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

KENGO SHIMIZU¹, KAZUHIRO HISHINUMA¹,* HIDEAKI IIDUKA²,**

¹Computer Science Course, Graduate School of Science and Technology, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa, 214-8571, Japan
²Department of Computer Science, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa, 214-8571, Japan

Abstract. We present a parallel computing proximal method for solving the problem of minimizing the sum of convex functions over the intersection of fixed point sets of quasi-nonexpansive mappings in a real Hilbert space. We also provide a convergence analysis of the method for constant and diminishing step sizes under certain assumptions as well as a convergence-rate analysis for a diminishing step size. Numerical comparisons show that the performance of the algorithm is comparable with existing subgradient methods.

Keywords. Fixed point; Nonsmooth convex optimization; Parallel computing; Proximal method; Quasi-nonexpansive mapping.

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

1. INTRODUCTION

In this paper, we consider the following problem [7, Problem 2.1] (see [3, 9, 10] for applications of Problem 1.1):

Problem 1.1. Let $H$ be a real Hilbert space. Suppose that

(A1) $Q_i: H \to H$ $(i \in \mathcal{I} := \{1, 2, \ldots, I\})$ is quasi-firmly nonexpansive;

(A2) $f_i: H \to \mathbb{R}$ $(i \in \mathcal{I})$ is convex and continuous with $\text{dom}(f_i) := \{x \in H: f_i(x) < +\infty\} = H$.

Then,

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

where one assumes that there exists a solution of Problem 1.1 (see Sections 2 and 4 for the details).

Algorithms for solving this problem have been proposed in [7, 9]. Reference [7] proposed parallel and incremental subgradient methods for solving Problem 1.1 and provided convergence as well as convergence-rate analyses. Reference [9, 10] proposed stochastic fixed point optimization algorithms for solving a convex stochastic optimization problem that is to minimize the expectation of $f_i$s over $\text{Fix}(Q_1)$. The stochastic fixed point optimization algorithms can be applied to the classifier ensemble problem.

There are methods for solving Problem 1.1 where $Q_i$ is taken to be a nonexpansive mapping, which is a stronger assumption than a quasi-nonexpansive mapping. Subgradient methods were presented in [4, 5, 6, 11], while proximal methods were presented in [8, 16].

*Research Fellow of the Japan Society for the Promotion of Science. **Corresponding author.
E-mail addresses: kengo@cs.meiji.ac.jp (K. Shimizu), kaz@cs.meiji.ac.jp (K. Hishinuma), iiduka@cs.meiji.ac.jp (H. Iiduka).
Received ?? ?, 20??; Accepted ?? ?, 20??.
In this paper, we present a parallel method for solving Problem 1.1. The method is obtained by combining the parallel method in [7] with the proximal method in [8]. 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 Problem 1.1 (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 proposed method with the existing subgradient methods.

This paper is organized as follows. Section 2 gives the mathematical preliminaries. Section 3 presents the parallel proximal method for solving Problem 1.1 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 the $N$-dimensional Euclidean space.

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

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

$Q$ is said to be quasi-nonexpansive [2, Definition 4.1(iii)] if $\|Q(x) - y\| \leq \|x - y\|$ for all $x \in H$ and for all $y \in \text{Fix}(Q)$. 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 $y \in \text{Fix}(Q)$,

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

where $\text{Id}(x) := x \ (x \in H)$. Any quasi-firmly nonexpansive mapping satisfies the quasi-nonexpansivity condition. Moreover, $Q$ is quasi-firmly nonexpansive if and only if $R := 2Q - \text{Id}$ is quasi-nonexpansive [2, Proposition 4.2], which implies that $(1/2)(\text{Id} + R)$ is quasi-firmly nonexpansive when $R$ is quasi-nonexpansive. Let $x, u \in H$ and $(x_n)_{n \in \mathbb{N}} \subset H$. $\text{Id} - Q$ is said to be demiclosed if a weak convergence of $(x_n)$ to $x$ and $\lim_{n \to +\infty} \|x_n - Q(x_n) - u\| = 0$ imply $x - Q(x) = u$. $\text{Id} - Q$ is demiclosed when $Q$ is nonexpansive, i.e., $\|Q(x) - Q(y)\| \leq \|x - y\| \ (x, y \in H)$ [2, Theorem 4.17]. The metric projection $P_C$ onto a nonempty, closed convex subset $C$ of $H$ is firmly nonexpansive, i.e., $\|P_C(x) - P_C(y)\|^2 + \|\text{Id} - P_C)(x) - (\text{Id} - P_C)(y)\|^2 \leq \|x - y\|^2 \ (x, y \in H)$. Moreover, $\text{Fix}(P_C) = C$ [2, Proposition 4.8, (4.8)].

2.2. Convexity, proximal point, and subdifferentiability. A function $f : H \to \mathbb{R}$ 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 [2, 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 $\beta$ [2, 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 $f : H \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex. Then the proximity operator of $f$ [2, Definition 12.23], [14], denoted by $\text{Prox}_f$, maps every $x \in H$ to the unique minimizer of $f(\cdot) + \langle \cdot, x \rangle$. 

\[(1/2)\|x - \cdot\|^2; \text{ i.e.,} \]

\[
\{\text{Prox}_f(x)\} = \arg\min_{y \in H} \left[ f(y) + \frac{1}{2}\|x - y\|^2 \right] (x \in H).
\]

The uniqueness and existence of \(\text{Prox}_f(x)\) are guaranteed for all \(x \in H\) [2, Definition 12.23], [13]. We call \(\text{Prox}_f(x)\) the proximal point of \(f\) at \(x\). Let \(\text{dom}(f) := \{x \in H: f(x) < +\infty\}\) be the domain of a function \(f : H \rightarrow (-\infty, +\infty)\).

