PARALLEL OPTIMIZATION ALGORITHM
FOR SMOOTH CONVEX OPTIMIZATION OVER
FIXED POINT SETS OF QUASI-NONEXPANSIVE MAPPINGS

Hideaki Iiduka
Meiji University

Abstract   Smooth convex optimization problems are solved over fixed point sets of quasi-nonexpansive mappings by using a distributed optimization technique. This is done for a networked system with an operator, who manages the system, and a finite number of users, by solving the problem of minimizing the sum of the operator's and users' differentiable, convex objective functions over the intersection of the operator's and users' fixed point sets of quasi-nonexpansive mappings. Under the assumption that the operator can communicate with all users, a parallel optimization algorithm can be devised that enables the operator to find a solution to the problem without using all user objective functions and quasi-nonexpansive mappings. This algorithm does not use proximity operators, in contrast to conventional parallel proximal algorithms. Moreover, it can optimize over fixed point sets of quasi-nonexpansive mappings, in contrast to conventional fixed point algorithms. Investigation of the algorithm's convergence properties for a constant step-size rule reveals that, with a small constant step size, it approximates the solution to the problem. Consideration of the case in which the step-size sequence is diminishing demonstrates that the algorithm converges to the problem solution. Application of the algorithm to network bandwidth allocation based on an operational policy is shown to make the network more stable and reliable.

Keywords: nonlinear programming, fixed point, network bandwidth allocation, parallel optimization algorithm, quasi-nonexpansive mapping, smooth convex optimization, strongly monotone operator

1.   Introduction

Optimization problems with a *fixed point constraint* (see, e.g., [7, 21, 24, 44]) enable consideration of constrained optimization problems in which the explicit form of the metric projection onto the constraint set is not always known; i.e., the constraint set is not simple in the sense that the projection cannot be easily calculated (e.g., the constraint set is the set of all minimizers of a convex function over a closed convex set [9, 44], the set of zeros of a set-valued, monotone operator [4, Proposition 23.38], or the level set of a nondifferentiable, convex function [3, Proposition 2.3]). These and related optimization problems include such practical problems as signal recovery [7], power control [16, 18, 40], bandwidth allocation [19, 20], storage allocation [23, 32], control optimization [25], beamforming [41], and minimal antenna-subset selection [46].

This paper focuses on a networked system consisting of an operator, who manage the system, and a finite number of participating users, and considers the problem of minimizing the sum of the operator's and users' differentiable, convex functions over the intersection of the operator's and users' fixed point constraint sets of *quasi-nonexpansive mappings*.

The motivations for considering this problem are to devise optimization algorithms that have a wider range of applications than previous algorithms for convex optimization over

fixed point sets of nonexpansive mappings [7, 21, 24, 44] and to solve the problem by using *parallel optimization techniques* [4, Chapter 27], [47, PART II].

Many parallel and optimizion algorithms have been presented for smooth or nonsmooth optimization. The *parallel proximal algorithms* [4, Proposition 27.8], [12, Algorithm 10.27], [38] are useful for minimizing the sum of nondifferentiable, convex functions over the whole space. They use the ideas of the *Dauglas-Rachford algorithm* [4, Chapters 25 and 27], [10, 12, 13, 29] and *forward-backward algorithm* [4, Chapters 25 and 27], [8, 11, 12], which use the *proximity operators* [4, Definition 12.23] of nondifferentiable, convex functions. The *incremental subgradient method* [5, Section 8.2] and *projected multi-agent algorithms* [30, 35–37] can minimize the sum of nondifferentiable, convex functions over simple constraint sets by using the *subgradients* [39, Section 23] of the nondifferentiable, convex functions instead of the proximity operators. The *fixed point optimization algorithms* [20, 23] can perform smooth convex distributed optimization over the fixed point sets of nonexpansive mappings. The centralized fixed point optimization algorithm [22] can optimize over two fixed point sets of a quasi-nonexpansive mapping and a nonexpansive mapping. There have been no reports, however, on distributed optimization algorithms for smooth convex optimization with fixed point constraints of quasi-nonexpansive mappings.

In this paper, we describe a parallel optimization algorithm for smooth convex optimization over fixed point sets of quasi-nonexpansive mappings. It is based on two well-known algorithms. The first is the *hybrid steepest descent algorithm* [44], a centralized algorithm for smooth convex optimization over fixed point sets of nonexpansive mappings. The operator and each user in the network can implement the hybrid steepest descent method. The second algorithm is the *parallel proximal algorithm* [4, Proposition 27.8], [12, Algorithm 10.27], [38] for nonsmooth convex optimization. From these two algorithms, a parallel optimization algorithm is formulated for smooth convex optimization with fixed point constraints. Since the operator can communicate with all users, the operator can find the solution to the main problem by using the information transmitted from all users.

This paper makes three contributions in relation to other work on convex optimization. The first is that the proposed parallel optimization algorithm does not use proximity operators, in contrast to several previous algorithms [8, 11, 12, 38, 46]. It uses the gradients of the operator's and users' convex functions.

The second is that the proposed algorithm can be applied to distributed smooth convex optimization over the fixed point sets of quasi-nonexpansive mappings while previous algorithms can only perform nonsmooth convex optimization over simple constraint sets [4, Subchapter 5.2],[8, 11, 12, 30, 35–38], centralized smooth convex optimization over fixed point sets of (quasi-)nonexpansive mappings [7, 21, 22, 24, 44], or distributed smooth convex optimization over fixed point sets of nonexpansive mappings [20, 23].

To clarify the advantages of dealing with distributed optimization over the fixed point sets of quasi-nonexpansive mappings, let us consider network resource allocation [6, 42], which is the sharing of available resources among users in the network so as to maximize the sum of their utilities subject to the feasible regions for allocating the resources. The problem of minimizing the sum of convex functions over the intersection of fixed point sets of nonexpansive mappings includes practical network resource allocation problems, such as power allocation [40], channel allocation [26], storage allocation [23, 32], and bandwidth allocation [19, 20, 27, 31, 34, 42]. Here, let us consider the network resource allocation problem with an operational constraint that makes the network more stable and reliable. When the operational constraint set can be expressed as the level set of a certain *nonsmooth*, convex function [22, subsection 1.2], it can be expressed as the fixed point set of the *subgradient pro-*

jection that satisfies the quasi nonexpansivity condition, *not* the nonexpansivity condition [3, Proposition 2.3] (see also subsection 2.1). This means that practical network resource allocation problems with operational constraints can be formulated as a convex optimization problem over the intersection of fixed point sets of quasi-nonexpansive mappings. The proposed parallel algorithm can thus be applied to network resource allocation with operational constraints. This paper focuses on bandwidth allocation with operational constraints and describes how the proposed parallel algorithm can solve it.

The third contribution is analysis of the proposed algorithm's convergence for different step-size rules. A small constant step size is shown to result in an approximate solution to the main problem. Thanks to the useful lemma of Maingé [33] (Proposition 2.6), it is also shown that the proposed algorithm with a diminishing step size converges to the solution to the problem.

This paper is organized as follows. Section 2 gives the mathematical preliminaries and states the main problem. Section 3 presents the proposed parallel optimization algorithm for solving the main problem and describes its convergence properties for a constant step size and for a diminishing step size. It also provides several application examples. Section 4 describes how application of the algorithm to network bandwidth allocation based on an operational policy makes the network stable and reliable. Section 5 concludes the paper with a brief summary and a mention of future work on distributed optimization over the fixed point sets of quasi-nonexpansive mappings.

2. Mathematical Preliminaries

Let $\mathbb{R}^N$ be an N-dimensional Euclidean space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$, and let $\mathbb{R}^N_+ := \{(x_i)^N_{i=1} \in \mathbb{R}^N \colon x_i \geq 0 \ (i = 1, 2, \ldots, N)\}$. Let $\mathbb{N}$ denote the set of all positive integers including zero. The identity mapping on $\mathbb{R}^N$ is denoted by Id, i.e., $\mathrm{Id}(x) := x \ (x \in \mathbb{R}^N)$.

2.1. Quasi nonexpansivity

A mapping $Q \colon \mathbb{R}^N \to \mathbb{R}^N$ is said to be *quasi-nonexpansive* [4, Definition 4.1(iii)] if $\|Q(x) - y\| \leq \|x - y\|$ for all $x \in \mathbb{R}^N$ and for all $y \in \mathrm{Fix}(Q)$.[*] The fixed point set of Q is denoted by $\mathrm{Fix}(Q) := \{x \in \mathbb{R}^N \colon Q(x) = x\}$. When a quasi-nonexpansive mapping has one fixed point, its fixed point set is closed and convex [3, Proposition 2.6]. $R \colon \mathbb{R}^N \to \mathbb{R}^N$ is called a *quasi-firmly nonexpansive* mapping [2] if a quasi-nonexpansive mapping $Q \colon \mathbb{R}^N \to \mathbb{R}^N$ exists such that $R = (1/2)(\mathrm{Id} + Q)$.

An important example of a quasi-firmly nonexpansive mapping is as follows. Let $f_0 \colon \mathbb{R}^N \to \mathbb{R}$ be a convex function with $\mathrm{lev}_{\leq 0} f_0 := \{x \in \mathbb{R}^N \colon f_0(x) \leq 0\} \neq \emptyset$. Then the *subdifferential* [4, Definition 16.1], [39, Section 23] of f_0 at $x \in \mathbb{R}^N$, denoted by

$$\partial f_0(x) := \left\{ z \in \mathbb{R}^N \colon f_0(y) \geq f_0(x) + \langle y - x, z \rangle \ \left(y \in \mathbb{R}^N \right) \right\},$$

has a point, and the *subgradient* of f_0 at x can be denoted by $f_0'(x) \in \partial f_0(x)$. The *subgradient projection relative to* f_0 [3, Proposition 2.3], [43, Subchapter 4.3], $Q_{\mathrm{sp}} \colon \mathbb{R}^N \to \mathbb{R}^N$, defined for all $x \in \mathbb{R}^N$ by

$$Q_{\mathrm{sp}}(x) := \begin{cases} x - \dfrac{f_0(x)}{\|f_0'(x)\|^2} f_0'(x) & \text{if } f_0(x) > 0, \\ x & \text{otherwise,} \end{cases}$$

---

[*]The definition of a quasi-nonexpansive mapping is that, if $x, y \in \mathbb{R}^N$ and if $y \in \mathrm{Fix}(Q) \ (\neq \emptyset)$, $\|Q(x) - y\| \leq \|x - y\|$. This proposition is always true if $\mathrm{Fix}(Q) = \emptyset$.

is quasi-firmly nonexpansive (i.e., $2Q_{\mathrm{sp}} - \mathrm{Id}$ is quasi-nonexpansive) and satisfies $\mathrm{Fix}(Q_{\mathrm{sp}}) = \mathrm{Fix}(2Q_{\mathrm{sp}} - \mathrm{Id}) = \mathrm{lev}_{\leq 0} f_0$. Moreover, Q_{sp} satisfies the following propositions.

Proposition 2.1.

(i) [2, Lemma 3.1] Q_{sp} *is fixed-point closed, i.e.,* $x \in \mathrm{Fix}(Q_{\mathrm{sp}})$ *whenever* $(x_n)_{n \in \mathbb{N}}$ $(\subset \mathbb{R}^N)$ *converges to* x $(\in \mathbb{R}^N)$ *and* $\lim_{n \to \infty} \|x_n - Q_{\mathrm{sp}}(x_n)\| = 0$.

(ii) $2Q_{\mathrm{sp}} - \mathrm{Id}$ *is fixed-point closed.*

It is obvious from $\mathrm{Fix}(Q_{\mathrm{sp}}) = \mathrm{Fix}(2Q_{\mathrm{sp}} - \mathrm{Id})$ and Proposition 2.1(i) that Proposition 2.1(ii) holds.

The following proposition indicates the properties of quasi-firmly nonexpansive mappings.

