SIAM J. OPTIM.
Vol. 0, No. 0, pp. 000–000

ACCELERATION METHOD COMBINING BROADCAST AND INCREMENTAL DISTRIBUTED OPTIMIZATION ALGORITHMS*

HIDEAKI IIDUKA† AND KAZUHIRO HISHINUMA†

Abstract. This paper considers a networked system consisting of an operator, which manages the system, and a finite number of subnetworks with all users, and studies the problem of minimizing the sum of the operator's and all users' objective functions over the intersection of the operator's and all users' constraint sets. When users in each subnetwork can communicate with each other, they can implement an incremental subgradient method that uses the transmitted information from their neighbor users. Since the operator can communicate with users in the subnetworks, it can implement a broadcast distributed algorithm that uses all available information in the subnetworks. We present an iterative method combining broadcast and incremental distributed optimization algorithms. Our method has faster convergence and a wider range of application compared with conventional distributed algorithms. We also prove that under certain assumptions our method converges to the solution to the problem in the sense of the strong topology of a Hilbert space. Moreover, we numerically compare our method with the conventional distributed algorithms in the case of a data storage system. The numerical results demonstrate the effectiveness and fast convergence of our method.

Key words. acceleration method, broadcast optimization algorithm, distributed optimization, fixed point, incremental subgradient method, nonexpansive mapping

AMS subject classifications. 49M37, 65K05, 90C25, 90C90

DOI. 10.1137/130939560

1. Introduction.

In this paper, we consider a networked system consisting of an operator, which manages the system, and a finite number of participating users. The main objective of this paper is to develop a novel distributed computing approach for resolving the following minimization problem (see [6, 8, 9, 11, 12, 13, 14, 16, 18, 19, 23, 24] and references therein for applications of problem (1.1)) on a Hilbert space H:

$$\text{Minimize} \sum_{i \in \{0\} \cup \mathcal{I}} f^{(i)}(x) \text{ subject to } x \in \bigcap_{i \in \{0\} \cup \mathcal{I}} \mathrm{Fix}\big(T^{(i)}\big), \tag{1.1}$$

where the operator has a convex function, $f^{(0)} \colon H \to \mathbb{R}$, and a nonexpansive mapping, $T^{(0)} \colon H \to H$; user i ($i \in \mathcal{I} := \{1, 2, \ldots, I\}$) has a convex function, $f^{(i)} \colon H \to \mathbb{R}$, and a nonexpansive mapping, $T^{(i)} \colon H \to H$; and $\mathrm{Fix}(T^{(i)})$ stands for the fixed point set of $T^{(i)}$ (i.e., $\mathrm{Fix}(T^{(i)}) := \{x \in H \colon T^{(i)}(x) = x\}$ $(i \in \{0\} \cup \mathcal{I})$).

Problem (1.1) enables us to discuss constrained optimization problems in which the explicit form of the metric projection onto the constraint set is not always known; i.e., the projection cannot be calculated explicitly. To describe this advantage more explicitly, let us consider the problem of minimizing a convex function f over the set of all minimizers of a convex function g with the Lipschitz gradient ∇g. We define $T := \mathrm{Id} - \alpha \nabla g$, where Id is the identity mapping on H, $\alpha \in (0, 2/L]$, and L (> 0) stands for the Lipschitz constant of ∇g. Accordingly, T is nonexpansive and

*Received by the editors October 2, 2013; accepted for publication (in revised form) July 21, 2014; published electronically DATE. This work was supported by the Japan Society for the Promotion of Science through a Grant-in-Aid for Scientific Research (C) (23500090).
http://www.siam.org/journals/siopt/x-x/93956.html
†Department of Computer Science, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214-8571, Japan (iiduka@cs.meiji.ac.jp, kaz@cs.meiji.ac.jp).

$\mathrm{Fix}(T) = \{x \in H : g(x) = \min_{y \in H} g(y)\}$ [10, Proposition 2.3]. Hence, problem (1.1) includes the minimization problem over $\{x \in H : g(x) = \min_{y \in H} g(y)\}$ onto which the projection cannot be calculated explicitly.

Next, let us consider the problem of minimizing a convex function f over the set of zeros of a set-valued, maximal monotone operator A. Since the resolvent of A, denoted by J_A, is firmly nonexpansive [2, Corollary 23.10] and $\mathrm{Fix}(J_A)$ coincides with the set of zeros of A, problem (1.1) includes the minimization problem over this set onto which the projection cannot be calculated explicitly. Section 4 will show that one application of problem (1.1) is storage allocation [19] in a peer-to-peer data system.

Distributed optimization methods can be implemented if all participants cooperate in the system even if each participant has its own private objective function and constraint set, and they enable each participant to solve problem (1.1) without using the private information of the other participants. Here, we describe the two distributed optimization algorithms that are useful for solving problem (1.1).

(I) Problem (1.1) can be solved under the assumption that the operator can communicate with all users because the operator manages the whole system. Accordingly, the operator can implement *broadcast optimization algorithms* (see [7, 9, 12, 22, 23] and references therein) which are given as follows: Suppose that the operator (user 0) has $x_n \in H$. Then it computes $x_n^{(0)} \in H$ by using x_n and its own private information $T^{(0)}$ and $f^{(0)}$; i.e., $x_n^{(0)} = x_n^{(0)}(x_n, T^{(0)}, f^{(0)})$. Moreover, user i ($i \in \mathcal{I}$) computes $x_n^{(i)} \in H$ by using the information x_n transmitted from the operator and user i's private information $T^{(i)}$ and $f^{(i)}$; i.e., $x_n^{(i)} = x_n^{(i)}(x_n, T^{(i)}, f^{(i)})$ ($i \in \mathcal{I}$), and it transmits $x_n^{(i)}$ to the operator. The operator computes $x_{n+1} \in H$ by using all $x_n^{(i)}$ ($i \in \{0\} \cup \mathcal{I}$); i.e., $x_{n+1} = x_{n+1}(x_n^{(0)}, x_n^{(1)}, \ldots, x_n^{(I)})$. Figure 1 illustrates the concept of broadcast distributed optimization algorithms when $I = 9$. It would be natural that user i ($i \in \{0\} \cup \mathcal{I}$) would try to choose $x_n^{(i)}$ so as to minimize $f^{(i)}$ over $\mathrm{Fix}(T^{(i)})$ as much as possible. When the operator uses the transmitted information from all users equally, $x_n^{(i)}$ ($i \in \{0\} \cup \mathcal{I}$) and x_{n+1} are defined as follows:

$$(1.2) \quad x_n^{(i)} := T^{(i)} \left(\mathrm{Id} - \lambda_n \nabla f^{(i)} \right) (x_n) \ (i \in \{0\} \cup \mathcal{I}), \quad x_{n+1} := \frac{1}{I+1} \sum_{i \in \{0\} \cup \mathcal{I}} x_n^{(i)},$$

where $(\lambda_n)_{n \in \mathbb{N}}$ is a step-size sequence.

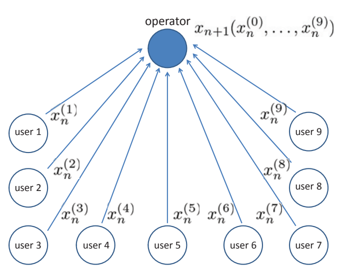


Fig. 1. *Broadcast optimization algorithm.*

(II) When a user communicates with a neighbor user via the network, it uses only its own private information and the information transmitted from the neighbor user.

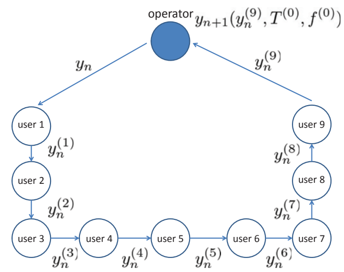


FIG. 2. *Incremental optimization algorithm.*

This enables each user to implement *incremental optimization algorithms* (see [4, Subchapter 8.2], [5, 12, 15, 17, 20] and references therein): Suppose that the operator has $y_n \in H$ with $y_n^{(0)} = y_n$ and user i ($i \in \mathcal{I}$) has the information $y_n^{(i-1)}$ transmitted from user $(i-1)$, which is one of user i's neighbors. Then, user i computes $y_n^{(i)} \in H$ by using $y_n^{(i-1)}$ and user i's private information $T^{(i)}$ and $f^{(i)}$; i.e., $y_n^{(i)} = y_n^{(i)}(y_n^{(i-1)}, T^{(i)}, f^{(i)})$ ($i \in \mathcal{I}$). The operator computes $y_{n+1} \in H$ as $y_{n+1} = y_{n+1}(y_n^{(I)}, T^{(0)}, f^{(0)})$. Figure 2 illustrates the concept of incremental distributed optimization algorithms when $I = 9$. The same discussion as in (1.2) leads us to

$$\begin{aligned} y_n^{(i)} &:= T^{(i)}\left(y_n^{(i-1)} - \lambda_n \nabla f^{(i)}\left(y_n^{(i-1)}\right)\right) \ (i \in \mathcal{I}), \\ y_{n+1} &:= T^{(0)}\left(y_n^{(I)} - \lambda_n \nabla f^{(0)}\left(y_n^{(I)}\right)\right), \end{aligned} \tag{1.3}$$

where one assumes that $f^{(i)}$ ($i \in \{0\} \cup \mathcal{I}$) is convex and Fréchet differentiable.[1]

Algorithms (1.2) and (1.3) suffer from two problems.

(i) The broadcast optimization algorithm (1.2) updates the next iteration x_{n+1} only after the operator has all the transmitted information $x_n^{(i)}$ ($i \in \{0\} \cup \mathcal{I}$). Hence, there is a possibility that algorithm (1.2) will be time-consuming when it is applied to large-scale networked systems with many users.

(ii) The incremental optimization algorithm (1.3) needs to go through all users to update the next iteration y_{n+1}. However, it would be physically difficult to go through all users in a large-scale complex network.

This paragraph describes a way of resolving issues (i) and (ii). Since user i ($i \in \mathcal{I}$) can communicate with its neighbor users, we can construct a subnetwork that consists of user i and its neighbors. This implies that the network is divided into S subnetworks (Figure 3 illustrates the network when $I = 9$ and $S = 3$). In this setting, each user in subnetwork s ($s \in \mathcal{S} := \{1, 2, \ldots, S\}$) can implement the incremental optimization algorithm. For each $s \in \mathcal{S}$, let $x_n^{(s)}$ be the point generated by $x_n = y_n^{(0)}$ and the incremental optimization algorithm. Then the operator can get all $x_n^{(s)}$ and compute $x_{n+1} = x_{n+1}(x_n^{(0)}, x_n^{(1)}, \ldots, x_n^{(S)}) = 1/(S+1)(x_n^{(0)} + \sum_{s \in \mathcal{S}} x_n^{(s)})$. This means the broadcast optimization algorithm can be implemented by the operator and subnetworks. Therefore, we can devise an algorithm combining the ideas of broadcast and incremental optimization algorithms. To update x_n to x_{n+1}, this algorithm needs

[1] The well-known incremental subgradient method [4, eqs. (8.9), (8.10), and (8.11)], [20] can be used when $f^{(i)}$ ($i \in \{0\} \cup \mathcal{I}$) is convex and nondifferentiable.

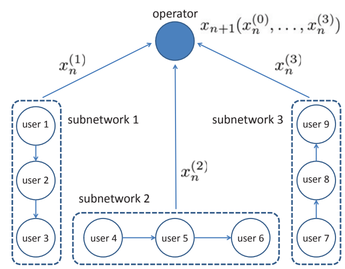


FIG. 3. *Proposed algorithm.*

to use S ($\leq I$) transmitted points from the subnetworks, while the broadcast optimization algorithm needs to use I transmitted points from all users. Hence, one can expect that it would resolve issue (i) (see section 4 for the fast convergence of the proposed algorithm). Moreover, it can resolve issue (ii) because it does not need to go through all users. Since the proposed algorithm coincides with the broadcast optimization algorithm when $S = I$, it can be considered to be a generalization of that algorithm. The proposed algorithm when $S = 1$ is similar to the incremental optimization algorithm implemented by all users.

The organization of the paper is as follows. Section 2 gives the necessary mathematical preliminaries. Section 3 presents an algorithm combining broadcast and incremental distributed optimization algorithms for solving problem (1.1) and its convergence analysis. Applications of this algorithm to storage allocation are described in section 4. Section 5 concludes by summarizing the key points of the paper.

2. Mathematical preliminaries.

2.1. Notation. Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Moreover, let Id be the identity mapping on H, and let $\mathbb{N}$ denote the set of all positive integers including zero. We define $\mathbb{R}_+ := \{x \in \mathbb{R} \colon x \geq 0\}$.

