{Prox}_{\gamma_n f_i}(x_{n,i-1})$. Moreover, user i has its own constraint set $\mathrm{Fix}(Q_i)$, with which it tries to find a fixed point of Q_i by using $x_{n,i}:=Q_i(y_{n,i})$. Accordingly, user I can compute $x_{n+1}=x_{n+1,0}=x_{n,I}$ and transmit it to user 1, which implies that Algorithm 1 is well-defined for the network system.

Algorithm 1 Incremental Proximal Method for Solving Problem 2.1

Require: $(\gamma_n)_{n\in\mathbb{N}} \subset (0,+\infty)$
$n \leftarrow 0$, $x_0 = x_{0,0} \in H$
loop
  for $i = 1$ to $i = I$ **do**
    $x_{n,i} := Q_i(\mathrm{Prox}_{\gamma_n f_i}(x_{n,i-1}))$
  end for
  $x_{n+1} = x_{n+1,0} = x_{n,I}$
  $n \leftarrow n+1$
end loop

Let us compare Algorithm 1 with the existing incremental subgradient method [8, Algorithm 4.1] for solving Problem 2.1 that is as follows:

$$\begin{aligned} &Q_{\alpha,i} := \alpha \mathrm{Id} + (1-\alpha)Q_i, \\ &g_{n,i} \in \partial f_i\left(Q_{\alpha,i}(x_{n,i-1})\right), \\ &x_{n,i} := Q_{\alpha,i}(x_{n,i-1}) - \lambda_n g_{n,i}, \\ &x_{n+1} = x_{n+1,0} = x_{n,I}. \end{aligned} \tag{3.1}$$

The difference between Algorithms 1 and (3.1) is the form of $x_{n,i}$; i.e., Algorithm 1 uses $x_{n,i} = Q_i(\mathrm{Prox}_{\gamma_n f_i}(x_{n,i-1}))$, while algorithm (3.1) uses $x_{n,i} := Q_{\alpha,i}(x_{n,i-1}) - \lambda_n g_{n,i}$. Section 4 compares the behaviors of Algorithm 1 and algorithm (3.1) for a concrete optimization problem.

First, we prove the following lemma:

Lemma 3.1. *Consider Problem 2.1 and define $y_{n,i} := \mathrm{Prox}_{\gamma_n f_i}(x_{n,i-1})$ for all $i \in \mathscr{I}$ and for all $n \in \mathbb{N}$. Then, Algorithm 1 satisfies that, for all $x \in X$ and for all $n \in \mathbb{N}$,*

$$\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - \sum_{i\in\mathscr{I}} \left\{ \|x_{n,i-1} - y_{n,i}\|^2 + \|x_{n,i} - y_{n,i}\|^2 \right\} + 2\gamma_n \sum_{i\in\mathscr{I}} (f_i(x) - f_i(y_{n,i})).$$

Proof. Let $x \in X$ and $n \in \mathbb{N}$ be fixed arbitrarily. The definition of $y_{n,i} := \mathrm{Prox}_{\gamma_n f_i}(x_{n,i-1})$ and Proposition 2.1(i) ensure that, for all $i \in \mathscr{I}$,

$$\langle x - y_{n,i}, x_{n,i-1} - y_{n,i} \rangle \leq \gamma_n (f_i(x) - f_i(y_{n,i})),$$

which, together with $2\langle x, y\rangle = \|x\|^2 + \|y\|^2 - \|x-y\|^2$ $(x, y \in H)$, implies that

$$2\gamma_n (f_i(x) - f_i(y_{n,i})) \geq \|x - y_{n,i}\|^2 + \|x_{n,i-1} - y_{n,i}\|^2 - \|x - x_{n,i-1}\|^2.$$

Accordingly, for all $i \in \mathscr{I}$,

$$\|y_{n,i} - x\|^2 \leq \|x_{n,i-1} - x\|^2 - \|x_{n,i-1} - y_{n,i}\|^2 + 2\gamma_n (f_i(x) - f_i(y_{n,i})). \tag{3.2}$$

The definition of $x_{n,i} := Q_i(y_{n,i})$ and (A1) guarantee that, for all $i \in \mathscr{I}$,

$$\|x_{n,i} - x\|^2 \leq \|y_{n,i} - x\|^2 - \|x_{n,i} - y_{n,i}\|^2. \tag{3.3}$$

Hence, (3.2) and (3.3) ensure that

$$\|x_{n,i} - x\|^2 \leq \|x_{n,i-1} - x\|^2 - \|x_{n,i-1} - y_{n,i}\|^2 - \|x_{n,i} - y_{n,i}\|^2 + 2\gamma_n (f_i(x) - f_i(y_{n,i})),$$

which implies that

$$\begin{aligned}\|x_{n+1}-x\|^2 &= \|x_{n,I}-x\|^2\\ &\leq \|x_{n,I-1}-x\|^2 - \left\{\|x_{n,I-1}-y_{n,I}\|^2 + \|x_{n,I}-y_{n,I}\|^2\right\} + 2\gamma_n(f_I(x)-f_I(y_{n,I}))\\ &\leq \|x_n-x\|^2 - \sum_{i\in\mathscr{I}}\left\{\|x_{n,i-1}-y_{n,i}\|^2 + \|x_{n,i}-y_{n,i}\|^2\right\} + 2\gamma_n\sum_{i\in\mathscr{I}}(f_i(x)-f_i(y_{n,i})).\end{aligned}$$

This completes the proof. □

The convergence analysis of Algorithm 1 depends on the following:

Assumpion 3.1. The sequence $(y_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ is bounded.

Assume that, for all $i \in \mathscr{I}$, $\operatorname{argmin}_{x\in H} f_i(x)(=\operatorname{Fix}(\operatorname{Prox}_{f_i})) \neq \emptyset$ and $\operatorname{Fix}(Q_i)$ is bounded. Then, we can choose a bounded, closed convex set C_i (e.g., C_i is a closed ball with a large enough radius) satisfying $C_i \supset \operatorname{Fix}(Q_i)$, and hence, we can compute

$$x_{n,i} := P_{C_i}[Q_i(y_{n,i})] \in C_i \tag{3.4}$$

instead of $x_{n,i}$ in Algorithm 1. The boundedness of C_i $(i \in \mathscr{I})$ implies that $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ is bounded. Proposition 2.1(ii) thus ensures that, for all $i \in \mathscr{I}$, for all $n \in \mathbb{N}$, and for all $x \in \operatorname{Fix}(\operatorname{Prox}_{f_i})$,

$$\|y_{n,i}-x\| \leq \|x_{n,i-1}-x\|,$$

which implies that $(y_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ is bounded. Hence, it can be assumed that $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ in Algorithm 1 is as in (3.4) in place of Assumption 3.1.