Proposition 2.2. *Suppose that* $Q \colon \mathbb{R}^N \to \mathbb{R}^N$ *is quasi-firmly nonexpansive with* $\mathrm{Fix}(Q) \neq \emptyset$ *and* $\alpha \in (0, 1]$ *and that* $Q_\alpha := \alpha \mathrm{Id} + (1 - \alpha)Q$. *Then the following hold:*

(i) $\mathrm{Fix}(Q) = \mathrm{Fix}(Q_\alpha)$.

(ii) Q_α *is quasi-nonexpansive.*

(iii) $\langle x - Q_\alpha(x), x - y \rangle \geq (1 - \alpha)\|x - Q(x)\|^2$ $(x \in \mathbb{R}^N, y \in \mathrm{Fix}(Q))$.

Proof. Items (i) and (ii) in Proposition 2.2 are deduced from [33, Remark 2.1(i0), (i1)]. From $\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$ $(x, y \in \mathbb{R}^N)$, it is found that, for all $x \in \mathbb{R}^N$ and for all $y \in \mathrm{Fix}(Q)$, $\langle x - Q(x), x - y \rangle = (1/2)(\|x - Q(x)\|^2 + \|x - y\|^2 - \|Q(x) - y\|^2)$. Since Q is quasi-firmly nonexpansive, $\|Q(x) - y\|^2 \leq \|x - y\|^2 - \|x - Q(x)\|^2$ $(x \in \mathbb{R}^N, y \in \mathrm{Fix}(Q))$ [2, section 3]. Accordingly, for all $x \in \mathbb{R}^N$ and for all $y \in \mathrm{Fix}(Q)$,

$$\langle x - Q(x), x - y \rangle = \frac{1}{2}\left(\|x - Q(x)\|^2 + \|x - y\|^2 - \|Q(x) - y\|^2\right) \geq \|x - Q(x)\|^2.$$

Hence, for all $x \in \mathbb{R}^N$ and for all $y \in \mathrm{Fix}(Q)$,

$$\langle x - Q_\alpha(x), x - y \rangle = (1 - \alpha)\langle x - Q(x), x - y \rangle \geq (1 - \alpha)\|x - Q(x)\|^2,$$

which means that item (iii) holds. This completes the proof. $\qquad\square$

2.2. Convex optimization problem and monotone variational inequality

An operator $A \colon \mathbb{R}^N \to \mathbb{R}^N$ is said to be *monotone* [4, Definition 20.1] if $\langle x - y, A(x) - A(y) \rangle \geq 0$ for all $x, y \in \mathbb{R}^N$. A is called a *strongly monotone* operator with $c > 0$ (c-strongly monotone operator) [4, Definition 22.1(iv)] if $\langle x - y, A(x) - A(y) \rangle \geq c\|x - y\|^2$ for all $x, y \in \mathbb{R}^N$. A is called a *Lipschitz continuous* operator with $L > 0$ (L-Lipschitz continuous) if $\|A(x) - A(y)\| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^N$.

Proposition 2.3. [44, Lemma 3.1] *Suppose that* $A \colon \mathbb{R}^N \to \mathbb{R}^N$ *is* c-*strongly monotone and* L-*Lipschitz continuous and that* $\mu \in (0, 2c/L^2)$. *For* $\lambda \in [0, 1]$, *define* $T_\lambda \colon \mathbb{R}^N \to \mathbb{R}^N$ *by* $T_\lambda(x) := x - \mu\lambda A(x)$ *for all* $x \in \mathbb{R}^N$. *Then for all* $x, y \in \mathbb{R}^N$,

$$\|T_\lambda(x) - T_\lambda(y)\| \leq (1 - \tau\lambda)\|x - y\|,$$

where $\tau := 1 - \sqrt{1 - \mu(2c - \mu L^2)} \in (0, 1]$.

The *variational inequality* problem [14, Chapter II], [28, Chapter I] for a monotone operator $A \colon \mathbb{R}^N \to \mathbb{R}^N$ over a nonempty, closed convex set $D \subset \mathbb{R}^N$ is to find a point in

$$\mathrm{VI}(D, A) := \{x^\star \in D \colon \langle y - x^\star, A(x^\star) \rangle \geq 0 \, (y \in D)\}.$$

Some properties of the solution set of the monotone variational inequality are as follows:

Proposition 2.4. *Suppose that D ($\subset \mathbb{R}^N$) is nonempty, closed, and convex, $A\colon \mathbb{R}^N \to \mathbb{R}^N$ is continuous, and $f\colon \mathbb{R}^N \to \mathbb{R}$ is convex and differentiable. Then*

(i) [14, Chapter 2, Proposition 2.1 (2.1) and (2.2)] $\mathrm{VI}(D, \nabla f) = \mathrm{argmin}_{x \in D} f(x) := \{x^\star \in D\colon f(x^\star) = \min_{x \in D} f(x)\}$.

(ii) [15, Corollary 2.2.5] $\mathrm{VI}(D, A) \neq \emptyset$ when D is compact.

(iii) [43, Theorem 2.31] $\mathrm{VI}(D, A)$ consists of one point if A is strongly monotone and Lipschitz continuous.

2.3. Main problem

The focus here is a networked system with an operator (denoted as user 0) and I users. Let

$$\mathcal{I} := \{1, 2, \ldots, I\} \text{ and } \overline{\mathcal{I}} := \{0\} \cup \mathcal{I}.$$

Suppose that user i ($i \in \overline{\mathcal{I}}$) has its own private objective function, denoted by $f^{(i)}\colon \mathbb{R}^N \to \mathbb{R}$, and a nonempty, closed convex constraint set, denoted by $C^{(i)}$ ($\subset \mathbb{R}^N$). Moreover, the following is assumed.

Assumption 2.1.

(A1) $Q^{(i)}\colon \mathbb{R}^N \to \mathbb{R}^N$ ($i \in \overline{\mathcal{I}}$) is quasi-firmly nonexpansive with $\mathrm{Fix}(Q^{(i)}) = C^{(i)}$ and $\bigcap_{i \in \overline{\mathcal{I}}} \mathrm{Fix}(Q^{(i)}) \neq \emptyset$.

(A2) $f^{(i)}\colon \mathbb{R}^N \to \mathbb{R}$ ($i \in \overline{\mathcal{I}}$) is convex and differentiable, and $\nabla f^{(i)}\colon \mathbb{R}^N \to \mathbb{R}^N$ is $c^{(i)}$-strongly monotone and $L^{(i)}$-Lipschitz continuous.

(A3) User i ($i \in \overline{\mathcal{I}}$) can use its own private $Q^{(i)}$ and $\nabla f^{(i)}$.

(A4) The operator can communicate with all users.

The following problem is discussed in this paper.

Problem 2.1.

$$\text{Minimize} \sum_{i \in \overline{\mathcal{I}}} f^{(i)}(x) \text{ subject to } x \in \bigcap_{i \in \overline{\mathcal{I}}} \mathrm{Fix}\left(Q^{(i)}\right).$$

The closedness and convexity [3, Proposition 2.6] of $\bigcap_{i \in \overline{\mathcal{I}}} \mathrm{Fix}(Q^{(i)})$ ($\neq \emptyset$) and (A2) guarantee the existence and uniqueness of the solution to Problem 2.1 (Proposition 2.4(i) and (iii)).

Problem 2.1 is closely related to network resource allocation [6, 42], which is a central issue in modern communication networks. The main objective of the allocation is to share the available resources among users in the network so as to maximize the sum of their utilities subject to the feasible regions for allocating the resources. Such a maximization problem, called the network resource allocation problem (see [6, 27, 34, 42], and references therein), includes future network resource allocation problems such as the channel allocation problem for a multi-carrier system [26], the storage allocation problem for a peer-to-peer network [23, 32], the power allocation problem for a wireless data network [40], and the bandwidth allocation problem [19, 20, 27, 31, 34, 42].

For example, the bandwidth allocation problem [27, 34, 42] is the problem of maximizing the sum of the utility function $\mathcal{U}^{(i)}$ of source i (user i) over the intersection of $\mathbb{R}_+^I \cap C_l$, where C_l stands for the capacity constraint of link l (for the definition of C_l, see (4.2)). When $C^{(i)}$ is defined by the intersection of the capacity constraints of links used by source i, $C^{(i)}$ can be expressed as the fixed point set of a certain nonexpansive mapping $Q^{(i)}$ (see (4.7)). Moreover, the condition $0 \in \bigcap_{i \in \mathcal{I}} C^{(i)} = \bigcap_{i \in \mathcal{I}} \mathrm{Fix}(Q^{(i)}) \neq \emptyset$ holds from $0 \in C_l$ for each link l. Since the operator manages the network, it knows the explicit form of $\bigcap_{i \in \mathcal{I}} C^{(i)}$. Accordingly, the operator can set $Q^{(0)}$ such that $\mathrm{Fix}(Q^{(0)}) = C^{(0)}$ and $C^{(0)} \cap \bigcap_{i \in \mathcal{I}} C^{(i)} = \bigcap_{i \in \overline{\mathcal{I}}} \mathrm{Fix}(Q^{(i)}) \neq \emptyset$.

The relationship between Problem 2.1 and the bandwidth allocation problem is described more fully in section 4. Other application examples are storage allocation [23, 32] and power allocation [40], and such allocation problems can be expressed as Problem 2.1.

The following notation is used.

$$Q_\alpha^{(i)} := \alpha^{(i)}\mathrm{Id} + \left(1 - \alpha^{(i)}\right)Q^{(i)}\left(\alpha^{(i)} \in (0,1)\right), \ X := \bigcap_{i\in\overline{\mathcal{I}}}\mathrm{Fix}\left(Q^{(i)}\right),$$

$$f := \sum_{i\in\overline{\mathcal{I}}} f^{(i)}, \ \{x^\star\} = \left\{x \in X \colon f(x) = f^\star := \min_{y\in X} f(y)\right\}$$

The following propositions are needed to prove one of the main theorems.

Proposition 2.5. *Suppose that Assumption 2.1 holds and $Q^{(i)}$ $(i \in \overline{\mathcal{I}})$ is fixed-point closed. Let (z_n) $(\subset \mathbb{R}^N)$ be a bounded sequence with $\lim_{n\to\infty}\|z_n - Q^{(i)}(z_n)\| = 0$ $(i \in \overline{\mathcal{I}})$. Then $\liminf_{n\to\infty}\langle z_n - x^\star, \nabla f(x^\star)\rangle \geq 0$.*

Proof. From the property of the limit inferior of $(\langle z_n - x^\star, \nabla f(x^\star)\rangle)_{n\in\mathbb{N}}$, there exists a subsequence (z_{n_i}) of $(z_n)_{n\in\mathbb{N}}$ such that

$$\liminf_{n\to\infty}\langle z_n - x^\star, \nabla f(x^\star)\rangle = \lim_{i\to\infty}\langle z_{n_i} - x^\star, \nabla f(x^\star)\rangle.$$

Since $(z_{n_i})_{i\in\mathbb{N}}$ is bounded, there exists $(z_{n_{i_j}})_{j\in\mathbb{N}}$ $(\subset (z_{n_i})_{i\in\mathbb{N}})$ converging to $\overline{x} \in \mathbb{R}^N$. It can be assumed without loss of generality that $(z_{n_i})_{i\in\mathbb{N}}$ converges to $\overline{x} \in \mathbb{R}^N$. Since $Q^{(i)}$ $(i \in \overline{\mathcal{I}})$ is fixed-point closed, $\overline{x} \in \mathrm{Fix}(Q^{(i)})$ $(i \in \overline{\mathcal{I}})$, and hence, $\overline{x} \in X$. Meanwhile, the quasi nonexpansivity of $Q^{(i)}$ $(i \in \overline{\mathcal{I}})$ and the nonempty condition of $\mathrm{Fix}(Q^{(i)})$ $(i \in \overline{\mathcal{I}})$ imply that X is nonempty, closed, and convex. Hence, the convexity and differentiability of $f^{(i)}$ $(i \in \overline{\mathcal{I}})$ and Proposition 2.4(i) guarantee that $\mathrm{VI}(X, \nabla f) = \{x^\star\}$. Therefore,

$$\liminf_{n\to\infty}\langle z_n - x^\star, \nabla f(x^\star)\rangle = \langle \overline{x} - x^\star, \nabla f(x^\star)\rangle \geq 0.$$

This completes the proof. $\square$

Proposition 2.6. [33, Lemma 2.1] *Let $(\Gamma_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ and suppose that $(\Gamma_{n_j})_{j\in\mathbb{N}}$ $(\subset (\Gamma_n)_{n\in\mathbb{N}})$ exists such that $\Gamma_{n_j} < \Gamma_{n_j+1}$ for all $j \in \mathbb{N}$. Then there exists $n_0 \in \mathbb{N}$ such that $(\tau(n))_{n\geq n_0}$ defined by $\tau(n) := \max\{k \leq n \colon \Gamma_k < \Gamma_{k+1}\}$ $(n \geq n_0)$ is increasing and $\lim_{n\to\infty}\tau(n) = \infty$. Moreover, $\Gamma_{\tau(n)} \leq \Gamma_{\tau(n)+1}$ and $\Gamma_n \leq \Gamma_{\tau(n)+1}$ for all $n \geq n_0$.*

3. Parallel Optimization Algorithm for Smooth Convex Optimization with Fixed Point Constraints of Quasi-Nonexpansive Mappings

This section presents the proposed parallel optimization algorithm for solving Problem 2.1.

