FIXED POINT OPTIMIZATION ALGORITHMS FOR DISTRIBUTED OPTIMIZATION IN NETWORKED SYSTEMS *

HIDEAKI IIDUKA †

Abstract. This paper considers a networked system with a finite number of users and deals with the problem of minimizing the sum of all users' objective functions over the intersection of all users' constraint sets, onto which the projection cannot be easily implemented. The main objective of this paper is to devise distributed optimization algorithms, which enable each user to find the solution of the problem without using other users' objective functions and constraint sets. To reach this goal, we first introduce easily implementable nonexpansive mappings of which the intersection of the fixed point sets is equal to the constraint set in the problem. We formulate the problem as a convex minimization problem over the intersection of the fixed point sets of the nonexpansive mappings. We then present an iterative algorithm, based on the conventional incremental subgradient methods which use the projection, for solving the problem. The algorithm can be implemented by using other nonexpansive mappings than the projection. We prove that the algorithm with slowly diminishing step-size sequences converges to a solution of the problem in the sense of weak topology of a Hilbert space. We also present a broadcast type of distributed optimization algorithm that weakly converges to a solution of the problem. Numerical examples for the bandwidth allocation demonstrate the convergence of these algorithms.

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

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

1. Introduction. This paper presents distributed optimization algorithms for solving the convex minimization problem,

$$\text{minimize } f(x) := \sum_{i\in I} f^{(i)}(x) \text{ subject to } x \in C := \bigcap_{i\in I} C^{(i)}, \tag{1.1}$$

where $f^{(i)}$ ($i \in I := \{1, 2, \ldots, K\}$) is a convex functional of a real Hilbert space H and $C^{(i)}$ ($\subset H$) ($i \in I$) is nonempty, closed, and convex.

We focus on Problem (1.1) in a networked system in which user i ($i \in I$) has its own private objective function, $f^{(i)}$, and constraint set, $C^{(i)}$, and cannot get the explicit forms of other users' objective functions and constraint sets. Problem (1.1) in this situation includes important and practical engineering problems, such as signal and image processing [9], channel allocation [20], bandwidth allocation [22], storage allocation [23], and power allocation [27] problems.

Distributed optimization algorithms for Problem (1.1) can be implemented through all users' cooperating, and they enable each user to find the optimal solution of Problem (1.1) without using the private information of other users such as their objective functions and constraint sets. A useful distributed algorithm for solving Problem (1.1) is the *incremental subgradient method* (see [4, Subchapter 8.2], [5, 19, 21, 24] and references therein). The sequence, $(x_n)_{n\in\mathbb{N}}$, is generated by the incremental subgradient

*This work was supported by the Japan Society for the Promotion of Science through a Grant-in-Aid for Young Scientists (B) (23760077), and in part by the Japan Society for the Promotion of Science through a Grant-in-Aid for Scientific Research (C) (23500090).

†Network Design Research Center, Kyushu Institute of Technology, Hibiya Kokusai Bldg. 1F 107, 2-2-3 Uchisaiwaicho, Chiyoda-ku, Tokyo 100-0011, Japan (iiduka@ndrc.kyutech.ac.jp).

method as follows: given $(\lambda_n)_{n\in\mathbb{N}} \subset (0,\infty)$ and $x_n := x_n^{(0)} \in H$,

$$\begin{cases} x_n^{(i)} := P_C\left(x_n^{(i-1)} - \lambda_n g_n^{(i)}\right), \ g_n^{(i)} \in \partial f^{(i)}\left(x_n^{(i-1)}\right) \ (i = 1,2,\ldots,K), \\ x_{n+1} := x_n^{(K)}, \end{cases} \tag{1.2}$$

where P_C is the metric projection onto C and $\partial f^{(i)}(x)$ stands for the subdifferential of $f^{(i)}$ at $x \in H$. Convergence analyses of Algorithm (1.2) have been done when $(\lambda_n)_{n\in\mathbb{N}}$ is a constant step-size [4, Subchapter 8.2.1], [5, 24], a diminishing step-size [4, Subchapter 8.2.1], [21, 24], or a dynamic step-size [4, Subchapter 8.2.2], [19, 24]. When $(\lambda_n)_{n\in\mathbb{N}}$ is a slowly diminishing step-size sequence, $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm (1.2) converges to a solution of Problem (1.1) [4, Proposition 8.2.6], [24, Proposition 2.4]. The analyses guarantee that, under the assumption that $C := \bigcap_{i\in I} C^{(i)}$ is known to the all users in advance and is simple enough so that the projection can be easily implemented, Algorithm (1.2) enables each user in the network to decide its own optimal solution by using only its own private objective function and the transmitted information from the neighbor user.

In this paper, we will discuss Problem (1.1) under the following assumptions:

(I) User i ($i \in I$) has its own private $f^{(i)}$ and $C^{(i)}$, and cannot get the explicit forms of other users' objective functions and constraint sets, i.e., none of users can use $P_C = P_{\bigcap_{i\in I} C^{(i)}}$.

(II) $C^{(i)}$ ($i \in I$) does not always have a simple form.[1]

(III) User i ($i \in I$) can use a firmly nonexpansive mapping[2], $T^{(i)} \colon H \to H$, satisfying $\mathrm{Fix}(T^{(i)}) := \{x \in H \colon T^{(i)}(x) = x\} = C^{(i)}$.

A particularly interesting application of Problem 1.1 under Assumptions (I), (II), and (III) is when user i ($i \in I$) has a nonempty, closed and convex constraint set,

$$C^{(i)} := \bigcap_{j\in J(i)} D_j^{(i)},$$

which is the intersection of simple, closed and convex sets $D_j^{(i)}$ ($j \in J(i) := \{1,2,\ldots,m(i)\}$) (e.g., $D_j^{(i)}$ is a closed ball, a closed cone, or a half-space). Note that the projection $P_{D_j^{(i)}}$ can be computed within a finite number of arithmetic operations. User i then can use $T^{(i)} \colon H \to H$ defined by

$$T^{(i)} := \frac{1}{2}\left(\mathrm{Id} + \prod_{j\in J(i)} P_{D_j^{(i)}}\right), \tag{1.3}$$

where Id stands for the identity mapping on H. The mapping $T^{(i)}$ satisfies the firm nonexpansivity condition because $P_{D_j^{(i)}}$ ($j \in J(i)$) is nonexpansive. Moreover,

$$\mathrm{Fix}\left(T^{(i)}\right) = \mathrm{Fix}\left(\prod_{j\in J(i)} P_{D_j^{(i)}}\right) = C^{(i)}.$$

[1] In this paper, D ($\subset H$) is said to be simple when D is, for example, a closed ball, a closed cone, or a half-space, onto which the projection can be easily implemented. In the case where $P_{C^{(i)}}$ can be implemented, user i can use $T^{(i)} = P_{C^{(i)}}$ satisfying the firm nonexpansivity and $\mathrm{Fix}(T^{(i)}) = C^{(i)}$.

[2] $T \colon H \to H$ is said to be *firmly nonexpansive* [1], [11, Chapter 12], [12, Chapter 1-11] if $\|T(x) - T(y)\|^2 \le \langle x-y, T(x)-T(y)\rangle$ ($x,y \in H$). Firm nonexpansivity is stronger than the nonexpansivity (i.e., $\|T(x)-T(y)\| \le \|x-y\|$ ($x,y \in H$)). $T := (1/2)(\mathrm{Id} + S)$ satisfies the firm nonexpansivity condition when S is nonexpansive [2, Definition 4.1, Proposition 4.2].

Let us consider the problem of network bandwidth allocation and treat the example of $T^{(i)}$ represented in (1.3). The objective of bandwidth allocation is to share the available bandwidth among K traffic sources so as to maximize all the network's utility subject to the capacity constraints for all links [28, Chapter 2]. The capacity constraints for all links are absolute constraints that are expressed as a finite number of inequalities. Hence, the constraint set in this problem can be expressed as a polyhedral set. In general, none of the sources can get the explicit form of the constraint set because there is no source who knows the utilization situation of all links. When source i has only the explicit forms of the capacity constraints for links used by source i, $C^{(i)}$ can be expressed as the intersection of $\mathbb{R}_+^K$ and the sets with these capacity constraints, i.e., $C^{(i)} := \mathbb{R}_+^K \cap \bigcap_{j\in J(i)} D_j^{(i)}$, where $\mathbb{R}_+^K := \{(x_1, x_2, \ldots, x_K) \in \mathbb{R}^K : x_i \geq 0 \ (i = 1, 2, \ldots, K)\}$ and $D_j^{(i)}$ $(\subset \mathbb{R}^K)$ $(j \in J(i) := \{1, 2, \ldots, m(i)\})$ are half-spaces with the capacity constraints for links used by source i (see Section 5 for examples of such sets). Source i then can use a firmly nonexpansive mapping $T^{(i)} := (1/2)(\mathrm{Id} + P_{\mathbb{R}_+^K} \prod_{j\in J(i)} P_{D_j^{(i)}})$, which satisfies $\mathrm{Fix}(T^{(i)}) = C^{(i)}$.

Another application is when $C^{(i)}$ $(i \in I)$ is the set of all minimizers of a Fréchet differentiable and convex functional $g^{(i)}$ over a simple, closed and convex set $D^{(i)}$, i.e.,

$$C^{(i)} := \left\{ x \in D^{(i)} : g^{(i)}(x) = \min_{y \in D^{(i)}} g^{(i)}(y) \right\}.$$

When $\nabla g^{(i)} : H \to H$ is Lipschitz continuous, $P_{D^{(i)}}(\mathrm{Id} - \lambda \nabla g^{(i)})$ with an adequate λ (> 0) is nonexpansive [14, Proposition 2.3]. Hence, user i can use $T^{(i)} : H \to H$ defined by

$$T^{(i)} := \frac{1}{2}\left(\mathrm{Id} + P_{D^{(i)}}\left(\mathrm{Id} - \lambda \nabla g^{(i)}\right)\right), \tag{1.4}$$

which satisfies the firm nonexpansivity condition (see [34, Theorem 46.C (1) and (2)]).

$$\mathrm{Fix}\left(T^{(i)}\right) = \mathrm{Fix}\left(P_{D^{(i)}}\left(\mathrm{Id} - \lambda \nabla g^{(i)}\right)\right) = C^{(i)}$$

Let us consider the case where user i has simple, closed and convex sets $D^{(i)}, D_j^{(i)}$ $(\subset H)$ $(j \in J(i) := \{1, 2, \ldots, m(i)\})$, and the intersection of these sets are empty. When $D^{(i)}$ is the absolute set, it would be reasonable to deal with a user i's constraint set $C^{(i)}$ as a subset of $D^{(i)}$ with the elements closest to the sets $D_j^{(i)}$ in terms of the mean square norm. This subset is referred to as a *generalized convex feasible set* (see [8, section I, Framework 2], [31, Definition 4.1]). When $\Phi^{(i)}(x)$ stands for the mean square value[3] of the distances from $x \in H$ to $D_j^{(i)}$, the generalized convex feasible set is defined by

$$C_{\Phi^{(i)}} := \left\{ x \in D^{(i)} : \Phi^{(i)}(x) = \min_{y \in D^{(i)}} \Phi^{(i)}(y) \right\}.$$

[3] $\Phi^{(i)} : H \to \mathbb{R}$ is defined as follows: for $w_j^{(i)}$ (> 0) $(j \in J(i))$ satisfying $\sum_{j\in J(i)} w_j^{(i)} = 1$, $\Phi^{(i)}(x) := (1/2) \sum_{j\in J(i)} w_j^{(i)} (\min_{y \in D_j^{(i)}} \|x - y\|)^2$ $(x \in H)$. $\Phi^{(i)}$ is Fréchet differentiable and convex, and $\nabla \Phi^{(i)} = \mathrm{Id} - \sum_{j\in J(i)} w_j^{(i)} P_{D_j^{(i)}}$ is Lipschitz continuous.

The fact that $C_{\Phi^{(i)}} \neq \emptyset$ is guaranteed when at least one of $D^{(i)}$ and $D_j^{(i)}$ is bounded [31, Remark 4.3(a)]. The set $C_{\Phi^{(i)}}$ is well defined because it is the set of all minimizers of $\Phi^{(i)}$ over $D^{(i)}$. Moreover, $C_{\Phi^{(i)}} = D^{(i)} \cap \bigcap_{j \in J(i)} D_j^{(i)}$ holds when $D^{(i)} \cap \bigcap_{j \in J(i)} D_j^{(i)} \neq \emptyset$, which implies that $C_{\Phi^{(i)}}$ is a generalization of $D^{(i)} \cap \bigcap_{j \in J(i)} D_j^{(i)}$. User i can use a firmly nonexpansive mapping $T^{(i)}$ defined by (1.4) when $\nabla g^{(i)} = \nabla \Phi^{(i)} = \mathrm{Id} - \sum_{j \in J(i)} w_j^{(i)} P_{D_j^{(i)}}$ (see Footnote 3 for the definition of $\Phi^{(i)}$), i.e., $T^{(i)} := (1/2)(\mathrm{Id} + P_{D^{(i)}}((1-\lambda)\mathrm{Id} + \lambda \sum_{j \in J(i)} w_j^{(i)} P_{D_j^{(i)}}))$, which satisfies $\mathrm{Fix}(T^{(i)}) = C_{\Phi^{(i)}}$.

When we consider a minimization problem in which the constraint set composed of the absolute set and the subsidiary sets is not feasible, we can provide a meaningful optimal solution by using a minimizer of an objective function over the generalized convex feasible set, i.e., a compromise solution that satisfies the absolute constraints and tries to satisfy the subsidiary constraints as much as possible. Generalized convex feasible sets have been used to discuss real-world optimization problems that can be formulated as such an infeasible optimization problem, including inconsistent signal feasibility problems [8], power control problems in which the constraints about the sufficient signal-to-interference-plus-noise ratio fall in the infeasible region [15], bandwidth allocation problems in which the constraints about the preferable transmission rate fall in the infeasible region [16], and optimal control problems given unsolvable stochastic algebraic Riccati equations [18].