We also have the following lemma:

Lemma 3.2. *Consider Problem 2.1 and suppose that Assumption 3.1 holds. Then,* $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ *is bounded.*

Proof. Assumption (A1) ensures that, for all $x \in X$, for all $i \in \mathscr{I}$, and for all $n \in \mathbb{N}$,

$$\|x_{n,i}-x\| \leq \|y_{n,i}-x\|,$$

which, together with Assumption 3.1, implies that $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ is bounded. □

3.1. **Constant step-size rule.** The following is a convergence analysis of Algorithm 1 with a constant step size, which can be proven by referring to the proofs in [8], [9], and [18].

Theorem 3.1. *Suppose that the assumptions in Lemma 3.2 hold. Then, Algorithm 1 with* $\gamma_n := \gamma > 0$ $(n \in \mathbb{N})$ *satisfies*

$$\liminf_{n\to+\infty}\sum_{i\in\mathscr{I}}\|y_{n,i}-Q_i(y_{n,i})\|^2 \leq M_1\gamma \quad \textit{and} \quad \liminf_{n\to+\infty}\sum_{i\in\mathscr{I}} f_i(y_{n,i}) \leq f^\star,$$

where $M_1 := \sup\{2\sum_{i\in\mathscr{I}}(f_i(x)-f_i(y_{n,i}))\colon n \in \mathbb{N}\} < +\infty$ *for some* $x \in X$ *and* $f^\star$ *is the optimal value of Problem 2.1.*

Proof. Let $x \in X$ be fixed arbitrarily. The definition of $\partial f_i(x)$ and the Cauchy-Schwarz inequality imply that, for all $i \in \mathscr{I}$, for all $n \in \mathbb{N}$, and for all $u_i \in \partial f_i(x)$,

$$f_i(x)-f_i(y_{n,i}) \leq \langle x-y_{n,i}, u_i\rangle \leq \|y_{n,i}-x\|\,\|u_i\|,$$

which, together with $\tilde{B} := \max_{i\in\mathscr{I}} \sup\{\|y_{n,i} - x\| \colon n \in \mathbb{N}\} < +\infty$ (by Assumption 3.1), implies that

$$\frac{M_1}{2} := \sup\left\{\sum_{i\in\mathscr{I}} (f_i(x) - f_i(y_{n,i})) \colon n \in \mathbb{N}\right\} \leq I\tilde{B} \max_{i\in\mathscr{I}} \|u_i\| < +\infty. \tag{3.5}$$

We first show that

$$\liminf_{n\to+\infty} \sum_{i\in\mathscr{I}} \left\{\|x_{n,i-1} - y_{n,i}\|^2 + \|x_{n,i} - y_{n,i}\|^2\right\} \leq M_1\gamma. \tag{3.6}$$

Define $X_{n,i} := \|x_{n,i-1} - y_{n,i}\|^2 + \|x_{n,i} - y_{n,i}\|^2$ $(i \in \mathscr{I}, n \in \mathbb{N})$. Assume that (3.6) does not hold. Then $\delta > 0$ exists such that

$$\liminf_{n\to+\infty} \sum_{i\in\mathscr{I}} X_{n,i} > M_1\gamma + 2\delta.$$

Accordingly, the property of the limit inferior of $(\sum_{i\in\mathscr{I}} X_{n,i})_{n\in\mathbb{N}}$ ensures that $n_0 \in \mathbb{N}$ exists such that, for all $n \geq n_0$,

$$\sum_{i\in\mathscr{I}} X_{n,i} > M_1\gamma + \delta. \tag{3.7}$$

Accordingly, Lemma 3.1 with $\gamma_n := \gamma$ $(n \in \mathbb{N})$ guarantees that, for all $n \geq n_0$,

$$\begin{aligned}
\|x_{n+1} - x\|^2 &\leq \|x_n - x\|^2 - \sum_{i\in\mathscr{I}} X_{n,i} + 2\gamma \sum_{i\in\mathscr{I}} (f_i(x) - f_i(y_{n,i})) \\
&< \|x_n - x\|^2 - (M_1\gamma + \delta) + M_1\gamma \\
&= \|x_n - x\|^2 - \delta \\
&< \|x_{n_0} - x\|^2 - \delta(n + 1 - n_0).
\end{aligned}$$

The right side of the above inequality approaches minus infinity as n diverges. Hence, we have a contradiction. This implies that (3.6) holds. Therefore,

$$\liminf_{n\to+\infty} \sum_{i\in\mathscr{I}} \|y_{n,i} - x_{n,i}\|^2 = \liminf_{n\to+\infty} \sum_{i\in\mathscr{I}} \|y_{n,i} - Q_i(y_{n,i})\|^2 \leq M_1\gamma.$$

Next, we show that

$$\liminf_{n\to+\infty} \sum_{i\in\mathscr{I}} f_i(y_{n,i}) \leq f^\star. \tag{3.8}$$

Assume that (3.8) does not hold. An argument similar to the one for obtaining (3.7) implies that there exist $\zeta > 0$ and $m_0 \in \mathbb{N}$ such that, for all $n \geq m_0$,

$$\sum_{i\in\mathscr{I}} f_i(y_{n,i}) - f^\star > \zeta.$$

Lemma 3.1 thus ensures that, for all $n \geq m_0$ and for all $x^\star \in X^\star := \{x^\star \in X \colon f(x^\star) = f^\star = \inf_{x\in X} f(x)\}$,

$$\begin{aligned}
\|x_{n+1} - x^\star\|^2 &\leq \|x_n - x^\star\|^2 + 2\gamma\left(f^\star - \sum_{i\in\mathscr{I}} f_i(y_{n,i})\right) \\
&< \|x_n - x^\star\|^2 - 2\gamma\zeta \\
&< \|x_{m_0} - x^\star\|^2 - 2\gamma\zeta(n + 1 - m_0),
\end{aligned}$$

which is a contradiction. Accordingly, (3.8) holds. This completes the proof. □

3.2. **Diminishing step-size rule.** The following is a convergence analysis of Algorithm 1 with a diminishing step size, which can be proven by referring to the proofs in [8], [9], and [18]. In particular, we can prove that Algorithm 1 strongly converges to the unique solution by referring to the proof of [8, Theorem 3.2]. Hence, we shall only prove the weak convergence of Algorithm 1.