We denote the fixed point set of a mapping $T \colon H \to H$ by $\mathrm{Fix}(T)$; i.e., $\mathrm{Fix}(T) := \{x \in H \colon T(x) = x\}$. The metric projection [2, Subchapter 4.2, Chapter 28] onto a nonempty, closed convex set C ($\subset H$) is denoted by P_C.

Let $\mathcal{I} := \{1, 2, \ldots, I\}$ be the set of users participating in the system, and let $\mathcal{S} := \{1, 2, \ldots, S\}$ be the set of subnetworks in the system. Let $\mathcal{I}_s$ denote the set of users in subnetwork s ($s \in \mathcal{S}$), and put $I_s := |\mathcal{I}_s|$ ($s \in \mathcal{S}$), where $|\mathcal{I}_s|$ stands for the number of elements in $\mathcal{I}_s$. Accordingly, we have that $\mathcal{I} = \bigcup_{s \in \mathcal{S}} \mathcal{I}_s$ and $I = \sum_{s \in \mathcal{S}} I_s$. User i ($i \in \{0\} \cup \mathcal{I}$) has its own private Fréchet differentiable, convex objective function, denoted by $f^{(i)} \colon H \to \mathbb{R}$, and closed convex constraint set, denoted by $C^{(i)}$ ($\subset H$), where user 0 stands for the operator.

2.2. Definitions and propositions. Let $f \colon H \to \mathbb{R}$ be Fréchet differentiable. Moreover, let $\nabla f \colon H \to H$ denote the gradient of f, and let $\alpha, L > 0$. ∇f is said to be α-*strongly monotone* [2, Definition 22.1(iv)] if $\langle x - y, \nabla f(x) - \nabla f(y) \rangle \geq \alpha \|x - y\|^2$ ($x, y \in H$). Suppose that f is α-strongly convex [2, Definition 10.5]; i.e., $f(\lambda x + (1 - \lambda) y) + (1/2)\alpha\lambda(1 - \lambda)\|x - y\|^2 \leq \lambda f(x) + (1 - \lambda) f(y)$ ($x, y \in H, \lambda \in [0, 1]$). Then, ∇f satisfies the strong monotonicity condition [2, Example 22.3]. ∇f is said to be

L-Lipschitz continuous [2, Definition 1.46] if $\|\nabla f(x) - \nabla f(y)\| \leq L\|x-y\|$ $(x, y \in H)$.

Proposition 2.1 (see [24, Lemma 3.1]). *Suppose that* $f\colon H \to \mathbb{R}$ *is Fréchet differentiable,* $\nabla f\colon H \to H$ *is* α*-strongly monotone and* L*-Lipschitz continuous, and* $\mu \in (0, 2\alpha/L^2)$. *Define* $T := \mathrm{Id} - \mu\lambda\nabla f$, *where* $\lambda \in [0,1]$. *Then* $\|T(x) - T(y)\| \leq (1-\tau\lambda)\|x-y\|$ $(x, y \in H)$, *where* $\tau := 1 - \sqrt{1-\mu(2\alpha - \mu L^2)} \in (0,1]$.

Proposition 2.2 (see [24, Proposition 2.7]). *Assume that* C $(\subset H)$ *is nonempty, closed, and convex,* $f\colon H \to \mathbb{R}$ *is Fréchet differentiable, and* $\nabla f\colon H \to H$ *is* α*-strongly monotone and* L*-Lipschitz continuous. Then there exists a unique minimizer of* f *over* C.

Proposition 2.3 (see [2, Proposition 17.10]). *Suppose that* $f\colon H \to \mathbb{R}$ *is Gâteaux differentiable and convex, and* $x \in H$. *Then* $f(y) \geq f(x) + \langle y - x, \nabla f(x)\rangle$ $(y \in H)$.

A mapping, $T\colon H \to H$, is said to be *nonexpansive* [2, Definition 4.1(ii)] if $\|T(x) - T(y)\| \leq \|x-y\|$ $(x, y \in H)$. T is referred to as a *firmly nonexpansive* mapping [2, Definition 4.1(i)] if $\|T(x) - T(y)\|^2 + \|(\mathrm{Id}-T)(x) - (\mathrm{Id}-T)(y)\|^2 \leq \|x-y\|^2$ $(x, y \in H)$. It is obvious that firm nonexpansivity means nonexpansivity.

Proposition 2.4 (see [2, Corollary 4.15, Remark 4.24(iii), Proposition 4.8]). *The following* (i)–(iii) *hold:*

(i) *Let* $T\colon H \to H$ *be nonexpansive. Then* $\mathrm{Fix}(T)$ *is closed and convex.*
(ii) $(1/2)(\mathrm{Id} + T)$ *is firmly nonexpansive when* $T\colon H \to H$ *is nonexpansive.*
(iii) *Let* C $(\subset H)$ *be nonempty, closed, and convex. Then* P_C *is firmly nonexpansive.*

The *resolvent* [2, Definition 23.1] of $A\colon H \to 2^H$ is defined by $J_A := (\mathrm{Id} + A)^{-1}$. A monotone operator $A\colon H \to 2^H$ is referred to as a *maximal monotone* operator [2, Definition 20.20] if there exists no monotone operator $B\colon H \to 2^H$ such that $\mathrm{gra}B := \{(x,u) \in H \times H \colon u \in B(x)\}$ properly contains $\mathrm{gra}A$.

Proposition 2.5 (see [2, Proposition 23.38, Corollary 23.8]). *The following* (i)–(ii) *hold:*

(i) *Let* $A\colon H \to 2^H$ *be monotone. Then* $\mathrm{Fix}(J_A) = \mathrm{zer}A := \{x \in H \colon 0 \in A(x)\}$.
(ii) $T\colon H \to H$ *is firmly nonexpansive if and only if it is the resolvent of a maximal monotone operator* $A\colon H \to 2^H$.

The following propositions will be also used to prove the main theorem.

Proposition 2.6 (see [3, Lemma 1.2]). *Assume that* $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+$ *satisfies* $a_{n+1} \leq (1-\alpha_n)a_n + \alpha_n\beta_n$ $(n \in \mathbb{N})$, *where* $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1]$ *and* $(\beta_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ *with* $\sum_{n=1}^{\infty} \alpha_n = \infty$ *and* $\limsup_{n\to\infty} \beta_n \leq 0$. *Then* $\lim_{n\to\infty} a_n = 0$.

Proposition 2.7 (see [21]). *Suppose that* $(x_n)_{n\in\mathbb{N}}$ $(\subset H)$ *weakly converges to* $x \in H$ *and* $y \in H$ *with* $y \neq x$. *Then* $\liminf_{n\to\infty} \|x_n - x\| < \liminf_{n\to\infty} \|x_n - y\|$.

2.3. Main problem. This paper deals with the following problem.

Problem 2.1. *Assume that*

(A1) $T^{(i)}\colon H \to H$ $(i \in \{0\} \cup \mathcal{I})$ *is firmly nonexpansive with* $\mathrm{Fix}(T^{(i)}) = C^{(i)}$, *and*
(A2) $\nabla f^{(i)}\colon H \to H$ $(i \in \{0\} \cup \mathcal{I})$ *is* $\alpha^{(i)}$*-strongly monotone and* $L^{(i)}$*-Lipschitz continuous.*

The main objective of this paper is to

$$\text{minimize} \sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}(x) \text{ subject to } x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}\left(T^{(i)}\right),$$

where one assumes that $\bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)}) \neq \emptyset$.

Assumption (A2) implies that $\nabla(\sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)})$ $(= \sum_{i\in\{0\}\cup\mathcal{I}} \nabla f^{(i)})$ is strongly monotone and Lipschitz continuous. Moreover, Proposition 2.4(i) and (A1) mean that

$\bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ is closed and convex. Hence, we can see from Proposition 2.2 that Problem 2.1 has a unique solution. Section 4 will show an application of Problem 2.1.

Proposition 2.5 implies that $T^{(i)}$ $(i \in \{0\}\cup\mathcal{I})$ in (A1) is the resolvent, denoted by $J_{A^{(i)}}$, of a maximal monotone operator $A^{(i)} \colon H \to 2^H$ $(i \in \{0\}\cup\mathcal{I})$ and $\mathrm{Fix}(T^{(i)}) = \mathrm{Fix}(J_{A^{(i)}}) = \mathrm{zer}A^{(i)}$ $(i \in \{0\}\cup\mathcal{I})$. Accordingly, Problem 2.1 is equivalent to the following problem: Given a maximal monotone operator $A^{(i)} \colon H \to 2^H$ $(i \in \{0\}\cup\mathcal{I})$ with $\mathrm{zer}A^{(i)} = C^{(i)}$ and a convex functional $f^{(i)}$ satisfying (A2),

$$\text{(2.1)} \quad \text{minimize} \sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}(x) \text{ subject to } x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{zer}A^{(i)} = \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}\left(J_{A^{(i)}}\right).$$

3. Proposed algorithm and its convergence analysis. To begin with, let us define the sets that are needed to describe our algorithm. Choose $s \in \mathcal{S}$ arbitrarily, and define $\mathcal{I}_s^{(j)}$ $(j = 1, 2, \ldots, I_s)$ as follows: First choose $i_s^{(1)} \in \mathcal{I}_s^{(1)} := \mathcal{I}_s$ randomly and set $\mathcal{I}_s^{(2)} := \mathcal{I}_s^{(1)} \backslash \{i_s^{(1)}\}$. Next, choose $i_s^{(2)} \in \mathcal{I}_s^{(2)}$ randomly. For $j = 3, 4, \ldots, I_s$, define $\mathcal{I}_s^{(j)} := \mathcal{I}_s^{(j-1)} \backslash \{i_s^{(j-1)}\}$ and choose $i_s^{(j)} \in \mathcal{I}_s^{(j)}$ randomly. The following algorithm combines the incremental and broadcast optimization algorithms. (Figure 3 illustrates the concept of Algorithm 3.1 when $I = 9$, $S = 3$, $I_s = 3$ $(s = 1, 2, 3)$, and $i_1^{(j)} = j$, $i_2^{(j)} = j + 3$, $i_3^{(j)} = j + 6$ $(j = 1, 2, 3)$.)

ALGORITHM 3.1.

Step 0. The operator and all users set $(\lambda_n)_{n\in\mathbb{N}}$. The operator chooses $x_0 \in H$ arbitrarily and transmits it to user $i_s^{(1)}$ $(s \in \mathcal{S})$. User $i_s^{(1)}$ $(s \in \mathcal{S})$ sets $x_0^{(i_s^{(0)})} := x_0$.

Step 1. Given x_n, the operator (user 0) computes $x_n^{(0)} \in H$:

$$x_n^{(0)} := T^{(0)}\left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right).$$

Given $x_n^{(i_s^{(0)})} = x_n$, user $i_s^{(j)}$ $(s \in \mathcal{S},\ j = 1, 2, \ldots, I_s)$ computes $x_n^{(i_s^{(j)})} \in H$ cyclically:

$$x_n^{\left(i_s^{(j)}\right)} := T^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)} - \lambda_n \nabla f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right) \ (j = 1, 2, \ldots, I_s).$$

User $i_s^{(I_s)}$ $(s \in \mathcal{S})$ transmits $x_n^{(i_s^{(I_s)})}$ to user $i_s^{(1)}$, and user $i_s^{(1)}$ $(s \in \mathcal{S})$ defines

$$x_n^{(s)} := x_n^{\left(i_s^{(I_s)}\right)}$$

and transmits $x_n^{(s)}$ to the operator.

Step 2. The operator computes $x_{n+1} \in H$ as

$$x_{n+1} := \frac{1}{S+1}\left(x_n^{(0)} + \sum_{s\in\mathcal{S}} x_n^{(s)}\right)$$

and transmits $x_{n+1}^{(i_s^{(0)})} := x_{n+1}$ to user $i_s^{(1)}$ $(s \in \mathcal{S})$. Put $n := n + 1$, and go to Step 1.

The convergence of Algorithm 3.1 depends on the following assumption.

ASSUMPTION 3.1. *The operator and all users have the common knowledge* $(\lambda_n)_{n\in\mathbb{N}}$ $(\subset (0,1))$ *satisfying*

$$\text{(C1)} \lim_{n\to\infty} \lambda_n = 0, \quad \text{(C2)} \sum_{n=0}^{\infty} \lambda_n = \infty, \quad \text{(C3)} \lim_{n\to\infty} \frac{1}{\lambda_{n+1}} \left| \frac{1}{\lambda_{n+1}} - \frac{1}{\lambda_n} \right| = 0$$

before they perform Algorithm 3.1.