The subdifferential [2, Definition 16.1] of \(f\) is defined by

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

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

**Proposition 2.1.** [2, Propositions 12.26, 12.27, 12.28, and 16.14] Let \(f : H \rightarrow (-\infty, \infty]\) be proper, lower semicontinuous, and convex. Then, the following hold:

(i) Let \(x, p \in H\). \(p = \text{Prox}_f(x)\) 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) \(\text{Prox}_f\) is firmly nonexpansive with \(\text{Fix}(\text{Prox}_f) = \arg\min_{x \in H} f(x)\).

(iii) If \(f\) is continuous at \(x \in \text{dom}(f)\), \(\partial f(x)\) 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 \(\delta\).

The following propositions will be used to prove the main theorems in this paper.

**Proposition 2.2.** [15, Lemma 3.1] Suppose that \((x_n)_{n \in \mathbb{N}} \subset H\) weakly converges to \(\hat{x} \in H\) and \(\hat{x} \neq \hat{x}\). Then, \(\liminf_{n \to +\infty} \|x_n - \hat{x}\| < \liminf_{n \to +\infty} \|x_n - \hat{x}\|\).

**Proposition 2.3.** [2, Theorem 9.1] When \(f : H \rightarrow \mathbb{R}\) is convex, \(f\) is weakly lower semicontinuous if and only if \(f\) is lower semicontinuous.

**Proposition 2.4.** [12, Lemma 2.1] Let \((\Gamma_n)_{n \in \mathbb{N}} \subset \mathbb{R}\) and suppose that \((\Gamma_n)_{j \in \mathbb{N}} \subset (\Gamma_n)_{n \in \mathbb{N}}\) exists such that \(\Gamma_n < \Gamma_{n+1}\) for all \(j \in \mathbb{N}\). Define \((\tau(n))_{n \geq n_0} \subset \mathbb{N}\) by \(\tau(n) := \max\{k \leq n : \Gamma_k < \Gamma_{k+1}\}\) 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 Parallel Proximal Method

Algorithm 1 is the proposed algorithm for solving Problem 1.1.

Let us consider a network system with \(J\) 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 \(\text{Fix}(Q_i)\). Moreover, let us assume that each user can communicate with other users. Then, at iteration \(n\), each user can have \(x_n\) in common. Since user \(i\) has its own objective function \(f_i\), it computes \(y_{n,i} := \text{Prox}_{p_i,f_i}(x_n)\). Moreover, user \(i\) has its own constraint set \(\text{Fix}(Q_i)\), with which it tries to find a fixed point of \(Q_i\) by using \(x_{n,i} := Q_i(y_{n,i})\). Since the users can communicate with each other, user \(i\) can receive all \(x_{n,i}\), and hence, user \(i\) can compute \(x_{n+1} := (1/J)\sum_{i \in G} x_{n,i}\).
The definition of $x_i^2$ which, together with Lemma 3.1, and algorithm (while algorithm (Algorithm 3.1) for concrete optimization problems.

First, we prove the following lemma:

**Lemma 3.1.** Suppose that (A1) and (A2) hold and define $y_{n,i} := \text{Prox}_{\gamma_i f_i}(x_n)$ for all $i \in \mathcal{I}$ and for all $n \in \mathbb{N}$. Then, Algorithm 1 satisfies that, for all $x \in X$ and for all $n \in \mathbb{N}$,

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

**Proof.** Let $x \in X$ and $n \in \mathbb{N}$ be fixed arbitrarily. The definition of $y_{n,i} := \text{Prox}_{\gamma_i f_i}(x_n)$ and Proposition 2.1(i) ensure that, for all $i \in \mathcal{I}$,

$$\langle x - y_{n,i}, x_n - y_{n,i} \rangle \leq \gamma_n (f_i(x) - f_i(y_{n,i})), \quad \text{which, together with } 2 \langle x, y \rangle = \|x\|^2 + \|y\|^2 - \|x - y\|^2 \quad (x, y \in H), \text{ implies that}$$

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

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

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

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

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

Let us compare Algorithm 1 with the existing parallel subgradient method [7, Algorithm 3.1] for solving Problem 1.1. The parallel subgradient method [7, Algorithm 3.1] is as follows:

$$Q_{\alpha,i} := \alpha I + (1 - \alpha) Q_i,$$

$$g_{n,i} \in \partial f_i(Q_{\alpha,i}(x_n)),$$

$$x_{n,i} := Q_{\alpha,i}(x_n) - \lambda_n g_{n,i}, \quad (3.1)$$

$$x_{n+1} := \frac{1}{T} \sum_{i \in \mathcal{I}} x_{n,i}.$$
Hence, (3.2) and (3.3) imply that
\[ \|x_{n,i} - x\|^2 \leq \|x_n - x\|^2 - \|x_n - y_{n,i}\|^2 - \|x_{n,i} - y_{n,i}\|^2 + 2\gamma_n (f_i(x) - f_i(y_{n,i})). \]
Summing the above inequality from \( i = 1 \) to \( i = I \) and the convexity of \( \| \cdot \|^2 \) ensure that
\[ I \|x_{n+1} - x\|^2 \leq \sum_{i \in \mathcal{I}} \|x_{n,i} - x\|^2 \leq I \|x_n - x\|^2 - \sum_{i \in \mathcal{I}} \left\{ \|x_n - y_{n,i}\|^2 + \|x_{n,i} - y_{n,i}\|^2 \right\} + 2\gamma_n \sum_{i \in \mathcal{I}} (f_i(x) - f_i(y_{n,i})), \]
which completes the proof.