Here, we formulate Problem (1.1) under Assumptions (I), (II), and (III) into the following convex optimization problem over the intersection of *fixed point sets of nonexpansive mappings*:

$$\text{Minimize } f(x) := \sum_{i \in I} f^{(i)}(x) \text{ subject to } x \in \bigcap_{i \in I} \mathrm{Fix}\left(T^{(i)}\right). \tag{1.5}$$

Centralized optimization algorithms [6, 17, 31], that use all $T^{(i)}$s and $f^{(i)}$s, have been developed for solving Problem (1.5). The first algorithm developed for solving Problem (1.5) works when $\nabla f \colon H \to H$ is strongly monotone and Lipschitz continuous. It is the hybrid steepest descent method [31, 32]: given $x_n \in H$,

$$\begin{cases} d_n := -\nabla \left(\sum_{i \in I} f^{(i)} \right)(x_n), \\ x_{n+1} := \prod_{i \in I} T^{(i)}\left(x_n + \lambda_n d_n\right). \end{cases}$$

The algorithm, with a slowly diminishing sequence $(\lambda_n)_{n \in \mathbb{N}}$, strongly converges to a solution of Problem (1.5) [32, Theorem 2.15, Remark 2.17 (a)]. Some algorithms [6, 17] have been proposed to accelerate the hybrid steepest descent method. Reference [6] presented an effective algorithm for solving the signal recovery problem, and this algorithm strongly converges to a solution of Problem (1.5) without using a diminishing sequence. Reference [17] presented the hybrid conjugate gradient method defined by $x_{n+1} := \prod_{i \in I} T^{(i)}(x_n + \lambda_n d_n)$ and the *conjugate gradient direction* [25, Chapter 5],

$$d_n := -\nabla \left(\sum_{i \in I} f^{(i)} \right)(x_n) + \beta_n d_{n-1}, \tag{1.6}$$

where $\beta_n \geq 0$ ($n \in \mathbb{N}$), and proved that the algorithm strongly converges to a solution of Problem (1.5) if $(\lambda_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ are slowly diminishing sequences [17, Theorem 4.1]. The numerical examples in [17] demonstrate that the hybrid conjugate gradient method converges faster than the hybrid steepest descent method.

The main goal of this paper is to devise an incremental gradient method for solving Problem (1.5). The framework of the proposed algorithm can be obtained by replacing $P_C := P_{\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})}$ in Algorithm (1.2) with $T^{(i)}$ and by replacing $\nabla(\sum_{i\in I} f^{(i)})$ in direction (1.6) with $\nabla f^{(i)}$. More precisely, $x_n := x_n^{(0)} \in H$, $d_{n-1}^{(i)} \in H$ ($i \in I$),

$$\begin{cases} d_n^{(i)} := -\nabla f^{(i)}\left(x_n^{(i-1)}\right) + \beta_n d_{n-1}^{(i)}, \\ x_n^{(i)} := T^{(i)}\left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right) \qquad (i = 1, 2, \ldots, K), \\ x_{n+1} := x_n^{(K)}. \end{cases} \tag{1.7}$$

To guarantee that Algorithm (1.7) converges to a solution of Problem (1.5), we will modify Algorithm (1.7) by using the idea of the fixed point algorithms in [13, 30] and prove that the algorithm with slowly diminishing step-size sequences weakly converges to a solution of Problem (1.5).

Broadcast optimization algorithms [7, 9, 27] have been proposed for solving Problem (1.1) and have been applied to practical problems such as signal and image processing, and power allocation. From such a viewpoint, we will also present a broadcast type of distributed optimization algorithm that weakly converges to a solution of Problem (1.5). This analysis of the algorithm will help us to resolve resource allocation problems in future wireless and wired networks.

The fixed point theory for nonexpansive mappings defined on infinite-dimensional spaces [1], [2, Chapter 4], [11, Chapter 3], [12, Chapter 1] is an important area of Nonlinear Analysis, and it has played a crucial role in resolving complex real-world problems in Hilbert spaces, such as inconsistent signal feasibility problems [8], signal recovery problems [6, 10], inverse problems in signal and image processing [9], and optimal control problems [18]. Thanks to it, we can perform convergence of the distributed optimization algorithms for Problem (1.5) that includes these problems in Hilbert spaces. We believe that our convergence analyses will help us to develop conventional algorithms [6, 8, 9, 10, 18] in the Hilbert space setting and to resolve unsolved optimization problems in Hilbert spaces. It would be desirable to perform convergence analyses of algorithms in the Banach space setting because optimization problems in Banach spaces include practical problems in various disciplines (e.g., optimal flow control problems [29, Chapter 1]). Our convergence analyses may provide useful hints on devising algorithms for solving optimization problems in Banach spaces because our analyses are done in the infinite-dimensional Hilbert space setting.

This paper is organized as follows. Section 2 gives the convex minimization problem over the intersection of the fixed point sets of nonexpansive mappings and mathematical preliminaries. Section 3 devises the *incremental fixed point optimization algorithm* for solving the problem. We also prove that the algorithm with slowly diminishing step-size sequences converges weakly to a solution of the problem. Section 4 describes the *broadcast fixed point optimization algorithm* that weakly converges to a solution of the problem. Section 5 applies the algorithms to a network bandwidth allocation problem and provides numerical examples for network bandwidth allocation. Section 6 concludes the paper.

2. Assumptions, Problem Formulation, and Mathematical Preliminaries. Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$, and let $\mathbb{N}$ denote the set of all positive integers including zero. Consider a networked system which consists of K users and suppose that the following assumptions are satisfied in the network.

Assumption 2.1.

(A1) $T^{(i)} \colon H \to H$ $(i \in I)$ *is firmly nonexpansive*[4] *with* $\bigcap_{i \in I} \mathrm{Fix}(T^{(i)}) := \bigcap_{i \in I} \{x \in H \colon T^{(i)}(x) = x\} \neq \emptyset$. *The explicit form of* $T^{(i)}$ *is its own private information; that is, other users cannot get the explicit form of* $T^{(i)}$.

(A2) $f^{(i)} \colon H \to \mathbb{R}$ $(i \in I)$ *is strictly convex*[5] *and Fréchet differentiable, and* $\nabla f^{(i)} \colon H \to H$ $(i \in I)$ *is* $(1/L^{(i)})$*-Lipschitz continuous*[6]*. The explicit form of* $f^{(i)}$ *is its own private information; that is, other users cannot get the explicit form of* $f^{(i)}$.

This paper discusses the following problem with information on the whole network:

Problem 2.1. *Under Assumption 2.1, we are interested in*

$$\textit{minimizing } f(x) := \sum_{i \in I} f^{(i)}(x) \textit{ subject to } x \in \bigcap_{i \in I} \mathrm{Fix}\left(T^{(i)}\right),$$

where

(A3) $\mathrm{Argmin}_X f := \{x^\star \in X := \bigcap_{i \in I} \mathrm{Fix}\left(T^{(i)}\right) : f(x^\star) = \min_{x \in X} f(x)\} \neq \emptyset$.

Assumption (A3) is guaranteed when at least one of the fixed point sets, $\mathrm{Fix}(T^{(i)})$, is bounded [33, Theorem 25.C]. Under $\mathrm{Argmin}_X f \neq \emptyset$, the strict convexity of f guarantees the uniqueness of the solution of Problem 2.1 [33, Corollary 25.15]. The *convexity* of $f^{(i)}$ ensures that *subsequences* of $(x_n)_{n \in \mathbb{N}}$ which are generated by Algorithms 3.1 and 4.1 weakly converge to a solution of Problem 2.1 (Lemmata 3.3(i), (ii) and 4.3(i), (ii)). To guarantee that Algorithms 3.1 and 4.1 weakly converge to the solution, we will need the *strict convexity* condition of $f^{(i)}$ (Lemma 3.3 (iii) and Lemma 4.3 (iii)). Reference [28, Chapter 2] provides examples of strictly convex objective functions in network resource allocation (see also Section 5).

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

Proposition 2.1. [14, Proposition 2.3] *Let* $f \colon H \to \mathbb{R}$ *be convex and Fréchet differentiable, and let* $\nabla f \colon H \to H$ *be* $(1/L)$*-Lipschitz continuous. For* $\lambda \in [0, 2L]$, *we define* $S_\lambda \colon H \to H$ *for all* $x \in H$ *by* $S_\lambda(x) := x - \lambda \nabla f(x)$. *Then* S_λ *is nonexpansive, i.e.,* $\|S_\lambda(x) - S_\lambda(y)\| \leq \|x - y\|$ *for all* $x, y \in H$.

Proposition 2.2. [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,* $\lim_{n \to \infty} a_n = 0$.

3. Incremental Fixed Point Optimization Algorithm for Distributed Optimization. This section presents the following incremental optimization algorithm.

Algorithm 3.1 (Incremental Fixed Point Optimization Algorithm).

---

[4]See Footnote 2 for the definition of a firmly nonexpansive mapping. $\mathrm{Fix}(T) := \{x \in H \colon T(x) = x\}$ is closed and convex when T is nonexpansive [12, Proposition 5.3].

[5]$f \colon H \to \mathbb{R}$ is said to be *strictly convex* [2, Definition 8.6] if, for all $x, y \in H$ with $x \neq y$ and for all $\lambda \in (0, 1)$, $f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y)$.

[6]$A \colon H \to H$ is said to be *Lipschitz continuous* with $L > 0$ (L-Lipschitz continuous) if $\|A(x) - A(y)\| \leq L \|x - y\|$ $(x, y \in H)$.

Step 0. User i ($i \in I$) sets $x^{(i)} \in H$ arbitrarily, and sets $d_{-1}^{(i)} := -\nabla f^{(i)}(x^{(i)})$. User K sets $x_0 \in H$ arbitrarily and transmits $x_0^{(0)} := x_0 \in H$ to user 1.

Step 1. Given $x_n = x_n^{(0)} \in H$ and $d_{n-1}^{(i)} \in H$ ($i \in I$), user i computes $x_n^{(i)} \in H$ cyclically by

$$\begin{cases} d_n^{(i)} := -\nabla f^{(i)}\left(x_n^{(i-1)}\right) + \beta_n d_{n-1}^{(i)}, \\ y_n^{(i)} := T^{(i)}\left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right), \\ x_n^{(i)} := \alpha_n x^{(i)} + (1-\alpha_n) y_n^{(i)} \qquad (i = 1, 2, \ldots, K). \end{cases}$$

Step 2. User K defines $x_{n+1} \in H$ by

$$x_{n+1} := x_n^{(K)}$$

and transmits $x_{n+1}^{(0)} := x_{n+1}$ to user 1. Put $n := n+1$, and go to Step 1.

We assume that all users participating in the network know the following common information before they execute the algorithm.

ASSUMPTION 3.1. *User i ($i \in I$) uses the decreasing sequences, $(\lambda_n)_{n\in\mathbb{N}} \subset (0, 2\min_{i\in I} L^{(i)}]$, $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1]$, and $(\beta_n)_{n\in\mathbb{N}} \subset (0,1]$, which converge to 0 and satisfy the following conditions*[7]*:*

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

$$\text{(C4)} \lim_{n\to\infty} \frac{\alpha_n}{\lambda_n} = 0, \ \text{(C5)} \frac{\lambda_n}{\lambda_{n+1}} \leq \sigma \text{ for some } \sigma \geq 1, \ \text{(C6)} \lim_{n\to\infty} \frac{\beta_n}{\alpha_{n+1}} = 0.$$

Our convergence result in this section depends on the following assumption.

ASSUMPTION 3.2. *The sequence $(y_n^{(i)})_{n\in\mathbb{N}}$, $i \in I$, which is generated by Algorithm 3.1 is bounded.*

User i can choose in advance a simple, bounded, closed and convex set, $X^{(i)}$ (e.g., $X^{(i)}$ is a closed ball with a large enough radius) satisfying $X^{(i)} \supset \mathrm{Fix}(T^{(i)})$. Then, user i can compute

$$y_n^{(i)} := P_{X^{(i)}}\left(T^{(i)}\left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right)\right) \tag{3.1}$$

instead of $y_n^{(i)}$ in Algorithm 3.1. Since $(y_n^{(i)})_{n\in\mathbb{N}} \subset X^{(i)}$ and $X^{(i)}$ is bounded, $(y_n^{(i)})_{n\in\mathbb{N}}$ ($i \in I$) is bounded. Hence, we may assume that $(y_n^{(i)})_{n\in\mathbb{N}}$ ($i \in I$) in Algorithm 3.1 is as in Equation (3.1) in place of Assumption 3.2. Let us show that Assumption 3.2 is satisfied by replacing $x_n^{(i)}$ in Algorithm 3.1 with $y_n^{(i)} := T^{(i)}(x_n^{(i-1)} + \lambda_n d_n^{(i)})$ and

$$x_n^{(i)} := P_{X^{(i)}}\left(\alpha_n x^{(i)} + (1-\alpha_n) y_n^{(i)}\right).$$

Indeed, $(x_n^{(i)})_{n\in\mathbb{N}} \subset X^{(i)}$ ($i \in I$) is bounded. The boundedness of $(x_n^{(i)})_{n\in\mathbb{N}}$, the Lipschitz continuity of $\nabla f^{(i)}$, and the convergence of $(\beta_n)_{n\in\mathbb{N}}$ to 0 guarantee that

[7] Examples of $(\lambda_n)_{n\in\mathbb{N}}$, $(\alpha_n)_{n\in\mathbb{N}}$, and $(\beta_n)_{n\in\mathbb{N}}$ are $\lambda_n := (2\min_{i\in I} L^{(i)})/(n+1)^a$, $\alpha_n := 1/(n+1)^b$, and $\beta_n := 1/(n+2)^c$ ($a \in (0,1/2), b \in (a, 1-a), b < c$). If $L^{(i)}$s are known from the beginning, we can choose $\lambda_n \in (0, 2\min_{i\in I} L^{(i)}]$ satisfying Conditions (C2)–(C5). From the convergence of $(\lambda_n)_{n\in\mathbb{N}}$ to 0, we see that even if $\lambda_0 > 2\min_{i\in I} L^{(i)}$, there exists $m \in \mathbb{N}$ such that $\lambda_n \leq 2\min_{i\in I} L^{(i)}$ for all $n \geq m$; that is, $(\lambda_n)_{n\geq m} \subset (0, 2\min_{i\in I} L^{(i)}]$ satisfies Conditions (C2)–(C5).

$(d_n^{(i)})_{n\in\mathbb{N}}$ is bounded (for details, see the proof of Lemma 3.1). The nonexpansivity of $T^{(i)}$ implies that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$, for all $i \in I$, and for all $n \in \mathbb{N}$, $\|y_n^{(i)} - x\| = \|T^{(i)}(x_n^{(i-1)} + \lambda_n d_n^{(i)}) - T^{(i)}(x)\| \leq \|x_n^{(i-1)} + \lambda_n d_n^{(i)} - x\| \leq \|x_n^{(i-1)} - x\| + \lambda_n \|d_n^{(i)}\| \leq \|x_n^{(i-1)} - x\| + 2\min_{i\in I} L^{(i)} \|d_n^{(i)}\| < \infty$, which means the boundedness of $(y_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$.