Theorem 3.2. *Suppose that the assumptions in Lemma 3.2 hold and* $\mathrm{Id} - Q_i$ $(i \in \mathscr{I})$ *is demiclosed. Let* $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ *be the sequence generated by Algorithm 1 with* $(\gamma_n)_{n\in\mathbb{N}}$ *satisfying* $\lim_{n\to+\infty} \gamma_n = 0$ *and* $\sum_{n=0}^{+\infty} \gamma_n = +\infty$*. Then, there exists a subsequence of each of* $(x_{n,i})_{n\in\mathbb{N}}$ *and* $(y_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ *that weakly converges to a solution of Problem 2.1. Moreover,* $(x_{n,i})_{n\in\mathbb{N}}$ *and* $(y_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ *strongly converge to a unique solution of Problem 2.1 if one of the following holds:*

(i) *One* f_i *is strongly convex;*
(ii) H *is finite-dimensional, and one* f_i *is strictly convex.*

Proof. We consider two cases.

Case 1: Suppose that there exists $m_0 \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$ and for all $x^\star \in X^\star$, $n \geq m_0$ implies $\|x_{n+1} - x^\star\| \leq \|x_n - x^\star\|$, where $X^\star := \{x^\star \in X : f(x^\star) = f^\star = \inf_{x\in X} f(x)\}$. Then, there exists $c := \lim_{n\to+\infty} \|x_n - x^\star\|$. Let $x^\star \in X^\star$ be fixed arbitrarily. Lemma 3.1, together with a discussion similar to that of (3.5), guarantees that there exists

$$\frac{M_2}{2} := \sup\left\{\sum_{i\in\mathscr{I}} (f_i(x^\star) - f_i(y_{n,i})) : n \in \mathbb{N}\right\} < +\infty$$

such that, for all $n \geq m_0$,

$$\sum_{i\in\mathscr{I}} X_{n,i} \leq \|x_n - x^\star\|^2 - \|x_{n+1} - x^\star\|^2 + M_2\gamma_n. \tag{3.9}$$

Accordingly, the conditions $\lim_{n\to+\infty} \gamma_n = 0$ and $c := \lim_{n\to+\infty} \|x_n - x^\star\|$ mean that

$$\lim_{n\to+\infty} \|x_{n,i-1} - y_{n,i}\| = 0 \text{ and } \lim_{n\to+\infty} \|x_{n,i} - y_{n,i}\| = 0 \ (i \in \mathscr{I}), \tag{3.10}$$

which, together with the triangle inequality, implies that

$$\lim_{n\to+\infty} \|x_{n,i-1} - x_{n,i}\| = 0 \ (i \in \mathscr{I}). \tag{3.11}$$

From Lemma 3.1, for all $x \in X$ and for all $k \in \mathbb{N}$,

$$2\gamma_k \sum_{i\in\mathscr{I}} (f_i(y_{k,i}) - f_i(x)) \leq \|x_k - x\|^2 - \|x_{k+1} - x\|^2. \tag{3.12}$$

Define $N_k(x) := \sum_{i\in\mathscr{I}} (f_i(y_{k,i}) - f_i(x))$ $(k \in \mathbb{N}, x \in X)$. We have that, for all $n \in \mathbb{N}$ and for all $x \in X$,

$$2\sum_{k=0}^{n} \gamma_k N_k(x) \leq \|x_0 - x\|^2 - \|x_{n+1} - x\|^2 \leq \|x_0 - x\|^2.$$

Accordingly, for all $x \in X$,

$$\sum_{k=0}^{+\infty} \gamma_k N_k(x) < +\infty. \tag{3.13}$$

Here, we show that, for all $x \in X$,

$$\liminf_{n\to+\infty} N_n(x) \leq 0. \tag{3.14}$$

Assume that (3.14) does not hold; i.e., there exists $x_0 \in X$ such that $\liminf_{n\to+\infty} N_n(x_0) > 0$. Then, $m_1 \in \mathbb{N}$ and $\theta > 0$ exist such that, for all $n \geq m_1$, $N_n(x_0) \geq \theta$. From (3.13) and $\sum_{n=0}^{+\infty} \gamma_n = +\infty$,

$$+\infty = \theta \sum_{k=m_1}^{+\infty} \gamma_k \leq \sum_{k=m_1}^{+\infty} \gamma_k N_k(x_0) < +\infty,$$

which is a contradiction. Hence, (3.14) holds, i.e., for all $x \in X$,

$$\liminf_{n\to+\infty} \sum_{i\in\mathscr{I}} f_i(y_{n,i}) \leq \sum_{i\in\mathscr{I}} f_i(x) =: f(x). \tag{3.15}$$

The definition of $u_{n,i} \in \partial f_i(x_n)$ and the Cauchy-Schwarz inequality ensure that, for all $i \in \mathscr{I}$ and for all $n \in \mathbb{N}$,

$$f_i(x_n) - f_i(y_{n,i}) \leq \langle x_n - y_{n,i}, u_{n,i}\rangle \leq B_1 \|x_{n,0} - y_{n,i}\|,$$

where $B_1 := \max_{i\in\mathscr{I}} \sup\{\|u_{n,i}\| \colon n \in \mathbb{N}\} < +\infty$ by Proposition 2.1(iii) and the boundedness of $(x_n)_{n\in\mathbb{N}}$ (see also Lemma 3.2). Accordingly, the triangle inequality implies that, for all $i \in \mathscr{I}$ and for all $n \in \mathbb{N}$,

$$f_i(x_n) - f_i(y_{n,i}) \leq B_1 \{\|x_{n,0} - x_{n,i}\| + \|x_{n,i} - y_{n,i}\|\} \leq B_1 \left\{ \sum_{j=0}^{i-1} \|x_{n,j} - x_{n,j+1}\| + \|x_{n,i} - y_{n,i}\| \right\}.$$