Algorithm 3.1.

Step 0. User i $(i \in \overline{\mathcal{I}})$ sets $\alpha^{(i)}$ $(\in (0,1))$, $Q_\alpha^{(i)} := \alpha^{(i)}\mathrm{Id} + (1 - \alpha^{(i)})Q^{(i)}$, μ $(\in (0, \min_{i\in\overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2}))$, and $(\lambda_n)_{n\in\mathbb{N}}$ $(\subset (0,1])$. The operator (user 0) sets x_0 $(\in \mathbb{R}^N)$ arbitrarily and transmits it to all users.

Step 1. User i $(i \in \overline{\mathcal{I}})$ computes $x_n^{(i)} \in \mathbb{R}^N$ using

$$x_n^{(i)} := Q_\alpha^{(i)}(x_n) - \mu\lambda_n\nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right).$$

User i $(i \in \mathcal{I})$ transmits $x_n^{(i)}$ to the operator.

Step 2. The operator computes $x_{n+1} \in \mathbb{R}^N$ as

$$x_{n+1} := \frac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} x_n^{(i)}$$

and transmits it to all users. The algorithm sets $n := n + 1$ and returns to Step 1.

Assumptions (A2) and (A3) ensure that user i ($i \in \overline{\mathcal{I}}$) knows $c^{(i)}$ and $L^{(i)}$, so user i can compute $2c^{(i)}/L^{(i)^2}$ before executing Algorithm 3.1. Since (A4) implies that user i ($i \in \mathcal{I}$) can transmit the value of $2c^{(i)}/L^{(i)^2}$ to the operator, the operator knows all the values of $2c^{(i)}/L^{(i)^2}$s, i.e., it can set $\mu \in (0, \min_{i \in \overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2})$. Since the operator can transmit this μ to all users, user i ($i \in \overline{\mathcal{I}}$) can obtain this μ before executing Algorithm 3.1. If the operator sets $(\lambda_n)_{n \in \mathbb{N}}$ ($\subset (0, 1]$) in advance, (A4) guarantees that the operator can inform all users of $(\lambda_n)_{n \in \mathbb{N}}$ before executing Algorithm 3.1. When $(\lambda_n)_{n \in \mathbb{N}}$ is a diminishing step size (subsection 3.2), the convergence of Algorithm 3.1 to the solution to Problem 2.1 is guaranteed regardless of the choice of μ (for details, see Remark 3.1).

Assumptions (A3) and (A4) imply that user i ($i \in \overline{\mathcal{I}}$) can compute *in parallel* $x_n^{(i)} = Q_\alpha^{(i)}(x_n) - \mu \lambda_n \nabla f^{(i)}(Q_\alpha^{(i)}(x_n))$ by using the information x_n transmitted from the operator and its own private information. The hybrid steepest descent method [44] is used to compute $x_n^{(i)}$. Moreover, (A4) ensures that the operator has access to all $x_n^{(i)}$ and can compute $x_{n+1} = (1/(I+1)) \sum_{i \in \overline{\mathcal{I}}} x_n^{(i)}$. This idea is based on the parallel proximal algorithm [4, Proposition 27.8].

The following is an important lemma that will be used to prove the main theorems.

Lemma 3.1. *Suppose that Assumption 2.1 holds. Then $(x_n)_{n \in \mathbb{N}}$ in Algorithm 3.1 satisfies the following properties:*
(i) $(x_n)_{n \in \mathbb{N}}$, $(x_n^{(i)})_{n \in \mathbb{N}}$, and $(\nabla f^{(i)}(Q_\alpha^{(i)}(x_n)))_{n \in \mathbb{N}}$ ($i \in \overline{\mathcal{I}}$) are bounded.
(ii) For all $n \in \mathbb{N}$,

$$\|x_{n+1} - x^\star\|^2 \leq \|x_n - x^\star\|^2 - \sum_{i \in \overline{\mathcal{I}}} \frac{2\alpha^{(i)}(1 - \alpha^{(i)})}{I+1} \|x_n - Q^{(i)}(x_n)\|^2 + \lambda_n M_n,$$

where

$$M_n := \frac{2\mu}{I+1} \sum_{i \in \overline{\mathcal{I}}} \left\{ \langle x^\star - x_n, \nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) \rangle + \mu \lambda_n \|\nabla f^{(i)}(Q_\alpha^{(i)}(x_n))\|^2 \right\}$$

satisfies $M_1 := \sup_{n \in \mathbb{N}} M_n < \infty$.

Proof. (i) Choose $x \in X \subset \text{Fix}(Q^{(i)})$ ($i \in \overline{\mathcal{I}}$) and $n \in \mathbb{N}$ arbitrarily. Then the triangle inequality and Proposition 2.3 guarantee that, for all $i \in \overline{\mathcal{I}}$,

$$\begin{aligned}
\|x_n^{(i)} - x\| &= \|Q_\alpha^{(i)}(x_n) - \mu \lambda_n \nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) - x\| \\
&\leq \|Q_\alpha^{(i)}(x_n) - \mu \lambda_n \nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) - (x - \mu \lambda_n \nabla f^{(i)}(x))\| \\
&\quad + \mu \lambda_n \|\nabla f^{(i)}(x)\| \\
&\leq (1 - \tau^{(i)} \lambda_n) \|Q_\alpha^{(i)}(x_n) - x\| + \mu \lambda_n \|\nabla f^{(i)}(x)\| \\
&\leq (1 - \tau \lambda_n) \|Q_\alpha^{(i)}(x_n) - x\| + \mu \lambda_n \|\nabla f^{(i)}(x)\|,
\end{aligned}$$

where $\tau^{(i)} := 1 - \sqrt{1 - \mu(2c^{(i)} - \mu L^{(i)^2})} \in (0,1]$ $(i \in \overline{\mathcal{I}})$ and $\tau := \min_{i \in \overline{\mathcal{I}}} \tau^{(i)}$. Hence, Proposition 2.2(i) and (ii) imply that, for all $i \in \overline{\mathcal{I}}$,

$$\left\| x_n^{(i)} - x \right\| \leq (1 - \tau\lambda_n) \left\| x_n - x \right\| + \mu\lambda_n \left\| \nabla f^{(i)}(x) \right\|. \tag{3.1}$$

Therefore, the triangle inequality and summing up the above inequality over all i ensure that

$$
\begin{aligned}
\|x_{n+1} - x\| &= \left\| \frac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} \left(x_n^{(i)} - x \right) \right\| \\
&\leq \frac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} \left\| x_n^{(i)} - x \right\| \\
&\leq (1 - \tau\lambda_n) \|x_n - x\| + \tau\lambda_n \frac{\mu}{\tau(I+1)} \sum_{i \in \overline{\mathcal{I}}} \left\| \nabla f^{(i)}(x) \right\|.
\end{aligned}
$$

Accordingly, induction leads to

$$\|x_n - x\| \leq \max \left\{ \|x_0 - x\|, \frac{\mu}{\tau(I+1)} \sum_{i \in \overline{\mathcal{I}}} \left\| \nabla f^{(i)}(x) \right\| \right\},$$

which means $(x_n)_{n \in \mathbb{N}}$ is bounded. Moreover, (3.1) and $(\lambda_n) \subset (0,1]$ $(n \in \mathbb{N})$ imply the boundedness of $(x_n^{(i)})_{n \in \mathbb{N}}$ $(i \in \overline{\mathcal{I}})$. From the quasi nonexpansivity of $Q_\alpha^{(i)}$ $(i \in \overline{\mathcal{I}})$, we have $\|Q_\alpha^{(i)}(x_n) - x\| \leq \|x_n - x\|$, which, together with the boundedness of $(x_n)_{n \in \mathbb{N}}$, means that $(Q_\alpha^{(i)}(x_n))_{n \in \mathbb{N}}$ $(i \in \overline{\mathcal{I}})$ is bounded. Since the Lipschitz continuity of $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ means that $\|\nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) - \nabla f^{(i)}(x)\| \leq L^{(i)}\|Q_\alpha^{(i)}(x_n) - x\|$, the boundedness of $(Q_\alpha^{(i)}(x_n))_{n \in \mathbb{N}}$ $(i \in \overline{\mathcal{I}})$ implies that $(\nabla f^{(i)}(Q_\alpha^{(i)}(x_n)))_{n \in \mathbb{N}}$ $(i \in \overline{\mathcal{I}})$ is bounded.

(ii) Choose $x \in X \subset \text{Fix}(Q^{(i)})$ $(i \in \overline{\mathcal{I}})$ and $n \in \mathbb{N}$ arbitrarily. From $-2\langle x, y \rangle = \|x - y\|^2 - \|x\|^2 - \|y\|^2$ $(x, y \in \mathbb{R}^N)$, we find that, for all $i \in \overline{\mathcal{I}}$,

$$
\begin{aligned}
&2 \left\langle x_n^{(i)} - x_n + \mu\lambda_n \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right), x_n - x \right\rangle \\
&= -2 \left\langle x_n - x_n^{(i)}, x_n - x \right\rangle + 2\mu\lambda_n \left\langle x_n - x, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right) \right\rangle \\
&= \left\| x_n^{(i)} - x \right\|^2 - \left\| x_n - x_n^{(i)} \right\|^2 - \|x_n - x\|^2 + 2\mu\lambda_n \left\langle x_n - x, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right) \right\rangle.
\end{aligned}
$$

Moreover, Proposition 2.2(iii) ensures that

$$2 \left\langle Q_\alpha^{(i)}(x_n) - x_n, x_n - x \right\rangle \leq -2 \left(1 - \alpha^{(i)} \right) \left\| x_n - Q^{(i)}(x_n) \right\|^2.$$

Accordingly, from $x_n^{(i)} := Q_\alpha^{(i)}(x_n) - \mu\lambda_n \nabla f^{(i)}(Q_\alpha^{(i)}(x_n))$,

$$
\begin{aligned}
&2 \left\langle x_n^{(i)} - x_n + \mu\lambda_n \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right), x_n - x \right\rangle \\
&= 2 \left\langle Q_\alpha^{(i)}(x_n) - x_n, x_n - x \right\rangle \\
&\leq -2 \left(1 - \alpha^{(i)} \right) \left\| x_n - Q^{(i)}(x_n) \right\|^2.
\end{aligned}
$$