Now let us show a convergence analysis of Algorithm 3.1.

THEOREM 3.1. *Under Assumptions 2.1, 3.1, and 3.2, the sequence* $(x_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$ *generated by Algorithm 3.1 weakly converges to the solution of Problem 2.1.*

Theorem 3.1 says that Algorithm 3.1 enables each user to find the solution of Problem 2.1 by using only its own private objective function and nonexpansive mapping and the transmitted information from the neighbor user.

Let us compare Algorithm 3.1 with the conventional incremental subgradient method. For convenience, we can rewrite the incremental subgradient method [4, Equations (8.9), (8.10), and (8.11)] (see also [24]): given $x_n := x_n^{(0)} \in \mathbb{R}^m$,

$$\begin{cases} x_n^{(i)} := P_C\left(x_n^{(i-1)} - \lambda_n g_n^{(i)}\right),\ g_n^{(i)} \in \partial f^{(i)}\left(x_n^{(i-1)}\right)\ (i = 1, 2, \ldots, K), \\ x_{n+1} := x_n^{(K)}. \end{cases} \tag{3.2}$$

Algorithm (3.2) can be used when $f^{(i)}$ $(i \in I)$ is convex and non-differentiable, and P_C can be easily implemented. References [4, Proposition 8.2.6] and [24, Proposition 2.4] describe that $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm (3.2) with $(\lambda_n)_{n\in\mathbb{N}}$ satisfying $\sum_{n=0}^{\infty} \lambda_n = \infty$ and $\sum_{n=0}^{\infty} \lambda_n^2 < \infty$ converges to a minimizer of f over C. For simplicity, we will consider Algorithm 3.1 with $\beta_n := 0$ $(n \in \mathbb{N})$: given $x_n := x_n^{(0)} \in H$,

$$\begin{cases} y_n^{(i)} := T^{(i)}\left(x_n^{(i-1)} - \lambda_n \nabla f^{(i)}\left(x_n^{(i-1)}\right)\right), \\ x_n^{(i)} := \alpha_n x^{(i)} + (1-\alpha_n) y_n^{(i)} \qquad (i = 1, 2, \ldots, K), \\ x_{n+1} := x_n^{(K)}. \end{cases} \tag{3.3}$$

Algorithm (3.3) is applicable when $f^{(i)}$ $(i \in I)$ is strictly convex and differentiable, and $P_{\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})}$ and $P_{\mathrm{Fix}(T^{(i)})}$ cannot be easily implemented. When user i has a complicated $C^{(i)}$, Algorithm (3.3) can be implemented by using the easily computable nonexpansive mapping $T^{(i)}$ such that $\mathrm{Fix}(T^{(i)}) = C^{(i)}$[8] (see Section 1 for examples of such mappings). Algorithm (3.3), in general, satisfies $y_n^{(i)}, x_n^{(i)} \notin \mathrm{Fix}(T^{(i)})$ $(n \in \mathbb{N})$, while Algorithm (3.2) satisfies $x_n^{(i)} = P_C(x_n^{(i-1)} - \lambda_n g_n^{(i)}) \in C$ $(n \in \mathbb{N})$. By using an iterative technique based on the convex combination of $x^{(i)}$ and $y_n^{(i)}$, which is used to solve fixed point problems [13, 30], we can prove that, if $(\lambda_n)_{n\in\mathbb{N}}$ and $(\alpha_n)_{n\in\mathbb{N}}$ satisfy Conditions (C1)–(C6), then $(x_n^{(i)})_{n\in\mathbb{N}}$ generated by Algorithm (3.3) weakly converges to an element in $\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$ and to the solution of Problem 2.1.

3.1. Proof of Theorem 3.1. We first prove the following lemma.

LEMMA 3.1. *Let* $(\lambda_n)_{n\in\mathbb{N}} \subset (0, \infty)$, $(\alpha_n)_{n\in\mathbb{N}} \subset (0, 1]$, *and* $(\beta_n)_{n\in\mathbb{N}} \subset (0, 1]$ *with* $\lim_{n\to\infty} \beta_n = 0$, *and suppose that* $\nabla f^{(i)} \colon H \to H$ $(i \in I)$ *is* $(1/L^{(i)})$*-Lipschitz continuous and Assumption 3.2 is satisfied. Then,* $(x_n^{(i)})_{n\in\mathbb{N}}$, $(\nabla f^{(i)}(x_n^{(i-1)}))_{n\in\mathbb{N}}$, $(d_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$, *and* $(x_n)_{n\in\mathbb{N}}$ *generated by Algorithm 3.1 are bounded.*

---

[8]When $C^{(i)}$ is simple, user i can use $T^{(i)} = P_{C^{(i)}}$ because $P_{C^{(i)}}$ is easily implemented and firmly nonexpansive and $\mathrm{Fix}(P_{C^{(i)}}) = C^{(i)}$.

Proof. Since $(y_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$ is bounded and $x_n^{(i)} := \alpha_n x^{(i)} + (1-\alpha_n) y_n^{(i)}$ $(i \in I, n \in \mathbb{N})$, $(x_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$ is also bounded. Moreover, $(x_n)_{n\in\mathbb{N}}$ is bounded from $x_{n+1} := x_n^{(K)}$ $(n \in \mathbb{N})$. The Lipschitz continuity of $\nabla f^{(i)}$ guarantees that $\|\nabla f^{(i)}(x_n^{(i-1)}) - \nabla f^{(i)}(x)\| \leq (1/L^{(i)})\|x_n^{(i-1)} - x\|$ for all $n \in \mathbb{N}$, for all $i \in I$, and for all $x \in H$. Hence, the boundedness of $(x_n^{(i-1)})_{n\in\mathbb{N}}$ ensures that $(\nabla f^{(i)}(x_n^{(i-1)}))_{n\in\mathbb{N}}$ $(i \in I)$ is bounded.

Since $\lim_{n\to\infty} \beta_n = 0$, there exists $n_1 \in \mathbb{N}$ such that $\beta_n \leq 1/2$ for all $n \geq n_1$. We put $M_1^{(i)} := \sup\{\|\nabla f^{(i)}(x_n^{(i-1)})\| \colon n \in \mathbb{N}\}$, $\bar{M}_1^{(i)} := \max\{M_1^{(i)}, \|d_{n_1}^{(i)}\|\}$ $(i \in I)$, and $\bar{M}_1 := \max_{i\in I} \bar{M}_1^{(i)} < \infty$. Then, we find that $\|d_{n_1}^{(i)}\| \leq 2\bar{M}_1$. The definition of $d_n^{(i)}$ implies that $\|d_{n+1}^{(i)}\| \leq \|\nabla f^{(i)}(x_{n+1}^{(i-1)})\| + \beta_{n+1}\|d_n^{(i)}\| \leq \bar{M}_1 + (1/2)\|d_n^{(i)}\|$ for all $n \geq n_1$ and for all $i \in I$. Fix $i \in I$ and suppose that $\|d_n^{(i)}\| \leq 2\bar{M}_1$ for some $n \geq n_1$. Then, we find that $\|d_{n+1}^{(i)}\| \leq \bar{M}_1 + (1/2)2\bar{M}_1 = 2\bar{M}_1$. Accordingly, induction guarantees that $\|d_n^{(i)}\| \leq 2\bar{M}_1$ for all $i \in I$ and for all $n \geq n_1$; i.e., $(d_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$ is bounded. □

Next, we have the following lemma.

LEMMA 3.2. *Suppose that $T^{(i)} \colon H \to H$ $(i \in I)$ is firmly nonexpansive with $\bigcap_{i\in I} \mathrm{Fix}(T^{(i)}) \neq \emptyset$, $f^{(i)} \colon H \to \mathbb{R}$ $(i \in I)$ is convex with $(1/L^{(i)})$-Lipschitz continuous gradient, and Assumption 3.2 is satisfied. Let $(\lambda_n)_{n\in\mathbb{N}} \subset (0, 2\min_{i\in I} L^{(i)}]$, $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1]$, and $(\beta_n)_{n\in\mathbb{N}} \subset (0,1]$ be the sequences in Assumption 3.1. Then the sequence $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 3.1 has the following properties:*

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

(ii) $\displaystyle\lim_{n\to\infty} \left\|x_n^{(i-1)} - y_n^{(i)}\right\| = 0$ $(i \in I)$;

(iii) $\displaystyle\lim_{n\to\infty} \left\|x_n - x_n^{(i-1)}\right\| = 0$ *and* $\displaystyle\lim_{n\to\infty} \left\|x_n - T^{(i)}(x_n)\right\| = 0$ $(i \in I)$.

Proof. (i) Lemma 3.1 ensures that $M_1 := \max_{i\in I}(\sup\{\|\nabla f^{(i)}(x_n^{(i-1)})\| \colon n \in \mathbb{N}\}) < \infty$ and $M_2 := \max_{i\in I}(\sup\{2\|d_n^{(i)}\| \colon n \in \mathbb{N}\}) < \infty$. Proposition 2.1 and the conditions, $\lambda_{n+1} \leq \lambda_n$ and $\beta_{n+1} \leq \beta_n$ $(n \in \mathbb{N})$, guarantee that, for all $i \in I$ and for all $n \geq 1$,

$$
\begin{aligned}
\left\|y_{n+1}^{(i)} - y_n^{(i)}\right\| &= \left\|T^{(i)}\left(x_{n+1}^{(i-1)} + \lambda_{n+1} d_{n+1}^{(i)}\right) - T^{(i)}\left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right)\right\| \\
&\leq \left\|\left(x_{n+1}^{(i-1)} + \lambda_{n+1} d_{n+1}^{(i)}\right) - \left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right)\right\| \\
&\leq \Big\| \left(x_{n+1}^{(i-1)} + \lambda_{n+1}\left(-\nabla f^{(i)}\left(x_{n+1}^{(i-1)}\right) + \beta_{n+1} d_n^{(i)}\right)\right) \\
&\quad - \left(x_n^{(i-1)} + \lambda_n\left(-\nabla f^{(i)}\left(x_n^{(i-1)}\right) + \beta_n d_{n-1}^{(i)}\right)\right) \Big\| \\
&= \Big\| \left(x_{n+1}^{(i-1)} - \lambda_{n+1}\nabla f^{(i)}\left(x_{n+1}^{(i-1)}\right)\right) - \left(x_n^{(i-1)} - \lambda_{n+1}\nabla f^{(i)}\left(x_n^{(i-1)}\right)\right) \\
&\quad + (\lambda_n - \lambda_{n+1})\nabla f^{(i)}\left(x_n^{(i-1)}\right) + \lambda_{n+1}\beta_{n+1} d_n^{(i)} - \lambda_n \beta_n d_{n-1}^{(i)} \Big\| \\
&\leq \left\|\left(x_{n+1}^{(i-1)} - \lambda_{n+1}\nabla f^{(i)}\left(x_{n+1}^{(i-1)}\right)\right) - \left(x_n^{(i-1)} - \lambda_{n+1}\nabla f^{(i)}\left(x_n^{(i-1)}\right)\right)\right\| \\
&\quad + |\lambda_n - \lambda_{n+1}| \left\|\nabla f^{(i)}\left(x_n^{(i-1)}\right)\right\| + \lambda_{n+1}\beta_{n+1}\left\|d_n^{(i)}\right\| + \lambda_n\beta_n\left\|d_{n-1}^{(i)}\right\| \\
&\leq \left\|x_{n+1}^{(i-1)} - x_n^{(i-1)}\right\| + M_1|\lambda_n - \lambda_{n+1}| + M_2\lambda_n\beta_n.
\end{aligned}
$$

Hence, for all $i \in I$ and for all $n \geq 1$,

$$
\begin{aligned}
&\left\| x_n^{(i)} - x_{n-1}^{(i)} \right\| = \left\| \left(\alpha_n x^{(i)} + (1-\alpha_n) y_n^{(i)} \right) - \left(\alpha_{n-1} x^{(i)} + (1-\alpha_{n-1}) y_{n-1}^{(i)} \right) \right\| \\
&= \left\| (1-\alpha_n) \left(y_n^{(i)} - y_{n-1}^{(i)} \right) + (\alpha_n - \alpha_{n-1}) \left(x^{(i)} - y_{n-1}^{(i)} \right) \right\| \\
&\leq (1-\alpha_n) \left\| y_n^{(i)} - y_{n-1}^{(i)} \right\| + |\alpha_n - \alpha_{n-1}| \left\| x^{(i)} - y_{n-1}^{(i)} \right\| \\
&\leq (1-\alpha_n) \left\{ \left\| x_n^{(i-1)} - x_{n-1}^{(i-1)} \right\| + M_1 |\lambda_n - \lambda_{n-1}| + M_2 \lambda_{n-1} \beta_{n-1} \right\} + M_3 |\alpha_n - \alpha_{n-1}| \\
&\leq (1-\alpha_n) \left\| x_n^{(i-1)} - x_{n-1}^{(i-1)} \right\| + M_1 |\lambda_n - \lambda_{n-1}| + M_2 \lambda_{n-1} \beta_{n-1} + M_3 |\alpha_n - \alpha_{n-1}|,
\end{aligned}
$$

where $M_3 := \max_{i \in I} (\sup\{\|x^{(i)} - y_n^{(i)}\| \colon n \in \mathbb{N}\}) < \infty$ from Assumption 3.2. Therefore, we find that, for all $n \geq 1$,

$$
\begin{aligned}
&\|x_{n+1} - x_n\| = \left\| x_n^{(K)} - x_{n-1}^{(K)} \right\| \\
&\leq (1-\alpha_n) \left\| x_n^{(K-1)} - x_{n-1}^{(K-1)} \right\| + M_1 |\lambda_n - \lambda_{n-1}| + M_2 \lambda_{n-1} \beta_{n-1} + M_3 |\alpha_n - \alpha_{n-1}| \\
&\leq (1-\alpha_n)^K \left\| x_n^{(0)} - x_{n-1}^{(0)} \right\| + K M_1 |\lambda_n - \lambda_{n-1}| + K M_2 \lambda_{n-1} \beta_{n-1} + K M_3 |\alpha_n - \alpha_{n-1}| \\
&\leq (1-\alpha_n) \|x_n - x_{n-1}\| + \hat{M}_1 |\lambda_n - \lambda_{n-1}| + \hat{M}_2 \lambda_{n-1} \beta_{n-1} + \hat{M}_3 |\alpha_n - \alpha_{n-1}|,
\end{aligned}
$$

where $\hat{M}_j := K M_j$ $(j = 1, 2, 3)$. Hence, Condition (C5) and $\lambda_n \leq 2L$ $(n \in \mathbb{N})$, where $L := \min_{i \in I} L^{(i)}$, guarantee that, for all $n \geq 1$,