Hence, for all $n \in \mathbb{N}$,

$$f(x_n) = \sum_{i\in\mathscr{I}} f_i(x_n) \leq B_1 \left\{ \sum_{i\in\mathscr{I}} \|x_{n,i} - y_{n,i}\| + \sum_{i\in\mathscr{I}} \sum_{j=0}^{i-1} \|x_{n,j} - x_{n,j+1}\| \right\} + \sum_{i\in\mathscr{I}} f_i(y_{n,i}).$$

Therefore, (3.10), (3.11), and (3.15) lead to the finding that, for all $x \in X$,

$$\liminf_{n\to+\infty} f(x_n) \leq f(x). \tag{3.16}$$

Accordingly, a subsequence $(x_{n_l})_{l\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ exists such that, for all $x \in X$,

$$\lim_{l\to\infty} f(x_{n_l}) = \liminf_{n\to\infty} f(x_n) \leq f(x). \tag{3.17}$$

Since $(x_{n_l})_{l\in\mathbb{N}}$ is bounded (see also Lemma 3.2), there exists $(x_{n_{l_m}})_{m\in\mathbb{N}}$ $(\subset (x_{n_l})_{l\in\mathbb{N}})$ such that $(x_{n_{l_m}})_{m\in\mathbb{N}}$ weakly converges to $x_* \in H$. From (3.10), $(y_{n_{l_m},i})$ $(i \in \mathscr{I})$ weakly converges to x_*. Hence, (3.10) and the demiclosedness of $\mathrm{Id} - Q_i$ ensure that $x_* \in \mathrm{Fix}(Q_i)$ $(i \in \mathscr{I})$, i.e., $x_* \in X$. Proposition 2.3 ensures that the continuity and convexity of f (by (A2)) imply that f is weakly lower semicontinuous, which means that $f(x_*) \leq \liminf_{m\to+\infty} f(x_{n_{l_m}})$. Therefore, (3.17) leads to the finding that, for all $x \in X$,

$$f(x_*) \leq \liminf_{m\to+\infty} f\left(x_{n_{l_m}}\right) = \lim_{m\to+\infty} f\left(x_{n_{l_m}}\right) \leq f(x),$$

that is, $x_* \in X^\star$. Let us take another subsequence $(x_{n_{l_k}})_{k\in\mathbb{N}}$ $(\subset (x_{n_l})_{l\in\mathbb{N}})$ such that $(x_{n_{l_k}})_{k\in\mathbb{N}}$ weakly converges to $x_{**} \in H$. A discussion similar to the one for obtaining $x_* \in X^\star$ guarantees that $x_{**} \in X^\star$. Here, it is proven that $x_* = x_{**}$. Let us assume that $x_* \neq x_{**}$. Then, the existence of $c := \lim_{n\to\infty} \|x_n - x^\star\|$ $(x^\star \in X^\star)$ and Proposition 2.2 imply that

$$\begin{aligned} c &= \lim_{m\to\infty} \left\|x_{n_{l_m}} - x_*\right\| < \lim_{m\to\infty} \left\|x_{n_{l_m}} - x_{**}\right\| \\ &= \lim_{n\to\infty} \|x_n - x_{**}\| = \lim_{k\to\infty} \left\|x_{n_{l_k}} - x_{**}\right\| < \lim_{k\to\infty} \left\|x_{n_{l_k}} - x_*\right\| \\ &= c, \end{aligned}$$

which is a contradiction. Hence, $x_* = x_{**}$. Accordingly, any subsequence of $(x_{n_l})_{l\in\mathbb{N}}$ converges weakly to $x_* \in X^\star$; i.e., $(x_{n_l})_{l\in\mathbb{N}}$ converges weakly to $x_* \in X^\star$. This means that x_* is a weak cluster point of $(x_n)_{n\in\mathbb{N}}$ and belongs to $X^\star$. A discussion similar to the one for obtaining $x_* = x_{**}$ guarantees that there is only one weak cluster point of $(x_n)_{n\in\mathbb{N}}$, so we can conclude that, in Case 1, $(x_n)_{n\in\mathbb{N}}$ weakly converges to a point in $X^\star$. Therefore, (3.11) ensures that $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathscr{I})$ weakly converges to a point in $X^\star$.

Case 2: Suppose that, for all $m \in \mathbb{N}$, there exist $n \in \mathbb{N}$ and $x_0^\star \in X^\star$ such that $n \geq m$ and $\|x_{n+1} - x_0^\star\| > \|x_n - x_0^\star\|$. This implies that $(x_{n_j})_{j\in\mathbb{N}}$ $(\subset (x_n)_{n\in\mathbb{N}})$ exists such that, for all $j \in \mathbb{N}$, $\|x_{n_j+1} - x_0^\star\| > \|x_{n_j} - x_0^\star\| =: \Gamma_{n_j}$. Proposition 2.4 thus guarantees that $m_1 \in \mathbb{N}$ exists such that, for all $n \geq m_1$, $\Gamma_{\tau(n)} \leq \Gamma_{\tau(n)+1}$, where $\tau(n)$ is defined as in Proposition 2.4. From Lemma 3.1 (see also (3.9)), for all $n \geq m_1$,

$$\sum_{i\in\mathscr{I}} \left\{ \left\|x_{\tau(n),i-1} - y_{\tau(n),i}\right\|^2 + \left\|x_{\tau(n),i} - y_{\tau(n),i}\right\|^2 \right\} \leq \Gamma_{\tau(n)}^2 - \Gamma_{\tau(n)+1}^2 + \tilde{M}_2\gamma_{\tau(n)} \leq \tilde{M}_2\gamma_{\tau(n)},$$

where

$$\frac{\tilde{M}_2}{2} := \sup\left\{ \sum_{i\in\mathscr{I}} \left(f_i(x^\star) - f_i(y_{\tau(n),i})\right) : n \in \mathbb{N} \right\}$$

is finite by Assumption 3.1 (see also (3.5)). Hence, the condition $\lim_{n\to+\infty} \gamma_{\tau(n)} = 0$ ensures that

$$\lim_{n\to+\infty} \left\|x_{\tau(n),i-1} - y_{\tau(n),i}\right\| = 0 \text{ and } \lim_{n\to+\infty} \left\|x_{\tau(n),i} - y_{\tau(n),i}\right\| = 0 \ (i \in \mathscr{I}), \tag{3.18}$$

which implies that

$$\lim_{n\to+\infty} \left\|x_{\tau(n),i-1} - x_{\tau(n),i}\right\| = 0 \ (i \in \mathscr{I}). \tag{3.19}$$