An example of $(\lambda_n)_{n\in\mathbb{N}}$ satisfying (C1)–(C3) is $\lambda_n := 1/(n+1)^a$ $(n \in \mathbb{N})$, where $a \in (0, 1/2)$. When the operator sets $(\lambda_n)_{n\in\mathbb{N}}$ satisfying (C1)–(C3), the operator can transmit the $(\lambda_n)_{n\in\mathbb{N}}$ to user $i_s^{(1)}$ $(s \in \mathcal{S})$ because the operator manages the whole system. Moreover, user $i_s^{(1)}$ can transmit the $(\lambda_n)_{n\in\mathbb{N}}$ to user $i_s^{(j)}$ $(j = 2, 3, \ldots, I_s)$ because user $i_s^{(j)}$ $(j = 1, 2, \ldots, I_s)$ can communicate with users in subnetwork s. Accordingly, Assumption 3.1 holds when the operator can set $(\lambda_n)_{n\in\mathbb{N}}$ satisfying (C1)–(C3).

Let us provide the property of $\mathrm{Id} - \lambda_n \nabla f^{(i)}$ $(i \in \{0\} \cup \mathcal{I}, n \in \mathbb{N})$ by using Proposition 2.1. Let $\mu^\star := \min_{i\in\{0\}\cup\mathcal{I}} 2\alpha^{(i)}/{L^{(i)}}^2$ and $(\lambda_n)_{n\in\mathbb{N}}$ $(\subset (0,1))$ satisfy (C1)–(C3), and choose $i \in \{0\} \cup \mathcal{I}$ and $x, y \in H$ arbitrarily. Proposition 2.1 indicates that, if $(\lambda_n)_{n\in\mathbb{N}} \subset (0, \mu^\star)$,

$$\begin{aligned}\left\| \left(\mathrm{Id} - \lambda_n \nabla f^{(i)}\right)(x) - \left(\mathrm{Id} - \lambda_n \nabla f^{(i)}\right)(y) \right\| &\leq \left(1 - \tau^{(i)}\lambda_n\right) \|x - y\| \\ &\leq (1 - \tau\lambda_n) \|x - y\| \end{aligned} \tag{3.1}$$

for all $n \in \mathbb{N}$, where $\tau^{(i)} := 1 - \sqrt{1 - \mu^\star(2\alpha^{(i)} - \mu^\star {L^{(i)}}^2)}$ $(i \in \{0\} \cup \mathcal{I})$ and $\tau := \min_{i\in\{0\}\cup\mathcal{I}} \tau^{(i)}$. Since (C1) ensures that $n_0 \in \mathbb{N}$ exists such that $(\lambda_n)_{n\geq n_0} \subset (0, \mu^\star)$, (3.1) is always true for all $n \geq n_0$. Therefore, we may assume without loss of generality that (3.1) holds for all $n \in \mathbb{N}$.

Let us perform a convergence analysis on Algorithm 3.1.

THEOREM 3.1. *Under Assumptions* (A1), (A2), *and* 3.1*, the sequences* $(x_n)_{n\in\mathbb{N}}$ *and* $(x_n^{(i_s^{(j)})})_{n\in\mathbb{N}}$ *$(s \in \mathcal{S},\ j = 1, 2, \ldots, I_s)$ generated by Algorithm* 3.1 *converge strongly to the solution to Problem* 2.1.

The discussion in subsection 2.3 and Theorem 3.1 describe that, under the assumptions in Theorem 3.1, $(x_n)_{n\in\mathbb{N}}$ and $(x_n^{(i_s^{(j)})})_{n\in\mathbb{N}}$ $(s \in \mathcal{S},\ j = 1, 2, \ldots, I_s)$ generated by Algorithm 3.1 when $T^{(i)} = J_{A^{(i)}}$ $(i \in \{0\} \cup \mathcal{I})$ converge strongly to the solution to problem (2.1).

3.1. Proof of Theorem 3.1.

We first prove the following lemma.

LEMMA 3.1. *The sequences* $(x_n)_{n\in\mathbb{N}}$ *and* $(x_n^{(i_s^{(j)})})_{n\in\mathbb{N}}$ *$(s \in \mathcal{S},\ j = 1, 2, \ldots, I_s)$ generated by Algorithm* 3.1 *are bounded.*

Proof. Choose $x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ and $n \in \mathbb{N}$ arbitrarily. Put $M_1 := \max\{\|\nabla f^{(0)}(x)\|, \max_{s\in\mathcal{S}, j=1,2,\ldots,I_s} \|\nabla f^{(i_s^{(j)})}(x)\|\}$. We have from the nonexpansivity of $T^{(0)}$, the triangle inequality, and (3.1) that

$$\begin{aligned}\left\| x_n^{(0)} - x \right\| &= \left\| T^{(0)}\left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - T^{(0)}(x) \right\| \\ &\leq \left\| \left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - x \right\| \\ &\leq \left\| \left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - \left(x - \lambda_n \nabla f^{(0)}(x)\right) \right\| + M_1\lambda_n \\ &\leq (1 - \tau\lambda_n) \|x_n - x\| + M_1\lambda_n. \end{aligned} \tag{3.2}$$

A discussion similar to the one for obtaining (3.2) guarantees that, for all $s \in \mathcal{S}$ and for all $j = 1, 2, \ldots, I_s$,

$$\left\| x_n^{(i_s^{(j)})} - x \right\| \leq (1 - \tau\lambda_n) \left\| x_n^{(i_s^{(j-1)})} - x \right\| + M_1\lambda_n. \tag{3.3}$$

Induction ensures that, for all $s \in \mathcal{S}$,

$$\begin{aligned}\left\| x_n^{(s)} - x \right\| &\leq (1 - \tau\lambda_n) \left\| x_n^{\left(i_s^{(I_s - 1)}\right)} - x \right\| + M_1 \lambda_n \\ &\leq (1 - \tau\lambda_n)^{I_s} \left\| x_n^{\left(i_s^{(0)}\right)} - x \right\| + I_s M_1 \lambda_n \\ &\leq (1 - \tau\lambda_n) \left\| x_n - x \right\| + I_s M_1 \lambda_n.\end{aligned} \tag{3.4}$$

The definition of x_{n+1}, the triangle inequality, (3.2), and (3.4) guarantee that

$$\begin{aligned}&\left\| x_{n+1} - x \right\| \\ &\leq \frac{1}{S+1} \left\{ \left\| x_n^{(0)} - x \right\| + \sum_{s \in \mathcal{S}} \left\| x_n^{(s)} - x \right\| \right\} \\ &\leq \frac{1}{S+1} \left\{ (1 - \tau\lambda_n) \left\| x_n - x \right\| + M_1 \lambda_n + S (1 - \tau\lambda_n) \left\| x_n - x \right\| + \sum_{s \in \mathcal{S}} I_s M_1 \lambda_n \right\} \\ &= (1 - \tau\lambda_n) \left\| x_n - x \right\| + \frac{M_1 \lambda_n}{S+1} \left(1 + \sum_{s \in \mathcal{S}} I_s \right).\end{aligned}$$

Therefore, we find

$$\left\| x_{n+1} - x \right\| \leq (1 - \tau\lambda_n) \left\| x_n - x \right\| + \frac{M_1}{\tau(S+1)} \left(1 + \sum_{s \in \mathcal{S}} I_s \right) \tau\lambda_n,$$

and hence, for all $n \in \mathbb{N}$ and for all $x \in \bigcap_{i \in \{0\} \cup \mathcal{I}} \mathrm{Fix}(T^{(i)})$,

$$\left\| x_{n+1} - x \right\| \leq \max \left\{ \left\| x_0 - x \right\|, \frac{M_1}{\tau(S+1)} \left(1 + \sum_{s \in \mathcal{S}} I_s \right) \right\}.$$

This means $(x_n)_{n \in \mathbb{N}}$ is bounded. Since (3.3) ensures that $\|x_n^{(i_s^{(1)})} - x\| \leq (1 - \tau\lambda_n)\|x_n - x\| + M_1\lambda_n$ ($n \in \mathbb{N}, s \in \mathcal{S}, x \in \bigcap_{i \in \{0\} \cup \mathcal{I}} \mathrm{Fix}(T^{(i)})$), $(x_n^{(i_s^{(1)})})_{n \in \mathbb{N}}$ ($s \in \mathcal{S}$) is also bounded. Hence, (3.3) and the boundedness of $(x_n)_{n \in \mathbb{N}}$ guarantee that $(x_n^{(i_s^{(j)})})_{n \in \mathbb{N}}$ ($s \in \mathcal{S}, j = 1, 2, \ldots, I_s$) is bounded. This proves Lemma 3.1. □

Lemma 3.2. *Algorithm* 3.1 *has the following properties:*

(i) $\lim_{n \to \infty} \frac{\|x_{n+1} - x_n\|}{\lambda_n} = 0$;

(ii) $\lim_{n \to \infty} \|x_n - x_n^{(0)}\| = 0$ *and* $\lim_{n \to \infty} \|x_n - x_n^{(i_s^{(j)})}\| = 0$ *(*$s \in \mathcal{S}$, $j = 1, 2, \ldots, I_s$*)*;

(iii) $\lim_{n \to \infty} \|x_n - T^{(0)}(x_n)\| = 0$ *and* $\lim_{n \to \infty} \|x_n - T^{(i_s^{(j)})}(x_n)\| = 0$ *(*$s \in \mathcal{S}$, $j = 1, 2, \ldots, I_s$*)*.

Proof. (i) Put $M_2 := \sup_{n \in \mathbb{N}} \|\nabla f^{(0)}(x_n)\| < \infty$ (M_2 is bounded by the Lipschitz continuity of $\nabla f^{(0)}$ and Lemma 3.1). The nonexpansivity of $T^{(0)}$, the triangle inequality, and (3.1) ensure that, for all $n \in \mathbb{N}$,

$$\begin{aligned}\left\| x_{n+1}^{(0)} - x_n^{(0)} \right\| &\leq \left\| \left(x_{n+1} - \lambda_{n+1} \nabla f^{(0)}(x_{n+1}) \right) - \left(x_n - \lambda_n \nabla f^{(0)}(x_n) \right) \right\| \\ &\leq \left\| \left(x_{n+1} - \lambda_{n+1} \nabla f^{(0)}(x_{n+1}) \right) - \left(x_n - \lambda_{n+1} \nabla f^{(0)}(x_n) \right) \right\| \\ &\quad + M_2 \left| \lambda_n - \lambda_{n+1} \right| \\ &\leq (1 - \tau\lambda_{n+1}) \left\| x_{n+1} - x_n \right\| + M_2 \left| \lambda_n - \lambda_{n+1} \right|.\end{aligned} \tag{3.5}$$

Lemma 3.1 and the Lipschitz continuity of $\nabla f^{(i_s^{(j)})}$ ($s \in \mathcal{S}$, $j = 1, 2, \ldots, I_s$) guarantee that $(\nabla f^{(i_s^{(j)})}(x_n^{(i_s^{(j-1)})}))_{n\in\mathbb{N}}$ ($s \in \mathcal{S}$, $j = 1, 2, \ldots, I_s$) is bounded, and hence, $M_3 := \max_{s\in\mathcal{S}, j=1,2,\ldots,I_s}(\sup_{n\in\mathbb{N}} \|\nabla f^{(i_s^{(j)})}(x_n^{(i_s^{(j-1)})})\|) < \infty$. A discussion similar to the one for obtaining (3.5) guarantees that, for all $n \in \mathbb{N}$, for all $s \in \mathcal{S}$, and for all $j = 1, 2, \ldots, I_s$,

$$\left\| x_{n+1}^{(i_s^{(j)})} - x_n^{(i_s^{(j)})} \right\| \leq (1 - \tau\lambda_{n+1}) \left\| x_{n+1}^{(i_s^{(j-1)})} - x_n^{(i_s^{(j-1)})} \right\| + M_3 |\lambda_n - \lambda_{n+1}| . \tag{3.6}$$

Hence, induction guarantees that, for all $n \in \mathbb{N}$ and for all $s \in \mathcal{S}$,

$$\begin{aligned} \left\| x_{n+1}^{(s)} - x_n^{(s)} \right\| &\leq (1 - \tau\lambda_{n+1}) \left\| x_{n+1}^{(i_s^{(I_s-1)})} - x_n^{(i_s^{(I_s-1)})} \right\| + M_3 |\lambda_n - \lambda_{n+1}| \\ &\leq (1 - \tau\lambda_{n+1}) \|x_{n+1} - x_n\| + I_s M_3 |\lambda_n - \lambda_{n+1}| . \end{aligned} \tag{3.7}$$