$$
\begin{aligned}
&\frac{\|x_{n+1} - x_n\|}{\lambda_n} \leq (1-\alpha_n) \frac{\|x_n - x_{n-1}\|}{\lambda_n} + \hat{M}_1 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n} + \hat{M}_2 \frac{\lambda_{n-1}}{\lambda_n} \beta_{n-1} + \hat{M}_3 \frac{|\alpha_n - \alpha_{n-1}|}{\lambda_n} \\
&\quad \leq (1-\alpha_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + (1-\alpha_n) \left\{ \frac{\|x_n - x_{n-1}\|}{\lambda_n} - \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} \right\} \\
&\qquad + \hat{M}_1 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n} + \sigma \hat{M}_2 \beta_{n-1} + \hat{M}_3 \frac{|\alpha_n - \alpha_{n-1}|}{\lambda_n} \\
&\quad \leq (1-\alpha_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + M_4 (1-\alpha_n) \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| \\
&\qquad + \hat{M}_1 \frac{|\lambda_n - \lambda_{n-1}|}{\lambda_n} + \sigma \hat{M}_2 \beta_{n-1} + \hat{M}_3 \frac{|\alpha_n - \alpha_{n-1}|}{\lambda_n} \\
&\quad \leq (1-\alpha_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + M_4 \alpha_n \frac{1}{\alpha_n} \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| \\
&\qquad + 2L \hat{M}_1 \alpha_n \frac{1}{\alpha_n} \frac{|\lambda_n - \lambda_{n-1}|}{2L \lambda_n} + \sigma \hat{M}_2 \alpha_n \frac{\beta_{n-1}}{\alpha_n} + \hat{M}_3 \alpha_n \frac{1}{\alpha_n} \frac{|\alpha_n - \alpha_{n-1}|}{\lambda_n} \\
&\quad \leq (1-\alpha_n) \frac{\|x_n - x_{n-1}\|}{\lambda_{n-1}} + M_4 \alpha_n \frac{1}{\alpha_n} \left| \frac{1}{\lambda_n} - \frac{1}{\lambda_{n-1}} \right| \\
&\qquad + 2L \hat{M}_1 \alpha_n \frac{1}{\alpha_n} \left| \frac{1}{\lambda_{n-1}} - \frac{1}{\lambda_n} \right| + \sigma \hat{M}_2 \alpha_n \frac{\beta_{n-1}}{\alpha_n} + \hat{M}_3 \alpha_n \frac{1}{\lambda_n} \left| 1 - \frac{\alpha_{n-1}}{\alpha_n} \right|,
\end{aligned}
$$

where $M_4 := \sup\{\|x_{n+1} - x_n\| \colon n \in \mathbb{N}\} < \infty$. Proposition 2.2 and Conditions (C1), (C2), (C3), and (C6) ensure that

$$
\lim_{n \to \infty} \frac{\|x_{n+1} - x_n\|}{\lambda_n} = 0. \tag{3.4}
$$

Equation (3.4) and $\lim_{n\to\infty} \lambda_n = 0$ imply that

$$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0. \tag{3.5}$$

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

$$\begin{aligned}
&2\left\|y_n^{(i)} - x\right\|^2 = 2\left\|T^{(i)}\left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right) - T^{(i)}(x)\right\|^2 \\
&\leq 2\left\langle \left(x_n^{(i-1)} + \lambda_n d_n^{(i)}\right) - x, y_n^{(i)} - x\right\rangle \\
&= \left\|\left(x_n^{(i-1)} - x\right) + \lambda_n d_n^{(i)}\right\|^2 + \left\|y_n^{(i)} - x\right\|^2 - \left\|\left(x_n^{(i-1)} - y_n^{(i)}\right) + \lambda_n d_n^{(i)}\right\|^2,
\end{aligned}$$

which implies that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$, for all $n \in \mathbb{N}$, and for all $i \in I$,

$$\begin{aligned}
&\left\|y_n^{(i)} - x\right\|^2 \leq \left\|\left(x_n^{(i-1)} - x\right) + \lambda_n d_n^{(i)}\right\|^2 - \left\|\left(x_n^{(i-1)} - y_n^{(i)}\right) + \lambda_n d_n^{(i)}\right\|^2 \\
&= \left\|x_n^{(i-1)} - x\right\|^2 + 2\lambda_n \left\langle x_n^{(i-1)} - x, d_n^{(i)}\right\rangle - \left\|x_n^{(i-1)} - y_n^{(i)}\right\|^2 - 2\lambda_n \left\langle x_n^{(i-1)} - y_n^{(i)}, d_n^{(i)}\right\rangle \\
&\leq \left\|x_n^{(i-1)} - x\right\|^2 - \left\|x_n^{(i-1)} - y_n^{(i)}\right\|^2 + M_5 \lambda_n,
\end{aligned}$$

where $M_5 := \max_{i\in I}(\sup\{2|\langle y_n^{(i)} - x, d_n^{(i)}\rangle| \colon n \in \mathbb{N}\}) < \infty$. Hence, the convexity of $\|\cdot\|^2$ guarantees that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$, for all $n \in \mathbb{N}$, and for all $i \in I$,

$$\begin{aligned}
&\left\|x_n^{(i)} - x\right\|^2 = \left\|\alpha_n \left(x^{(i)} - x\right) + (1 - \alpha_n)\left(y_n^{(i)} - x\right)\right\|^2 \\
&\leq \alpha_n \left\|x^{(i)} - x\right\|^2 + (1 - \alpha_n)\left\|y_n^{(i)} - x\right\|^2 \\
&\leq \alpha_n \left\|x^{(i)} - x\right\|^2 + (1 - \alpha_n)\left\{\left\|x_n^{(i-1)} - x\right\|^2 - \left\|x_n^{(i-1)} - y_n^{(i)}\right\|^2 + M_5 \lambda_n\right\} \\
&\leq \alpha_n \left\|x^{(i)} - x\right\|^2 + \left\|x_n^{(i-1)} - x\right\|^2 - (1 - \alpha_n)\left\|x_n^{(i-1)} - y_n^{(i)}\right\|^2 + M_5 \lambda_n.
\end{aligned} \tag{3.6}$$

Accordingly, we find that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$ and for all $n \in \mathbb{N}$,

$$\begin{aligned}
&\|x_{n+1} - x\|^2 = \left\|x_n^{(K)} - x\right\|^2 \\
&\leq \alpha_n \left\|x^{(K)} - x\right\|^2 + \left\|x_n^{(K-1)} - x\right\|^2 - (1 - \alpha_n)\left\|x_n^{(K-1)} - y_n^{(K)}\right\|^2 + M_5 \lambda_n \\
&\leq \alpha_n \sum_{i\in I}\left\|x^{(i)} - x\right\|^2 + \left\|x_n^{(0)} - x\right\|^2 - (1 - \alpha_n)\sum_{i\in I}\left\|x_n^{(i-1)} - y_n^{(i)}\right\|^2 + K M_5 \lambda_n \\
&= \alpha_n \sum_{i\in I}\left\|x^{(i)} - x\right\|^2 + \|x_n - x\|^2 - (1 - \alpha_n)\sum_{i\in I}\left\|x_n^{(i-1)} - y_n^{(i)}\right\|^2 + \hat{M}_5 \lambda_n,
\end{aligned}$$

where $\hat{M}_5 := K M_5$. This inequality means that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$ and for all

$n \in \mathbb{N}$,

$$(1-\alpha_n)\sum_{i\in I}\left\|x_n^{(i-1)}-y_n^{(i)}\right\|^2 \leq \alpha_n \sum_{i\in I}\left\|x^{(i)}-x\right\|^2+\|x_n-x\|^2-\|x_{n+1}-x\|^2+\hat{M}_5\lambda_n$$

$$=\alpha_n \sum_{i\in I}\left\|x^{(i)}-x\right\|^2+(\|x_n-x\|+\|x_{n+1}-x\|)(\|x_n-x\|-\|x_{n+1}-x\|)+\hat{M}_5\lambda_n$$

$$\leq\alpha_n \sum_{i\in I}\left\|x_0^{(i)}-x\right\|^2+(\|x_n-x\|+\|x_{n+1}-x\|)\|x_n-x_{n+1}\|+\hat{M}_5\lambda_n.$$

Equation (3.5), the boundedness of $(x_n)_{n\in\mathbb{N}}$, and $\lim_{n\to\infty}\alpha_n=\lim_{n\to\infty}\lambda_n=0$ ensure that

$$\lim_{n\to\infty}\sum_{i\in I}\left\|x_n^{(i-1)}-y_n^{(i)}\right\|^2=0;\ \text{i.e.,}\ \lim_{n\to\infty}\left\|x_n^{(i-1)}-y_n^{(i)}\right\|=0\ (i\in I).$$

(iii) Since $\|x_n^{(i)}-y_n^{(i)}\|=\alpha_n\|x^{(i)}-y_n^{(i)}\|$ $(i\in I, n\in\mathbb{N})$ and $\lim_{n\to\infty}\alpha_n=0$, we find that $\lim_{n\to\infty}\|x_n^{(i)}-y_n^{(i)}\|=0$ $(i\in I)$. From $\|x_n-x_n^{(i-1)}\|=\|x_n^{(0)}-x_n^{(i-1)}\|\leq \|x_n^{(0)}-y_n^{(1)}\|+\|y_n^{(1)}-x_n^{(1)}\|+\cdots+\|x_n^{(i-2)}-y_n^{(i-1)}\|+\|y_n^{(i-1)}-x_n^{(i-1)}\|$, $\lim_{n\to\infty}\|x_n^{(i-1)}-y_n^{(i)}\|=0$, and $\lim_{n\to\infty}\|x_n^{(i)}-y_n^{(i)}\|=0$ $(i\in I)$, we have

$$\lim_{n\to\infty}\left\|x_n-x_n^{(i-1)}\right\|=0\ (i\in I). \tag{3.7}$$

From $\|x_n-y_n^{(i)}\|\leq\|x_n-x_n^{(i-1)}\|+\|x_n^{(i-1)}-y_n^{(i)}\|$, we find that

$$\lim_{n\to\infty}\left\|x_n-y_n^{(i)}\right\|=0\ (i\in I).$$

Moreover, since $\|y_n^{(i)}-T^{(i)}(x_n)\|=\|T^{(i)}(x_n^{(i-1)}+\lambda_n d_n^{(i)})-T^{(i)}(x_n)\|\leq\|x_n^{(i-1)}-x_n\|+\lambda_n\|d_n^{(i)}\|$, (3.7), and $\lim_{n\to\infty}\lambda_n=0$, we also have

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

Therefore, from $\|x_n-T^{(i)}(x_n)\|\leq\|x_n-y_n^{(i)}\|+\|y_n^{(i)}-T^{(i)}(x_n)\|$, $\lim_{n\to\infty}\|x_n-y_n^{(i)}\|=0$, and $\lim_{n\to\infty}\|y_n^{(i)}-T^{(i)}(x_n)\|=0$ $(i\in I)$, we find that

$$\lim_{n\to\infty}\left\|x_n-T^{(i)}(x_n)\right\|=0\ (i\in I). \tag{3.8}$$

This proves Lemma 3.2. □

Lemma 3.2 and the strict convexity of f lead us to the weak convergence of $(x_n^{(i)})_{n\in\mathbb{N}}$ $(i\in I)$ to the solution of Problem 2.1.

LEMMA 3.3. *Suppose that the assumptions in Lemma 3.2 are satisfied. Then, the following hold.*

(i) *There exists 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 $x^*\in\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$;*

(ii) *$x^*\in\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$ is the solution of Problem 2.1*[9]*;*

[9]Equation (3.7) and Items (i) and (ii) in Lemma 3.3 imply that $(x_{n_k}^{(i)})_{k\in\mathbb{N}}$ $(\subset (x_n^{(i)})_{n\in\mathbb{N}})$ $(i\in I)$ weakly converges to the minimizer of f over $\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$.

(iii) *if f is strictly convex, $(x_n^{(i)})_{n\in\mathbb{N}}$ $(i \in I)$ generated by Algorithm 3.1 weakly converges to x^*.*

Proof. (i) The boundedness of $(x_n)_{n\in\mathbb{N}}$ guarantees the existence of $(x_{n_k})_{k\in\mathbb{N}}$ $(\subset (x_n)_{n\in\mathbb{N}})$ such that $(x_{n_k})_{k\in\mathbb{N}}$ weakly converges to $x^* \in H$. Fix $i \in I$ arbitrarily and assume that $x^* \notin \mathrm{Fix}(T^{(i)})$. Then, the Opial's condition[10], (3.8), and the nonexpansivity of $T^{(i)}$ imply that

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

This is a contradiction. Therefore, $x^* \in \mathrm{Fix}(T^{(i)})$ $(i \in I)$; that is, $x^* \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$.