From (3.12), for all $n \geq m_1$,

$$2\gamma_{\tau(n)} N_{\tau(n)}(x_0^\star) \leq \Gamma_{\tau(n)}^2 - \Gamma_{\tau(n)+1}^2 \leq 0,$$

which, together with $\gamma_{\tau(n)} \geq 0$ $(n \geq m_1)$, implies that $N_{\tau(n)}(x_0^\star) \leq 0$. Accordingly,

$$\limsup_{n\to+\infty} \sum_{i\in\mathscr{I}} f_i\left(y_{\tau(n),i}\right) \leq f^\star.$$

An argument similar to the one for obtaining (3.16), together with (3.18) and (3.19), implies that

$$\limsup_{n\to+\infty} f\left(x_{\tau(n)}\right) \leq f^\star.$$

Choose a subsequence $(x_{\tau(n_k)})_{k\in\mathbb{N}}$ of $(x_{\tau(n)})_{n\geq m_1}$ arbitrarily. Then,

$$\limsup_{k\to+\infty} f\left(x_{\tau(n_k)}\right) \leq \limsup_{n\to+\infty} f\left(x_{\tau(n)}\right) \leq f^\star. \tag{3.20}$$

The boundedness of $(x_{\tau(n_k)})_{k\in\mathbb{N}}$ ensures that $(x_{\tau(n_{k_l})})_{l\in\mathbb{N}}$ $(\subset (x_{\tau(n_k)})_{k\in\mathbb{N}})$ exists such that $(x_{\tau(n_{k_l})})_{l\in\mathbb{N}}$ weakly converges to $x_\star \in H$. Then, (3.18) and the demiclosedness of $\mathrm{Id} - Q_i$ ensure that $x_\star \in X$. Moreover, Proposition 2.3 and (3.20) guarantee that

$$f(x_\star) \leq \liminf_{l\to+\infty} f\left(x_{\tau\left(n_{k_l}\right)}\right) \leq \limsup_{l\to+\infty} f\left(x_{\tau\left(n_{k_l}\right)}\right) \leq f^\star,$$

that is, $x_\star \in X^\star$. Therefore, $(x_{\tau(n_{k_l})})_{l\in\mathbb{N}}$ weakly converges to $x_\star \in X^\star$. From (3.19), $(x_{\tau(n_{k_l}),i})_{l\in\mathbb{N}}$ $(i \in \mathscr{I})$ weakly converges to $x_\star \in X^\star$. From Cases 1 and 2, there exists a subsequence of $(x_n)_{n\in\mathbb{N}}$ that weakly converges to a point in $X^\star$. This completes the proof. □

The following is a convergence-rate analysis of Algorithm 1 with a diminishing step size.

Theorem 3.3. *Suppose that the assumptions in Theorem 3.2 hold and a monotone decreasing sequence* $(\gamma_n)_{n\in\mathbb{N}}$ *satisfies* $\lim_{n\to+\infty}(n\gamma_n)^{-1} = 0$ *and* $\lim_{n\to+\infty} n^{-1}\sum_{k=0}^{n-1}\gamma_k = 0$. *Then, Algorithm 1 satisfies that, for all* $n \geq 1$,

$$\sum_{i\in\mathscr{I}}\left(\frac{1}{n}\sum_{k=0}^{n-1}\|y_{k,i} - Q_i(y_{k,i})\|^2\right) \leq \frac{\|x_0 - x\|^2}{n} + \frac{\tilde{M}_1}{n}\sum_{k=0}^{n-1}\gamma_k \quad \textit{and}$$

$$\sum_{i\in\mathscr{I}} f_i\left(\frac{1}{n}\sum_{k=0}^{n-1} y_{k,i}\right) \leq f^\star + \frac{B}{2n\gamma_n},$$

where $x^\star$ *is a solution of Problem 2.1,* $\tilde{M}_1 := \sup\{2\sum_{i\in\mathscr{I}}(f_i(x^\star) - f_i(y_{n,i}))\colon n \in \mathbb{N}\} < +\infty$, *and* $B := \sup\{\|x_n - x^\star\|^2\colon n \in \mathbb{N}\} < +\infty$.

Proof. Let $x^\star \in X^\star$. Lemma 3.1 implies that, for all $n \geq 1$,

$$\sum_{i\in\mathscr{I}}\sum_{k=0}^{n-1}\left\{\|x_{k,i-1} - y_{k,i}\|^2 + \|x_{k,i} - y_{k,i}\|^2\right\} \leq \|x_0 - x\|^2 + \tilde{M}_1\sum_{k=0}^{n-1}\gamma_k,$$

which in turn implies that

$$\begin{aligned}\sum_{i\in\mathscr{I}}\left(\frac{1}{n}\sum_{k=0}^{n-1}\|x_{k,i} - y_{k,i}\|^2\right) &\leq \frac{1}{n}\sum_{i\in\mathscr{I}}\sum_{k=0}^{n-1}\left\{\|x_{k,i} - y_{k,i}\|^2 + \|x_{k,i} - y_{k,i}\|^2\right\}\\ &\leq \frac{\|x_0 - x\|^2}{n} + \frac{\tilde{M}_1}{n}\sum_{k=0}^{n-1}\gamma_k.\end{aligned}$$

Lemma 3.1 indicates that, for all $k \in \mathbb{N}$,

$$\sum_{i\in\mathscr{I}} f_i(y_{k,i}) - f^\star \leq \frac{1}{2\gamma_k}\left\{\|x_k - x^\star\|^2 - \|x_{k+1} - x^\star\|^2\right\}.$$