The convergence analysis of Algorithm 1 depends on the following:

**Assumption 3.1.** The sequence \((y_{n,i})_{n \in \mathbb{N}}\) \((i \in \mathcal{I})\) is bounded.

Assume that, for all \( i \in \mathcal{I} \), \( \arg\min_{x \in H} f_i(x) = \text{Fix}(\text{Prox}_{f_i}) \neq \emptyset \) and \( \text{Fix}(Q_i) \) is bounded. Then, we can choose in advance of running the algorithm a bounded, closed convex set \( C_i \) (e.g., \( C_i \) is a closed ball with a large enough radius) satisfying \( C_i \supset \text{Fix}(Q_i) \). Accordingly, we can compute
\[ x_{n,i} := P_{C_i}(Q_i(y_{n,i})) \in C_i \]  \hspace{1cm} (3.4)
instead of \( x_{n,i} \) in Algorithm 1. The boundedness of \( C_i \) \((i \in \mathcal{I})\) implies that \( (x_{n,i})_{n \in \mathbb{N}} \) \((i \in \mathcal{I})\) is bounded. Accordingly, \( (x_{n,i})_{n \in \mathbb{N}} \) is also bounded. Moreover, Proposition 2.1(ii) ensures that, for all \( i \in \mathcal{I} \), for all \( n \in \mathbb{N} \), and for all \( x \in \text{Fix}(\text{Prox}_{f_i}) \), \( \|y_{n,i} - x\| \leq \|x_n - x\| \). Hence, the boundedness of \( (x_{n,i})_{n \in \mathbb{N}} \) guarantees that \( (y_{n,i})_{n \in \mathbb{N}} \) \((i \in \mathcal{I})\) is bounded. Hence, it can be assumed that \( (x_{n,i})_{n \in \mathbb{N}} \) \((i \in \mathcal{I})\) in Algorithm 1 is as in (3.4) in place of Assumption 3.1.

We also have the following lemma:

**Lemma 3.2.** Suppose that (A1), (A2), and Assumption 3.1 hold. Then, \((x_{n,i})_{n \in \mathbb{N}}\) \((i \in \mathcal{I})\) and \((x_n)_{n \in \mathbb{N}}\) are bounded.

**Proof.** Assumption (A1) ensures that, for all \( x \in X \), for all \( i \in \mathcal{I} \), and for all \( n \in \mathbb{N} \),
\[ \|x_{n,i} - x\| \leq \|y_{n,i} - x\|, \]
which, together with Assumption 3.1, implies that \( (x_{n,i})_{n \in \mathbb{N}} \) \((i \in \mathcal{I})\) is bounded. Hence, the definition of \( x_n \) implies that \( (x_n)_{n \in \mathbb{N}} \) is also bounded. \( \square \)

### 3.1. Constant step-size rule

The following is a convergence analysis of Algorithm 1 with a constant step size, which indicates that Algorithm 1 with a small constant step size may approximate a solution of Problem 1.1.

**Theorem 3.1.** Suppose that (A1), (A2), and Assumption 3.1 hold. Then, Algorithm 1 with \( \gamma_n := \gamma > 0 \) satisfies that
\[ \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} \|y_{n,i} - Q_i(y_{n,i})\|^2 \leq IM_1 \gamma \quad \text{and} \quad \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} f_i(y_{n,i}) \leq f^*, \]
where \( M_1 := \sup \{(2/I)\sum_{i \in \mathcal{I}} (f_i(x) - f_i(y_{n,i})) : n \in \mathbb{N}\} < +\infty \) for some \( x \in X \) and \( f^* \) is the optimal value of Problem 1.1.
Proof. Let \( x \in X \) be fixed arbitrarily. The definition of \( \partial f_i(x) \) and the Cauchy-Schwarz inequality imply that, for all \( i \in \mathcal{I} \), 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 \( \tilde{B} := \max_{i \in \mathcal{I}} \sup \{\|y_{n,i} - x\| : n \in \mathbb{N}\} < +\infty \) (by Assumption 3.1), implies that

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

We first show that

\[
    \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} \left\{ \|x_n - y_{n,i}\|^2 + \|x_n - y_{n,i}\|_2^2 \right\} \leq IM_1 \gamma. \tag{3.6}
\]

If (3.6) does not hold, there exists \( \delta > 0 \) such that

\[
    \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} X_{n,i} > IM_1 \gamma + 2\delta.
\]

Accordingly, the property of the limit inferior of \( \left( \sum_{i \in \mathcal{I}} \{\|x_n - y_{n,i}\|^2 + \|x_n - y_{n,i}\|_2^2\} \right)_{n \in \mathbb{N}} \) ensures that

\[
    \sum_{i \in \mathcal{I}} X_{n,i} > IM_1 \gamma + \delta. \tag{3.7}
\]

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

\[
    \|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - \frac{1}{I} \sum_{i \in \mathcal{I}} X_{n,i} + \frac{2}{I} \gamma \sum_{i \in \mathcal{I}} (f_i(x) - f_i(y_{n,i}))
    < \|x_n - x\|^2 - \frac{1}{I} (IM_1 \gamma + \delta) + M_1 \gamma
    = \|x_n - x\|^2 - \frac{\delta}{I}
    < \|x_{n_0} - x\|^2 - \frac{\delta}{I} (n + 1 - n_0).
\]