Therefore, for all $i \in \overline{\mathcal{I}}$,

$$
\begin{aligned}
\left\| x_n^{(i)} - x \right\|^2 \leq{}& \|x_n - x\|^2 + \left\| x_n - x_n^{(i)} \right\|^2 - 2\mu\lambda_n \left\langle x_n - x, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right) \right\rangle \\
&- 2 \left(1 - \alpha^{(i)} \right) \left\| x_n - Q^{(i)}(x_n) \right\|^2.
\end{aligned}
$$

Moreover, from $\|x - y\|^2 \leq 2\|x\|^2 + 2\|y\|^2 \ (x, y \in \mathbb{R}^N)$,

$$
\begin{aligned}
\left\| x_n - x_n^{(i)} \right\|^2 &= \left\| \left(x_n - Q_\alpha^{(i)}(x_n) \right) + \mu\lambda_n \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\|^2 \\
&\leq 2 \left\| x_n - Q_\alpha^{(i)}(x_n) \right\|^2 + 2\mu^2\lambda_n^2 \left\| \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\|^2 \\
&= 2 \left(1 - \alpha^{(i)} \right)^2 \left\| x_n - Q^{(i)}(x_n) \right\|^2 + 2\mu^2\lambda_n^2 \left\| \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\|^2 .
\end{aligned}
$$

Hence, for all $i \in \overline{\mathcal{I}}$,

$$
\begin{aligned}
\left\| x_n^{(i)} - x \right\|^2 &\leq \| x_n - x \|^2 - 2\alpha^{(i)} \left(1 - \alpha^{(i)} \right) \left\| x_n - Q^{(i)}(x_n) \right\|^2 \\
&\quad + 2\mu^2\lambda_n^2 \left\| \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\|^2 \\
&\quad - 2\mu\lambda_n \left\langle x_n - x, \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\rangle ,
\end{aligned}
$$

which, together with the convexity of $\| \cdot \|^2$ and putting $x := x^\star$, implies that

$$
\begin{aligned}
& \| x_{n+1} - x^\star \|^2 \\
&\leq \frac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} \left\| x_n^{(i)} - x^\star \right\|^2 \\
&\leq \| x_n - x^\star \|^2 - \frac{2}{I+1} \sum_{i \in \overline{\mathcal{I}}} \alpha^{(i)} \left(1 - \alpha^{(i)} \right) \left\| x_n - Q^{(i)}(x_n) \right\|^2 \\
&\quad - \frac{2\mu\lambda_n}{I+1} \sum_{i \in \overline{\mathcal{I}}} \left\{ \left\langle x_n - x^\star, \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\rangle - \mu\lambda_n \left\| \nabla f^{(i)} \left(Q_\alpha^{(i)}(x_n) \right) \right\|^2 \right\} \\
&= \| x_n - x^\star \|^2 - \frac{2}{I+1} \sum_{i \in \overline{\mathcal{I}}} \alpha^{(i)} \left(1 - \alpha^{(i)} \right) \left\| x_n - Q^{(i)}(x_n) \right\|^2 + \lambda_n M_n,
\end{aligned}
$$

where $M_n := (2\mu/(I+1)) \sum_{i \in \overline{\mathcal{I}}} \{ \langle x^\star - x_n, \nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) \rangle + \mu\lambda_n \| \nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) \|^2 \}$ and Lemma 3.1(i) guarantee $M_1 := \sup_{n \in \mathbb{N}} M_n < \infty$. This completes the proof. $\qquad \square$

3.1. Constant step-size rule

The discussion in this subsection is based on the following assumption.

Assumption 3.1. *User i ($i \in \overline{\mathcal{I}}$) has $(\lambda_n)_{n \in \mathbb{N}}$ satisfying*

$$
\text{(C1)} \ \lambda_n := \lambda \in (0, 1] \ (n \in \mathbb{N}).
$$

Let us perform a convergence analysis on Algorithm 3.1 under Assumption 3.1.

Theorem 3.1. *Suppose that Assumptions 2.1 and 3.1 hold. Then $(x_n)_{n \in \mathbb{N}}$ in Algorithm 3.1 satisfies the relations*

$$
\liminf_{n \to \infty} \left\| x_n - Q^{(i)}(x_n) \right\|^2 \leq \frac{(I+1) M_1 \lambda}{2\alpha^{(i)} \left(1 - \alpha^{(i)} \right)} \ \left(i \in \overline{\mathcal{I}} \right),
$$

$$
\liminf_{n \to \infty} f(x_n) \leq f^\star + \mu M_2 \lambda + M_3 \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \sqrt{\frac{\left(1 - \alpha^{(i)} \right)(I+1) M_1 \lambda}{2\alpha^{(i)}}},
$$

where M_1 is as in Lemma 3.1, $M_2 := \sup_{n \in \mathbb{N}} \sum_{i \in \overline{\mathcal{I}}} \| \nabla f^{(i)}(Q_\alpha^{(i)}(x_n)) \|^2 < \infty$, and $M_3 := \sup_{n \in \mathbb{N}} \| x_n - x^\star \| < \infty$.

If $\lambda > 0$ can be chosen such that $\mu M_2 \lambda \approx 0$, $(I + 1)M_1\lambda/(2\alpha^{(i)}(1 - \alpha^{(i)})) \approx 0$, and $L^{(i)}\sqrt{(1 - \alpha^{(i)})(I + 1)M_1\lambda/(2\alpha^{(i)})} \approx 0$ $(i \in \overline{\mathcal{I}})$, Theorem 3.1 says that

$$\liminf_{n\to\infty} \left\| x_n - Q^{(i)}(x_n) \right\|^2 \approx 0 \; \left(i \in \overline{\mathcal{I}} \right) \text{ and } \liminf_{n\to\infty} f(x_n) \approx f^\star.$$

Therefore, Theorem 3.1 indicates that Algorithm 3.1 with a small enough λ may approximate the solution to Problem 2.1.

Proof. First, let us show that

$$\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}\left(1 - \alpha^{(i)}\right) \left\| x_n - Q^{(i)}(x_n) \right\|^2 \leq \frac{(I + 1)M_1\lambda}{2}. \tag{3.2}$$

Assume that (3.2) does not hold. Accordingly, $\delta \, (> 0)$ can be chosen such that

$$\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}\left(1 - \alpha^{(i)}\right) \left\| x_n - Q^{(i)}(x_n) \right\|^2 > \frac{(I + 1)M_1\lambda}{2} + 2\delta.$$

The property of the limit inferior of $(\sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}(1 - \alpha^{(i)})\|x_n - Q^{(i)}(x_n)\|^2)_{n\in\mathbb{N}}$ guarantees that there exists $n_0 \in \mathbb{N}$ such that $\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}(1 - \alpha^{(i)})\|x_n - Q^{(i)}(x_n)\|^2 - \delta \leq \sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}(1 - \alpha^{(i)})\|x_n - Q^{(i)}(x_n)\|^2$ for all $n \geq n_0$. Accordingly, for all $n \geq n_0$,

$$\sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}\left(1 - \alpha^{(i)}\right) \left\| x_n - Q^{(i)}(x_n) \right\|^2 > \frac{(I + 1)M_1\lambda}{2} + \delta.$$

Hence, Lemma 3.1(ii) leads to the finding that, for all $n \geq n_0$,

$$\begin{aligned} \left\| x_{n+1} - x^\star \right\|^2 &\leq \left\| x_n - x^\star \right\|^2 - \sum_{i\in\overline{\mathcal{I}}} \frac{2\alpha^{(i)}\left(1 - \alpha^{(i)}\right)}{I + 1} \left\| x_n - Q^{(i)}(x_n) \right\|^2 + M_1\lambda \\ &< \left\| x_n - x^\star \right\|^2 - \frac{2}{I + 1}\left\{ \frac{(I + 1)M_1\lambda}{2} + \delta \right\} + M_1\lambda \\ &= \left\| x_n - x^\star \right\|^2 - \frac{2}{I + 1}\delta. \end{aligned}$$

Therefore, induction ensures that, for all $n \geq n_0$,

$$0 \leq \left\| x_{n+1} - x^\star \right\|^2 < \left\| x_{n_0} - x^\star \right\|^2 - \frac{2}{I + 1}\delta\left(n + 1 - n_0\right).$$

Since the right side of the above inequality approaches minus infinity as n diverges, there is a contradiction. Therefore, (3.2) holds. Since $\liminf_{n\to\infty} \alpha^{(i)}(1 - \alpha^{(i)})\|x_n - Q^{(i)}(x_n)\|^2 \leq \liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \alpha^{(i)}(1 - \alpha^{(i)})\|x_n - Q^{(i)}(x_n)\|^2$ $(i \in \overline{\mathcal{I}})$, there is also another finding:

$$\liminf_{n\to\infty} \left\| x_n - Q^{(i)}(x_n) \right\|^2 \leq \frac{(I + 1)M_1\lambda}{2\alpha^{(i)}\left(1 - \alpha^{(i)}\right)} \; \left(i \in \overline{\mathcal{I}} \right). \tag{3.3}$$

Let $i \in \overline{\mathcal{I}}$ be fixed arbitrarily. Inequality (3.3) and the property of the limit inferior of $(\|x_n - Q^{(i)}(x_n)\|^2)_{n\in\mathbb{N}}$ guarantee the existence of a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \left\| x_{n_k} - Q^{(i)}(x_{n_k}) \right\|^2 = \liminf_{n\to\infty} \left\| x_n - Q^{(i)}(x_n) \right\|^2 \leq \frac{(I + 1)M_1\lambda}{2\alpha^{(i)}\left(1 - \alpha^{(i)}\right)}.$$

Therefore, for all $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that, for all $k \geq k_0$,

$$\left\| x_{n_k} - Q^{(i)}\left(x_{n_k}\right) \right\| \leq \sqrt{\frac{(I+1)\,M_1\lambda}{2\alpha^{(i)}\left(1-\alpha^{(i)}\right)} + \epsilon}. \tag{3.4}$$

Here, it is proven that, for all $k \geq k_0$,

$$\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}\left(x_n\right) \right\rangle \leq M_3 \sum_{i\in\overline{\mathcal{I}}} L^{(i)}\left(1-\alpha^{(i)}\right)\left\|x_{n_k} - Q^{(i)}\left(x_{n_k}\right)\right\|$$
$$+ \mu M_2\lambda + 2\epsilon. \tag{3.5}$$

Now, let us assume that (3.5) does not hold for all $k \geq k_0$, i.e., there exists $n_1 \in \mathbb{N}$ such that, for all $n \geq n_1$,

$$\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}\left(x_n\right) \right\rangle > M_3 \sum_{i\in\overline{\mathcal{I}}} L^{(i)}\left(1-\alpha^{(i)}\right)\left\|x_n - Q^{(i)}\left(x_n\right)\right\|$$
$$+ \mu M_2\lambda + 2\epsilon.$$

Since the property of the limit inferior of $(\sum_{i\in\overline{\mathcal{I}}}\langle x_n - x^\star, \nabla f^{(i)}(x_n)\rangle)_{n\in\mathbb{N}}$ implies the existence of $n_2 \in \mathbb{N}$ such that $\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}}\langle x_n - x^\star, \nabla f^{(i)}(x_n)\rangle - \epsilon \leq \sum_{i\in\overline{\mathcal{I}}}\langle x_n - x^\star, \nabla f^{(i)}(x_n)\rangle$ for all $n \geq n_2$, it is found that, for all $n \geq n_3 := \max\{n_1, n_2\}$,

$$\sum_{i\in\overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}\left(x_n\right) \right\rangle > M_3 \sum_{i\in\overline{\mathcal{I}}} L^{(i)}\left(1-\alpha^{(i)}\right)\left\|x_n - Q^{(i)}\left(x_n\right)\right\|$$
$$+ \mu M_2\lambda + \epsilon. \tag{3.6}$$