Therefore, we find from the triangle inequality, (3.5), and (3.7) that, for all $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \frac{1}{S+1} \Bigg\{ (1 - \tau\lambda_n) \|x_n - x_{n-1}\| + M_2 |\lambda_n - \lambda_{n-1}| \\ &\quad + S (1 - \tau\lambda_n) \|x_n - x_{n-1}\| + \sum_{s\in\mathcal{S}} I_s M_3 |\lambda_n - \lambda_{n-1}| \Bigg\} \\ &= (1 - \tau\lambda_n) \|x_n - x_{n-1}\| + M_4 |\lambda_n - \lambda_{n-1}| , \end{aligned}$$

where $M_4 := (M_2 + M_3 \sum_{s\in\mathcal{S}} I_s)/(S+1) < \infty$. Accordingly, we have that, for all $n \in \mathbb{N}$,

$$\begin{aligned} \frac{\|x_{n+1} - x_n\|}{\lambda_n} &\leq (1 - \tau\lambda_n) \frac{\|x_n - x_{n-1}\|}{\lambda_n} + M_4 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n} \\ &= (1 - \tau\lambda_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + (1 - \tau\lambda_n) \left\{ \frac{\|x_n - x_{n-1}\|}{\lambda_n} - \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} \right\} \\ &\quad + M_4 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n} \\ &\leq (1 - \tau\lambda_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + M_5 \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| + M_4 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n}, \end{aligned} \tag{3.8}$$

where $M_5 := \sup_{n\in\mathbb{N}}(1 - \tau\lambda_n)\|x_n - x_{n-1}\|$ ($M_5 < \infty$ holds from (C1) and the boundedness of $(x_n)_{n\in\mathbb{N}}$). We also have from $1 \leq 1/\lambda_n$ ($n \in \mathbb{N}$) that, for all $n \in \mathbb{N}$,

$$\begin{aligned} M_5 \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| + M_4 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n} &\leq M_5 \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| + M_4 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n \lambda_{n-1}} \\ &= \frac{M_4 + M_5}{\tau} \tau\lambda_n \frac{1}{\lambda_n} \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| . \end{aligned} \tag{3.9}$$

Inequalities (3.8) and (3.9) lead one to deduce that, for all $n \in \mathbb{N}$,

$$\frac{\|x_{n+1} - x_n\|}{\lambda_n} \leq (1 - \tau\lambda_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + \frac{M_4 + M_5}{\tau} \tau\lambda_n \frac{1}{\lambda_n} \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| .$$

Since (C2) and (C3) imply that $\sum_{n=0}^{\infty} \tau\lambda_n = \infty$ and $\lim_{n\to\infty}((M_4+M_5)/\tau)(1/\lambda_n)|1/\lambda_n - 1/\lambda_{n-1}| = 0$, Proposition 2.6 leads us to

$$\lim_{n\to\infty} \frac{\|x_{n+1} - x_n\|}{\lambda_n} = 0.$$

Moreover, from (C1), we have

$$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0. \tag{3.10}$$

(ii) Choose $x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ arbitrarily. From the firm nonexpansivity of $T^{(0)}$, we have that, for all $n \in \mathbb{N}$,

$$\left\|x_n^{(0)} - x\right\|^2 \le \left\|\left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - x\right\|^2 - \left\|\left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - x_n^{(0)}\right\|^2,$$

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

$$\begin{aligned} \left\|x_n^{(0)} - x\right\|^2 &\le \|x_n - x\|^2 - 2\lambda_n \left\langle x_n - x, \nabla f^{(0)}(x_n)\right\rangle \\ &\quad - \left\|x_n - x_n^{(0)}\right\|^2 + 2\lambda_n \left\langle x_n - x_n^{(0)}, \nabla f^{(0)}(x_n)\right\rangle \\ &= \|x_n - x\|^2 - \left\|x_n - x_n^{(0)}\right\|^2 + 2\lambda_n \left\langle x - x_n^{(0)}, \nabla f^{(0)}(x_n)\right\rangle \\ &\le \|x_n - x\|^2 - \left\|x_n - x_n^{(0)}\right\|^2 + M_6\lambda_n, \end{aligned} \tag{3.11}$$

where $M_6 := 2\sup_{n\in\mathbb{N}} |\langle x - x_n^{(0)}, \nabla f^{(0)}(x_n)\rangle| < \infty$. A calculation similar to (3.11) guarantees that, for all $n \in \mathbb{N}$, for all $s \in \mathcal{S}$, and for all $j = 1, 2, \ldots, I_s$,

$$\left\|x_n^{\left(i_s^{(j)}\right)} - x\right\|^2 \le \left\|x_n^{\left(i_s^{(j-1)}\right)} - x\right\|^2 - \left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\|^2 + M_7\lambda_n,$$

where $M_7 := 2\max_{s\in\mathcal{S}, j=1,2,\ldots,I_s}(\sup_{n\in\mathbb{N}} |\langle x - x_n^{(i_s^{(j)})}, \nabla f^{(i_s^{(j)})}(x_n^{(i_s^{(j-1)})})\rangle|) < \infty$. Induction shows that, for all $n \in \mathbb{N}$ and for all $s \in \mathcal{S}$,

$$\begin{aligned} \left\|x_n^{(s)} - x\right\|^2 &\le \left\|x_n^{\left(i_s^{(I_s-1)}\right)} - x\right\|^2 - \left\|x_n^{\left(i_s^{(I_s-1)}\right)} - x_n^{\left(i_s^{(I_s)}\right)}\right\|^2 + M_7\lambda_n \\ &\le \|x_n - x\|^2 - \sum_{j=1}^{I_s} \left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\|^2 + I_s M_7 \lambda_n. \end{aligned} \tag{3.12}$$

Summing up (3.12) over all s means that, for all $n \in \mathbb{N}$,

$$\sum_{s\in\mathcal{S}} \left\|x_n^{(s)} - x\right\|^2 \le S\|x_n - x\|^2 - \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s} \left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\|^2 + \sum_{s\in\mathcal{S}} I_s M_7 \lambda_n. \tag{3.13}$$

Since the convexity of $\|\cdot\|^2$ ensures that, for all $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - x\|^2 &= \left\|\frac{1}{S+1}\left\{\left(x_n^{(0)} - x\right) + \sum_{s\in\mathcal{S}}\left(x_n^{(s)} - x\right)\right\}\right\|^2 \\ &\le \frac{1}{S+1}\left\{\left\|x_n^{(0)} - x\right\|^2 + \sum_{s\in\mathcal{S}} \left\|x_n^{(s)} - x\right\|^2\right\}, \end{aligned} \tag{3.14}$$

inequalities (3.11), (3.13), and (3.14) lead us to

$$\begin{aligned}\|x_{n+1}-x\|^2 &\le \frac{1}{S+1}\Bigg\{ \|x_n - x\|^2 - \left\|x_n - x_n^{(0)}\right\|^2 + M_6\lambda_n \\ &\qquad + S\|x_n-x\|^2 - \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\|^2 + \sum_{s\in\mathcal{S}} I_s M_7\lambda_n\Bigg\} \\ &= \|x_n-x\|^2 - \frac{1}{S+1}\left\{\left\|x_n - x_n^{(0)}\right\|^2 + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\|^2\right\} \\ &\qquad + \frac{M_6 + \sum_{s\in\mathcal{S}} I_s M_7}{S+1}\lambda_n.\end{aligned}$$

Accordingly, we find that, for all $n \in \mathbb{N}$,

$$\begin{aligned} 0 &\le \frac{1}{S+1}\left\{\left\|x_n - x_n^{(0)}\right\|^2 + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\|^2\right\} \\ &\le \|x_n-x\|^2 - \|x_{n+1}-x\|^2 + \frac{M_6 + \sum_{s\in\mathcal{S}} I_s M_7}{S+1}\lambda_n \\ &= (\|x_n - x\| + \|x_{n+1}-x\|)(\|x_n-x\| - \|x_{n+1}-x\|) + \frac{M_6 + \sum_{s\in\mathcal{S}} I_s M_7}{S+1}\lambda_n \\ &\le (\|x_n - x\| + \|x_{n+1}-x\|)\|x_n - x_{n+1}\| + \frac{M_6 + \sum_{s\in\mathcal{S}} I_s M_7}{S+1}\lambda_n. \end{aligned}$$

The boundedness of $(x_n)_{n\in\mathbb{N}}$, (3.10), and (C1) mean that the right-hand side of the above inequality converges to 0. Therefore, we have

$$\lim_{n\to\infty}\left\|x_n - x_n^{(0)}\right\| = 0 \tag{3.15}$$

and

$$\lim_{n\to\infty}\left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right\| = 0 \quad (s\in\mathcal{S},\ j = 1,2,\ldots,I_s). \tag{3.16}$$

Moreover, from

$$\begin{aligned}\left\|x_n - x_n^{\left(i_s^{(j)}\right)}\right\| &= \left\|\left(x_n^{\left(i_s^{(0)}\right)} - x_n^{\left(i_s^{(1)}\right)}\right) + \cdots + \left(x_n^{\left(i_s^{(j-1)}\right)} - x_n^{\left(i_s^{(j)}\right)}\right)\right\| \\ &\le \sum_{k=1}^{j}\left\|x_n^{\left(i_s^{(k-1)}\right)} - x_n^{\left(i_s^{(k)}\right)}\right\| \quad (n\in\mathbb{N},\ s\in\mathcal{S},\ j=1,2,\ldots,I_s)\end{aligned}$$

and (3.16), we have

$$\lim_{n\to\infty}\left\|x_n - x_n^{\left(i_s^{(j)}\right)}\right\| = 0 \quad (s\in\mathcal{S},\ j=1,2,\ldots,I_s). \tag{3.17}$$

(iii) The nonexpansivity of $T^{(0)}$ guarantees that, for all $n\in\mathbb{N}$,

$$\left\|x_n^{(0)} - T^{(0)}(x_n)\right\| \le \left\|\left(x_n - \lambda_n\nabla f^{(0)}(x_n)\right) - x_n\right\| \le \lambda_n\left\|\nabla f^{(0)}(x_n)\right\|.$$

Hence, (C1) and the boundedness of $(\nabla f^{(0)}(x_n))_{n\in\mathbb{N}}$ imply that

$$\lim_{n\to\infty}\left\|x_n^{(0)} - T^{(0)}(x_n)\right\| = 0. \tag{3.18}$$

Accordingly, from $\|x_n - T^{(0)}(x_n)\| \leq \|x_n - x_n^{(0)}\| + \|x_n^{(0)} - T^{(0)}(x_n)\|$ $(n \in \mathbb{N})$, (3.15), and (3.18), we have that

$$\lim_{n\to\infty}\left\|x_n - T^{(0)}(x_n)\right\| = 0.$$

The nonexpansivity of $T^{(i_s^{(j)})}$ implies that, for all $n \in \mathbb{N}$, for all $s \in \mathcal{S}$, and for all $j = 1, 2, \ldots, I_s$,

$$\begin{aligned}\left\|x_n^{\left(i_s^{(j)}\right)} - T^{\left(i_s^{(j)}\right)}(x_n)\right\| &\leq \left\|\left(x_n^{\left(i_s^{(j-1)}\right)} - \lambda_n \nabla f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right) - x_n\right\| \\ &\leq \left\|x_n^{\left(i_s^{(j-1)}\right)} - x_n\right\| + M_3\lambda_n.\end{aligned}$$

Hence, (3.17) and (C1) guarantee that

$$\lim_{n\to\infty}\left\|x_n^{\left(i_s^{(j)}\right)} - T^{\left(i_s^{(j)}\right)}(x_n)\right\| = 0 \quad (s \in \mathcal{S},\ j = 1, 2, \ldots, I_s). \tag{3.19}$$

Since the triangle inequality ensures that $\|x_n - T^{(i_s^{(j)})}(x_n)\| \leq \|x_n - x_n^{(i_s^{(j)})}\| + \|x_n^{(i_s^{(j)})} - T^{(i_s^{(j)})}(x_n)\|$ $(n \in \mathbb{N},\ s \in \mathcal{S},\ j = 1, 2, \ldots, I_s)$, we find from (3.17) and (3.19) that

$$\lim_{n\to\infty}\left\|x_n - T^{\left(i_s^{(j)}\right)}(x_n)\right\| = 0 \quad (s \in \mathcal{S},\ j = 1, 2, \ldots, I_s).$$

This proves Lemma 3.2. □

Regarding the processing order within each subnetwork and the proof of Lemma 3.2, we can make the following remark.