(ii) Let $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$ be fixed arbitrarily. We find from $\{\nabla f^{(i)}(x_n^{(i-1)})\} = \partial f^{(i)}(x_n^{(i-1)})$ that $f^{(i)}(x) \geq f^{(i)}(x_n^{(i-1)}) + \langle x - x_n^{(i-1)}, \nabla f^{(i)}(x_n^{(i-1)})\rangle$ $(n \in \mathbb{N}, i \in I)$. So, the nonexpansivity of $T^{(i)}$ guarantees that, for all $i \in I$ and for all $n \geq 1$,

$$\begin{aligned}
&\left\| y_n^{(i)} - x \right\|^2 \leq \left\| \left(x_n^{(i-1)} - x \right) + \lambda_n d_n^{(i)} \right\|^2 \\
&= \left\| x_n^{(i-1)} - x \right\|^2 + 2\lambda_n \left\langle x_n^{(i-1)} - x, d_n^{(i)} \right\rangle + \lambda_n^2 \left\| d_n^{(i)} \right\|^2 \\
&\leq \left\| x_n^{(i-1)} - x \right\|^2 + 2\lambda_n \left\langle x_n^{(i-1)} - x, -\nabla f^{(i)}\left(x_n^{(i-1)} \right) + \beta_n d_{n-1}^{(i)} \right\rangle + M_2 \lambda_n^2 \\
&\leq \left\| x_n^{(i-1)} - x \right\|^2 + 2\lambda_n \left\langle x - x_n^{(i-1)}, \nabla f^{(i)}\left(x_n^{(i-1)} \right) \right\rangle + M_6 \lambda_n \beta_n + M_2 \lambda_n^2 \\
&\leq \left\| x_n^{(i-1)} - x \right\|^2 + 2\lambda_n \left(f^{(i)}(x) - f^{(i)}\left(x_n^{(i-1)} \right) \right) + M_6 \lambda_n \beta_n + M_2 \lambda_n^2,
\end{aligned}$$

where $M_6 := \max_{i\in I}(\sup\{2|\langle x_{n+1}^{(i-1)} - x, d_n^{(i)}\rangle| \colon n \in \mathbb{N}\}) < \infty$. Hence, from (3.6) we have, for all $i \in I$ and for all $n \geq 1$,

$$\begin{aligned}
&\left\| x_n^{(i)} - x \right\|^2 \leq \alpha_n \left\| x^{(i)} - x \right\|^2 + (1 - \alpha_n) \left\| y_n^{(i)} - x \right\|^2 \\
&\leq \alpha_n \left\| x^{(i)} - x \right\|^2 + (1 - \alpha_n) \left\{ \left\| x_n^{(i-1)} - x \right\|^2 + 2\lambda_n \left(f^{(i)}(x) - f^{(i)}\left(x_n^{(i-1)} \right) \right) \right. \\
&\quad \left. + M_6 \lambda_n \beta_n + M_2 \lambda_n^2 \right\} \\
&\leq \alpha_n \left\| x^{(i)} - x \right\|^2 + \left\| x_n^{(i-1)} - x \right\|^2 + 2(1 - \alpha_n)\lambda_n \left(f^{(i)}(x) - f^{(i)}\left(x_n^{(i-1)} \right) \right) \\
&\quad + M_6 \lambda_n \beta_n + M_2 \lambda_n^2.
\end{aligned}$$

[10]Suppose that $(x_n)_{n\in\mathbb{N}}$ $(\subset H)$ weakly converges to $\hat{x} \in H$ and $\bar{x} \neq \hat{x}$. Then, the following condition, called the Opial's condition [26], is satisfied: $\liminf_{n\to\infty} \|x_n - \hat{x}\| < \liminf_{n\to\infty} \|x_n - \bar{x}\|$.

Therefore, from $x_{n+1} = x_n^{(K)} = x_{n+1}^{(0)}$ $(n \in \mathbb{N})$ we find that, for all $n \geq 1$,

$$\begin{aligned}\|x_{n+1} - x\|^2 \leq & \alpha_n \sum_{i \in I} \left\| x^{(i)} - x \right\|^2 + \|x_n - x\|^2 \\ & + 2(1-\alpha_n)\lambda_n \left(f(x) - \sum_{i \in I} f^{(i)}\left(x_n^{(i-1)}\right)\right) + \hat{M}_6 \lambda_n \beta_n + \hat{M}_2 \lambda_n^2 \\ = & \alpha_n \sum_{i \in I} \left\| x^{(i)} - x \right\|^2 + \|x_n - x\|^2 + \hat{M}_6 \lambda_n \beta_n + \hat{M}_2 \lambda_n^2 \\ & + 2(1-\alpha_n)\lambda_n \left(f(x) - f(x_n) + \sum_{i \in I} \left[f^{(i)}(x_n) - f^{(i)}\left(x_n^{(i-1)}\right)\right]\right),\end{aligned}$$

where $\hat{M}_6 := K M_6$. This inequality means that

$$\begin{aligned} 2(1-\alpha_n)(f(x_n) - f(x)) \leq & \frac{\alpha_n}{\lambda_n} \sum_{i \in I} \left\| x^{(i)} - x \right\|^2 + \frac{\|x_n - x\|^2 - \|x_{n+1} - x\|^2}{\lambda_n} \\ & + \hat{M}_6 \beta_n + \hat{M}_2 \lambda_n + 2(1-\alpha_n) \sum_{i \in I} \left[f^{(i)}(x_n) - f^{(i)}\left(x_n^{(i-1)}\right)\right] \\ = & \frac{\alpha_n}{\lambda_n} \sum_{i \in I} \left\| x^{(i)} - x \right\|^2 + \frac{(\|x_n - x\| + \|x_{n+1} - x\|)(\|x_n - x\| - \|x_{n+1} - x\|)}{\lambda_n} \\ & + \hat{M}_6 \beta_n + \hat{M}_2 \lambda_n + 2(1-\alpha_n) \sum_{i \in I} \left[f^{(i)}(x_n) - f^{(i)}\left(x_n^{(i-1)}\right)\right] \\ \leq & \frac{\alpha_n}{\lambda_n} \sum_{i \in I} \left\| x^{(i)} - x \right\|^2 + \frac{(\|x_n - x\| + \|x_{n+1} - x\|)\|x_n - x_{n+1}\|}{\lambda_n} \\ & + \hat{M}_6 \beta_n + \hat{M}_2 \lambda_n + 2(1-\alpha_n) \sum_{i \in I} \left[f^{(i)}(x_n) - f^{(i)}\left(x_n^{(i-1)}\right)\right]. \end{aligned} \tag{3.9}$$

On the other hand, for all $i \in I$ and for all $n \in \mathbb{N}$, we have

$$f^{(i)}(x_n) - f^{(i)}\left(x_n^{(i-1)}\right) \leq \left\langle x_n - x_n^{(i-1)}, \nabla f^{(i)}(x_n) \right\rangle \leq \left\| \nabla f^{(i)}(x_n) \right\| \left\| x_n - x_n^{(i-1)} \right\|,$$

which from the boundedness of $(\nabla f(x_n))_{n \in \mathbb{N}}$ and (3.7) implies that

$$\limsup_{n \to \infty} \left(f^{(i)}(x_n) - f^{(i)}\left(x_n^{(i-1)}\right)\right) \leq 0 \ (i \in I).$$

Hence, (3.9), (3.4), Condition (C4), and the convergence of $(\lambda_n)_{n \in \mathbb{N}}$, $(\alpha_n)_{n \in \mathbb{N}}$, and $(\beta_n)_{n \in \mathbb{N}}$ to 0 ensure that, for all $x \in \bigcap_{i \in I} \mathrm{Fix}(T^{(i)})$,

$$\limsup_{n \to \infty} (f(x_n) - f(x)) \leq 0.$$

This inequality, the weak convergence of $(x_{n_k})_{k \in \mathbb{N}}$ to $x^* \in \bigcap_{i \in I} \mathrm{Fix}(T^{(i)})$, and the convexity and continuity of f[11] guarantee that, for all $x \in \bigcap_{i \in I} \mathrm{Fix}(T^{(i)})$,

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

[11] When $f\colon H \to \mathbb{R}$ is convex, f is weakly lower semicontinuous if and only if f is lower semicontinuous [2, Theorem 9.1]. Since f in Problem 2.1 is convex and continuous, and $(x_{n_k})_{k \in \mathbb{N}}$ weakly converges to x^*, we have $f(x^*) \leq \liminf_{k \to \infty} f(x_{n_k})$.

i.e., $x^* \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$ is the solution of Problem 2.1.

(iii) Since $f\colon H \to \mathbb{R}$ is strictly convex, the uniqueness of the solution, denoted by $x^\star$, of Problem 2.1 is guaranteed. Hence, Lemma 3.3(i), (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 Lemma 3.3(i), (ii), we can prove that $(x_{n_l})_{l\in\mathbb{N}}$ also weakly converges to $x^\star$ with $\langle x^\star, w\rangle = \lim_{k\to\infty}\langle x_{n_k}, w\rangle$ $(w \in H)$; that is, any subsequence of $(x_n)_{n\in\mathbb{N}}$ weakly converges to $x^\star$. Hence, $(x_n)_{n\in\mathbb{N}} = (x_{n-1}^{(K)})_{n\in\mathbb{N}}$ weakly converges to $x^\star$. This implies from (3.7) that $(x_n^{(i-1)})_{n\in\mathbb{N}}$ $(i \in I)$ also converges weakly to $x^\star$. Therefore, we can conclude that $(x_n^{(i)})_{n\in\mathbb{N}}$ $(i\in I)$ weakly converges to the solution of Problem 2.1. □

3.2. Analysis of random incremental fixed point optimization algorithm. Let us analyze an incremental fixed point optimization algorithm where one user is randomly chosen from $I^{(j)}$ $(j = 0, 1, \ldots, K-1)$ defined as follows: we first choose $i^{(0)} \in I^{(0)} := I$ randomly. We then define $I^{(1)} := I^{(0)}\setminus\{i^{(0)}\}$ and choose $i^{(1)} \in I^{(1)}$ randomly. For $j = 2, 3, \ldots, K-1$, we set $I^{(j)} := I^{(j-1)}\setminus\{i^{(j-1)}\}$ and choose $i^{(j)} \in I^{(j)}$ randomly. We define $i^{(K)} := i^{(0)}$.

Algorithm 3.2 (Random Incremental Fixed Point Optimization Algorithm).

Step 0. User i $(i \in I)$ sets $x^{(i)} \in H$ arbitrarily, and sets $d_{-1}^{(i)} := -\nabla f^{(i)}(x^{(i)})$. User $i^{(0)}$ sets $x_0 \in H$ arbitrarily and transmits $x_0^{(i^{(0)})} := x_0$ to user $i^{(1)}$.

Step 1. Given $x_n := x_n^{(i^{(0)})} \in H$ and $d_{n-1}^{(i)} \in H$ $(i \in I)$, user $i^{(j)}$ $(j \in I)$ computes $x_n^{(i^{(j)})} \in H$ cyclically by

$$\begin{cases} d_n^{(i^{(j)})} := -\nabla f^{(i^{(j)})}\left(x_n^{(i^{(j-1)})}\right) + \beta_n d_{n-1}^{(i^{(j)})}, \\ y_n^{(i^{(j)})} := T^{(i^{(j)})}\left(x_n^{(i^{(j-1)})} + \lambda_n d_n^{(i^{(j)})}\right), \\ x_n^{(i^{(j)})} := \alpha_n x^{(i^{(j)})} + (1-\alpha_n) y_n^{(i^{(j)})} \qquad (j = 1, 2, \ldots, K). \end{cases}$$

Step 2. User $i^{(0)}$ $(= i^{(K)})$ defines $x_{n+1} \in H$ by

$$x_{n+1} := x_n^{(i^{(0)})}$$

and transmits $x_{n+1}^{(i^{(0)})} := x_{n+1}$ to user $i^{(1)}$. Put $n := n+1$, and go to Step 1.

Algorithm 3.2 when $i^{(j)} = j$ $(j \in I)$ coincides with Algorithm 3.1. We can prove a convergence analysis of Algorithm 3.2 by referring to the proof of Theorem 3.1 (Subsection 3.1).[12]

Theorem 3.2. *Assume that Assumptions 2.1 and 3.1 are satisfied, and the sequence $(y_n^{(i^{(j)})})_{n\in\mathbb{N}}$, $i^{(j)} \in I^{(j)}$, which is generated by Algorithm 3.2 is bounded. Then the sequence $(x_n^{(i^{(j)})})_{n\in\mathbb{N}}$ $(i^{(j)} \in I^{(j)})$ generated by Algorithm 3.2 weakly converges to the solution of Problem 2.1.*

When one user is randomly chosen from $I_n^{(j)} := I_n^{(j-1)}\setminus\{i_n^{(j-1)}\}$ which depends on n, in general, $I_n^{(j)} \ni i_n^{(j)} \neq i_{n+1}^{(j)} \in I_{n+1}^{(j)}$ holds, which implies $T^{(i_n^{(j)})} \neq T^{(i_{n+1}^{(j)})}$.

[12] We can obtain all formulas (e.g., (3.9)) in Subsection 3.1 that do not depend on $i \in I$ by using firmly nonexpansive mappings $T^{(i^{(j)})}$ $(j \in I)$. Therefore, we can prove Theorem 3.2 by referring to the proof of Theorem 3.1.

We cannot show in this case that Algorithm 3.2 weakly converges to the solution of Problem 2.1 because the proof of Theorem 3.1 uses essentially nonexpansive mappings, $T^{(i^{(j)})} = T^{(i_n^{(j)})}$ $(n \in \mathbb{N})$, which do not depend on n and satisfy $T^{(i_n^{(j)})} = T^{(i_{n+1}^{(j)})}$ $(n \in \mathbb{N})$. Hence, in the future, we should try to devise random incremental gradient methods that can be applied when user $i_n^{(j)}$ does not always coincide with user $i_{n+1}^{(j)}$.

4. Broadcast Fixed Point Optimization Algorithm for Distributed Optimization. In this section, we present the following broadcast type of distributed optimization algorithm which can be implemented under the assumption that each user can directly communicate with other users.[13]

ALGORITHM 4.1 (Broadcast Fixed Point Optimization Algorithm).

Step 0. User i ($i \in I$) transmits an arbitrarily chosen $x_0^{(i)} \in H$ to the all users, and computes $x_0 := (1/K)\sum_{i\in I} x_0^{(i)}$. User i sets $d_0^{(i)} := -\nabla f^{(i)}(x_0)$.

Step 1. Given $x_n, d_n^{(i)} \in H$, user i computes $x_{n+1}^{(i)} \in H$ by

$$\begin{cases} y_n^{(i)} := T^{(i)}\left(x_n + \lambda_n d_n^{(i)}\right), \\ x_{n+1}^{(i)} := \alpha_n x_0^{(i)} + (1-\alpha_n) y_n^{(i)} \end{cases}$$

and transmits $x_{n+1}^{(i)}$ to the all users.

Step 2. User i computes $x_{n+1} \in H$ and $d_{n+1}^{(i)} \in H$ by

$$\begin{aligned} x_{n+1} &:= \frac{1}{K}\sum_{i\in I} x_{n+1}^{(i)}, \\ d_{n+1}^{(i)} &:= -\nabla f^{(i)}(x_{n+1}) + \beta_{n+1} d_n^{(i)}. \end{aligned}$$

Put $n := n+1$, and go to Step 1.

In this section, we assume the following:

ASSUMPTION 4.1. *The sequence $(y_n^{(i)})_{n\in\mathbb{N}}$, $i \in I$, which is generated by Algorithm 4.1 is bounded.*

The same discussion as in Assumption 3.2 describing the existence of a simple, bounded, closed and convex set, $X^{(i)}$ $(i \in I)$, satisfying $X^{(i)} \supset \mathrm{Fix}(T^{(i)})$, leads us to Assumption 4.1 (for details, see Section 3).

THEOREM 4.1. *Under Assumptions 2.1, 3.1, and 4.1, the sequence $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 4.1 weakly converges to the solution of Problem 2.1.*

We can see from Theorem 4.1 that Algorithm 4.1 enables each user to solve Problem 2.1 by using only its private objective function and nonexpansive mapping and the transmitted information from all users.