On the other hand, Lemma 3.1(ii) and the Cauchy-Schwarz inequality guarantee that, for all $n \geq n_3$,

$$\|x_{n+1} - x^\star\|^2$$
$$\leq \|x_n - x^\star\|^2 + \frac{2\mu\lambda}{I+1} \sum_{i\in\overline{\mathcal{I}}} \left\langle x^\star - x_n, \nabla f^{(i)}\left(Q_\alpha^{(i)}\left(x_n\right)\right)\right\rangle + \frac{2\mu^2 M_2\lambda^2}{I+1}$$
$$\leq \|x_n - x^\star\|^2 + \frac{2\mu\lambda}{I+1} \sum_{i\in\overline{\mathcal{I}}} \left\langle x^\star - x_n, \nabla f^{(i)}\left(x_n\right)\right\rangle + \frac{2\mu^2 M_2\lambda^2}{I+1}$$
$$+ \frac{2\mu\lambda}{I+1} \sum_{i\in\overline{\mathcal{I}}} \|x^\star - x_n\|\left\|\nabla f^{(i)}\left(Q_\alpha^{(i)}\left(x_n\right)\right) - \nabla f^{(i)}\left(x_n\right)\right\|,$$

which, together with the Lipschitz continuity of $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ and the definition of $Q_\alpha^{(i)}$ $(i \in \overline{\mathcal{I}})$, implies that, for all $n \geq n_3$,

$$\|x_{n+1} - x^\star\|^2 \leq \|x_n - x^\star\|^2 + \frac{2\mu\lambda}{I+1} \sum_{i\in\overline{\mathcal{I}}} \left\langle x^\star - x_n, \nabla f^{(i)}\left(x_n\right)\right\rangle + \frac{2\mu^2 M_2\lambda^2}{I+1}$$
$$+ \frac{2\mu M_3\lambda}{I+1} \sum_{i\in\overline{\mathcal{I}}} L^{(i)}\left(1-\alpha^{(i)}\right)\left\|Q^{(i)}\left(x_n\right) - x_n\right\|.$$

Hence, from (3.6), for all $n \geq n_3$,

$$\|x_{n+1} - x^\star\|^2$$
$$< \|x_n - x^\star\|^2 + \frac{2\mu M_3 \lambda}{I+1} \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \left(1 - \alpha^{(i)}\right) \left\|Q^{(i)}(x_n) - x_n\right\|$$
$$- \frac{2\mu\lambda}{I+1} \left\{ M_3 \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \left(1 - \alpha^{(i)}\right) \left\|x_n - Q^{(i)}(x_n)\right\| + \mu M_2 \lambda + \epsilon \right\} + \frac{2\mu^2 M_2 \lambda^2}{I+1}$$
$$= \|x_n - x^\star\|^2 - \frac{2\mu\lambda}{I+1}\epsilon$$
$$\leq \|x_{n_3} - x^\star\|^2 - \frac{2\mu\lambda}{I+1}\epsilon\,(n + 1 - n_3).$$

Since the right side of the above inequality approaches minus infinity as n diverges, there is a contradiction. Thus, (3.5) holds for all $k \geq k_0$. Therefore, (3.4) and (3.5) lead to the deduction that, for all $\epsilon > 0$,

$$\liminf_{n \to \infty} \sum_{i \in \overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}(x_n) \right\rangle \leq M_3 \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \left(1 - \alpha^{(i)}\right) \sqrt{\frac{(I+1)M_1\lambda}{2\alpha^{(i)}\left(1 - \alpha^{(i)}\right)} + \epsilon}$$
$$+ \mu M_2 \lambda + 2\epsilon.$$

Since $\epsilon\,(> 0)$ is arbitrary, for all $x \in X$,

$$\liminf_{n \to \infty} \sum_{i \in \overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}(x_n) \right\rangle \leq M_3 \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \sqrt{\frac{\left(1 - \alpha^{(i)}\right)(I+1)M_1\lambda}{2\alpha^{(i)}}} + \mu M_2 \lambda.$$

From $f^{(i)}(x^\star) \geq f^{(i)}(x_n) + \left\langle x^\star - x_n, \nabla f^{(i)}(x_n) \right\rangle$ $(n \in \mathbb{N}, i \in \overline{\mathcal{I}})$ (by $\{\nabla f^{(i)}(x_n)\} = \partial f^{(i)}(x_n)$), $f := \sum_{i \in \overline{\mathcal{I}}} f^{(i)}$, and $f^\star := f(x^\star)$,

$$\liminf_{n \to \infty} f(x_n) - f^\star = \liminf_{n \to \infty} \left(f(x_n) - f^\star\right)$$
$$\leq \liminf_{n \to \infty} \sum_{i \in \overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}(x_n) \right\rangle$$
$$\leq \mu M_2 \lambda + M_3 \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \sqrt{\frac{\left(1 - \alpha^{(i)}\right)(I+1)M_1\lambda}{2\alpha^{(i)}}}.$$

This completes the proof. $\qquad\square$

3.2. Diminishing step-size rule

The discussion in this subsection is based on the following assumption.

Assumption 3.2. $Q^{(i)} \colon \mathbb{R}^N \to \mathbb{R}^N$ $(i \in \overline{\mathcal{I}})$ *is fixed-point closed.*[†] *User i $(i \in \overline{\mathcal{I}})$ has $(\lambda_n)_{n \in \mathbb{N}}$ satisfying*

$$\text{(C2)} \ \lim_{n \to \infty} \lambda_n = 0 \ \text{and} \ \text{(C3)} \ \sum_{n=0}^{\infty} \lambda_n = \infty.$$

[†]Let $g^{(i)} \colon \mathbb{R}^N \to \mathbb{R}$ $(i \in \overline{\mathcal{I}})$ be convex and nondifferentiable and define $Q^{(i)}$ $(i \in \overline{\mathcal{I}})$ by the subgradient projection relative to $g^{(i)}$. Then $Q^{(i)}$ is quasi-firmly nonexpansive and fixed-point closed with $\text{Fix}(Q^{(i)}) = \text{lev}_{\leq 0} g^{(i)}$ (see subsection 2.1).

An example of $(\lambda_n)_{n \in \mathbb{N}}$ is $\lambda_n := 1/(n+1)^a$ $(n \in \mathbb{N})$, where $a \in (0, 1]$.

Let us perform a convergence analysis on Algorithm 3.1 under Assumption 3.2.

Theorem 3.2. *Suppose that Assumptions 2.1 and 3.2 hold. Then the sequence $(x_n)_{n \in \mathbb{N}}$ generated by Algorithm 3.1 converges to $x^\star$.*

Regarding Assumption 3.2 and Theorem 3.2, the following remark can be made.

Remark 3.1. The condition $\mu \in (0, \min_{i \in \overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2})$ is needed for $\mathrm{Id} - \mu\lambda_n \nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ to be contractive (see Proposition 2.3). Condition (C2) implies that, if $\mu \geq \min_{i \in \overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2}$, there exists $m \in \mathbb{N}$ such that $\mu\lambda_n < \min_{i \in \overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2}$ for all $n \geq m$. Accordingly, for obtaining a sufficiently large n, $\bar{\mu} \in (0, \min_{i \in \overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2})$ and $\bar{\lambda}_n \in (0, 1)$ can be chosen with $\mu\lambda_n = \bar{\mu}\bar{\lambda}_n$. Since $(\bar{\lambda}_n = (\mu/\bar{\mu})\lambda_n)_{n \geq m}$ also satisfies (C2) and (C3), Theorem 3.2 ensures that Algorithm 3.1 with an initial point x_m converges to $x^\star$. Therefore, the convergence of Algorithm 3.1 is guaranteed regardless of the choice of μ. This implies that Algorithm 3.1 with $\mu := 1$, i.e., $(x_n)_{n \in \mathbb{N}}$ generated by

$$\begin{cases} x_0 \in \mathbb{R}^N, \\ x_n^{(i)} := Q_\alpha^{(i)}(x_n) - \lambda_n \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right) \ \ (i \in \overline{\mathcal{I}}), \\ x_{n+1} := \dfrac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} x_n^{(i)} \end{cases}$$

converges to $x^\star$ under the assumptions in Theorem 3.2.

Proof. We distinguish two cases.

Case 1: Suppose that there exists $m_0 \in \mathbb{N}$ such that $\|x_{n+1} - x^\star\| \leq \|x_n - x^\star\|$ for all $n \geq m_0$. In this case, the existence of $\lim_{n \to \infty} \|x_n - x^\star\|$ is guaranteed. Since Lemma 3.1(ii) implies that, for all $n \geq m_0$,

$$\sum_{i \in \overline{\mathcal{I}}} \frac{2\alpha^{(i)}\left(1 - \alpha^{(i)}\right)}{I+1} \left\|x_n - Q^{(i)}(x_n)\right\|^2 \leq \|x_n - x^\star\|^2 - \|x_{n+1} - x^\star\|^2 + M_1\lambda_n,$$

$\lim_{n \to \infty}(1/(I+1)) \sum_{i \in \overline{\mathcal{I}}} 2\alpha^{(i)}(1 - \alpha^{(i)})\|x_n - Q^{(i)}(x_n)\|^2 = 0$, i.e.,

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

Hence, Lemma 3.1(i) and Proposition 2.5 ensure that

$$\liminf_{n \to \infty} \langle x_n - x^\star, \nabla f(x^\star) \rangle \geq 0. \tag{3.7}$$

Furthermore, from Lemma 3.1(ii), $\lambda_n(-M_n) \leq \|x_n - x^\star\|^2 - \|x_{n+1} - x^\star\|^2$ $(n \in \mathbb{N})$. Summing up this inequality from $n = 0$ to $n = m$ $(m \in \mathbb{N})$ implies that $\sum_{n=0}^{m} \lambda_n(-M_n) \leq \|x_0 - x^\star\|^2 - \|x_{m+1} - x^\star\|^2 \leq \|x_0 - x^\star\|^2 < \infty$, so

$$\sum_{n=0}^{\infty} \lambda_n(-M_n) < \infty.$$

Now, under the assumption that $\liminf_{n \to \infty}(-M_n) > 0$, $m_1 \in \mathbb{N}$ and $\gamma > 0$ can be chosen such that $-M_n \geq \gamma$ for all $n \geq m_1$. Accordingly, (C3) means that

$$\infty = \gamma \sum_{n=m_1}^{\infty} \lambda_n \leq \sum_{n=m_1}^{\infty} \lambda_n(-M_n) < \infty,$$

which is a contradiction. Therefore, $\liminf_{n\to\infty}(-M_n) \leq 0$, i.e.,

$$\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \left\{ \left\langle x_n - x^\star, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right)\right\rangle - \mu\lambda_n \left\|\nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right)\right\|^2 \right\} \leq 0,$$

which, together with (C2) and Lemma 3.1(i) (the boundedness of $(\nabla f^{(i)}(Q_\alpha^{(i)}(x_n)))_{n\in\mathbb{N}}$ $(i \in \overline{\mathcal{I}})$), imply that

$$\liminf_{n\to\infty} \sum_{i\in\overline{\mathcal{I}}} \left\langle x_n - x^\star, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right)\right\rangle \leq 0. \tag{3.8}$$

Moreover, the Cauchy-Schwarz inequality and the Lipschitz continuity of $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ mean that, for all $n \geq m_0$ and for all $i \in \overline{\mathcal{I}}$,

$$\begin{aligned}
&\left\langle x_n - x^\star, \nabla f^{(i)}(x_n)\right\rangle \\
&= \left\langle x_n - x^\star, \nabla f^{(i)}(x_n) - \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right)\right\rangle + \left\langle x_n - x^\star, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right)\right\rangle \\
&\leq L^{(i)} M_3 \left\|x_n - Q_\alpha^{(i)}(x_n)\right\| + \left\langle x_n - x^\star, \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right)\right\rangle.
\end{aligned} \tag{3.9}$$

From $\lim_{n\to\infty}\|x_n - Q_\alpha^{(i)}(x_n)\| = (1-\alpha^{(i)})\lim_{n\to\infty}\|x_n - Q^{(i)}(x_n)\| = 0$ $(i \in \overline{\mathcal{I}})$ and (3.8),

$$\liminf_{n\to\infty}\left\langle x_n - x^\star, \nabla f(x_n)\right\rangle = \liminf_{n\to\infty}\sum_{i\in\overline{\mathcal{I}}}\left\langle x_n - x^\star, \nabla f^{(i)}(x_n)\right\rangle \leq 0. \tag{3.10}$$

Since the strong monotonicity of $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ guarantees that, for all $n \geq m_0$,

$$\left\langle x_n - x^\star, \nabla f(x^\star)\right\rangle \leq -\sum_{i\in\overline{\mathcal{I}}} c^{(i)}\|x_n - x^\star\|^2 + \left\langle x_n - x^\star, \nabla f(x_n)\right\rangle,$$

(3.7) and (3.10) lead to the deduction that

$$\begin{aligned}
0 &\leq \liminf_{n\to\infty}\left\langle x_n - x^\star, \nabla f(x^\star)\right\rangle \\
&\leq \liminf_{n\to\infty}\left\{ -\sum_{i\in\overline{\mathcal{I}}} c^{(i)}\|x_n - x^\star\|^2 + \left\langle x_n - x^\star, \nabla f(x_n)\right\rangle \right\} \\
&= -\sum_{i\in\overline{\mathcal{I}}} c^{(i)}\lim_{n\to\infty}\|x_n - x^\star\|^2 + \liminf_{n\to\infty}\left\langle x_n - x^\star, \nabla f(x_n)\right\rangle \\
&\leq -\sum_{i\in\overline{\mathcal{I}}} c^{(i)}\lim_{n\to\infty}\|x_n - x^\star\|^2,
\end{aligned}$$

i.e., $\lim_{n\to\infty}\|x_n - x^\star\| = 0$. This means that $(x_n)_{n\in\mathbb{N}}$ converges to $x^\star$.