Remark 3.1. Let $s \in \mathcal{S}$ be fixed arbitrarily. When one user is randomly chosen from $\mathcal{I}_{s,n}^{(j)} := \mathcal{I}_{s,n}^{(j-1)} \setminus \{i_{s,n}^{(j-1)}\}$ which depends on n, in general, $\mathcal{I}_{s,n}^{(j)} \ni i_{s,n}^{(j)} \neq i_{s,n+1}^{(j)} \in \mathcal{I}_{s,n+1}^{(j)}$ holds, which implies $T^{(i_{s,n}^{(j)})} \neq T^{(i_{s,n+1}^{(j)})}$. We cannot show in this case that Algorithm 3.1 strongly converges to the solution to Problem 2.1 because the proof of Lemma 3.2(i) uses essentially nonexpansive mappings, $T^{(i_s^{(j)})} = T^{(i_{s,n}^{(j)})}$ $(n \in \mathbb{N})$; more specifically, (3.6) is given by using $T^{(i_s^{(j)})} = T^{(i_{s,n}^{(j)})} = T^{(i_{s,n+1}^{(j)})}$ $(n \in \mathbb{N})$ and the nonexpansivity of $T^{(i_s^{(j)})}$. Hence, in the future, we should try to devise distributed optimization methods that can be applied when user $i_{s,n}^{(j)}$ does not always coincide with user $i_{s,n+1}^{(j)}$.

We can prove the following lemma by using Lemmas 3.1 and 3.2. This lemma leads one to deduce the weak convergence of $(x_n)_{n\in\mathbb{N}}$ in Algorithm 3.1 to the solution to Problem 2.1.

LEMMA 3.3. *Algorithm* 3.1 *has the following properties:*

(i) *There exists* $(x_{n_k})_{k\in\mathbb{N}}$ $(\subset (x_n)_{n\in\mathbb{N}})$ *which converges weakly to a point* $x^* \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$;

(ii) $x^* \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ *coincides with the unique solution to Problem* 2.1;

(iii) $(x_n)_{n\in\mathbb{N}}$ *weakly converges to* x^*.

Proof. (i) Lemma 3.1 guarantees the existence of a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $(x_{n_k})_{k\in\mathbb{N}}$ weakly converges to a point $x^* \in H$. Assume that $x^* \notin \mathrm{Fix}(T^{(0)})$. Proposition 2.7, the nonexpansivity of $T^{(0)}$, and Lemma 3.2(iii) mean that

$$\begin{aligned}\liminf_{k\to\infty} \|x_{n_k} - x^*\| &< \liminf_{k\to\infty} \left\|x_{n_k} - T^{(0)}(x^*)\right\| \\ &= \liminf_{k\to\infty} \left\|\left(x_{n_k} - T^{(0)}\left(x_{n_k}\right)\right) + \left(T^{(0)}\left(x_{n_k}\right) - T^{(0)}(x^*)\right)\right\| \\ &= \liminf_{k\to\infty} \left\|T^{(0)}\left(x_{n_k}\right) - T^{(0)}(x^*)\right\| \\ &\le \liminf_{k\to\infty} \|x_{n_k} - x^*\|.\end{aligned}$$

This is a contradiction. Therefore, $x^* \in \mathrm{Fix}(T^{(0)})$. We shall prove that $x^* \in \bigcap_{i\in\mathcal{I}} \mathrm{Fix}(T^{(i)})$. Let $s \in \mathcal{S}$ be fixed arbitrarily, let $j \in \{1, 2, \ldots, I_s\}$ be chosen arbitrarily, and assume that $x^* \notin \mathrm{Fix}(T^{(i_s^{(j)})})$. A discussion similar to the one above produces a contradiction:

$$\liminf_{k\to\infty} \|x_{n_k} - x^*\| < \liminf_{k\to\infty} \left\|x_{n_k} - T^{\left(i_s^{(j)}\right)}(x^*)\right\| \le \liminf_{k\to\infty} \|x_{n_k} - x^*\|.$$

Hence, $x^* \in \mathrm{Fix}(T^{(i_s^{(j)})})$; i.e., $x^* \in \bigcap_{s\in\mathcal{S}} \bigcap_{j=1}^{I_s} \mathrm{Fix}(T^{(i_s^{(j)})}) = \bigcap_{i\in\mathcal{I}} \mathrm{Fix}(T^{(i)})$. Therefore, we find that $x^* \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$.

(ii) Let $x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ be chosen arbitrarily. The nonexpansivity of $T^{(0)}$ and the equation $\|x-y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2$ $(x, y \in H)$ mean that, for all $n \in \mathbb{N}$,

$$\begin{aligned}\left\|x_n^{(0)} - x\right\|^2 &\le \left\|(x_n - x) - \lambda_n \nabla f^{(0)}(x_n)\right\|^2 \\ &\le \|x_n - x\|^2 - 2\lambda_n \left\langle x_n - x, \nabla f^{(0)}(x_n)\right\rangle + M_2^2 \lambda_n^2,\end{aligned}$$

where $M_2 := \sup_{n\in\mathbb{N}} \|\nabla f^{(0)}(x_n)\| < \infty$. Since Proposition 2.3 ensures that $\langle x_n - x, \nabla f^{(0)}(x_n)\rangle \ge f^{(0)}(x_n) - f^{(0)}(x)$ $(n \in \mathbb{N})$, we have that, for all $n \in \mathbb{N}$,

$$\left\|x_n^{(0)} - x\right\|^2 \le \|x_n - x\|^2 + 2\lambda_n \left(f^{(0)}(x) - f^{(0)}(x_n)\right) + M_2^2 \lambda_n^2. \tag{3.20}$$

A discussion similar to the one for obtaining (3.20) guarantees that, for all $n \in \mathbb{N}$, for all $s \in \mathcal{S}$, and for all $j = 1, 2, \ldots, I_s$,

$$\left\|x_n^{\left(i_s^{(j)}\right)} - x\right\|^2 \le \left\|x_n^{\left(i_s^{(j-1)}\right)} - x\right\|^2 + 2\lambda_n \left(f^{\left(i_s^{(j)}\right)}(x) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right) + M_3^2 \lambda_n^2,$$

where $M_3 := \max_{s\in\mathcal{S}, j=1,2,\ldots,I_s} (\sup_{n\in\mathbb{N}} \|\nabla f^{(i_s^{(j)})}(x_n^{(i_s^{(j-1)})})\|) < \infty$. Accordingly, for all $n \in \mathbb{N}$ and for all $s \in \mathcal{S}$, we have

$$\left\|x_n^{(s)} - x\right\|^2 \le \|x_n - x\|^2 + 2\lambda_n \sum_{j=1}^{I_s} \left(f^{\left(i_s^{(j)}\right)}(x) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right) + I_s M_3^2 \lambda_n^2.$$

Summing up the above inequality over all s means that, for all $n \in \mathbb{N}$,

$$\begin{aligned}\sum_{s\in\mathcal{S}} \left\|x_n^{(s)} - x\right\|^2 \leq S \|x_n - x\|^2 + \sum_{s\in\mathcal{S}} I_s M_3^2 \lambda_n^2 \\ + 2\lambda_n \sum_{s\in\mathcal{S}} \sum_{j=1}^{I_s} \left(f^{\left(i_s^{(j)}\right)}(x) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right).\end{aligned} \tag{3.21}$$

Inequalities (3.14), (3.20), and (3.21) lead one to deduce that, for all $n \in \mathbb{N}$,

$$\begin{aligned}&\|x_{n+1} - x\|^2 \\ &\leq \frac{1}{S+1}\left\{\left\|x_n^{(0)} - x\right\|^2 + \sum_{s\in\mathcal{S}}\left\|x_n^{(s)} - x\right\|^2\right\} \\ &\leq \frac{1}{S+1}\Bigg\{ \|x_n - x\|^2 + 2\lambda_n\left(f^{(0)}(x) - f^{(0)}(x_n)\right) + M_2^2\lambda_n^2 + S\|x_n - x\|^2 \\ &\quad + \sum_{s\in\mathcal{S}} I_s M_3^2\lambda_n^2 + 2\lambda_n \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left(f^{\left(i_s^{(j)}\right)}(x) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right)\Bigg\} \\ &= \|x_n - x\|^2 + \frac{M_2^2 + \sum_{s\in\mathcal{S}} I_s M_3^2}{S+1}\lambda_n^2 \\ &\quad + \frac{2\lambda_n}{S+1}\left\{\left(f^{(0)}(x) - f^{(0)}(x_n)\right) + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left(f^{\left(i_s^{(j)}\right)}(x) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right)\right\},\end{aligned}$$

which, together with $f^{(0)} + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s} f^{\left(i_s^{(j)}\right)} = f^{(0)} + \sum_{i\in\mathcal{I}} f^{(i)} = \sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)} =: f$, implies that

$$\begin{aligned}&\|x_{n+1} - x\|^2 \\ &\leq \|x_n - x\|^2 + \frac{M_2^2 + \sum_{s\in\mathcal{S}} I_s M_3^2}{S+1}\lambda_n^2 \\ &\quad + \frac{2\lambda_n}{S+1}\left\{\sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}(x) - f^{(0)}(x_n) - \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s} f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right\} \\ &= \|x_n - x\|^2 + \frac{M_2^2 + \sum_{s\in\mathcal{S}} I_s M_3^2}{S+1}\lambda_n^2 \\ &\quad + \frac{2\lambda_n}{S+1}\left\{f(x) - f(x_n) + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left[f^{\left(i_s^{(j)}\right)}(x_n) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right]\right\}.\end{aligned}$$

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

$$\begin{aligned}\frac{2}{S+1}\left(f(x_n) - f(x)\right) \leq \frac{\|x_n - x\|^2 - \|x_{n+1} - x\|^2}{\lambda_n} + \frac{M_2^2 + \sum_{s\in\mathcal{S}} I_s M_3^2}{S+1}\lambda_n \\ + \frac{2}{S+1}\sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left[f^{\left(i_s^{(j)}\right)}(x_n) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right)\right].\end{aligned} \tag{3.22}$$

On the other hand, Proposition 2.3 and the Cauchy–Schwarz inequality guarantee that, for all $n \in \mathbb{N}$, for all $s \in \mathcal{S}$, and for all $j = 1, 2, \ldots, I_s$,

$$\begin{aligned} f^{\left(i_s^{(j)}\right)}(x_n) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right) &\leq \left\langle x_n - x_n^{\left(i_s^{(j-1)}\right)}, \nabla f^{\left(i_s^{(j)}\right)}(x_n) \right\rangle \\ &\leq \left\| x_n - x_n^{\left(i_s^{(j-1)}\right)} \right\| \left\| \nabla f^{\left(i_s^{(j)}\right)}(x_n) \right\|. \end{aligned}$$

Accordingly, Lemmas 3.2(ii) and 3.1 and (A2) (the Lipschitz continuity of $f^{(i)}$ $(i \in \mathcal{I})$) ensure that $\limsup_{n\to\infty}(f^{(i_s^{(j)})}(x_n) - f^{(i_s^{(j)})}(x_n^{(i_s^{(j-1)})})) \leq 0$ $(s \in \mathcal{S}, j = 1, 2, \ldots, I_s)$; i.e.,

$$\begin{aligned} &\limsup_{n\to\infty} \sum_{s\in\mathcal{S}} \sum_{j=1}^{I_s} \left[f^{\left(i_s^{(j)}\right)}(x_n) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right) \right] \\ &\leq \sum_{s\in\mathcal{S}} \sum_{j=1}^{I_s} \left(\limsup_{n\to\infty} \left[f^{\left(i_s^{(j)}\right)}(x_n) - f^{\left(i_s^{(j)}\right)}\left(x_n^{\left(i_s^{(j-1)}\right)}\right) \right] \right) \leq 0. \end{aligned} \tag{3.23}$$

Moreover, since $(\|x_n - x\|^2 - \|x_{n+1} - x\|^2)/\lambda_n = (\|x_n - x\| + \|x_{n+1} - x\|)(\|x_n - x\| - \|x_{n+1} - x\|)/\lambda_n \leq (\|x_n - x\| + \|x_{n+1} - x\|)(\|x_n - x_{n+1}\|)/\lambda_n$ $(n \in \mathbb{N})$, Lemmas 3.1 and 3.2(i) imply that

$$\limsup_{n\to\infty} \frac{\|x_n - x\|^2 - \|x_{n+1} - x\|^2}{\lambda_n} \leq 0. \tag{3.24}$$

Therefore, (3.22), (3.23), (3.24), and (C1) guarantee that, for all $x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$,

$$\limsup_{n\to\infty} \left(f(x_n) - f(x)\right) = \limsup_{n\to\infty} \left(\sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}(x_n) - \sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}(x) \right) \leq 0.$$

Since (A2) implies that $f := \sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}$ is convex and continuous, f is weakly lower continuous [2, Theorem 9.1]; i.e.,

$$f(x^*) \leq \liminf_{k\to\infty} f\left(x_{n_k}\right),$$

where $(x_{n_k})_{k\in\mathbb{N}}$ is a subsequence in Lemma 3.3(i) that converges weakly to $x^* \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$. Accordingly, we find that, for all $x \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$,

$$\sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}\left(x^*\right) = f(x^*) \leq \liminf_{k\to\infty} f\left(x_{n_k}\right) \leq \limsup_{k\to\infty} f\left(x_{n_k}\right) \leq f(x) = \sum_{i\in\{0\}\cup\mathcal{I}} f^{(i)}(x).$$

Since Problem 2.1 has a unique solution, denoted by $x^\star$, this inequality implies that $x^* \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ coincides with $x^\star$.