Let us compare Algorithms 3.1 and 4.1. In Algorithm 4.1, when user i ($i \in I$) has $x_n^{(i)}$, each point is broadcast to all users. Then, user i computes $y_n^{(i)} := T^{(i)}(x_n + \lambda_n d_n^{(i)})$ by using $x_n := (1/K)\sum_{i\in I} x_n^{(i)}$. All users have $(x_n)_{n\in\mathbb{N}}$, which weakly converges to the solution of Problem 2.1 (see Theorem 4.1). Therefore, all users can solve Problem 2.1. Meanwhile, in Algorithm 3.1, $y_n^{(i)} := T^{(i)}(x_n^{(i-1)} + \lambda_n d_n^{(i)})$ uses $x_n^{(i-1)}$, which is the transmitted information from user $(i-1)$. User i in this case only computes $(x_n^{(i)})_{n\in\mathbb{N}}$, which weakly converges to the solution of Problem 2.1 (see Theorem 3.1). Hence, all users using Algorithm 3.1 can also solve Problem 2.1.

[13]This implies that all users have access to all information and can execute all steps in Algorithm 4.1.

4.1. Proof of Theorem 4.1. We omit the proof of the following result since it is similar to the proof of Lemma 3.1.

LEMMA 4.1. *Let $(\lambda_n)_{n\in\mathbb{N}} \subset (0,\infty)$, $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1]$, $(\beta_n)_{n\in\mathbb{N}} \subset (0,1]$ with $\lim_{n\to\infty}\beta_n = 0$, and suppose that $\nabla f^{(i)}\colon H \to H$ $(i \in I)$ is $(1/L^{(i)})$-Lipschitz continuous and Assumption 4.1 is satisfied. Then, the sequences $(x_n^{(i)})_{n\in\mathbb{N}}$, $(\nabla f^{(i)}(x_n))_{n\in\mathbb{N}}$, $(d_n^{(i)})_{n\in\mathbb{N}}$, $i \in I$, and $(x_n)_{n\in\mathbb{N}}$ which are generated by Algorithm 4.1 are bounded.*

LEMMA 4.2. *Suppose that $T^{(i)}\colon H \to H$ $(i \in I)$ is firmly nonexpansive with $\bigcap_{i\in I}\mathrm{Fix}(T^{(i)}) \neq \emptyset$, $f^{(i)}\colon H \to \mathbb{R}$ $(i \in I)$ is convex with $(1/L^{(i)})$-Lipschitz continuous gradient, and Assumption 4.1 is satisfied. Let $(\lambda_n)_{n\in\mathbb{N}} \subset (0, 2\min_{i\in I} L^{(i)}]$, $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1]$, and $(\beta_n)_{n\in\mathbb{N}} \subset (0,1]$ be the sequences in Assumption 3.1. Then, Algorithm 4.1 has the following properties:*

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

(ii) $\displaystyle\lim_{n\to\infty}\left\|x_n - y_n^{(i)}\right\| = 0$ $(i \in I)$;

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

Proof. (i) From Proposition 2.1 and the conditions, $\lambda_{n+1} \le \lambda_n$ and $\beta_{n+1} \le \beta_n$ $(n \in \mathbb{N})$ we find that, for all $i \in I$ and for all $n \ge 1$,

$$
\begin{aligned}
&\left\|y_{n+1}^{(i)} - y_n^{(i)}\right\| = \left\|T^{(i)}\left(x_{n+1} + \lambda_{n+1}d_{n+1}^{(i)}\right) - T^{(i)}\left(x_n + \lambda_n d_n^{(i)}\right)\right\|\\
\le&\left\|\left(x_{n+1} + \lambda_{n+1}d_{n+1}^{(i)}\right) - \left(x_n + \lambda_n d_n^{(i)}\right)\right\|\\
=&\left\|\left(x_{n+1} + \lambda_{n+1}\left(-\nabla f^{(i)}(x_{n+1}) + \beta_{n+1}d_n^{(i)}\right)\right) - \left(x_n + \lambda_n\left(-\nabla f^{(i)}(x_n) + \beta_n d_{n-1}^{(i)}\right)\right)\right\|\\
=&\Big\|\left(x_{n+1} - \lambda_{n+1}\nabla f^{(i)}(x_{n+1})\right) - \left(x_n - \lambda_{n+1}\nabla f^{(i)}(x_n)\right)\\
&+ (\lambda_n - \lambda_{n+1})\nabla f^{(i)}(x_n) + \lambda_{n+1}\beta_{n+1}d_n^{(i)} - \lambda_n\beta_n d_{n-1}^{(i)}\Big\|\\
\le&\left\|\left(x_{n+1} - \lambda_{n+1}\nabla f^{(i)}(x_{n+1})\right) - \left(x_n - \lambda_{n+1}\nabla f^{(i)}(x_n)\right)\right\|\\
&+ |\lambda_n - \lambda_{n+1}|\left\|\nabla f^{(i)}(x_n)\right\| + \lambda_{n+1}\beta_{n+1}\left\|d_n^{(i)}\right\| + \lambda_n\beta_n\left\|d_{n-1}^{(i)}\right\|\\
\le&\|x_{n+1} - x_n\| + N_1|\lambda_n - \lambda_{n+1}| + N_2\lambda_n\beta_n,
\end{aligned}
$$

where $N_1 := \max_{i\in I}(\sup\{\|\nabla f^{(i)}(x_n)\|\colon n \in \mathbb{N}\}) < \infty$ and $N_2 := \max_{i\in I}(\sup\{2\|d_n^{(i)}\|\colon n \in \mathbb{N}\}) < \infty$. Hence, for all $i \in I$ and for all $n \ge 1$,

$$
\begin{aligned}
&\left\|x_{n+1}^{(i)} - x_n^{(i)}\right\| = \left\|\left(\alpha_n x_0^{(i)} + (1-\alpha_n)y_n^{(i)}\right) - \left(\alpha_{n-1}x_0^{(i)} + (1-\alpha_{n-1})y_{n-1}^{(i)}\right)\right\|\\
=&\left\|(1-\alpha_n)\left(y_n^{(i)} - y_{n-1}^{(i)}\right) + (\alpha_n - \alpha_{n-1})\left(x_0^{(i)} - y_{n-1}^{(i)}\right)\right\|\\
\le&(1-\alpha_n)\left\|y_n^{(i)} - y_{n-1}^{(i)}\right\| + |\alpha_n - \alpha_{n-1}|\left\|x_0^{(i)} - y_{n-1}^{(i)}\right\|\\
\le&(1-\alpha_n)\left\{\|x_n - x_{n-1}\| + N_1|\lambda_n - \lambda_{n-1}| + N_2\lambda_{n-1}\beta_{n-1}\right\} + N_3|\alpha_n - \alpha_{n-1}|\\
\le&(1-\alpha_n)\|x_n - x_{n-1}\| + N_1|\lambda_n - \lambda_{n-1}| + N_2\lambda_{n-1}\beta_{n-1} + N_3|\alpha_n - \alpha_{n-1}|,
\end{aligned}
\tag{4.1}
$$

where $N_3 := \max_{i\in I}(\sup\{\|x_0^{(i)} - y_n^{(i)}\|\colon n \in \mathbb{N}\}) < \infty$. Moreover, for all $n \ge 1$, we have $\|x_{n+1} - x_n\| = \|(1/K)\sum_{i\in I}(x_{n+1}^{(i)} - x_n^{(i)})\| \le (1/K)\sum_{i\in I}\|x_{n+1}^{(i)} - x_n^{(i)}\|$, and hence, $K\|x_{n+1} - x_n\| \le \sum_{i\in I}\|x_{n+1}^{(i)} - x_n^{(i)}\|$. Accordingly, summing up (4.1) over all

i means that, for all $n \geq 1$, we have

$$\|x_{n+1} - x_n\| \leq (1-\alpha_n)\|x_n - x_{n-1}\| + N_1|\lambda_n - \lambda_{n-1}| + N_2\lambda_{n-1}\beta_{n-1} + N_3|\alpha_n - \alpha_{n-1}|.$$

A similar argument as in the proof of Lemma 3.2(i) leads us to

$$\lim_{n\to\infty} \frac{\|x_{n+1} - x_n\|}{\lambda_n} = 0. \tag{4.2}$$

From (4.2) and the convergence of $(\lambda_n)_{n\in\mathbb{N}}$ to 0 we obtain that

$$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0. \tag{4.3}$$

(ii) From the firm nonexpansivity of $T^{(i)}$ $(i \in I)$ and the equation, $2\langle x, y\rangle = \|x\|^2 + \|y\|^2 - \|x-y\|^2$ $(x, y \in H)$, we find that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$, for all $n \in \mathbb{N}$, and for all $i \in I$,

$$\begin{aligned} &2\left\|y_n^{(i)} - x\right\|^2 = 2\left\|T^{(i)}\left(x_n + \lambda_n d_n^{(i)}\right) - T^{(i)}(x)\right\|^2 \leq 2\left\langle \left(x_n + \lambda_n d_n^{(i)}\right) - x, y_n^{(i)} - x\right\rangle \\ &= \left\|(x_n - x) + \lambda_n d_n^{(i)}\right\|^2 + \left\|y_n^{(i)} - x\right\|^2 - \left\|\left(x_n - y_n^{(i)}\right) + \lambda_n d_n^{(i)}\right\|^2, \end{aligned}$$

which implies that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$, for all $n \in \mathbb{N}$, and for all $i \in I$, we have

$$\begin{aligned} &\left\|y_n^{(i)} - x\right\|^2 \leq \left\|(x_n - x) + \lambda_n d_n^{(i)}\right\|^2 - \left\|\left(x_n - y_n^{(i)}\right) + \lambda_n d_n^{(i)}\right\|^2 \\ &= \|x_n - x\|^2 + 2\lambda_n \left\langle x_n - x, d_n^{(i)}\right\rangle - \left\|x_n - y_n^{(i)}\right\|^2 - 2\lambda_n\left\langle x_n - y_n^{(i)}, d_n^{(i)}\right\rangle \\ &\leq \|x_n - x\|^2 - \left\|x_n - y_n^{(i)}\right\|^2 + N_4\lambda_n, \end{aligned} \tag{4.4}$$

where $N_4 := \max_{i\in I}(\sup\{2|\langle y_n^{(i)} - x, d_n^{(i)}\rangle| \colon n \in \mathbb{N}\}) < \infty$. Hence, the convexity of $\|\cdot\|^2$ guarantees that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$, for all $n \in \mathbb{N}$, and for all $i \in I$,

$$\begin{aligned} &\left\|x_{n+1}^{(i)} - x\right\|^2 = \left\|\alpha_n\left(x_0^{(i)} - x\right) + (1-\alpha_n)\left(y_n^{(i)} - x\right)\right\|^2 \\ &\leq \alpha_n\left\|x_0^{(i)} - x\right\|^2 + (1-\alpha_n)\left\|y_n^{(i)} - x\right\|^2 \\ &\leq \alpha_n\left\|x_0^{(i)} - x\right\|^2 + (1-\alpha_n)\left\{\|x_n - x\|^2 - \left\|x_n - y_n^{(i)}\right\|^2 + N_4\lambda_n\right\} \\ &\leq \alpha_n\left\|x_0^{(i)} - x\right\|^2 + \|x_n - x\|^2 - (1-\alpha_n)\left\|x_n - y_n^{(i)}\right\|^2 + N_4\lambda_n. \end{aligned} \tag{4.5}$$

On the other hand, the convexity of $\|\cdot\|^2$ means that, for all $x \in H$ and for all $n \in \mathbb{N}$,

$$\|x_{n+1} - x\|^2 = \left\|\frac{1}{K}\sum_{i\in I}\left(x_{n+1}^{(i)} - x\right)\right\|^2 \leq \frac{1}{K}\sum_{i\in I}\left\|x_{n+1}^{(i)} - x\right\|^2. \tag{4.6}$$

Summing up (4.5) over all i we get from (4.6) that, for all $x \in \bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$ and for all $n \in \mathbb{N}$,

$$\|x_{n+1} - x\|^2 \leq \frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + \|x_n - x\|^2 - \frac{1-\alpha_n}{K}\sum_{i\in I}\left\|x_n - y_n^{(i)}\right\|^2 + N_4\lambda_n,$$

which implies that

$$\begin{aligned}\frac{1-\alpha_n}{K}\sum_{i\in I}\left\|x_n - y_n^{(i)}\right\|^2 &\le \frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + \|x_n - x\|^2 - \|x_{n+1}-x\|^2 + N_4\lambda_n\\ =&\frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + (\|x_n - x\| + \|x_{n+1}-x\|)(\|x_n - x\| - \|x_{n+1}-x\|) + N_4\lambda_n\\ \le&\frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + (\|x_n - x\| + \|x_{n+1}-x\|)\|x_n - x_{n+1}\| + N_4\lambda_n.\end{aligned}$$

From (4.3), the boundedness of $(x_n)_{n\in\mathbb{N}}$, and $\lim_{n\to\infty}\alpha_n = \lim_{n\to\infty}\lambda_n = 0$ ensure us that

$$\lim_{n\to\infty}\sum_{i\in I}\left\|x_n - y_n^{(i)}\right\|^2 = 0;\ \text{i.e.,}\ \lim_{n\to\infty}\left\|x_n - y_n^{(i)}\right\| = 0\ (i\in I).$$

(iii) The nonexpansivity of $T^{(i)}$ implies that, for all $i\in I$ and for all $n\in\mathbb{N}$, $\|y_n^{(i)} - T^{(i)}(x_n)\| = \|T^{(i)}(x_n + \lambda_n d_n^{(i)}) - T^{(i)}(x_n)\| \le \lambda_n\|d_n^{(i)}\|$. Hence, the boundedness of $(d_n^{(i)})_{n\in\mathbb{N}}$ $(i\in I)$ and $\lim_{n\to\infty}\lambda_n = 0$ guarantee that $\lim_{n\to\infty}\|y_n^{(i)} - T^{(i)}(x_n)\| = 0$. Accordingly, we find from $\lim_{n\to\infty}\|x_n - y_n^{(i)}\| = 0$ and $\|x_n - T^{(i)}(x_n)\| \le \|x_n - y_n^{(i)}\| + \|y_n^{(i)} - T^{(i)}(x_n)\|$ $(i\in I, n\in\mathbb{N})$ that

$$\lim_{n\to\infty}\left\|x_n - T^{(i)}(x_n)\right\| = 0\ (i\in I). \tag{4.7}$$

This proves Lemma 4.2. □

LEMMA 4.3. *Suppose that the assumptions in Lemma 4.2 are satisfied. Then, the following hold.*

(i) *There exists 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 $x^*\in\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$;*

(ii) *$x^*\in\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$ is a solution of Problem 2.1;*

(iii) *if f is strictly convex, $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 4.1 weakly converges to x^*.*

Proof. (i) The boundedness of $(x_n)_{n\in\mathbb{N}}$ guarantees the existence of $(x_{n_k})_{k\in\mathbb{N}}$ $(\subset (x_n)_{n\in\mathbb{N}})$ such that $(x_{n_k})_{k\in\mathbb{N}}$ weakly converges to $x^*\in H$. Fix $i\in I$ arbitrarily and assume that $x^*\notin\mathrm{Fix}(T^{(i)})$. In the same manner as in the proof of Lemma 3.3(i), Opial's condition, (4.7), and the nonexpansivity of $T^{(i)}$ lead us to Lemma 4.3(i).