Summing the above inequality from $k = 0$ to $k = n-1$ implies that, for all $n \geq 1$,

$$\frac{1}{n}\sum_{k=0}^{n-1}\sum_{i\in\mathscr{I}} f_i(y_{k,i}) - f^\star \leq \frac{1}{2n}\sum_{k=0}^{n-1}\frac{1}{\gamma_k}\left\{\|x_k - x^\star\|^2 - \|x_{k+1} - x^\star\|^2\right\}.$$

The definition of $X_n := \sum_{k=0}^{n-1}\gamma_k^{-1}\{\|x_k - x^\star\|^2 - \|x_{k+1} - x^\star\|^2\}$ means that

$$X_n = \frac{\|x_0 - x^\star\|}{\gamma_0} + \sum_{k=1}^{n-1}\left\{\frac{\|x_k - x^\star\|^2}{\gamma_k} - \frac{\|x_k - x^\star\|^2}{\gamma_{k-1}}\right\} - \frac{\|x_n - x^\star\|^2}{\gamma_{n-1}},$$

which, together with $\gamma_n \leq \gamma_{n-1}$ $(n \geq 1)$ and $B := \sup\{\|x_n - x^\star\|^2 \colon n \in \mathbb{N}\} < +\infty$ (by Lemma 3.2), implies that

$$X_n \leq \frac{B}{\gamma_0} + B\sum_{k=1}^{n-1}\left(\frac{1}{\gamma_k} - \frac{1}{\gamma_{k-1}}\right) = \frac{B}{\gamma_{n-1}} \leq \frac{B}{\gamma_n}.$$

The convexity of f_i thus ensures that, for all $n \geq 1$,

$$\sum_{i \in \mathscr{I}} f_i\left(\frac{1}{n}\sum_{k=0}^{n-1} y_{k,i}\right) - f^\star \leq \frac{B}{2n\gamma_n},$$

which completes the proof. □

Theorem 3.3 leads to the following corollary:

Corollary 3.1. *Suppose that the assumptions in Theorem 3.1 hold. Then, Algorithm 1 with $\gamma_n := n^{-1/2}$ $(n \geq 1)$ satisfies that, for all $n \geq 1$,*

$$\sum_{i \in \mathscr{I}}\left(\frac{1}{n}\sum_{k=0}^{n-1}\|y_{k,i} - Q_i(y_{k,i})\|^2\right) = \mathscr{O}\left(\sqrt{\frac{1+\ln n}{n}}\right) \text{ and } \sum_{i \in \mathscr{I}} f_i\left(\frac{1}{n}\sum_{k=0}^{n-1} y_{k,i}\right) \leq f^\star + \frac{B}{2\sqrt{n}},$$

where B is defined as in Theorem 3.3 and $\mathscr{O}$ stands for the Landau notation.

Proof. The step size $(\gamma_n)_{n\geq 1}$ defined by $\gamma_n := n^{-1/2}$ $(n \geq 1)$ is monotone decreasing and satisfies $\lim_{n\to+\infty}\gamma_n = 0$, $\lim_{n\to+\infty}(n\gamma_n)^{-1} = 0$, and $\sum_{n=0}^{+\infty}\gamma_n = +\infty$. Moreover, the Cauchy-Schwarz inequality and $\sum_{k=0}^{n-1} k^{-1} \leq 1 + \ln n$ mean that

$$\frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{\sqrt{k}} \leq \frac{\sqrt{n}}{n}\sqrt{\sum_{k=0}^{n-1}\frac{1}{k}} \leq \sqrt{\frac{1+\ln n}{n}},$$

which implies that $\lim_{n\to+\infty} n^{-1}\sum_{k=0}^{n-1}\gamma_k = 0$. Theorem 3.3 indicates that Algorithm 1 with $\gamma_n := n^{-1/2}$ satisfies the assertion in Corollary 3.1. □

Here, let us compare Algorithm 1 with the existing parallel proximal method [18, Algorithm 1] for solving Problem 2.1 that is as follows:

$$\begin{aligned} x_{n,i} &:= Q_i(\mathrm{Prox}_{\gamma_n f_i}(x_n)), \\ x_{n+1} &:= \frac{1}{I}\sum_{i \in \mathscr{I}} x_{n,i}. \end{aligned} \tag{3.21}$$

From Theorem 3.3 in [18] and Theorem 3.3 in this paper, Algorithm 1 has the same convergence rate as algorithm (3.21) (see [11] for a convergence rate analysis of stochastic approximation methods).

4. Numerical Comparisons

Let us compare the performance of Algorithm 1 with the one of the existing parallel proximal methods (PPM) [18, Algorithm 1] (see (3.21)), parallel subgradient method (PSM) [8, Algorithm 3.1], and incremental subgradient method (ISM) [8, Algorithm 4.1] (see (3.1)) for the following problem (4.2) (see also [8, Problem 5.1] and [18, Problem (4.1)]): Let $a_{i,j} > 0$, $b_{i,j}, d_i \in \mathbb{R}$ $(i \in \mathscr{I}, j = 1, 2, \ldots, N)$, and $c_i := (c_{i,j})_{j=1}^N \in \mathbb{R}^N$ $(i \in \mathscr{I})$

with $c_{i,j} > 0$. A function $g_i\colon \mathbb{R}^N \to \mathbb{R}$ and a mapping $Q_i\colon \mathbb{R}^N \to \mathbb{R}^N$ are defined for all $x \in \mathbb{R}^N$ by

$$g_i(x) := \begin{cases} \langle c_i, x\rangle + d_i & \text{if } \langle c_i, x\rangle > -d_i, \\ 0 & \text{otherwise} \end{cases}$$

and

$$(4.1) \qquad Q_i(x) := \begin{cases} x - \dfrac{g_i(x)}{\|z_i(x)\|^2} z_i(x) & \text{if } g_i(x) > 0, \\ x & \text{if } x \in \mathrm{lev}_{\leq 0} g_i := \{x \in \mathbb{R}^N \colon g_i(x) \leq 0\}, \end{cases}$$

where $z_i(x)$ is any vector in $\partial g_i(x)$. The mapping Q_i defined by (4.1) is called the *subgradient projection* related to g_i. Q_i satisfies quasi-firm nonexpansivity, and $\mathrm{Id} - Q_i$ satisfies the demiclosedness condition [1, Lemma 3.1]. Moreover, $\mathrm{Fix}(Q_i)$ is equal to the level set of g_i defined by $\mathrm{lev}_{\leq 0} g_i$. A function $f_i\colon \mathbb{R}^N \to \mathbb{R}$ is defined for all $x := (x_j)_{j=1}^N \in \mathbb{R}^N$ by

$$f_i(x) := \sum_{j=1}^N a_{i,j} \left| x_j - b_{i,j} \right|.$$

Then, we would like to

$$(4.2) \qquad \text{minimize } f(x) := \sum_{i \in \mathscr{I}} f_i(x) \text{ subject to } x \in X := \bigcap_{i \in \mathscr{I}} \mathrm{Fix}(Q_i) = \bigcap_{i \in \mathscr{I}} \mathrm{lev}_{\leq 0} g_i.$$

Obviously, problem (4.2) is an example of Problem 2.1.