(iii) Lemmas 3.3(i) and 3.3(ii) ensure that $(x_{n_k})_{k\in\mathbb{N}}$ $(\subset (x_n)_{n\in\mathbb{N}})$ exists such that $(x_{n_k})_{k\in\mathbb{N}}$ weakly converges to $x^\star$. Let us take another weakly converging subsequence, $(x_{n_l})_{l\in\mathbb{N}}$, of $(x_n)_{n\in\mathbb{N}}$. Then, from Lemmas 3.2(i) and 3.2(ii), we can prove that $(x_{n_l})_{l\in\mathbb{N}}$ also weakly converges to $x^\star$; i.e., any subsequence of $(x_n)_{n\in\mathbb{N}}$ weakly converges to the unique solution $x^\star$. Hence, we can conclude that $(x_n)_{n\in\mathbb{N}}$ weakly converges to $x^\star$. This proves Lemma 3.3. ∎

Now we are in a position to prove Theorem 3.1.

Proof. Let $x^\star \in \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}(T^{(i)})$ be the solution to Problem 2.1. The nonexpansivity of $T^{(0)}$ and the inequality $\|x-y\|^2 \leq \|x\|^2 + 2\langle y-x, y\rangle$ $(x,y\in H)$ mean that, for all $n\in\mathbb{N}$,

$$\begin{aligned}\left\|x_n^{(0)} - x^\star\right\|^2 &\leq \left\|\left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - \left(x^\star - \lambda_n \nabla f^{(0)}\left(x^\star\right)\right) - \lambda_n \nabla f^{(0)}\left(x^\star\right)\right\|^2 \\ &\leq \left\|\left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right) - \left(x^\star - \lambda_n \nabla f^{(0)}\left(x^\star\right)\right)\right\|^2 \\ &\quad + 2\lambda_n \left\langle x^\star - \left(x_n - \lambda_n \nabla f^{(0)}(x_n)\right), \nabla f^{(0)}\left(x^\star\right)\right\rangle,\end{aligned}$$

which, together with (3.1) and the Cauchy–Schwarz inequality, implies that

$$\begin{aligned}\left\|x_n^{(0)} - x^\star\right\|^2 &\leq (1-\tau\lambda_n)\left\|x_n - x^\star\right\|^2 + 2\lambda_n \left\langle x^\star - x_n, \nabla f^{(0)}\left(x^\star\right)\right\rangle \\ &\quad + 2\lambda_n^2 \left\langle \nabla f^{(0)}(x_n), \nabla f^{(0)}\left(x^\star\right)\right\rangle \\ &\leq (1-\tau\lambda_n)\left\|x_n - x^\star\right\|^2 + 2\lambda_n \left\langle x^\star - x_n, \nabla f^{(0)}\left(x^\star\right)\right\rangle \\ &\quad + 2M_2\lambda_n^2 \left\|\nabla f^{(0)}\left(x^\star\right)\right\| \\ &= (1-\tau\lambda_n)\left\|x_n - x^\star\right\|^2 \\ &\quad + \tau\lambda_n \frac{2}{\tau}\left\{\left\langle x^\star - x_n, \nabla f^{(0)}\left(x^\star\right)\right\rangle + M_2 \left\|\nabla f^{(0)}\left(x^\star\right)\right\| \lambda_n\right\}.\end{aligned} \tag{3.25}$$

A calculation similar to (3.25) means that, for all $n\in\mathbb{N}$, for all $s\in\mathcal{S}$, and for all $j=1,2,\ldots,I_s$,

$$\begin{aligned}\left\|x_n^{\left(i_s^{(j)}\right)} - x^\star\right\|^2 &\leq (1-\tau\lambda_n)\left\|x_n^{\left(i_s^{(j-1)}\right)} - x^\star\right\|^2 \\ &\quad + \tau\lambda_n\frac{2}{\tau}\left\{\left\langle x^\star - x_n^{\left(i_s^{(j-1)}\right)}, \nabla f^{\left(i_s^{(j)}\right)}\left(x^\star\right)\right\rangle + M_3\left\|\nabla f^{\left(i_s^{(j)}\right)}\left(x^\star\right)\right\|\lambda_n\right\}.\end{aligned}$$

Accordingly, for all $n\in\mathbb{N}$ and for all $s\in\mathcal{S}$,

$$\begin{aligned}\left\|x_n^{(s)} - x^\star\right\|^2 \leq (1-\tau\lambda_n)\left\|x_n - x^\star\right\|^2 + \tau\lambda_n\frac{2}{\tau}\Bigg\{ & M_3 \sum_{j=1}^{I_s}\left\|\nabla f^{\left(i_s^{(j)}\right)}\left(x^\star\right)\right\|\lambda_n \\ & + \sum_{j=1}^{I_s}\left\langle x^\star - x_n^{\left(i_s^{(j-1)}\right)}, \nabla f^{\left(i_s^{(j)}\right)}\left(x^\star\right)\right\rangle\Bigg\}.\end{aligned}$$

Summing up the above inequality over all s leads to, for all $n\in\mathbb{N}$,

$$\begin{aligned}\sum_{s\in\mathcal{S}}\left\|x_n^{(s)} - x^\star\right\|^2 \leq S(1-\tau\lambda_n)\left\|x_n - x^\star\right\|^2 + \tau\lambda_n\frac{2}{\tau}\Bigg\{ & M_3\sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\|\nabla f^{\left(i_s^{(j)}\right)}\left(x^\star\right)\right\|\lambda_n \\ & + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\langle x^\star - x_n^{\left(i_s^{(j-1)}\right)}, \nabla f^{\left(i_s^{(j)}\right)}\left(x^\star\right)\right\rangle\Bigg\}.\end{aligned} \tag{3.26}$$

Therefore, from (3.14), (3.25), and (3.26), we have, for all $n\in\mathbb{N}$,

$$\left\|x_{n+1} - x^\star\right\|^2 \leq (1-\tau\lambda_n)\left\|x_n - x^\star\right\|^2 + \tau\lambda_n X_n,$$

where

$$X_n := \frac{1}{S+1}\left[\frac{2}{\tau}\left\{\left\langle x^\star - x_n, \nabla f^{(0)}(x^\star)\right\rangle + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\langle x^\star - x_n^{\left(i_s^{(j-1)}\right)}, \nabla f^{\left(i_s^{(j)}\right)}(x^\star)\right\rangle \right.\right.$$
$$\left.\left. + \left(M_2\left\|\nabla f^{(0)}(x^\star)\right\| + M_3\sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\|\nabla f^{\left(i_s^{(j)}\right)}(x^\star)\right\|\right)\lambda_n\right\}\right] \ (n\in\mathbb{N}).$$

Since Lemmas 3.3(iii) and 3.2(ii) guarantee that $(x_n)_{n\in\mathbb{N}}$ and $(x_n^{(i_s^{(j)})})_{n\in\mathbb{N}}$ ($s\in\mathcal{S}$, $j = 1,2,\ldots,I_s$) converge weakly to $x^\star$, we have

$$\lim_{n\to\infty}\left\{\left\langle x^\star - x_n, \nabla f^{(0)}(x^\star)\right\rangle + \sum_{s\in\mathcal{S}}\sum_{j=1}^{I_s}\left\langle x^\star - x_n^{\left(i_s^{(j-1)}\right)}, \nabla f^{\left(i_s^{(j)}\right)}(x^\star)\right\rangle\right\} = 0.$$

Therefore, (C1) ensures that $\lim_{n\to\infty} X_n = 0$. Hence, Proposition 2.6 and (C2) lead one to deduce that

$$\lim_{n\to\infty}\|x_n - x^\star\| = 0. \tag{3.27}$$

This implies that $(x_n)_{n\in\mathbb{N}}$ converges strongly to $x^\star$. Moreover, since $\|x_n^{(i_s^{(j)})} - x^\star\| \leq \|x_n^{(i_s^{(j)})} - x_n\| + \|x_n - x^\star\|$ ($n\in\mathbb{N}$, $s\in\mathcal{S}$, $j = 1,2,\ldots,I_s$), Lemma 3.2(ii) and (3.27) lead to

$$\lim_{n\to\infty}\left\|x_n^{\left(i_s^{(j)}\right)} - x^\star\right\| = 0 \quad (s\in\mathcal{S},\ j=1,2,\ldots,I_s);$$

i.e., $(x_n^{(i_s^{(j)})})_{n\in\mathbb{N}}$ ($s\in\mathcal{S}$, $j = 1,2,\ldots,I_s$) also converges strongly to $x^\star$. This proves Theorem 3.1. □

4. Application of Algorithm 3.1 to storage allocation.

4.1. Storage allocation problem. One application of Problem 2.1 is storage allocation [19] in a peer-to-peer (P2P) data system. Here, we consider a P2P data storage system network in which peer i ($i\in\mathcal{I}$) offers a storage capacity $c_o^{(i)}$ that is to be shared with other peers and demands a storage capacity $c_s^{(i)}$ that is to be used for storing its own data.

The supply and demand functions of peer i are defined as follows [19, Definition 2 and Assumption A]: $s^{(i)}(p) := a^{(i)}[p - p_{\min}^{(i)}]^+$, $d^{(i)}(p) := b^{(i)}[p_{\max}^{(i)} - p]^+$ ($p \geq 0$), where $x^+ := \max\{0, x\}$ ($x\in\mathbb{R}$), $a^{(i)}$, $b^{(i)}$, $p_{\max}^{(i)}$ (> 0), and $p_{\min}^{(i)}$ (≥ 0). Peer i is entirely described by four parameters, $a^{(i)}$, $b^{(i)}$, $p_{\max}^{(i)}$, and $p_{\min}^{(i)}$. The two price parameters $p_{\min}^{(i)}$ and $p_{\max}^{(i)}$, respectively, represent the minimum value of the unit price p_o at which peer i will sell some of its own disk space and the maximum value of the unit price p_s that it will pay for storage space, and $a^{(i)}$ and $b^{(i)}$, respectively, correspond to the increase in sold capacity with the unit price p_o ($\geq p_{\min}^{(i)}$) and the decrease in bought storage space with the unit price p_s ($\leq p_{\max}^{(i)}$). For a given p (≥ 0), $s^{(i)}(p)$ (resp., $d^{(i)}(p)$) is the amount of storage capacity that peer i would choose to sell (resp., buy) if peer i were paid (resp., charged) a unit price p for it.

The utility function $U^{(i)}$ of peer i is of the following form [19, section II]:

$$U^{(i)}\left(c_s^{(i)}, c_o^{(i)}, \varepsilon^{(i)}\right) := V^{(i)}\left(c_s^{(i)}\right) - P^{(i)}\left(c_o^{(i)}\right) - \varepsilon^{(i)}, \tag{4.1}$$

where $x \wedge y := \min\{x, y\}$ $(x, y \in \mathbb{R})$,

$$V^{(i)}\left(c_s^{(i)}\right) := \frac{1}{b^{(i)}}\left[-\frac{\left(c_s^{(i)} \wedge b^{(i)} p_{\max}^{(i)}\right)^2}{2} + b^{(i)} p_{\max}^{(i)}\left(c_s^{(i)} \wedge b^{(i)} p_{\max}^{(i)}\right)\right]$$

is peer i's valuation obtained when it uses $c_s^{(i)}$, $O^{(i)}(c_o^{(i)}) := (1/a^{(i)})(c_o^{(i)})^2/2$ is the opportunity cost of offering $c_o^{(i)}$ for other peers without using $c_o^{(i)}$ for itself, $p_{\min}^{(i)} c_o^{(i)}$ is the data transfer cost,

$$P^{(i)}\left(c_o^{(i)}\right) := O^{(i)}\left(c_o^{(i)}\right) + p_{\min}^{(i)} c_o^{(i)}$$

stands for the overall nonmonetary cost of peer i offering $c_o^{(i)}$ in the system, and

$$\varepsilon^{(i)} := p_s c_s^{(i)} - p_o c_o^{(i)}$$

is the monetary price paid by peer i.