Case 2: Suppose that there exists $(x_{n_j})_{j\in\mathbb{N}}$ such that $\|x_{n_j} - x^\star\| < \|x_{n_j+1} - x^\star\|$ for all $j \in \mathbb{N}$. Define $\Gamma_n := \|x_n - x^\star\|^2$ $(n \in \mathbb{N})$. Proposition 2.6 guarantees the existence of $m_2 \in \mathbb{N}$ such that $\Gamma_{\tau(n)} < \Gamma_{\tau(n)+1}$ for all $n \geq m_2$, where $\tau(n)$ is as in Proposition 2.6. Lemma 3.1(ii) implies that, for all $n \geq m_2$,

$$\sum_{i\in\overline{\mathcal{I}}} \frac{2\alpha^{(i)}\left(1-\alpha^{(i)}\right)}{I+1}\left\|x_{\tau(n)} - Q^{(i)}\left(x_{\tau(n)}\right)\right\|^2 \leq \Gamma_{\tau(n)} - \Gamma_{\tau(n)+1} + \lambda_{\tau(n)}M_{\tau(n)}$$

$$< \lambda_{\tau(n)}M_{\tau(n)} \leq M_1\lambda_{\tau(n)}, \tag{3.11}$$

which, together with (C2), means that

$$\lim_{n\to\infty} \left(\frac{1}{I+1}\right) \sum_{i\in\overline{\mathcal{I}}} 2\alpha^{(i)} \left(1 - \alpha^{(i)}\right) \left\|x_{\tau(n)} - Q^{(i)}\left(x_{\tau(n)}\right)\right\|^2 = 0,$$

i.e.,

$$\lim_{n\to\infty} \left\|x_{\tau(n)} - Q^{(i)}\left(x_{\tau(n)}\right)\right\| = 0 \ \left(i\in\overline{\mathcal{I}}\right). \tag{3.12}$$

Hence, from Lemma 3.1(i) and Proposition 2.5, $\liminf_{n\to\infty}\langle x_{\tau(n)} - x^\star, \nabla f(x^\star)\rangle \geq 0$, i.e.,

$$\limsup_{n\to\infty} \left\langle x^\star - x_{\tau(n)}, \nabla f\left(x^\star\right)\right\rangle \leq 0. \tag{3.13}$$

Moreover, from (3.11), $M_{\tau(n)} > 0$ $(n \geq m_2)$, which implies that, for all $n \geq m_2$,

$$\sum_{i\in\overline{\mathcal{I}}} \left\langle x_{\tau(n)} - x^\star, \nabla f^{(i)}\left(Q_\alpha^{(i)}\left(x_{\tau(n)}\right)\right)\right\rangle < \mu M_2 \lambda_{\tau(n)}. \tag{3.14}$$

A discussion similar to the one for obtaining (3.9) implies that, for all $n \geq m_2$ and for all $i \in \overline{\mathcal{I}}$,

$$\begin{aligned}&\left\langle x_{\tau(n)} - x^\star, \nabla f^{(i)}\left(x_{\tau(n)}\right)\right\rangle \\ &\leq L^{(i)} M_3 \left\|x_{\tau(n)} - Q_\alpha^{(i)}\left(x_{\tau(n)}\right)\right\| + \left\langle x_{\tau(n)} - x^\star, \nabla f^{(i)}\left(Q_\alpha^{(i)}\left(x_{\tau(n)}\right)\right)\right\rangle.\end{aligned}$$

Summing up the above inequality over all i and (3.14) guarantee that, for all $n \geq m_2$,

$$\left\langle x_{\tau(n)} - x^\star, \nabla f\left(x_{\tau(n)}\right)\right\rangle < M_3 \sum_{i\in\overline{\mathcal{I}}} L^{(i)} \left(1 - \alpha^{(i)}\right) \left\|x_{\tau(n)} - Q^{(i)}\left(x_{\tau(n)}\right)\right\| + \mu M_2 \lambda_{\tau(n)},$$

which, together with (3.12) and (C2), implies that

$$\limsup_{n\to\infty} \left\langle x_{\tau(n)} - x^\star, \nabla f\left(x_{\tau(n)}\right)\right\rangle \leq 0. \tag{3.15}$$

Accordingly, the strong monotonicity of $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$, (3.13), and (3.15) lead to the deduction that

$$\begin{aligned}0 &\leq \sum_{i\in\overline{\mathcal{I}}} c^{(i)} \limsup_{n\to\infty} \Gamma_{\tau(n)} \\ &\leq \limsup_{n\to\infty} \left\{\left\langle x_{\tau(n)} - x^\star, \nabla f\left(x_{\tau(n)}\right)\right\rangle + \left\langle x^\star - x_{\tau(n)}, \nabla f\left(x^\star\right)\right\rangle\right\} \\ &\leq \limsup_{n\to\infty} \left\langle x_{\tau(n)} - x^\star, \nabla f\left(x_{\tau(n)}\right)\right\rangle + \limsup_{n\to\infty} \left\langle x^\star - x_{\tau(n)}, \nabla f\left(x^\star\right)\right\rangle \\ &\leq 0,\end{aligned}$$

i.e., $\lim_{n\to\infty} \Gamma_{\tau(n)} = 0$. Since Proposition 2.6 ensures that $0 \leq \|x_n - x^\star\|^2 =: \Gamma_n \leq \Gamma_{\tau(n)+1}$ $(n \geq m_2)$, $\lim_{n\to\infty} \Gamma_n = 0$. That is, $(x_n)_{n\in\mathbb{N}}$ converges to $x^\star$. This completes the proof. $\square$

3.3. Examples of Algorithm 3.1

Problem 2.1 is first considered for $\mathcal{I} = \emptyset$, i.e., the problem of finding

$$\{x^\star\} = \underset{x \in \text{Fix}(Q)}{\operatorname{argmin}} f(x), \tag{3.16}$$

where $\nabla f \colon \mathbb{R}^N \to \mathbb{R}^N$ is c-strongly monotone and L-Lipschitz continuous, and $Q \colon \mathbb{R}^N \to \mathbb{R}^N$ is quasi-firmly nonexpansive with $\text{Fix}(Q) \neq \emptyset$.

The following discussion proceeds from Theorems 3.1 and 3.2.

Corollary 3.1. *Suppose that* $\mu \in (0, 2c/L^2)$, $(\lambda_n)_{n \in \mathbb{N}} \subset (0, 1]$, $\alpha \in (0, 1)$, *and* $Q_\alpha := \alpha\text{Id} + (1 - \alpha)Q$. *Let* $(x_n)_{n \in \mathbb{N}}$ *be a sequence generated by an arbitrarily chosen initial point* $x_0 \ (\in \mathbb{R}^N)$ *and*

$$x_{n+1} := Q_\alpha(x_n) - \mu\lambda_n \nabla f(Q_\alpha(x_n)) \quad (n \in \mathbb{N}). \tag{3.17}$$

Then the following hold:

(i) *If* $(\lambda_n)_{n \in \mathbb{N}}$ *satisfies (C1),*

$$\liminf_{n \to \infty} \|x_n - Q(x_n)\|^2 \le \frac{M_1 \lambda}{2\alpha(1 - \alpha)},$$

$$\liminf_{n \to \infty} f(x_n) \le f^\star + \mu M_2 \lambda + M_3 L \sqrt{\frac{(1 - \alpha)M_1 \lambda}{2\alpha}},$$

where $M_1 := \sup_{n \in \mathbb{N}} 2\mu\{\langle x^\star - x_n, \nabla f(Q_\alpha(x_n))\rangle + \mu\lambda\|\nabla f(Q_\alpha(x_n))\|^2\} < \infty$, $M_2 := \sup_{n \in \mathbb{N}} \|\nabla f(Q_\alpha(x_n))\|^2 < \infty$, *and* $M_3 := \sup_{n \in \mathbb{N}} \|x_n - x^\star\| < \infty$.

(ii) *If* Q *is fixed-point closed and if* $(\lambda_n)_{n \in \mathbb{N}}$ *satisfies (C2) and (C3),* $(x_n)_{n \in \mathbb{N}}$ *converges to the unique solution to problem* (3.16).

Let us compare the results of Yamada and Ogura [45] with Corollary 3.1. Their theorem 4 guarantees that the algorithm in (3.17) with (C2) and (C3) converges to the solution $x^\star$ if $\overline{x} \in \text{Fix}(Q)$ and $\overline{\mu} \in (0, 2c/L^2)$ exist such that Q is quasi-shrinking[‡] on $B_{\overline{x}}(\rho(x_0)) := \{x \in \mathbb{R}^N \colon \|x - \overline{x}\| \le \rho(x_0)\}$, where $\rho(x_0) := \max\{\|\overline{\mu}\nabla f(\overline{x})\|/\tau, \|x_0 - \overline{x}\|, \max_{\alpha_n > \overline{\mu}} \|x_n - \overline{x}\|\}$ and $\tau := 1 - \sqrt{1 - \overline{\mu}(2c - \overline{\mu}L^2)} \in (0, 1]$. It would be difficult to check for the existence of $B_{\overline{x}}(\rho(x_0))$ on which Q is quasi-shrinking before executing algorithm (3.17). Meanwhile, Corollary 3.1(ii) guarantees that algorithm (3.17) does not require checking in advance whether complicated assumptions, such as the existence of $B_{\overline{x}}(\rho(x_0))$, are satisfied, and converges to $x^\star$ when Q is quasi-firmly nonexpansive and fixed-point closed.

The next case considered is one in which $Q^{(i)}$ $(i \in \overline{\mathcal{I}})$ is the subgradient projection relative to a convex functional $g^{(i)}$, which is defined for all $x \in \mathbb{R}^N$ as follows (see also subsection 2.1):

$$Q_{\text{sp}}^{(i)}(x) := \begin{cases} x - \dfrac{g^{(i)}(x)}{\left\|g^{(i)'}(x)\right\|^2} g^{(i)'}(x) & \text{if } g^{(i)}(x) > 0, \\ x & \text{otherwise,} \end{cases}$$

where $g^{(i)'}(x) \in \partial g^{(i)}(x)$ $(i \in \overline{\mathcal{I}}, x \in \mathbb{R}^N)$. The mapping $Q_{\text{sp}}^{(i)}$ $(i \in \overline{\mathcal{I}})$ is quasi-firmly nonexpansive and fixed-point closed (see subsection 2.1). Problem 2.1 in this case is to find

$$\{x^\star\} = \underset{x \in \bigcap_{i \in \overline{\mathcal{I}}} \text{lev}_{\le 0} g^{(i)}}{\operatorname{argmin}} \sum_{i \in \overline{\mathcal{I}}} f^{(i)}(x). \tag{3.18}$$

Therefore, Theorems 3.1 and 3.2 lead to the following.