(ii) Let $x\in\bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$ be arbitrarily chosen. From (4.4) and the differentiability of $f^{(i)}$ $(i\in I)$ we have, for all $i\in I$ and for all $n\ge 1$, that

$$\begin{aligned}\left\|y_n^{(i)} - x\right\|^2 &\le \left\|(x_n - x) + \lambda_n d_n^{(i)}\right\|^2 = \|x_n - x\|^2 + 2\lambda_n\left\langle x_n - x, d_n^{(i)}\right\rangle + \lambda_n^2\left\|d_n^{(i)}\right\|^2\\ \le&\|x_n - x\|^2 + 2\lambda_n\left\langle x_n - x, -\nabla f^{(i)}(x_n) + \beta_n d_{n-1}^{(i)}\right\rangle + N_2\lambda_n^2\\ \le&\|x_n - x\|^2 + 2\lambda_n\left\langle x - x_n, \nabla f^{(i)}(x_n)\right\rangle + N_5\lambda_n\beta_n + N_2\lambda_n^2\\ \le&\|x_n - x\|^2 + 2\lambda_n\left(f^{(i)}(x) - f^{(i)}(x_n)\right) + N_5\lambda_n\beta_n + N_2\lambda_n^2,\end{aligned}$$

where $N_5 := \max_{i\in I}(\sup\{2|\langle x_{n+1} - x, d_n^{(i)}\rangle| \colon n \in \mathbb{N}\}) < \infty$. Hence, for all $n \geq 1$, that

$$\frac{1}{K}\sum_{i\in I}\left\|y_n^{(i)} - x\right\|^2 \leq \|x_n - x\|^2 + \frac{2\lambda_n}{K}\sum_{i\in I}\left(f^{(i)}(x) - f^{(i)}(x_n)\right) + N_5\lambda_n\beta_n + N_2\lambda_n^2$$
$$\leq \|x_n - x\|^2 + \frac{2\lambda_n}{K}\left(f(x) - f(x_n)\right) + N_5\lambda_n\beta_n + N_2\lambda_n^2.$$

Accordingly, from (4.5) and (4.6) we obtain, for all $n \geq 1$, that

$$\|x_{n+1} - x\|^2 \leq \frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + \frac{1-\alpha_n}{K}\sum_{i\in I}\left\|y_n^{(i)} - x\right\|^2$$
$$\leq \frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + (1-\alpha_n)\left\{\|x_n - x\|^2 + \frac{2\lambda_n}{K}\left(f(x) - f(x_n)\right) + N_5\lambda_n\beta_n + N_2\lambda_n^2\right\}$$
$$\leq \frac{\alpha_n}{K}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + \|x_n - x\|^2 + \frac{2\lambda_n(1-\alpha_n)}{K}\left(f(x) - f(x_n)\right) + N_5\lambda_n\beta_n + N_2\lambda_n^2.$$

Therefore, we get, for all $n \geq 1$, that

$$\frac{2(1-\alpha_n)}{K}\left(f(x_n) - f(x)\right)$$
$$\leq \frac{\alpha_n}{K\lambda_n}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + \frac{\|x_n - x\|^2 - \|x_{n+1} - x\|^2}{\lambda_n} + N_5\beta_n + N_2\lambda_n$$
$$\leq \frac{\alpha_n}{K\lambda_n}\sum_{i\in I}\left\|x_0^{(i)} - x\right\|^2 + \frac{(\|x_n - x\| + \|x_{n+1} - x\|)\|x_n - x_{n+1}\|}{\lambda_n} + N_5\beta_n + N_2\lambda_n.$$

Hence, Condition (C4), (4.2), and $\lim_{n\to\infty}\alpha_n = \lim_{n\to\infty}\lambda_n = \lim_{n\to\infty}\beta_n = 0$ ensure us that, for all $x \in \bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$,

$$\limsup_{n\to\infty}\left(f(x_n) - f(x)\right) \leq 0. \tag{4.8}$$

From (4.8), the weak convergence of $(x_{n_k})_{k\in\mathbb{N}}$ to $x^* \in \bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$, and the convexity and continuity of f guarantee that, for all $x \in \bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$, we deduce that

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

i.e., $x^* \in \bigcap_{i\in I}\mathrm{Fix}(T^{(i)})$ is a solution of Problem 2.1.

(iii) Since f is strictly convex, the solution of Problem 2.1 is unique. Therefore, in the same manner as in the proof of Lemma 3.3(iii), we can prove the weak convergence of $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 4.1 to the solution of Problem 2.1. □

5. Numerical Examples. Let us apply Algorithms 3.1, 3.2, and 4.1 to the network bandwidth allocation problem. The objective of utility-based bandwidth allocation is to share the available bandwidth among traffic sources so as to maximize the overall utility subject to the capacity constraints [28, Chapter 2]. The utility function of source i (user i) is defined for all $x \in \mathbb{R}_+$ as follows [28, Equation (2.4)]: given $w^{(i)} > 0$ and $a^{(i)} > 0$, we define

$$\mathcal{U}^{(i)}(x) := \begin{cases} w^{(i)}\log x \ \left(a^{(i)} = 1\right), \\ w^{(i)}\dfrac{x^{1-a^{(i)}}}{1-a^{(i)}} \ \left(a^{(i)} \neq 1\right). \end{cases}$$

Since the utility function is strictly concave, $f^{(i)} := -\mathcal{U}^{(i)}$ is strictly convex and continuously differentiable. We assume that source i has its own private $f^{(i)} := -\mathcal{U}^{(i)}$ and $C^{(i)}$ with the capacity constraints for links used by source i.

Consider the following network bandwidth allocation problem on a network [28, Fig.2.2] (see Figure 5.1) that consists of three links and four sources:

$$\text{Maximize } \mathcal{U}(x) := \sum_{i\in I} \mathcal{U}^{(i)}(x_i) \text{ subject to } x \in \bigcap_{i \in I} C^{(i)}, \tag{5.1}$$

where $I := \{1,2,3,4\}$, $\mathcal{U}^{(1)}(x) := \log x$, $\mathcal{U}^{(2)}(x) := 2\log x$, $\mathcal{U}^{(3)}(x) := 0.5^{-1}x^{0.5}$, $\mathcal{U}^{(4)}(x) := 0.8^{-1}x^{0.8}$ $(x \in \mathbb{R}_+)$, $D^{(1)} := \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \colon x_1 + x_3 \leq c_1\}$[14], $D^{(2)} := \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \colon x_2 + x_3 \leq c_2\}$, $D^{(3)} := \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \colon x_2 + x_4 \leq c_3\}$, $C^{(1)} := \mathbb{R}^4_+ \cap D^{(1)}$, $C^{(2)} := \mathbb{R}^4_+ \cap D^{(2)} \cap D^{(3)}$, $C^{(3)} := \mathbb{R}^4_+ \cap D^{(1)} \cap D^{(2)}$, $C^{(4)} := \mathbb{R}^4_+ \cap D^{(3)}$.

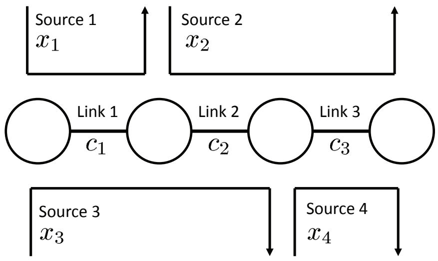

FIG. 5.1. *Network with three links and four sources*

To apply Algorithms 3.1, 3.2, and 4.1 to Problem (5.1), we define $T^{(i)} \colon \mathbb{R}^4 \to \mathbb{R}^4$ $(i \in I)$ by

$$T^{(1)} := \frac{1}{2}\left(\mathrm{Id} + P_{\mathbb{R}^4_+} P_{D^{(1)}}\right), \; T^{(2)} := \frac{1}{2}\left(\mathrm{Id} + P_{\mathbb{R}^4_+} P_{D^{(2)}} P_{D^{(3)}}\right),$$
$$T^{(3)} := \frac{1}{2}\left(\mathrm{Id} + P_{\mathbb{R}^4_+} P_{D^{(1)}} P_{D^{(2)}}\right), \; T^{(4)} := \frac{1}{2}\left(\mathrm{Id} + P_{\mathbb{R}^4_+} P_{D^{(3)}}\right),$$

which satisfy

$$\begin{aligned}
&\bigcap_{i\in I} \mathrm{Fix}\left(T^{(i)}\right) \\
=&\mathrm{Fix}\left(P_{\mathbb{R}^4_+} P_{D^{(1)}}\right) \cap \mathrm{Fix}\left(P_{\mathbb{R}^4_+} P_{D^{(2)}} P_{D^{(3)}}\right) \cap \mathrm{Fix}\left(P_{\mathbb{R}^4_+} P_{D^{(1)}} P_{D^{(2)}}\right) \cap \mathrm{Fix}\left(P_{\mathbb{R}^4_+} P_{D^{(3)}}\right) \\
=&\left(\mathbb{R}^4_+ \cap D^{(1)}\right) \cap \left(\mathbb{R}^4_+ \cap D^{(2)} \cap D^{(3)}\right) \cap \left(\mathbb{R}^4_+ \cap D^{(1)} \cap D^{(2)}\right) \cap \left(\mathbb{R}^4_+ \cap D^{(3)}\right) \\
=&\mathbb{R}^4_+ \cap \bigcap_{i=1}^{3} D^{(i)} = \bigcap_{i\in I} C^{(i)} \neq \emptyset.
\end{aligned}$$

We can see from the fact that $\bigcap_{i\in I} \mathrm{Fix}(T^{(i)}) = \mathbb{R}^4_+ \cap \bigcap_{i=1}^{3} D^{(i)}$ that any point in $\bigcap_{i\in I} \mathrm{Fix}(T^{(i)})$ satisfies the capacity constraints for all links. We set a closed ball, X

[14] The projection onto $D := \{x \in \mathbb{R}^K : \langle a, x\rangle \leq c\}$, where $a \ (\neq 0) \in \mathbb{R}^K$ and $c \in \mathbb{R}$, is expressed as follows [1, p.406], [2, Subchapter 28.3]: $P_D(x) := x + [(c - \langle a, x\rangle)/\|a\|^2]x$ $(x \notin D)$, or x $(x \in D)$.

($\supset \mathrm{Fix}(T^{(i)})$ $(i \in I)$), with a large enough radius and use Algorithms 3.1, 3.2, and 4.1 with $y_n^{(i)} := P_X(T^{(i)}(x_n^{(i-1)} + \lambda_n d_n^{(i)}))$, $y_n^{(i^{(j)})} := P_X(T^{(i^{(j)})}(x_n^{(i^{(j-1)})} + \lambda_n d_n^{(i^{(j)})}))$, and $y_n^{(i)} := P_X(T^{(i)}(x_n + \lambda_n d_n^{(i)}))$ $(n \in \mathbb{N}, i, j \in I, i^{(1)} := 2, i^{(2)} := 4, i^{(3)} := 3, i^{(0)} = i^{(4)} := 1)$, respectively. We use $\lambda_n := 10^{-3}/(n+1)^a$, $\alpha_n := 1/(n+1)^b$, and $\beta_n := 1/(n+2)^c$ $(a \in (0, 1/2), b \in (a, 1-a), c > b)$ which satisfying Conditions (C1)–(C6). Theorems 3.1, 3.2, and 4.1 guarantee that Algorithms 3.1, 3.2, and 4.1 in the above case converge to the solution of Problem (5.1). To compare the distributed optimization algorithms with a centralized optimization algorithm[15], we use the hybrid conjugate gradient method (HCGM) [17] defined by $x_{n+1} := T^{(1)}T^{(2)}T^{(3)}T^{(4)}(x_n + \lambda_n d_n)$ and $d_{n+1} := -\nabla(\sum_{i \in I} f^{(i)})(x_{n+1}) + \beta_{n+1} d_n$ $(n \in \mathbb{N})$, where $x_0 \in \mathbb{R}^4$ and $d_0 := -\nabla(\sum_{i \in I} f^{(i)})(x_0)$.

We set $c_1 := 5$, $c_2 := 4$, $c_3 := 5$, and $x := x_0 = x^{(i)} = x_0^{(i)} = x_0^{(0)} = x_0^{(i^{(0)})}$ $(i \in I)$ in Algorithms 3.1, 3.2, and 4.1, and HCGM. We selected one hundred random points $x = x(k)$ $(k = 1, 2, \ldots, 100)$ and executed Algorithms 3.1, 3.2, and 4.1, and HCGM for these points. Let $x(k)$ be one of the randomly selected points and let $(x_n(k))_{n \in \mathbb{N}}$ be the sequence generated by $x(k)$ and one of Algorithms 3.1, 3.2, 4.1, and HCGM. To check whether Algorithms 3.1, 3.2, and 4.1, and HCGM converge to a point in $\bigcap_{i \in I} C^{(i)} = \bigcap_{i \in I} \mathrm{Fix}(T^{(i)})$, we employed the following evaluation functions[16]: $D_n(k) := \sum_{i \in I} \|x_n(k) - T^{(i)}(x_n(k))\|$ $(k = 1, 2, \ldots, 100, n \in \mathbb{N})$ and $D_n := (1/100) \sum_{k=1}^{100} D_n(k)$ $(n \in \mathbb{N})$. We also employed $x_{n,j} := (1/100) \sum_{k=1}^{100} x_n(k)_j$ $(j \in I, n \in \mathbb{N})$, where $x_n(k) = (x_n(k)_j)_{j=1,2,3,4}$. The computer used in the experiment had an Intel Boxed Core i7 i7-870 2.93 GHz 8 M CPU and 8 GB of memory. The language was MATLAB 7.13.