On the other hand, the operator (denoted by peer 0), which manages the P2P data storage system, tries to maximize its revenue, which is the total amount that the peers are charged. Since the monetary price paid by peer i is $\varepsilon^{(i)} = p_s c_s^{(i)} - p_o c_o^{(i)}$, $c_s^{(i)} = d^{(i)}(p_s)$, and $c_o^{(i)} = s^{(i)}(p_o)$, the utility function of the operator can be represented by

$$U^{(0)}(p_s, p_o) := \sum_{i \in \mathcal{I}} \varepsilon^{(i)} = \sum_{i \in \mathcal{I}} \left[p_s d^{(i)}(p_s) - p_o s^{(i)}(p_o)\right]. \tag{4.2}$$

We define a performance measure, called *social welfare* [19, Definition 3], in the whole system as the sum of the utility functions of all peers and the operator. From (4.1) and (4.2), social welfare can be expressed as follows: For all $c_s := (c_s^{(1)}, c_s^{(2)}, \ldots, c_s^{(I)})^T$, $c_o := (c_o^{(1)}, c_o^{(2)}, \ldots, c_o^{(I)})^T \in \mathbb{R}^I$,

$$\begin{aligned} W(c_s, c_o) &:= \sum_{i \in \mathcal{I}} U^{(i)}\left(c_s^{(i)}, c_o^{(i)}, \varepsilon^{(i)}\right) + U^{(0)}(p_s, p_o) \\ &= \sum_{i \in \mathcal{I}} \left[V^{(i)}\left(c_s^{(i)}\right) - P^{(i)}\left(c_o^{(i)}\right)\right], \end{aligned} \tag{4.3}$$

where x^T denotes the transpose of the vector x. It is desirable to maximize W defined by (4.3) because it makes the whole system stable and reliable. We call $W^{(i)} \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined for all $(c_s^{(i)}, c_o^{(i)}) \in \mathbb{R} \times \mathbb{R}$ by

$$W^{(i)}\left(c_s^{(i)}, c_o^{(i)}\right) := V^{(i)}\left(c_s^{(i)}\right) - P^{(i)}\left(c_o^{(i)}\right)$$

the *welfare* of peer i.

A payment-based management scheme is based on monetary exchanges where peers can buy storage space in the system for a unit price p_s and sell some of their disk capacity for a unit price p_o. In the profit-oriented pricing scheme, the operator strives to extract the maximum profit out of the business by buying and selling storage spaces. Assuming that the operator knows that peer i ($i \in \mathcal{I}$) will sell $s^{(i)}(p_o)$ and buy $d^{(i)}(p_s)$, it tries to choose p_s and p_o so as to maximize its profit $U^{(0)}(p_s, p_o)$. Accordingly, the constrained set and objective function of the operator (peer 0) are defined as follows [19, section III.C]:

$$C^{(0)} := \mathbb{R}_+ \times \mathbb{R}_+ \cap \left\{ (p_s, p_o) \in \mathbb{R} \times \mathbb{R} \colon \sum_{i \in \mathcal{I}} s^{(i)}(p_o) \geq \sum_{i \in \mathcal{I}} d^{(i)}(p_s) \right\}, \tag{4.4}$$

$$f^{(0)}(p_s, p_o) := -U^{(0)}(p_s, p_o) = -\left[p_s \sum_{i \in \mathcal{I}} d^{(i)}(p_s) - p_o \sum_{i \in \mathcal{I}} s^{(i)}(p_o) \right] \tag{4.5}$$

for all $(p_s, p_o) \in \mathbb{R} \times \mathbb{R}$. The operator must have $C^{(0)}$ defined in (4.4) because $\sum_{i \in \mathcal{I}} c_s^{(i)} = \sum_{i \in \mathcal{I}} d^{(i)}(p_s)$, which is used for storing data, must not exceed $\sum_{i \in \mathcal{I}} c_o^{(i)} = \sum_{i \in \mathcal{I}} s^{(i)}(p_o)$ offered by the peers. Here, let us define a mapping $T^{(0)} \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ for all $(p_s, p_o) \in \mathbb{R} \times \mathbb{R}$ by

$$T^{(0)}(p_s, p_o) := \frac{1}{2} \left[(p_s, p_o) + P_{\mathbb{R}_+ \times \mathbb{R}_+} \left\{ P_{\hat{C}^{(0)}}(p_s, p_o) \right\} \right], \tag{4.6}$$

where $\hat{C}^{(0)} := \{(p_s, p_o) \in \mathbb{R} \times \mathbb{R} \colon \sum_{i \in \mathcal{I}} s^{(i)}(p_o) \geq \sum_{i \in \mathcal{I}} d^{(i)}(p_s)\}$. Since $s^{(i)}$ and $d^{(i)}$ are affine, $\hat{C}^{(0)}$ is a half-space, which means that $P_{\hat{C}^{(0)}}$ can be easily computed within a finite number of arithmetic operations [1, p. 406], [2, Subchapter 28.3]. $T^{(0)}$ defined in (4.6) satisfies the firm nonexpansivity condition (Propositions 2.4(ii) and 2.4(iii)), and

$$\mathrm{Fix}\left(T^{(0)}\right) := \left\{ (p_s, p_o) \in \mathbb{R} \times \mathbb{R} \colon T^{(0)}(p_s, p_o) = (p_s, p_o) \right\} = C^{(0)}$$

because $\mathrm{Fix}(T^{(0)}) = \mathrm{Fix}(P_{\mathbb{R}_+ \times \mathbb{R}_+} P_{\hat{C}^{(0)}}) = \mathbb{R}_+ \times \mathbb{R}_+ \cap \hat{C}^{(0)} =: C^{(0)}$. Moreover, since $s^{(i)}$ and $d^{(i)}$ are affine, $f^{(0)}$ in (4.5) satisfies the strong convexity condition. Hence, we can see that $\nabla f^{(0)}$ is strongly monotone and Lipschitz continuous.

Meanwhile, peer i ($i \in \mathcal{I}$) selfishly chooses strategies that maximize its welfare $W^{(i)}$. Accordingly, the constrained set and objective function of peer i ($i \in \mathcal{I}$) can be expressed as

$$C^{(i)} := \left[p_{\min}^{(i)}, p_{\max}^{(i)} \right] \times \left[p_{\min}^{(i)}, p_{\max}^{(i)} \right] = \mathrm{Fix}\left(P_{C^{(i)}}\right) =: \mathrm{Fix}\left(T^{(i)}\right), \tag{4.7}$$

$$f^{(i)}(p_s, p_o) := -\left[V^{(i)}\left(d^{(i)}(p_s)\right) - P^{(i)}\left(s^{(i)}(p_o)\right) \right] \tag{4.8}$$

for all $(p_s, p_o) \in \mathbb{R} \times \mathbb{R}$. Since $s^{(i)}$ and $d^{(i)}$ are affine and $V^{(i)}$ and $P^{(i)}$ have quadratic forms, $f^{(i)}$ ($i \in \mathcal{I}$) in (4.8) satisfies the strong convexity condition. Hence, we have that $\nabla f^{(i)}$ ($i \in \mathcal{I}$) is strongly monotone and Lipschitz continuous. $T^{(i)} := P_{C^{(i)}}$ ($i \in \mathcal{I}$) in (4.7) is easily computable and firmly nonexpansive (Proposition 2.4(iii)) [1, p. 406], [2, Subchapter 28.3].

The main objective of the profit-oriented pricing scheme is to determine optimal prices p_s and p_o so as to maximize the operator's profit $U^{(0)}$. Meanwhile, it is desirable to maximize the social welfare W to make the whole system stable and reliable. Therefore, we can formulate the storage allocation problem for the profit-oriented pricing scheme as one of maximizing the weighted mean of the operator's profit and social welfare, $\lambda U^{(0)} + (1-\lambda)W$, for some weight parameter λ ($\in (0,1)$).

PROBLEM 4.1 (storage allocation problem for profit-oriented pricing scheme).

$$\text{Maximize } \lambda U^{(0)}(p_s,p_o) + (1-\lambda)W(p_s,p_o) = -\left[\lambda f^{(0)} + (1-\lambda)\sum_{i\in\mathcal{I}} f^{(i)}\right](p_s,p_o)$$

$$\text{subject to } (p_s,p_o) \in \left\{(p_s,p_o)\in\mathbb{R}_+\times\mathbb{R}_+ \colon \sum_{i\in\mathcal{I}} s^{(i)}(p_o) \geq \sum_{i\in\mathcal{I}} d^{(i)}(p_s)\right\}$$
$$\cap \bigcap_{i\in\mathcal{I}} \left[p_{\min}^{(i)}, p_{\max}^{(i)}\right] \times \left[p_{\min}^{(i)}, p_{\max}^{(i)}\right] = \bigcap_{i\in\{0\}\cup\mathcal{I}} \mathrm{Fix}\left(T^{(i)}\right),$$

where $\lambda \in (0,1)$ *is a parameter chosen in advance and* $f^{(i)} \colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ *and* $T^{(i)}\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}\times\mathbb{R}$ *(*$i\in\{0\}\cup\mathcal{I}$*) are defined as in* (4.5)*,* (4.6)*,* (4.7)*, and* (4.8)*.*

Therefore, we can conclude that Problem 4.1 can be formulated as Problem 2.1.

4.2. Experimental results. Let us apply Algorithm 3.1 to Problem 4.1 with $\lambda = 1/2$. We used a MacBook Air 11-inch, Mid 2013. The computer had a 1.30GHz Intel Core i5-4250U processor and 4GB 1600MHz DDR3 memory. Algorithm 3.1 was written in C++. The experiment used random numbers in the range of $(0,5]$ for $a^{(i)}$ and $b^{(i)}$, random numbers in the range of $[0,10]$ for $p_{\min}^{(i)}$, and random numbers in the range of $[90,100]$ for $p_{\max}^{(i)}$. The random numbers were generated using the function `random-real` in the srfi-27 library of Gauche.[2] In the experiment, we set $\lambda_n := 10^{-3}/(n+1)^a$ ($a = 0.10, 0.25, 0.45$), $I := 100$, $S := 2, 5, 10, 20, 50, 100$, and $I_s = I/S$, and performed 100 samplings, each starting from different random initial points.[3] We averaged the results of the 100 samplings.[4] Note that Algorithm 3.1 when $S = I$ (i.e., $I_s = 1$) coincides with the broadcast optimization algorithm, and Algorithm 3.1 when $S = 1$ (i.e., $I = I_1$) is similar to the incremental optimization algorithm implemented by all peers.

Figure 4 shows the behavior of p_s and p_o when $\lambda_n := 10^{-3}/(n+1)^{0.10}$. The plots show that although Algorithm 3.1 behaves differently depending on the choice of S, it converges to the same point for $S = 2, 5, 10, 20, 50, 100$. In particular, we can see from these graphs that the required numbers of iterations for $S = 2, 5, 10, 20, 50$ (Figures 4(a)–4(e)) are less than those for $S = I = 100$ (Figure 4(f)). This means that Algorithm 3.1 with $S < 100$ has fewer iterations compared with the conventional broadcast optimization algorithm (Algorithm 3.1 with $S = I = 100$).

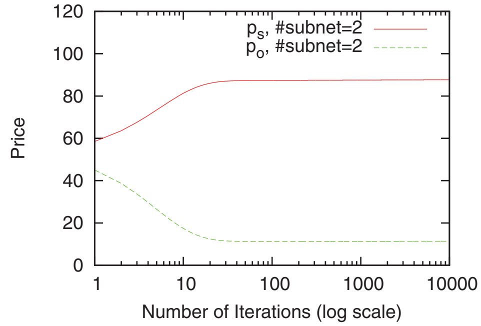
Figure 5 describes the behaviors of p_s and p_o when $\lambda_n := 10^{-3}/(n+1)^a$ ($a = 0.10, 0.25, 0.45$) and $S = 20$. The graphs show that Algorithm 3.1 converges to the same point with $\lambda_n := 10^{-3}/(n+1)^a$ ($a = 0.10, 0.25, 0.45$) and that it converges faster with $a = 0.10$ than with $a = 0.25, 0.45$. A similar trend was observed in the numerical results for $S = 2, 5, 10, 50, 100$.[5] Figures 4 and 5 indicate that the optimal p_s (denoted by $p_s^\star$) is larger than the optimal p_o (denoted by $p_o^\star$). Since Algorithm 3.1 converges in $C^{(0)}$ defined by (4.4), $C^\star := \sum_{i\in\mathcal{I}} s^{(i)}(p_o^\star) \approx \sum_{i\in\mathcal{I}} d^{(i)}(p_s^\star)$ holds. Hence,

[2]We used Gauche Scheme Shell, Version 0.9.3.3 [utf-8,pthreads], x86_64-apple-darwin12.