[‡]See [45] for the definition of a quasi-shrinking mapping.

Corollary 3.2. *Suppose that* $\mu \in (0, \min_{i \in \overline{\mathcal{I}}} 2c^{(i)}/L^{(i)^2})$, $(\lambda_n)_{n \in \mathbb{N}} \subset (0, 1]$, $Q_\alpha^{(i)} := (1/2)(\mathrm{Id} + Q_{\mathrm{sp}}^{(i)})$ $(i \in \overline{\mathcal{I}})$, *and* $(x_n)_{n \in \mathbb{N}}$ *is defined by an arbitrary initial point* x_0 $(\in \mathbb{R}^N)$ *and, for all* $n \in \mathbb{N}$,

$$x_n^{(i)} := Q_\alpha^{(i)}(x_n) - \mu \lambda_n \nabla f^{(i)}\left(Q_\alpha^{(i)}(x_n)\right) \ (i \in \overline{\mathcal{I}}),$$

$$x_{n+1} := \frac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} x_n^{(i)}.$$

Then the following hold:
(i) *If* $(\lambda_n)_{n \in \mathbb{N}}$ *satisfies (C1),*

$$\liminf_{n \to \infty} \left\| x_n - Q_{\mathrm{sp}}^{(i)}(x_n) \right\|^2 \leq 2(I+1) M_1 \lambda \ \left(i \in \overline{\mathcal{I}}\right),$$

$$\liminf_{n \to \infty} f(x_n) \leq f^\star + \mu M_2 \lambda + M_3 \sum_{i \in \overline{\mathcal{I}}} L^{(i)} \sqrt{\frac{(I+1) M_1 \lambda}{2}},$$

where $M_1 := \sup_{n \in \mathbb{N}}(2\mu/(I+1)) \sum_{i \in \overline{\mathcal{I}}} \{\langle x^\star - x_n, \nabla f^{(i)}(Q_\alpha^{(i)}(x_n))\rangle + \mu\lambda\|\nabla f^{(i)}(Q_\alpha^{(i)}(x_n))\|^2\}$ $< \infty$, $M_2 := \sup_{n \in \mathbb{N}} \sum_{i \in \overline{\mathcal{I}}} \|\nabla f^{(i)}(Q_\alpha^{(i)}(x_n))\|^2 < \infty$, *and* $M_3 := \sup_{n \in \mathbb{N}} \|x_n - x^\star\| < \infty$.
(ii) *If* $(\lambda_n)_{n \in \mathbb{N}}$ *satisfies (C2) and (C3),* $(x_n)_{n \in \mathbb{N}}$ *converges to the unique solution to problem* (3.18).

Remark 3.1, Corollary 3.1(ii) and Corollary 3.2(ii) say that the algorithm with $\mu := 1$ in Corollary 3.1 (resp. Corollary 3.2) converges to the solution to problem (3.16) (resp. problem (3.18)).

4. Application of Algorithm 3.1 to Bandwidth Allocation

The objective of bandwidth allocation [27, 34, 42] is to share the available bandwidth among traffic sources so as to maximize the overall utility under the capacity constraints. The utility function of source i (user i) $(i \in \mathcal{I})$ is defined for all $x \in \mathbb{R}_+$ as follows [42, (2.4)]: given $w^{(i)} > 0$ and $v^{(i)} > 0$,[§]

$$\mathcal{U}^{(i)}(x) := \begin{cases} w^{(i)} \log(x+1) & \text{if } v^{(i)} = 1, \\ w^{(i)} \dfrac{(x+1)^{1-v^{(i)}}}{1-v^{(i)}} & \text{if } v^{(i)} \neq 1. \end{cases} \tag{4.1}$$

The values of the parameters $w^{(i)}$ and $v^{(i)}$ are source i's private information.

The capacity constraint for each link is an inequality constraint in which the sum of the transmission rates of all sources sharing the link is less than or equal to the capacity of the link. Hence, the capacity constraint set for each link l $(\in \mathcal{L} := \{1, 2, \ldots, L\})$ is expressed as $\mathbb{R}_+^I \cap C_l$, where

$$C_l := \left\{ x := (x_1, x_2, \ldots, x_I) \in \mathbb{R}^I : \sum_{i \in \mathcal{I}} x_i I_{i,l} \leq c_l \right\}, \tag{4.2}$$

[§]$\mathcal{U}_{\mathrm{pf}}(x) := \sum_{i \in \mathcal{I}} w^{(i)} \log x_i$ is called the *weighted proportionally fair* function [27, 34, 42]. To enable us to define the weighted proportionally fair function on $\mathbb{R}_+$, we define $\mathcal{U}^{(i)}(x) := w^{(i)} \log(x+1)$ $(x \in \mathbb{R}_+)$ when $v^{(i)} := 1$. The same discussion as in the case where $v^{(i)} = 1$ leads to $\mathcal{U}^{(i)}(x) := w^{(i)}(x+1)^{1-v^{(i)}}/(1-v^{(i)})$ $(x \in \mathbb{R}_+)$ when $v^{(i)} \neq 1$.

c_l (> 0) stands for the capacity of link l, and $I_{i,l}$ takes the value 1 if l is the link used by source i, and 0 otherwise. Let $\mathcal{L}^{(i)}$ $(i \in \mathcal{I})$ be the set of all links used by source i. When source i $(i \in \mathcal{I})$ knows only the capacity constraints for links used by source i, it has the constraint set defined by

$$C^{(i)} := \mathbb{R}_+^I \cap \bigcap_{l \in \mathcal{L}^{(i)}} C_l. \tag{4.3}$$

This section discusses a bandwidth allocation problem subject to not only the capacity constraints but also an *operational constraint* [22, section 1]. The operator has an operational policy to make the network more stable and reliable.

For example, when sources exist in the network such that they get a low (resp. high) degree of satisfaction, the operator attempts to re-allocate bandwidth so as to enable them to get a high (resp. low) degree of satisfaction. When the available bandwidth is limited in the network, the operator needs to control the sum of the transmission rates of all sources. When the network is controlled by using a certain indicator function representing the network's performance, the operator tries to design the network so as to satisfy a constraint incorporating the indicator function.

The operational constraint set representing such operational policies can be written as

$$C^{(0)} := \left\{ x \in \mathbb{R}^I \colon \mathcal{P}(x) \le p \right\}, \tag{4.4}$$

where $\mathcal{P} \colon \mathbb{R}^I \to \mathbb{R}$ is convex and is not always differentiable, and $p \in \mathbb{R}$.

The operator can set $C^{(0)} = \{x \in \mathbb{R}^I \colon x_{i_0} \le p\}$ when it tries to limit the transmission rate of source i_0 and $C^{(0)} = \{x \in \mathbb{R}^I \colon \sum_{i \in \mathcal{I}} x_i \le p\}$ when it tries to limit the transmission rates of all sources. It can also set $C^{(0)} = \{x \in \mathbb{R}^I \colon \sum_{i \in \mathcal{I}} \omega^{(i)} \mathcal{P}^{(i)}(x_i) \le p\}$ $(\omega^{(i)} \ge 0$ $(i \in \mathcal{I})$, $\mathcal{P}^{(i)} \colon \mathbb{R} \to \mathbb{R}$ is nondifferentiable) when the network is controlled by $\mathcal{P}(x) := \sum_{i \in \mathcal{I}} \omega^{(i)} \mathcal{P}^{(i)}(x_i)$. If the network's performance increases when source i's transmission rate is more than a certain value x^0 (> 0), $\mathcal{P}^{(i)}(x)$ can be, for example, expressed as 0 $(0 \le x \le x^0)$ or $x - x^0$ $(x \ge x^0)$. Given the network's performance measure $\mathcal{P}(x)$ and its optimal value $\mathcal{P}^\star$, the operator attempts to allocate bandwidth to all sources so as to satisfy $\mathcal{P}(x) = \mathcal{P}^\star$ as much as possible. When it is sufficient to only satisfy $\mathcal{P}(x) - \mathcal{P}^\star \le \epsilon$ for some $\epsilon \in \mathbb{R}$ to make the network stable, the operator determines an appropriate ϵ and sets $C^{(0)} := \{x \in \mathbb{R}^I \colon \mathcal{P}(x) \le \mathcal{P}^\star + \epsilon\}$.

The operator (user 0) can define its utility as a function of the transmission rates allocated to all the sources, i.e.,

$$\mathcal{U}^{(0)}(x_1, x_2, \ldots, x_I). \tag{4.5}$$

The operator can set $\mathcal{U}^{(0)}$ to allocate the bandwidth fairly and effectively (e.g., for all $x := (x_1, x_2, \ldots, x_I) \in \mathbb{R}^I$, $\mathcal{U}^{(0)}(x) := (1/I) \sum_{i \in \mathcal{I}} x_i$).

Therefore, the objective in bandwidth allocation is to solve the following problem:¶

$$\text{Maximize } \mathcal{U}^{(0)}(x) + \sum_{i \in \mathcal{I}} \mathcal{U}^{(i)}(x_i) \text{ subject to } x \in \mathbb{R}_+^I \cap C^{(0)} \cap \bigcap_{i \in \mathcal{I}} C^{(i)}, \tag{4.6}$$

where $\mathcal{U}^{(i)}$, $C^{(i)}$ $(i \in \mathcal{I})$ $(i \in \mathcal{I})$, $C^{(0)}$, and $\mathcal{U}^{(0)}$ are defined as in (4.1), (4.3), (4.4), and (4.5).

¶Conventional bandwidth allocation problem is to maximize $\sum_{i \in \mathcal{I}} \mathcal{U}^{(i)}(x_i)$ subject to $x \in \mathbb{R}_+^I \cap \bigcap_{i \in \mathcal{I}} C^{(i)}$ [42].