The experiment was conducted on a MacBook Air (11-inch, Early 2015) with a 1.6GHz Intel Core i5 CPU processor, 8 GB, 1600 MHz DDR3 memory, and macOS Mojave 10.14.5 operating system. PPM, PSM, ISM, and Algorithm 1 were written in Python 3.6.8 with the NumPy 1.15.0 package. We set $I = 256$ and $N = 1000$ and randomly chose $a_{i,j} \in (0, 100]$, $b_{i,j} \in [-100, 100)$, $d_i \in [-1, 0)$, and $c_{i,j} \in [-0.5, 0.5)$. The stopping condition was $n = 10000$. The step sizes were as follows:

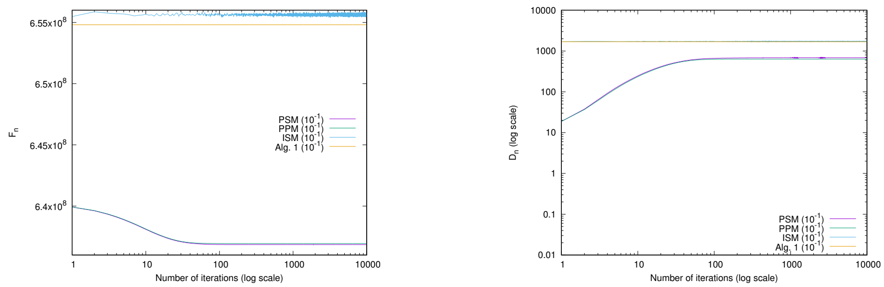
$$\text{Constant step sizes: } \gamma_n = \lambda_n := 10^{-1}, 10^{-3},$$
$$\text{Diminishing step sizes: } \gamma_n = \lambda_n := \frac{10^{-1}}{n+1}, \frac{10^{-3}}{n+1}.$$

The performance measures were as follows: for $n \in \mathbb{N}$,

$$F_n := \frac{1}{10} \sum_{s=1}^{10} \sum_{i \in \mathscr{I}} f_i(x_n(s)) \text{ and } D_n := \frac{1}{10} \sum_{s=1}^{10} \sum_{i \in \mathscr{I}} \|x_n(s) - Q_i(x_n(s))\|,$$

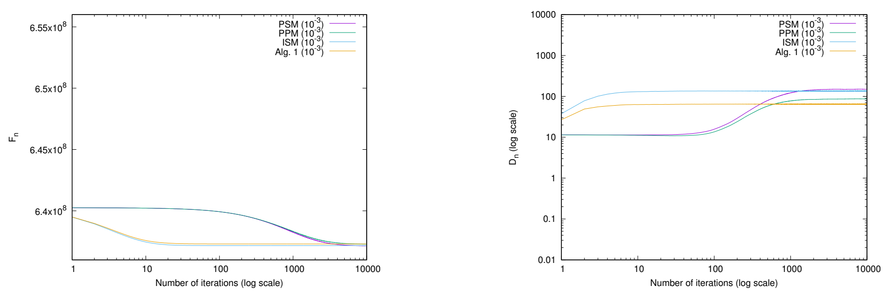
where $(x_n(s))_{n \in \mathbb{N}}$ is the sequence generated by each of the four methods with a randomly chosen initial point $x_0(s) \in [0, 1)^N$ $(s = 1, 2, \ldots, 10)$. If $(D_n)_{n \in \mathbb{N}}$ converges to 0, the methods converge to a point in X.

Figures 1 and 2 show the behaviors of F_n and D_n for the methods with constant step sizes 10^{-1} and 10^{-3}. These figures indicate that they did not converge to a point in X. These results suggest that the methods with a constant step size are not always able to solve Problem 2.1 (see, e.g., Theorem 3.1 for a convergence analysis of Algorithm 1 with a constant step size).



(A) Evaluation of F_n (B) Evaluation of D_n

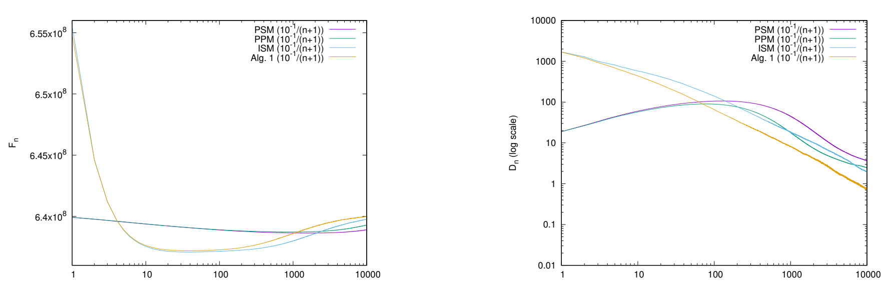
FIGURE 1. Behaviors of F_n and D_n for Algorithm 1, PPM, PSM, and ISM with $\gamma_n = \lambda_n = 10^{-1}$



(A) Evaluation of F_n (B) Evaluation of D_n

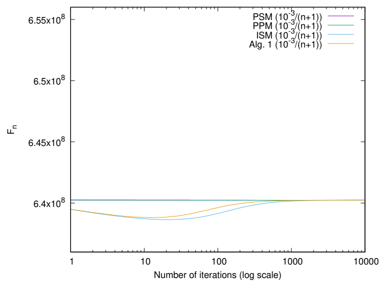
FIGURE 2. Behaviors of F_n and D_n for Algorithm 1, PPM, PSM, and ISM with $\gamma_n = \lambda_n = 10^{-3}$

Figures 3 and 4 show the behaviors of F_n and D_n for the methods with diminishing step sizes $10^{-1}/(n+1)$ and $10^{-3}/(n+1)$. These figures show that Algorithm 1 with these diminishing step sizes converged to a point in X, as guaranteed by Theorem 3.2. These figures also show that F_n remains stable. In particular, we can see that Algorithm 1 converged fastest.