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,

\[
    \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} \|y_{n,i} - x_{n,i}\|^2 = \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} \|y_{n,i} - Q_i(y_{n,i})\|^2 \leq IM_1 \gamma.
\]

Next, we show that

\[
    \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} f_i(y_{n,i}) \leq f^*. \tag{3.8}
\]

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 \( n \geq m_0 \),

\[
    \sum_{i \in \mathcal{I}} f_i(y_{n,i}) - f^* > \zeta.
\]
Lemma 3.1 thus ensures that, for all \( n \geq m_0 \) and for all \( x^* \in X^* := \{ x^* \in X : f(x^*) = f^* = \inf_{x \in X} f(x) \} \),

\[
\|x_{n+1} - x^*\|^2 \leq \|x_n - x^*\|^2 + \frac{2}{I} \gamma \left( f^* - \sum_{i \in I} f_i(y_{n,i}) \right) < \|x_n - x^*\|^2 - \frac{2}{I} \gamma \xi
\]

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

3.2. **Diminishing step-size rule.** The following is a convergence analysis of Algorithm 1 with a diminishing step size.

**Theorem 3.2.** Suppose that (A1), (A2), and Assumption 3.1 hold and \( \text{Id} - Q_i \) is demiclosed.* Let \( (x_n)_{n \in \mathbb{N}} \) 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 \( (x_n)_{n \in \mathbb{N}} \), \( (x_{n,i})_{n \in \mathbb{N}} \), and \( (y_{n,i})_{n \in \mathbb{N}} \) (i.e., \( i \in I \)) that weakly converges to a solution of Problem 1.1. Moreover, \( (x_n)_{n \in \mathbb{N}} \), \( (x_{n,i})_{n \in \mathbb{N}} \), and \( (y_{n,i})_{n \in \mathbb{N}} \) (i.e., \( i \in I \)) strongly converge to a unique solution of Problem 1.1 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 \( x^* \in X^* \), \( n \geq m_0 \) implies \( \|x_{n+1} - x^*\| \leq \|x_n - x^*\| \), where \( X^* := \{ x^* \in X : f(x^*) = f^* = \inf_{x \in X} f(x) \} \). Then, there exists \( c := \lim_{n \to +\infty} \|x_n - x^*\| \). Let \( x^* \in X^* \) be fixed arbitrarily. Lemma 3.1, together with a discussion similar to that of (3.5), guarantees that there exists

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

such that, for all \( n \geq m_0 \),

\[
\frac{1}{I} \sum_{i \in I} \left\{ \|x_n - y_{n,i}\|^2 + \|x_{n,i} - y_{n,i}\|^2 \right\} \leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + M_2 \gamma_n. \quad (3.9)
\]

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

\[
\lim_{n \to +\infty} \|x_n - y_{n,i}\| = 0 \quad \text{and} \quad \lim_{n \to +\infty} \|x_{n,i} - y_{n,i}\| = 0 \quad (i \in I). \quad (3.10)
\]

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

\[
\frac{2}{I} \gamma_k \sum_{i \in I} f_i(y_{k,i}) - f_i(x) \leq \|x_k - x\|^2 - \|x_{k+1} - x\|^2, \quad (3.11)
\]

which implies that, for all \( n \in \mathbb{N} \) and for all \( x \in X \),

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

* See Section 4 for an example in which \( Q_i \) is quasi-firmly nonexpansive and \( \text{Id} - Q_i \) is demiclosed.
Accordingly, for all \( x \in X \),
\[
2 \sum_{k=0}^{+\infty} \gamma_N(x) < +\infty.
\] (3.12)

Here, we show that, for all \( x \in X \),
\[
\liminf_{n \to +\infty} N_n(x) \leq 0.
\] (3.13)

Assume that (3.13) 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 (3.12) and \( \sum_{n=0}^{+\infty} \gamma_n = +\infty \),
\[
+\infty = \frac{2\theta}{I} \sum_{k=m_1}^{+\infty} \gamma_k \leq \frac{2}{I} \sum_{k=m_1}^{+\infty} \gamma_N(x_0) < +\infty,
\]
which is a contradiction. Hence, (3.13) holds, i.e., for all \( x \in X \),
\[
\liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} f_i(y_{n,i}) \leq \sum_{i \in \mathcal{I}} f_i(x) =: f(x).
\] (3.14)

The definition of \( u_{n,i} \in \partial f_i(x_n) \) and the Cauchy-Schwarz inequality ensure that, for all \( i \in \mathcal{I} \) and for all \( n \in \mathbb{N} \),
\[
f_i(x_n) - f_i(y_{n,i}) \leq \langle x_n - y_{n,i}, u_{n,i} \rangle \leq \|x_n - y_{n,i}\| \|u_{n,i}\|.
\]

Proposition 2.1(iii) and the boundedness of \( (x_n)_{n \in \mathbb{N}} \) (see also Lemma 3.2) guarantee that there exists \( B_1 := \max_{i \in \mathcal{I}} \sup \{ \|u_{n,i}\| : n \in \mathbb{N} \} < +\infty \) such that, for all \( n \in \mathbb{N} \),
\[
f(x_n) = \sum_{i \in \mathcal{I}} f_i(x_n) \leq B_1 \sum_{i \in \mathcal{I}} \|x_n - y_{n,i}\| + \sum_{i \in \mathcal{I}} f_i(y_{n,i}).
\]

Therefore, (3.10) and (3.14) lead to the finding that, for all \( x \in X \),
\[
\liminf_{n \to +\infty} f(x_n) \leq B_1 \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} \|x_n - y_{n,i}\| + \liminf_{n \to +\infty} \sum_{i \in \mathcal{I}} f_i(y_{n,i}) \leq f(x).
\] (3.15)