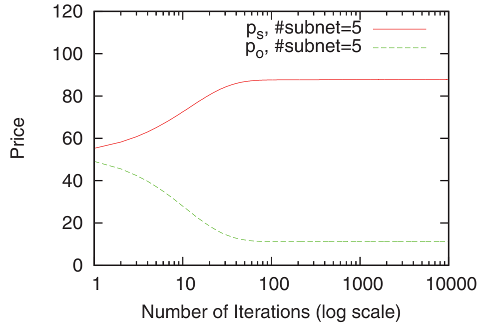
[3]Random values were generated by the `rand` function in the C Standard Library.

[4]We used gnuplot Version 4.6 (patchlevel 3) to make the graphs in this paper from experiment results.

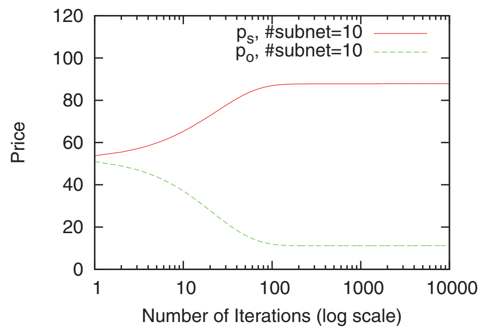
[5]We omitted the details of the results for $a = 0.10, 0.25, 0.45$ and $S = 2, 5, 10, 50, 100$ because of lack of space.



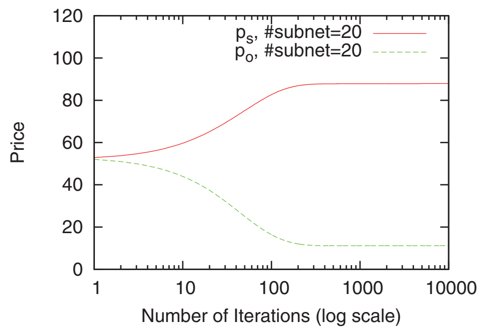
(a) The number of subnetworks is 2.



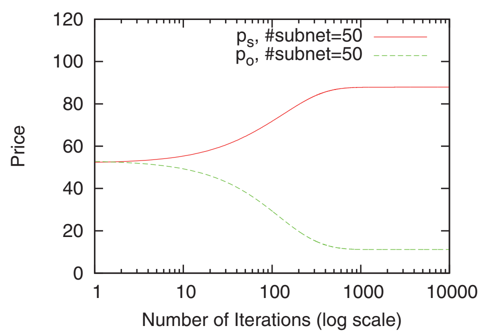
(b) The number of subnetworks is 5.



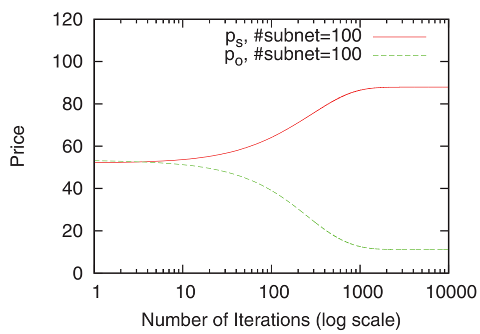
(c) The number of subnetworks is 10.



(d) The number of subnetworks is 20.



(e) The number of subnetworks is 50.

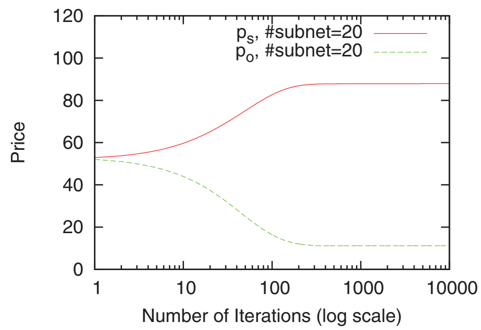


(f) The number of subnetworks is 100.

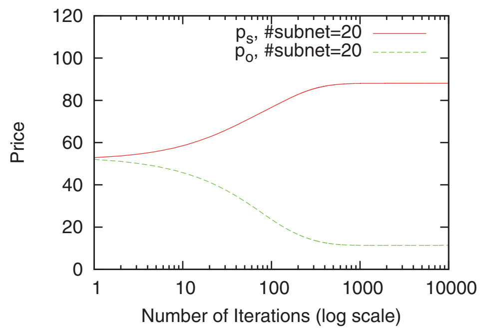
FIG. 4. *Behavior of p_s and p_o when $\lambda_n := 10^{-3}/(n+1)^{0.10}$.*

the operator's revenue $U^{(0)}(p_s^\star, p_o^\star)$ is approximately $(p_s^\star - p_o^\star)C^\star > 0$; i.e., the operator makes a profit in the situation depicted in the experiment.

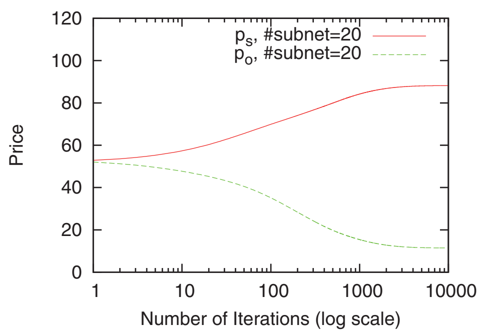
Figure 6 is the relation between the number of subnetworks and running time until $n = 10^4$ in the case of $\lambda_n := 10^{-3}/(n+1)^{0.10}$. Here, Algorithm 3.1 with $S = 2, 5, 10$ dramatically reduces the running time compared with the conventional broadcast optimization algorithm (Algorithm 3.1 with $S = I = 100$); see also issue (i) in section 1. Therefore, we can conclude from Figures 4 and 6 that Algorithm 3.1 with $S < I$ converges to the solution to Problem 4.1 faster than the broadcast optimization



(a) $\lambda_n := 10^{-3}/(n+1)^{0.10}$.



(b) $\lambda_n := 10^{-3}/(n+1)^{0.25}$.



(c) $\lambda_n := 10^{-3}/(n+1)^{0.45}$.

FIG. 5. *Behavior of p_s and p_o when the number of subnetworks is* 20.

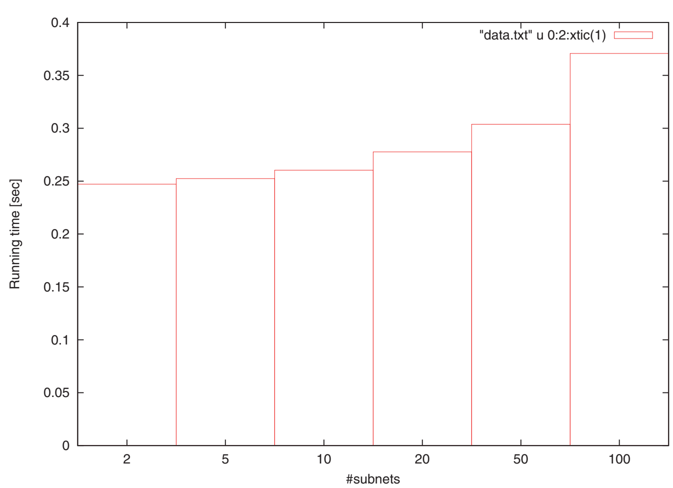


FIG. 6. *Relation between the number of subnetworks and running time until $n = 10^4$ in the case of $\lambda_n := 10^{-3}/(n+1)^{0.10}$.*

algorithm can. Moreover, Figures 4 and 6 indicate that the smaller S is, the fewer the required iterations become and the shorter the running time becomes. This implies that if there are subnetworks in which as many peers as possible participate and if the incremental optimization algorithm can be implemented by them, the operator can quickly find the solution to Problem 4.1, thanks to full cooperation from many peers. However, it would be physically difficult for many peers to implement the incremental optimization algorithm because real networked systems are complex and composed of a number of subnetworks (see issue (ii)). Meanwhile, Algorithm 3.1 can be applied when each peer communicates with its neighbor peers, and the applications of Algorithm 3.1 do not depend on the network topology. Therefore, we can conclude that Algorithm 3.1 is a good way of solving convex optimization problems in large-scale and complex networked systems.

5. Conclusion and future work. We discussed the problem of minimizing the sum of convex objective functions over the intersection of fixed point sets of nonexpansive mappings in a Hilbert space and presented a novel distributed optimization algorithm for solving the problem and its convergence analysis. The algorithm combines the conventional broadcast and incremental optimization algorithms. The convergence analysis guarantees that the algorithm, with a slowly diminishing step-size sequence, converges strongly to the solution to the problem. Finally, we described a numerical experiment that used the algorithm for storage allocation. The numerical results and discussions showed that our algorithm converges to the solution to the storage allocation problem faster than the conventional broadcast optimization algorithm can, and has a wider range of application compared with the conventional incremental optimization algorithm.

The convergence analysis ensures that our algorithm works when the processing order within each subnetwork is deterministic at all times. It would be desirable to devise distributed optimization algorithms which work when the processing order is randomized at each iteration because they have a wider range of application compared with our algorithm. Therefore, in the future, we need to devise such distributed optimization algorithms.

Acknowledgments. We are sincerely grateful to Associate Editor Patrick L. Combettes and the two anonymous referees for helping us improve the original manuscript.

REFERENCES

[1] H. H. Bauschke and J. M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev., 38 (1996), pp. 367–426.
[2] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
[3] V. Berinde, *Iterative Approximation of Fixed Points*, Springer, Berlin, 2007.
[4] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, MA, 2003.
[5] D. Blatt, A. O. Hero, and H. Gauchman, *A convergent incremental gradient method with a constant step size*, SIAM J. Optim., 18 (2007), pp. 29–51.
[6] P. L. Combettes, *A block-iterative surrogate constraint splitting method for quadratic signal recovery*, IEEE Trans. Signal Process., 51 (2003), pp. 1771–1782.
[7] P. L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators*, J. Convex Anal., 16 (2009), pp. 727–748.
[8] P. L. Combettes and P. Bondon, *Hard-constrained inconsistent signal feasibility problems*, IEEE Trans. Signal Process., 47 (1999), pp. 2460–2468.

[9] P. L. Combettes and J. C. Pesquet, *A proximal decomposition method for solving convex variational inverse problems*, Inverse Problems, 24 (2008), 065014.
[10] H. Iiduka, *Fixed point optimization algorithm and its application to network bandwidth allocation*, J. Comput. Appl. Math., 236 (2012), pp. 1733–1742.
[11] H. Iiduka, *Fixed point optimization algorithm and its application to power control in CDMA data networks*, Math. Program., 133 (2012), pp. 227–242.
[12] H. Iiduka, *Fixed point optimization algorithms for distributed optimization in networked systems*, SIAM J. Optim., 23 (2013), pp. 1–26.
[13] H. Iiduka, *Acceleration method for convex optimization over the fixed point set of a nonexpansive mapping*, Math. Program., to appear.
[14] H. Iiduka and I. Yamada, *A use of conjugate gradient direction for the convex optimization problem over the fixed point set of a nonexpansive mapping*, SIAM J. Optim., 19 (2009), pp. 1881–1893.
[15] B. Johansson, M. Rabi, and M. Johansson, *A randomized incremental subgradient method for distributed optimization in networked systems*, SIAM J. Optim., 20 (2009), pp. 1157–1170.
[16] M. Kaneko, P. Popovski, and J. Dahl, *Proportional fairness in multi-carrier system: Upper bound and approximation algorithms*, IEEE Commun. Lett., 10 (2006), pp. 462–464.
[17] K. C. Kiwiel, *Convergence of approximate and incremental subgradient methods for convex optimization*, SIAM J. Optim., 14 (2004), pp. 807–840.
[18] S. H. Low and D. E. Lapsley, *Optimization flow control*—I: *Basic algorithm and convergence*, IEEE/ACM Trans. Netw., 7 (1999), pp. 861–874.
[19] P. Maillé and L. Toka, *Managing a peer-to-peer data storage system in a selfish society*, IEEE J. Sel. Areas Commun., 26 (2008), pp. 1295–1301.
[20] A. Nedić and D. P. Bertsekas, *Incremental subgradient methods for nondifferentiable optimization*, SIAM J. Optim., 12 (2001), pp. 109–138.
[21] Z. Opial, *Weak convergence of the sequence of successive approximation for nonexpansive mappings*, Bull. Amer. Math. Soc., 73 (1967), pp. 591–597.
[22] J. C. Pesquet and N. Pustelnik, *A parallel inertial proximal optimization method*, Pac. J. Optim., 8 (2012), pp. 273–306.
[23] S. Sharma and D. Teneketzis, *An externalities-based decentralized optimal power allocation algorithm for wireless networks*, IEEE/ACM Trans. Netw., 17 (2009), pp. 1819–1831.
[24] I. Yamada, *The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings*, in Inherently Parallel Algorithms for Feasibility and Optimization and Their Applications, D. Butnariu, Y. Censor, and S. Reich, eds., North-Holland, Amsterdam, 2001, pp. 473–504.