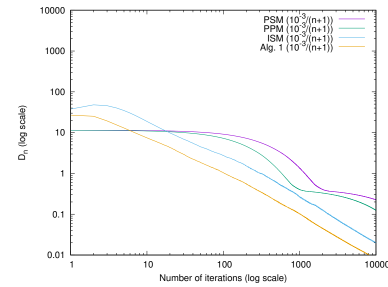


(A) Evaluation of F_n (B) Evaluation of D_n

FIGURE 3. Behaviors of F_n and D_n for Algorithm 1, PPM, PSM, and ISM with $\gamma_n = \lambda_n = 10^{-1}/(n+1)$



(A) Evaluation of F_n



(B) Evaluation of D_n

FIGURE 4. Behaviors of F_n and D_n for Algorithm 1, PPM, PSM, and ISM with $\gamma_n = \lambda_n = 10^{-3}/(n+1)$

5. CONCLUSION

This paper presented an incremental proximal method for solving the minimization problem of the sum of convex functions over the intersection of fixed point sets of quasi-nonexpansive mappings in a real Hilbert space. It also provided convergence and convergence-rate analyses. A numerical experiment compared the performance of the proposed method with those of the existing parallel and incremental methods for a concrete nonsmooth convex optimization problem over the intersection of level sets of convex functions. It showed that, whereas none of the methods with constant step sizes converged to a solution of the problem, the methods with diminishing step sizes solved the problem. In particular, the proposed method with diminishing step sizes performed better than the existing methods with the same step sizes.

Acknowledgments The authors would like to thank Professor Wataru Takahashi for giving us a chance to submit our paper to this journal. This work was supported by a JSPS KAKENHI Grant, Number JP18K11184.

REFERENCES

[1] H. H. Bauschke, J. Chen, A projection method for approximating fixed points of quasinonexpansive mappings without the usual demiclosedness condition, J. Nonlinear Convex Anal. 15 (2014), 129–135.
[2] H. H. Bauschke, P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert space, Math. Oper. Res. 26 (2001), 248–264.
[3] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, (2011), Springer, New York.
[4] Y. Hayashi, H. Iiduka, Optimality and convergence for convex ensemble learning with sparsity and diversity based on fixed point optimization, Neurocomputing 273 (2018), 367–372.
[5] H. Iiduka, Fixed point optimization algorithms for distributed optimization in networked systems, SIAM J. Optim. 23 (2013), 1–26.
[6] H. Iiduka, Parallel computing subgradient method for nonsmooth convex optimization over the intersection of fixed point sets of nonexpansive mappings, Fixed Point Theory Appl. 2015 (2015), Article ID 72.
[7] H. Iiduka, Incremental subgradient method for nonsmooth convex optimization with fixed point constraints, Optim. Method Softw. 31 (2016), 931–951.
[8] H. Iiduka, Convergence analysis of iterative methods for nonsmooth convex optimization over fixed point sets of quasi-nonexpansive mappings, Math. Program. 159 (2016), 509–538.

[9] H. Iiduka, Proximal point algorithms for nonsmooth convex optimization with fixed point constraints, Eur. J. Oper. Res. 253 (2016), 503–513.
[10] H. Iiduka, Stochastic fixed point optimization algorithm for classifier ensemble, IEEE Trans. Cybernetics (2019).
[11] H. Iiduka, Stochastic approximation methods using diagonal positive definite matrices for convex optimization with fixed point constraints, submitted.
[12] H. Iiduka, K. Hishinuma, Acceleration method combining broadcast and incremental distributed optimization algorithms, SIAM J. Optim. 24 (2014), 1840–1863.
[13] P. E. Maingé, The viscosity approximation process for quasi-nonexpansive mappings in Hilbert spaces, Comput. Math. Appl. 59 (2010), 74–79.
[14] G. J. Minty, A theorem on maximal monotonic sets in Hilbert space, J. Math. Anal. Appl. 11 (1965), 434–439.
[15] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, C. R. Acad. Sci. Paris Sér. A Math. 255 (1962), 2897–2899.
[16] Z. Opial, Weak convergence of the sequence of successive approximation for nonexpansive mappings, Bull. Amer. Math. Soc. 73 (1967), 591–597.
[17] K. Sakurai, T. Jimba, H. Iiduka, Iterative methods for parallel convex optimization with fixed point constraints, Journal of Nonlinear and Variational Analysis 3 (2019), 115–126.
[18] K. Shimizu, K. Hishinuma, H. Iiduka, Parallel computing proximal method for nonsmooth convex optimization with fixed point constraints of quasi-nonexpansive mappings, submitted.

(H. Oishi) COMPUTER SCIENCE COURSE, GRADUATE SCHOOL OF SCIENCE AND TECHNOLOGY, MEIJI UNIVERSITY, 1-1-1 HIGASHIMITA, TAMA-KU, KAWASAKI-SHI, KANAGAWA, 214-8571, JAPAN
E-mail address: `_pin_haru@cs.meiji.ac.jp`

(Y. Kobayashi) COMPUTER SCIENCE COURSE, GRADUATE SCHOOL OF SCIENCE AND TECHNOLOGY, MEIJI UNIVERSITY, 1-1-1 HIGASHIMITA, TAMA-KU, KAWASAKI-SHI, KANAGAWA, 214-8571, JAPAN
E-mail address: `yuukbys@cs.meiji.ac.jp`

(H. Iiduka) DEPARTMENT OF COMPUTER SCIENCE, MEIJI UNIVERSITY, 1-1-1 HIGASHIMITA, TAMA-KU, KAWASAKI-SHI, KANAGAWA, 214-8571, JAPAN
E-mail address: `iiduka@cs.meiji.ac.jp`