Accordingly, a subsequence \( (x_{n_k})_{k \in \mathbb{N}} \) of \( (x_n)_{n \in \mathbb{N}} \) exists such that, for all \( x \in X \),
\[
\lim f(x_{n_k}) = \liminf_{n \to +\infty} f(x_n) \leq f(x).
\] (3.16)

Since \( (x_{n_k})_{k \in \mathbb{N}} \) is bounded (see also Lemma 3.2), there exists \( (x_{n_{l_m}})_{m \in \mathbb{N}} \subset (x_{n_k})_{k \in \mathbb{N}} \) such that \( (x_{n_{l_m}})_{m \in \mathbb{N}} \) weakly converges to \( x_s \in H \). From (3.10), \( (y_{n_{l_m},i}) \ (i \in \mathcal{I}) \) weakly converges to \( x_s \). Hence, (3.10) and the demiclosedness of \( \text{Id} - Q_i \) ensure that \( x_s \in \text{Fix}(Q_i) \ (i \in \mathcal{I}) \), i.e., \( x_s \in X \). Proposition 2.3 ensures that the continuity and convexity of \( f \) (by (A2)) imply that \( f \) is weakly lower semicontinuous, which means that
\[
f(x_s) \leq \liminf_{m \to +\infty} f(x_{n_{l_m}}).
\]

Therefore, (3.16) leads to the finding that, for all \( x \in X \),
\[
f(x_s) \leq \liminf_{m \to +\infty} f(x_{n_{l_m}}) = \lim_{m \to +\infty} f(x_{n_{l_m}}) \leq f(x),
\]
that is, \( x_s \in X^* \). Let us take another subsequence \( (x_{n_{k_l}})_{k \in \mathbb{N}} \subset (x_{n_k})_{k \in \mathbb{N}} \) such that \( (x_{n_{k_l}})_{k \in \mathbb{N}} \) weakly converges to \( x_{ss} \in H \). A discussion similar to the one for obtaining \( x_s \in X^* \) guarantees that \( x_{ss} \in X^* \).
Here, it is proven that $x_* = x_{ss}$. Now, let us assume that $x_* \neq x_{ss}$. Then, the existence of $c := \lim_{n \to \infty} \| x_n - x^* \| (x^* \in X^*)$ and Proposition 2.2 imply that

$$c = \lim_{n \to \infty} \| x_n - x_* \| < \lim_{n \to \infty} \| x_n - x_{ss} \| = \lim_{n \to \infty} \| x_n - x_{ss} \| < \lim_{k \to \infty} \| x_{nk} - x_* \| = c,$$

which is a contradiction. Hence, $x_* = x_{ss}$. Accordingly, any subsequence of $(x_n)_{n \in \mathbb{N}}$ converges weakly to $x_* \in X^*$; i.e., $(x_n)_{n \in \mathbb{N}}$ converges weakly to $x_* \in X^*$. This means that $x_*$ is a weak cluster point of $(x_n)_{n \in \mathbb{N}}$ and belongs to $X^*$. A discussion similar to the one for obtaining $x_* = x_{ss}$ 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 $X^*$.

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

$$\frac{1}{I} \sum_{i \in I} \left\{ \| x^{(n)} - y^{(n),i} \|^2 + \| x^{(n),i} - y^{(n),i} \| \right\} \leq \Gamma^2(n) - \Gamma^2(n+1) + \tilde{M}_2 \gamma(n) \leq \tilde{M}_2 \gamma(n),$$

where

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

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

$$\lim_{n \to +\infty} \| x^{(n)} - y^{(n),i} \| = 0 \quad \text{and} \quad \lim_{n \to +\infty} \| x^{(n),i} - y^{(n),i} \| = 0 \quad (i \in I). \quad (3.17)$$

From (3.11), for all $n \geq m_1$,

$$\frac{2}{I} \gamma(n) N_{\tau(n)}(x_0^*) \leq \Gamma^2(n) - \Gamma^2(n+1) \leq 0,$$

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

$$\limsup_{n \to +\infty} \sum_{i \in I} f_i(y^{(n),i}) \leq f^*.$$  

An argument similar to the one for obtaining (3.15), together with (3.17), implies that

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

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

$$\limsup_{k \to +\infty} f(x_{\tau(n_k)}) \leq \limsup_{n \to +\infty} f(x_{\tau(n)}) \leq f^*.$$  

(3.18)

The boundedness of $(x_{\tau(n_k)})_{k \in \mathbb{N}}$ ensures that $(x_{\tau(n_k)})_{k \in \mathbb{N}} \subset (x_{\tau(k)})_{k \in \mathbb{N}}$ exists such that $(x_{\tau(n_k)})_{k \in \mathbb{N}}$ weakly converges to $x_* \in H$. Then, (3.17) and the demiclosedness of $\text{Id} - Q_1$ ensure that $x_* \in X$. Moreover, Proposition 2.3 and (3.18) guarantee that

$$f(x_*) \leq \liminf_{l \to +\infty} f\left(x_{\tau(n_l)}\right) \leq \limsup_{l \to +\infty} f\left(x_{\tau(n_l)}\right) \leq f^*.$$
that is, \( x_* \in X^* \). Therefore, \( (x_{(n)}^i)_{i \in \mathbb{N}} \) weakly converges to \( x_* \in X^* \). From Cases 1 and 2, there exists a subsequence of \( (x_n)_{n \in \mathbb{N}} \) that weakly converges to a point in \( X^* \).