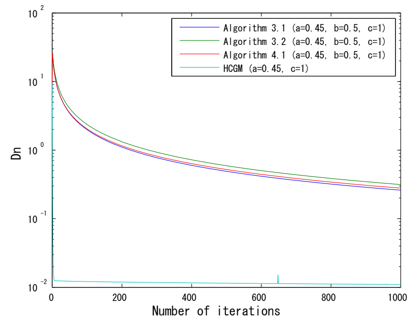

FIG. 5.2. *Behavior of D_n for Algorithms 3.1, 3.2, and 4.1, and HCGM when $a = 0.45$, $b = 0.5$, and $c = 1$*

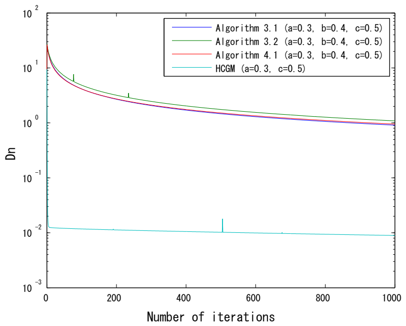

FIG. 5.3. *Behavior of D_n for Algorithms 3.1, 3.2, and 4.1, and HCGM when $a = 0.3$, $b = 0.4$, and $c = 0.5$*

Figures 5.2 and 5.3 indicate the behaviors of D_n for Algorithms 3.1, 3.2, and 4.1, and HCGM. These figures shows that the $(D_n)_{n \in \mathbb{N}}$s generated by these algorithms converge to 0; i.e., the algorithms converge to a point in $\bigcap_{i \in I} C^{(i)} = \bigcap_{i \in I} \mathrm{Fix}(T^{(i)})$, HCGM converges to the point fastest, and Algorithms 3.1, 3.2, and 4.1 when $a = 0.45$, $b = 0.5$, and $c = 1$ converge to the point faster than they do when $a = 0.3$, $b = 0.4$, and

[15] Although there are well known centralized optimization algorithms [6, 31, 32] for Problem (5.1), we apply HCGM, which is the basis for devising Algorithms 3.1, 3.2, and 4.1 (see Section 1), to Problem (5.1), and see how Algorithms 3.1, 3.2, and 4.1, and HCGM with the same $(\lambda_n)_{n \in \mathbb{N}}$ and $(\beta_n)_{n \in \mathbb{N}}$ work.

[16] $x \in \mathbb{R}^4$ satisfies $\sum_{i \in I} \|x - T^{(i)}(x)\| = 0$ if and only if $x \in \mathrm{Fix}(T^{(i)})$ $(i \in I)$, i.e., $x \in \bigcap_{i \in I} \mathrm{Fix}(T^{(i)}) = \mathbb{R}_+^4 \cap \bigcap_{i=1}^3 D^{(i)} = \bigcap_{i \in I} C^{(i)}$.

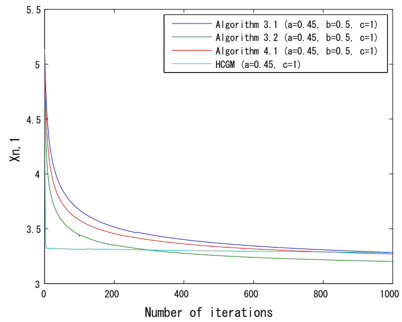

FIG. 5.4. *Behavior of* $x_{n,1}$ *for Algorithms 3.1, 3.2, and 4.1, and HCGM when* $a = 0.45$, $b = 0.5$, *and* $c = 1$

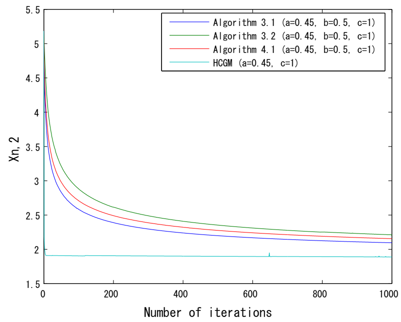

FIG. 5.5. *Behavior of* $x_{n,2}$ *for Algorithms 3.1, 3.2, and 4.1, and HCGM when* $a = 0.45$, $b = 0.5$, *and* $c = 1$

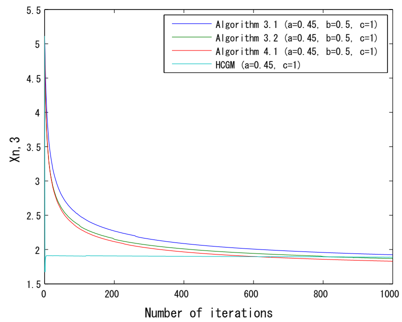

FIG. 5.6. *Behavior of* $x_{n,3}$ *for Algorithms 3.1, 3.2, and 4.1, and HCGM when* $a = 0.45$, $b = 0.5$, *and* $c = 1$

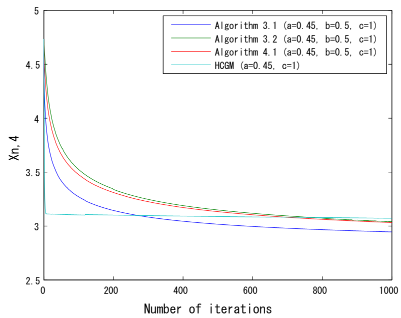

FIG. 5.7. *Behavior of* $x_{n,4}$ *for Algorithms 3.1, 3.2, and 4.1, and HCGM when* $a = 0.45$, $b = 0.5$, *and* $c = 1$

$c = 0.5$. Figures 5.4–5.7 show the behaviors of $x_{n,j}$ $(j \in I)$ when $a = 0.45$, $b = 0.5$, and $c = 1$. Figures 5.8–5.11 show the behaviors of $x_{n,j}$ $(j \in I)$ when $a = 0.3$, $b = 0.4$, and $c = 0.5$. These figures show that, although the behaviors of the distributed optimization algorithms differ depending on the choice of the step-size sequences, the different $(x_{n,j})_{n\in\mathbb{N}}$ $(j \in I)$ generated by the algorithms converge to the same point.

6. Conclusion and Future Work. We discussed the problem of minimizing the sum of all users' objective functions over the intersection of all users' constraint sets in a Hilbert space and presented two distributed fixed point optimization algorithms for solving the problem. One algorithm is based on conventional incremental subgradient methods, and the other is a broadcast type of distributed optimization algorithm. The algorithms use easily implementable nonexpansive mappings of which the intersection of the fixed point sets is equal to the intersection of all users' constraint sets. They can be applied to the problem when the projection onto each user's constraint set cannot be easily implemented. We showed that the algorithms with slowly diminishing step-size sequences weakly converge to the solution of the problem. Finally, we gave numerical results to support the convergence analyses on the algorithms.

In the future, we should consider developing distributed optimization algorithms for solving minimization problems in which all users' objective functions are nonconvex

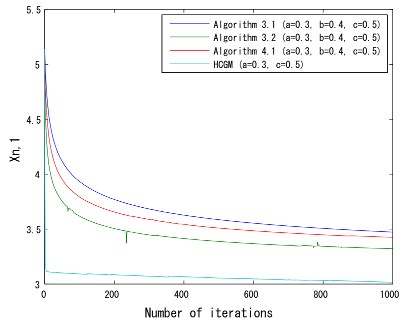

Fig. 5.8. *Behavior of $x_{n,1}$ for Algorithms 3.1, 3.2, and 4.1, and HCGM when $a = 0.3$, $b = 0.4$, and $c = 0.5$*

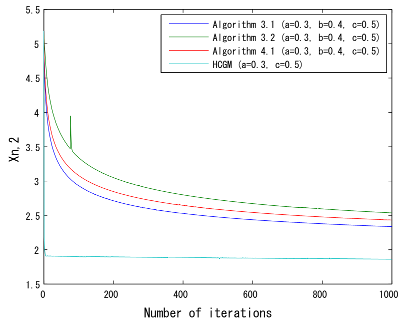

Fig. 5.9. *Behavior of $x_{n,2}$ for Algorithms 3.1, 3.2, and 4.1, and HCGM when $a = 0.3$, $b = 0.4$, and $c = 0.5$*

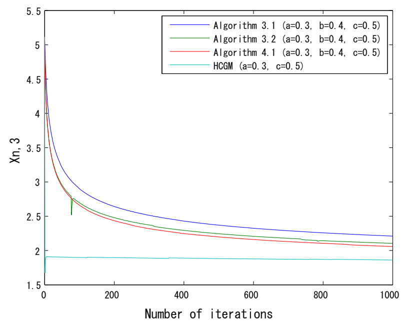

Fig. 5.10. *Behavior of $x_{n,3}$ for Algorithms 3.1, 3.2, and 4.1, and HCGM when $a = 0.3$, $b = 0.4$, and $c = 0.5$*

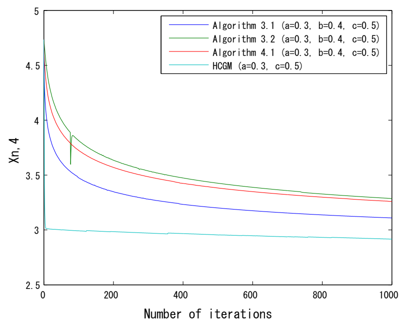

Fig. 5.11. *Behavior of $x_{n,4}$ for Algorithms 3.1, 3.2, and 4.1, and HCGM when $a = 0.3$, $b = 0.4$, and $c = 0.5$*

(for example, the signal-to-interference-plus-noise ratio, which is used to evaluate the performance of each user in a wireless network, is not concave). We also need to devise incremental fixed point optimization algorithms which work where one user is randomly chosen at any time.

Acknowledgments I am sincerely grateful to Associate Editor Patrick L. Combettes and the two anonymous referees for helping me improve the original manuscript.

REFERENCES

[1] Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Review **38**(3), 367–426 (1996)
[2] Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer (2011)
[3] Berinde, V.: Iterative Approximation of Fixed Points. Springer (2007)
[4] Bertsekas, D.P., Nedić, A., Ozdaglar, A.E.: Convex Analysis and Optimization. Athena Scientific (2003)
[5] Blatt, D., Hero, A.O., Gauchman, H.: A convergent incremental gradient method with a constant step size. SIAM Journal on Optimization **18**, 29–51 (2007)
[6] Combettes, P.L.: A block-iterative surrogate constraint splitting method for quadratic signal recovery. IEEE Transactions on Signal Processing **51**(7), 1771–1782 (2003)
[7] Combettes, P.L.: Iterative construction of the resolvent of a sum of maximal monotone opera-

tors. Journal of Convex Analysis **16**, 727–748 (2009)
[8] Combettes, P.L., Bondon, P.: Hard-constrained inconsistent signal feasibility problems. IEEE Transactions on Signal Processing **47**(9), 2460–2468 (1999)
[9] Combettes, P.L., Pesquet, J.C.: A proximal decomposition method for solving convex variational inverse problems. Inverse Problems **24**, article ID 065,014 (2008)
[10] Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. Multiscale Modeling and Simulation **4**, 1168–1200 (2005)
[11] Goebel, K., Kirk, W.A.: Topics in Metric Fixed Point Theory. Cambridge Studies in Advanced Mathematics. Cambridge University Press (1990)
[12] Goebel, K., Reich, S.: Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings. Dekker (1984)
[13] Halpern, B.: Fixed points of nonexpansive maps. Bulletin of the American Mathematical Society **73**, 957–961 (1967)
[14] Iiduka, H.: Fixed point optimization algorithm and its application to network bandwidth allocation. Journal of Computational and Applied Mathematics **236**, 1733–1742 (2012)
[15] Iiduka, H.: Fixed point optimization algorithm and its application to power control in cdma data networks. Mathematical Programming **133**, 227–242 (2012)
[16] Iiduka, H.: Iterative algorithm for triple-hierarchical constrained nonconvex optimization problem and its application to network bandwidth allocation. SIAM Journal on Optimization **22**(3), 862–878 (2012)
[17] Iiduka, H., Yamada, I.: A use of conjugate gradient direction for the convex optimization problem over the fixed point set of a nonexpansive mapping. SIAM Journal on Optimization **19**(4), 1881–1893 (2009)
[18] Iiduka, H., Yamada, I.: Computational method for solving a stochastic linear-quadratic control problem given an unsolvable stochastic algebraic riccati equation. SIAM Journal on Control and Optimization (to appear)
[19] Johansson, B., Rabi, M., Johansson, M.: A randomized incremental subgradient method for distributed optimization in networked systems. SIAM Journal on Optimization **20**, 1157–1170 (2009)
[20] Kaneko, M., Popovski, P., Dahl, J.: Proportional fairness in multi-carrier system: upper bound and approximation algorithms. IEEE Communications Letters **10**(6), 462–464 (2006)
[21] Kiwiel, K.C.: Convergence of approximate and incremental subgradient methods for convex optimization. SIAM Journal on Optimization **14**, 807–840 (2004)
[22] Low, S., Lapsley, D.E.: Optimization flow control, i: Basic algorithm and convergence. IEEE/ACM Transactions on Networking **7**(6), 861–874 (1999)
[23] Maillé, P., Toka, L.: Managing a peer-to-peer data storage system in a selfish society. IEEE Journal on Selected Areas in Communications **26**(7), 1295–1301 (2008)
[24] Nedić, A., Bertsekas, D.P.: Incremental sugradient methods for nondifferentiable optimization. SIAM Journal on Optimization **12**, 109–138 (2001)
[25] Nocedal, J., Wright, S.J.: Numerical Optimization, 2nd edn. Springer Series in Operations Research and Financial Engineering. Springer (2006)
[26] Opial, Z.: Weak convergence of the sequence of successive approximation for nonexpansive mappings. Bulletin of the American Mathematical Society **73**, 591–597 (1967)
[27] Sharma, S., Teneketzis, D.: An externalities-based decentralized optimal power allocation algorithm for wireless networks. IEEE/ACM Transactions on Networking **17**, 1819–1831 (2009)
[28] Srikant, R.: Mathematics of Internet Congestion Control. Birkhauser (2004)
[29] Ulbrich, M.: Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces. MOS-SIAM Series on Optimization (2011)
[30] Wittmann, R.: Approximation of fixed points of nonexpansive mappings. Archiv der Mathematik **58**(5), 486–491 (1992)
[31] Yamada, I.: The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. In: D. Butnariu, Y. Censor, S. Reich (eds.) Inherently Parallel Algorithms for Feasibility and Optimization and Their Applications, pp. 473–504. Elsevier (2001)
[32] Yamada, I., Ogura, N., Shirakawa, N.: A numerical robust hybrid steepest descent method for the convexly constrained generalized inverse problems. Contemporary Mathematics **313**, 269–305 (2002)
[33] Zeidler, E.: Nonlinear Functional Analysis ans Its Applications II/B. Nonlinear Monotone Operators. Springer (1985)
[34] Zeidler, E.: Nonlinear Functional Analysis ans Its Applications III. Variational Methods and Optimization. Springer (1985)