Suppose that assumption (i) in Theorem 3.2 holds. The strong convexity of \( f := \sum_{i \in \mathcal{F}} f^{(i)} \) implies that \( X^* \) consists of one point, denoted by \( x^* \). In Case 1, the strong convexity of \( f \) guarantees that there exists \( \beta > 0 \) such that, for all \( \alpha \in (0, 1) \) and for all \( l \in \mathbb{N} \), \( (\beta/2)(1 - \alpha)\|x_{n} - x^*\|^2 \leq \alpha f(x_{n}) + (1 - \alpha)f^* - f(\alpha x_{n} + (1 - \alpha)x^*) \). Accordingly, from the existence of \( c := \lim_{n \to +\infty} \|x_{n} - x^*\| \) and (3.16), we have
\[
\frac{\beta}{2} (1 - \alpha) \lim_{l \to +\infty} \|x_{n} - x^*\|^2 \leq \alpha \lim_{l \to +\infty} (\alpha f(x_{n}) + (1 - \alpha)f^*)
+ \limsup_{l \to +\infty} (-f(\alpha x_{n} + (1 - \alpha)x^*)
\leq f^* - \liminf_{l \to +\infty} f(\alpha x_{n} + (1 - \alpha)x^*) ,
\]
which, together with the weak convergence of \( (x_{n})_{l \in \mathbb{N}} \) to \( x^* \) and Proposition 2.3, implies that
\[
\frac{\beta}{2} (1 - \alpha) \lim_{l \to +\infty} \|x_{n} - x^*\|^2 \leq f^* - f(\alpha x^* + (1 - \alpha)x^*) = 0.
\]
Hence, \( (x_{n})_{l \in \mathbb{N}} \) strongly converges to \( x^* \). Therefore, from [2, Theorem 5.11], the whole sequence \( (x_{n})_{n \in \mathbb{N}} \) strongly converges to \( x^* \). From (3.10), \( (x_{n}, i \in \mathbb{N} \) and \( (y_{n}, i \in \mathcal{F}) \) strongly converge to \( x^* \). In Case 2, the strong convexity of \( f \) leads to the deduction that, for all \( \alpha \in (0, 1) \) and for all \( l \in \mathbb{N} \),
\[
\frac{\beta}{2} (1 - \alpha) \limsup_{l \to +\infty} \|x_{(n)}^i - x^*\|^2 \leq \alpha \limsup_{l \to +\infty} f(x_{(n)}^i) + (1 - \alpha)f^*
- \liminf_{l \to +\infty} f(\alpha x_{(n)}^i + (1 - \alpha)x^*).
\]
The weak convergence of \( (x_{(n)}^i)_{i \in \mathbb{N}} \) to \( x^* \), the weakly lower semicontinuity of \( f \) (by Proposition 2.3), and (3.18) imply that
\[
\frac{\beta}{2} (1 - \alpha) \limsup_{l \to +\infty} \|x_{(n)}^i - x^*\|^2 \leq f^* - f(\alpha x^* + (1 - \alpha)x^*) = 0,
\]
which implies that \( (x_{(n)}^i)_{i \in \mathbb{N}} \) strongly converges to \( x^* \). When another subsequence \( (x_{(n_{m})}^i)_{m \in \mathbb{N}} \subset (x_{(n)}^i)_{k \in \mathbb{N}} \) can be chosen, a discussion similar to the one for showing the weak convergence of \( (x_{(n)}^i)_{i \in \mathbb{N}} \) to a point in \( X^* \) guarantees that \( (x_{(n)}^i)_{m \in \mathbb{N}} \) also weakly converges to a point in \( X^* \). Furthermore, a discussion similar to the one for showing the strong convergence of \( (x_{(n)}^i)_{i \in \mathbb{N}} \) to \( x^* \) ensures that \( (x_{(n_{m})}^i)_{m \in \mathbb{N}} \) strongly converges to the same \( x^* \). Hence, it is guaranteed that \( (x_{(n)}^i)_{i \in \mathbb{N}} \) strongly converges to \( x^* \). Since \( (x_{(n)}^i)_{k \in \mathbb{N}} \) is an arbitrary subsequence of \( (x_{(n)}^i)_{n \geq m_{1}}, (x_{(n)}^i)_{n \geq m_{1}} \) strongly converges to \( x^* \); i.e., \( \lim_{n \to +\infty} \Gamma_{(n)} = \lim_{n \to +\infty} \|x_{(n)} - x^*\| = 0 \). Accordingly, Proposition 2.4 ensures that
\[
\limsup_{n \to +\infty} \|x_{n} - x^*\| \leq \limsup_{n \to +\infty} \Gamma_{(n)+1} = 0,
\]
which implies that, in Case 2, the whole sequence \( (x_{n})_{n \in \mathbb{N}} \) converges to \( x^* \). Moreover, Lemma 3.1 and \( \lim_{n \to +\infty} y_{n} = 0 \) imply that \( \lim_{n \to +\infty} \|x_{n} - y_{n,i}\| = \lim_{n \to +\infty} \|x_{n,i} - y_{n,i}\| = 0 \) (\( i \in \mathcal{F} \)). Therefore, \( (x_{n,i})_{n \in \mathbb{N}} \) and \( (y_{n,i})_{n \in \mathcal{F}} \) converges to \( x^* \).

Suppose that assumption (ii) in Theorem 3.2 holds. Let \( x^* \in X^* \) be the unique solution to Problem 1.1. In Case 1, it is guaranteed that \( (x_{n})_{n \in \mathbb{N}} \) converges to \( x^* \in X^* \). From (3.10), \( (x_{n,i})_{n \in \mathbb{N}} \) and \( (y_{n,i})_{n \in \mathcal{F}} \) strongly converge to \( x^* \). Moreover, in Case 2, the convergence of \( (x_{(n)}^i)_{i \in \mathbb{N}} \) to \( x^* \) is guaranteed. A discussion similar to the one for showing the strong convergence of \( (x_{(n)}^i)_{n \geq m_{1}} \) to \( x^* \) ensures that
Proposition 2.4 thus guarantees that the whole sequence \((x_n)_{n \in \mathbb{N}}\) converges to \(x^* \in X^*\). Lemma 3.1 and \(\lim_{n \to +\infty} \gamma_n = 0\) imply that \(\lim_{n \to +\infty} \|x_n - y_{n,i}\| = \lim_{n \to +\infty} \|x_{n,i} - y_{n,i}\| = 0\) \((i \in \mathcal{I})\). Therefore, \((x_{n,i})_{n \in \mathbb{N}}\) and \((y_{n,i})_{n \in \mathbb{N}}\) \((i \in \mathcal{I})\) converge to \(x^*\). This completes the proof. \(\square\)

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

**Theorem 3.3.** Suppose that the assumptions in Theorem 3.1 hold and a monotone decreasing sequence \((\gamma_n)_{n \in \mathbb{N}}\) satisfies \(\lim_{n \to +\infty} \gamma_n = 0\), \(\lim_{n \to +\infty} (n \gamma_n)^{-1} = 0\), \(\sum_{n=0}^{+\infty} \gamma_n = +\infty\), and \(\lim_{n \to +\infty} n^{-1} \sum_{k=0}^{n-1} \gamma_k = 0\). Then, Algorithm 1 satisfies that, for all \(n \geq 1\),

\[
\sum_{i \in \mathcal{I}} \left( \frac{1}{n} \sum_{k=0}^{n-1} \|y_{k,i} - Q_i(y_{k,i})\|^2 \right) \leq \frac{I}{n} \|x_0 - x\|^2 + \frac{\tilde{M}_1}{n} \sum_{k=0}^{n-1} \gamma_k \quad \text{and} \quad \sum_{i \in \mathcal{I}} f_i \left( \frac{1}{n} \sum_{k=0}^{n-1} y_{k,i} \right) \leq f^* + \frac{IB}{2n \gamma_n},
\]

where \(x^*\) is a solution of Problem 1.1, \(\tilde{M}_1 := \sup \{2 \sum_{i \in \mathcal{I}} (f_i(x^*) - f_i(y_{n,i})) : n \in \mathbb{N} \} < +\infty\), and \(B := \sup \{\|x_n - x^*\|^2 : n \in \mathbb{N} \} < +\infty\).

**Proof.** Let \(x^* \in X^*\). Lemma 3.1 implies that, for all \(n \geq 1\),

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

which in turn implies that

\[
\sum_{i \in \mathcal{I}} \left( \frac{1}{n} \sum_{k=0}^{n-1} \|y_{k,i} - x_{k,i}\|^2 \right) \leq \frac{1}{n} \sum_{i \in \mathcal{I}} \sum_{k=0}^{n-1} \left\{ \|y_{k,i} - x_{k,i}\|^2 + \|x_{k,i} - y_{k,i}\|^2 \right\} \leq \frac{I}{n} \|x_0 - x\|^2 + \frac{\tilde{M}_1}{n} \sum_{k=0}^{n-1} \gamma_k.
\]

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

\[
\sum_{i \in \mathcal{I}} f_i(y_{k,i}) - f^* \leq \frac{I}{2 \gamma_k} \left\{ \|x_k - x^*\|^2 \right\}.
\]

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

\[
\frac{1}{n} \sum_{k=0}^{n-1} f_i(y_{k,i}) - f^* \leq \frac{I}{2n} \sum_{k=0}^{n-1} \frac{1}{\gamma_k} \left\{ \|x_k - x^*\|^2 \right\}.
\]

The definition of \(X_n\) means that

\[
X_n = \frac{\|x_0 - x^*\|}{\gamma_0} + \sum_{k=0}^{n-1} \left\{ \frac{\|x_k - x^*\|^2}{\gamma_k} - \frac{\|x_k - x^*\|^2}{\gamma_{k-1}} \right\} = \frac{\|x_n - x^*\|^2}{\gamma_{n-1}},
\]

which, together with \(\gamma_n \leq \gamma_{n-1} \quad (n \geq 1)\) and \(B := \sup \{\|x_n - x^*\|^2 : n \in \mathbb{N} \} < +\infty\) (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 \(n \geq 1\),

\[
\sum_{i \in \mathcal{I}} f_i \left( \frac{1}{n} \sum_{k=0}^{n-1} y_{k,i} \right) - f^* \leq \frac{IB}{2n \gamma_n},
\]

which completes the proof. \(\square\)
Let us consider the rate of convergence of Algorithm 1 with \( \gamma_n := n^{-1/2} \) \((n \geq 1)\). The step size \((\gamma_n)_{n \geq 1}\) 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

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

which implies that \(\lim_{n \to +\infty} n^{-1} \sum_{k=0}^{n-1} k = 0\). Theorem 3.3 indicates that Algorithm 1 with \( \gamma_n := n^{-1/2} \) satisfies that, for all \( n \geq 1 \),

\[
\sum_{i \in \mathcal{I}} \left( \frac{1}{n} \sum_{k=0}^{n-1} \|y_{k,i} - Q_i(y_{k,i})\|^2 \right) = O\left( \sqrt{\frac{1 + \ln n}{n}} \right) \quad \text{and} \quad \sum_{i \in \mathcal{I}} f_i \left( \frac{1}{n} \sum_{k=0}^{n-1} y_{k,i} \right) \leq f^* + \frac{IB}{2\sqrt{n}},
\]

where \(O\) stands for the Landau notation (see [10] 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 subgradient method (PSM) [7, Algorithm 3.1] (see (3.1)) and incremental subgradient method (ISM) [7, Algorithm 4.1] for the following problem (see also [7, Problem 5.1]): Let \(a_{i,j} > 0, b_{i,j}, d_i \in \mathbb{R} \) \((i \in \mathcal{I}, j = 1, 2, \ldots, N)\), and \(c_i := (c_{i,j})_{j=1}^{N} \in \mathbb{R}^{N} \) \((i \in \mathcal{I})\) with \(c_{i,j} > 0\). Then,

\[
\text{minimize } f(x) := \sum_{i \in \mathcal{I}} f_i(x) \text{ subject to } x \in X := \bigcap_{i \in \mathcal{I}} \text{Fix}(Q_i) = \bigcap_{i \in \mathcal{I}} \text{lev}_{\leq 0} g_i,
\]

where \(f_i: \mathbb{R}^N \to \mathbb{R} \) and \(Q_i: \mathbb{R}^N \to \mathbb{R}^N \) are defined for all \( x := (x_j)_{j=1}^{N} \in \mathbb{R}^N \) by

\[
f_i(x) := \sum_{j=1}^{N} a_{i,j} |x_j - b_{i,j}| \quad \text{and} \quad Q_i(x) := \begin{cases} x - \frac{g_i(x)}{\|z_i(x)\|^2} z_i(x) & \text{if } g_i(x) > 0, \\ x & \text{if } x \in \text{lev}_{\leq 0} g_i := \{ x \in \mathbb{R}^N : g_i(x) \leq 0 \}, \end{cases}
\]

\(g_i: \mathbb{R}^N \to \mathbb{R} \) is 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 \(z_i(x)\) is any vector in \(\partial g_i(x)\). The above mapping \(Q_i\) is called the subgradient projection related to \(g_i\). \(Q_i\) satisfies quasi-firm nonexpansivity, and \(\text{Id} - Q_i\) satisfies the demiclosedness condition [1, Lemma 3.1].

The experiment was conducted on a MacBook Air (13-inch, 2017) with a 1.8 GHz Intel (R) Core (TM) i5 CPU processor, 8 GB, 1600 MHz DDR3 memory, and Mac OS Catalina (Version 10.15) operating system. PSM, ISM, and Algorithm 1 were written in Python 3.7.4 with the NumPy 1.17.2 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:

- Constant step sizes: \(\gamma_n := 10^{-1}, 10^{-3}\),
- Diminishing step sizes: \(\gamma_n := \frac{10^{-1}}{n+1}, \frac{10^{-3}}{n+1}\).
The performance measures were as follows: for \( n \in \mathbb{N} \),
\[
F_n := \frac{1}{10} \sum_{s=1}^{10} \sum_{l \in S} f_l(x_n(s)) \quad \text{and} \quad D_n := \frac{1}{10} \sum_{s=1}^{10} \sum_{l \in S} \|x_n(s) - Q_l(x_n(s))\|,
\]
where \((x_n(s))_{n \in \mathbb{N}}\) is the sequence generated by each of the three algorithms with the 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 algorithms converge to a point in \( X \).

Figure 1 shows that the algorithms with \( \gamma_n = \lambda_n = 10^{-1} \) did not converge to a point in \( X \). Figure 2 indicates that, although the values of \( D_n \) generated by the algorithms with \( \gamma_n = \lambda_n = 10^{-3} \) were less than those generated by the algorithms with \( \gamma_n = \lambda_n = 10^{-1} \), the algorithms with \( \gamma_n = \lambda_n = 10^{-3} \) did not converge to a point in \( X \). These results imply that it would be difficult to set an appropriate constant step size in advance.

**Figure 1.** Behaviors of \( F_n \) and \( D_n \) for Algorithm 1, PSM, and ISM with \( \gamma_n = \lambda_n = 10^{-1} \)

**Figure 2.** Behaviors of \( F_n \) and \( D_n \) for Algorithm 1, PSM, and ISM with \( \gamma_n = \lambda_n = 10^{-3} \)

Meanwhile, Figures 3 and 4 show that Algorithm 1 with diminishing step sizes \( \gamma_n = 10^{-1}/(n + 1), 10^{-3}/(n + 1) \) converged to a point in \( X \), as guaranteed by Theorem 3.2. These figures also show
that $F_n$ remains stable. Accordingly, from Theorem 3.2, Algorithm 1 converged to a solution of problem (4.1). Figures 3 and 4 also indicate that Algorithm 1 performs comparably to PSM and ISM.

![Algorithm 1, PSM, and ISM with $\gamma_n = \lambda_n = 10^{-1}/(n+1)$](image1)

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

![Algorithm 1, PSM, and ISM with $\gamma_n = \lambda_n = 10^{-3}/(n+1)$](image2)

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

5. **Conclusion**

This paper presented a parallel 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. Numerical comparisons showed that the performance of the algorithm is almost the same as those of the existing methods.

**Acknowledgments**

The authors would like to thank Professor Xiaolong Qin for giving us a chance to submit our paper to this journal. This work was supported by JSPS KAKENHI Grant Numbers, JP18K11184 and JP17J09220.
REFERENCES