Here, a mapping $Q^{(i)} \colon \mathbb{R}^I \to \mathbb{R}^I$ $(i \in \mathcal{I})$ is defined by

$$Q^{(i)}(x) := P_{\mathbb{R}_+^I} \prod_{l \in \mathcal{L}^{(i)}} P_{C_l}, \tag{4.7}$$

where P_D stands for the *metric projection* onto a nonempty, closed convex set D $(\subset \mathbb{R}^I)$. Since $\mathbb{R}_+^I$ and C_l $(l \in \mathcal{L})$ are half-space, $P_{\mathbb{R}_+^I}$ and P_{C_l} can be computed within a finite number of arithmetic operations [1, p. 406], [4, Subchapter 28.3],[||] which means $Q^{(i)}$ can be easily calculated. Moreover, from [1, Proposition 2.10, Theorem 4.17, Corollary 4.18], $Q^{(i)}$ $(i \in \mathcal{I})$ in (4.7) satisfies the nonexpansivity condition and the fixed point closedness condition, and

$$\mathrm{Fix}\left(Q^{(i)}\right) = C^{(i)} := \mathbb{R}_+^I \cap \bigcap_{l \in \mathcal{L}^{(i)}} C_l.$$

The subgradient projection relative to $\mathcal{P}(\cdot) - p$ is defined by

$$Q^{(0)}(x) := \begin{cases} x - \dfrac{\mathcal{P}(x) - p}{\|\mathcal{P}'(x)\|^2} \mathcal{P}'(x) & \text{if } \mathcal{P}(x) > p, \\ x & \text{otherwise,} \end{cases} \tag{4.8}$$

where $\mathcal{P}'(x) \in \partial \mathcal{P}(x) := \{z \in \mathbb{R}^I \colon (\mathcal{P}(y) - p) \geq (\mathcal{P}(x) - p) + \langle y - x, z \rangle (y \in \mathbb{R}^I)\}$. From subsection 2.1, $Q^{(0)}$ in (4.8) satisfies the quasi-firm nonexpansivity condition, and

$$\mathrm{Fix}\left(Q^{(0)}\right) = C^{(0)} := \left\{x \in \mathbb{R}^I \colon \mathcal{P}(x) \leq p\right\}.$$

The set $C^{(0)} := \{x \in \mathbb{R}^I \colon \mathcal{P}(x) \leq p\}$ with a nonsmooth $\mathcal{P}$ can be represented as the fixed point set of a quasi-nonexpansive mapping, not a nonexpansive mapping [3, Proposition 2.3]. This means that optimization problems with nonsmooth constraints cannot be formulated as optimization problems with fixed point constraints of nonexpansive mappings.[**]

The function $f^{(i)} := -\mathcal{U}^{(i)}$ $(i \in \mathcal{I})$ defined by (4.1) is convex on $\mathbb{R}_+^I$, and $\nabla f^{(i)}$ $(i \in \mathcal{I})$ is strongly monotone and Lipschitz continuous on a compact, convex set $X^{(i)}$ $(\subset \mathbb{R}_+^I)$. Since source i knows the explicit form of $C^{(i)}$, source i can set $X^{(i)}$ $(\supset \mathrm{Fix}(Q^{(i)}) = C^{(i)})$ onto which the projection can be easily computed (e.g., $X^{(i)}$ is a closed ball with a large enough radius or a box constraint set). Assume that the gradient of $f^{(0)} := -\mathcal{U}^{(0)}$ satisfies the strong monotonicity and Lipschitz continuity conditions (e.g., $\mathcal{U}^{(0)}(x_1, x_2, \ldots, x_I) := (1/I) \sum_{i \in \mathcal{I}} x_i$). Problem (4.6) can thus be expressed as Problem 2.1 with $f^{(i)} := -\mathcal{U}^{(i)}$ $(i \in \overline{\mathcal{I}})$ defined by (4.1) and (4.5) and with $Q^{(i)}$ $(i \in \overline{\mathcal{I}})$ defined by (4.7) and (4.8).

Assumption (A2) requires the strong monotonicity and Lipschitz continuity conditions of $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ on the whole space. However, since the constraint set of Problem 2.1 is $X := \bigcap_{i \in \overline{\mathcal{I}}} \mathrm{Fix}(Q^{(i)})$, it is sufficient that $\nabla f^{(i)}$ $(i \in \overline{\mathcal{I}})$ is strongly monotone and Lipschitz continuous on $X^{(i)}$ $(\supset \mathrm{Fix}(Q^{(i)}) \supset X)$. This means that $Q_\alpha^{(i)}(x_n)$ $(i \in \overline{\mathcal{I}}, n \in \mathbb{N})$ in Algorithm 3.1 must be is in $X^{(i)}$. Therefore, Algorithm 3.1 for problem (4.6) can be represented as

$$\begin{cases} x_0 \in \mathbb{R}^I, \\ x_n^{(i)} := P_{X^{(i)}} \left[Q_\alpha^{(i)}(x_n)\right] - \mu\lambda_n \nabla f^{(i)} \left(P_{X^{(i)}} \left[Q_\alpha^{(i)}(x_n)\right]\right) \ \ (i \in \overline{\mathcal{I}}), \\ x_{n+1} := \dfrac{1}{I+1} \sum_{i \in \overline{\mathcal{I}}} x_n^{(i)}. \end{cases} \tag{4.9}$$

[||]The metric projection onto a half-space $H := \{x \in \mathbb{R}^I \colon \langle a, x \rangle \leq b\}$, where a $(\neq 0) \in \mathbb{R}^I$ and $b \in \mathbb{R}$, is expressed as $P_H(x) := x - [\max\{0, \langle a, x \rangle - b\}/\|a\|^2]a$ $(x \in \mathbb{R}^I)$.
[**]Suppose that C is a nonempty, closed convex set of $\mathbb{R}^I$, $\mathcal{P} \colon \mathbb{R}^I \to \mathbb{R}$ is convex and differentiable, and $\nabla\mathcal{P} \colon \mathbb{R}^I \to \mathbb{R}^I$ is L-Lipschitz continuous. Thus, $T := P_C(\mathrm{Id} - \lambda\nabla\mathcal{P})$ $(\lambda \in [0, 2/L])$ is nonexpansive and $\mathrm{Fix}(T) = \mathrm{argmin}_{x \in C} \mathcal{P}(x)$ [17, subsection 2.1, Proposition 2.3].

Algorithm (4.9) can be used to solve problem (4.6), as indicated by the results in section 3.[††]

5. Summary and Future Work

This paper proposed the parallel optimization algorithm minimizes the sum of convex functions over the intersection of the fixed point sets of quasi-nonexpansive mappings. An investigation of the convergence properties for a constant step-size rule and a diminishing step-size rule showed that, with a small constant step size, the algorithm may give an approximate solution to the minimization problem and that, with a diminishing sequence, it converges to the solution. A potential application of the algorithm is to network bandwidth allocation with operational constraints, while the existing bandwidth allocation algorithms in [31, section III] and [42, Chapters 2 and 3] can be applied to only conventional bandwidth allocation problems (without operational constraints).

The numerical experiments and theoretical analyses for bandwidth allocation in [31, 42] were discussed for networks with single-digit numbers of links and sources. (The networks they discussed [31, Figure 2] and [42, Figure 2.1] consisted of two links and three sources.) Future work includes application of the algorithm to specific bandwidth allocation problems in such networks and numeric evaluation of the convergence and stability of the algorithm. It also includes numerical comparison of the proposed parallel algorithm with a centralized algorithm [22] for bandwidth allocation with operational constraints and evaluation of the performance of the algorithm.

The problem of minimizing the sum of nondifferentiable functions or nonconvex functions over the intersection of the fixed point sets of quasi-nonexpansive mappings includes important and practical engineering problems. For example, the power control problem is one of maximizing the sum of differentiable, nonconvex functions over the fixed point sets of certain nonexpansive mappings [18], and the minimal antenna-subset selection problem is one of minimizing the sum of nondifferentiable, convex functions over the fixed point sets of certain quasi-nonexpansive mappings [46]. Therefore, we need to devise distributed optimization algorithms for nonsmooth convex optimization and for smooth nonconvex optimization with fixed point constraints of quasi-nonexpansive mappings.

Acknowledgments

References

[1] H.H. Bauschke and J.M. Borwein: On projection algorithms for solving convex feasibility problems. *SIAM Review*, **38** (1996), 367–426.

[2] H.H. Bauschke, J. Chen, and X. Wang: A projection method for approximating fixed points of quasi nonexpansive mappings without the usual demiclosedness condition. *Journal of Nonlinear and Convex Analysis*, **15** (2014), 129–135.

[3] H.H. Bauschke and P.L. Combettes: A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert space. *Mathematics of Operations Research*, **26** (2001),

---

[††]The proofs of Theorems 3.1 and 3.2 are discussed for the domain in which $\nabla f^{(i)}$ ($i \in \overline{\mathcal{I}}$) is strongly monotone and Lipschitz continuous. Hence, convergence of algorithm (4.9) can be proved by referring to the proofs of Theorems 3.1 and 3.2 and using the nonexpansivity of $P_{X^{(i)}}$ with $X^{(i)} \supset \text{Fix}(Q^{(i)})$ ($i \in \overline{\mathcal{I}}$).

248–264.

[4] H.H. Bauschke and P.L. Combettes: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces* (Springer, New York, NY, 2011).

[5] D.P. Bertsekas, A. Nedić, and A.E. Ozdaglar: *Convex Analysis and Optimization* (Athena Scientific, Belmont, MA, 2003).

[6] D. Bertsekas and R. Gallager: *Data Networks* (Prentice Hall, Upper Saddle River, NJ, 1987).

[7] P.L. Combettes: A block-iterative surrogate constraint splitting method for quadratic signal recovery. *IEEE Transactions on Signal Processing*, **51** (2003), 1771–1782.

[8] P.L. Combettes: Iterative construction of the resolvent of a sum of maximal monotone operators. *Journal of Convex Analysis*, **16** (2009), 727–748.

[9] P.L. Combettes and P. Bondon: Hard-constrained inconsistent signal feasibility problems. *IEEE Transactions on Signal Processing*, **47** (1999), 2460–2468.

[10] P.L. Combettes and J.C. Pesquet: A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, **1** (2007), 564–574.

[11] P.L. Combettes and J.C. Pesquet: A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems*, **24** (2008), article ID 065014.

[12] P.L. Combettes and J.C. Pesquet: Proximal splitting methods in signal processing. In H.H. Bauschke, R.S. Burachik, P.L. Combettes, V. Elser, D.R. Luke, and H. Wolkowicz (eds.): *Fixed-Point Algorithms for Inverse Problems in Science and Engineering* (Springer, New York, 2011), 185–212.

[13] J. Eckstein and D.P. Bertsekas: On the Douglas-Rachfold splitting method and proximal point algorithm for maximal monotone operators. *Mathematical Programming*, **55** (1992), 293–318.

[14] I. Ekeland and R. Témam: *Convex Analysis and Variational Problems* (Classics In Applied Mathematics, 28. SIAM, Philadelphia, PA, 1999).

[15] F. Facchinei and J.S. Pang: *Finite-Dimensional Variational Inequalities and Complementarity Problems I* (Springer, New York, 2003).

[16] F. Facchinei, J.S. Pang, G. Scutari, and L. Lampariello: VI-constrained hemivariational inequalities: distributed algorithms and power control in ad-hoc networks. *Mathematical Programming*, **145** (2014), 59–96.

[17] H. Iiduka: Iterative algorithm for solving triple-hierarchical constrained optimization problem. *Journal of Optimization Theory and Applications*, **148** (2011), 580–592.

[18] H. Iiduka: Fixed point optimization algorithm and its application to power control in CDMA data networks. *Mathematical Programming*, **133** (2012), 227–242.

[19] H. Iiduka: Iterative algorithm for triple-hierarchical constrained nonconvex optimization problem and its application to network bandwidth allocation. *SIAM Journal on Optimization*, **22** (2012), 862–878.

[20] H. Iiduka: Fixed point optimization algorithms for distributed optimization in networked systems. *SIAM Journal on Optimization*, **23** (2013), 1–26.

[21] H. Iiduka: Acceleration method for convex optimization over the fixed point set of a nonexpansive mapping. *Mathematical Programming*, **149** (2015), 131–165.

[22] H. Iiduka: Convex optimization over fixed point sets of quasi-nonexpansive and nonexpansive mappings in utility-based bandwidth allocation problems with operational constraints. *Journal of Computational and Applied Mathematics*, **282** (2015), 225–236.

[23] H. Iiduka and K. Hishinuma: Acceleration method combining broadcast and incremental distributed optimization algorithms. *SIAM Journal on Optimization*, **24** (2014), 1840–1863.

[24] 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 Journal on Optimization*, **19** (2009), 1881–1893.

[25] H. Iiduka and I. Yamada: Computational method for solving a stochastic linear-quadratic control problem given an unsolvable stochastic algebraic Riccati equation. *SIAM Journal on Control and Optimization*, **50** (2012), 2173–2192.

[26] M. Kaneko, P. Popovski, and J. Dahl: Proportional fairness in multi-carrier system: upper bound and approximation algorithms. *IEEE Communications Letters*, **10** (2006), 462–464.

[27] F.P. Kelly: Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, **8** (1997), 33–37.

[28] D. Kinderlehrer and G. Stampacchia: *An Introduction to Variational Inequalities and Their Applications* (Classics In Applied Mathematics, 31. SIAM, Philadelphia, PA, 2000).

[29] P.L. Lions and B. Mercier: Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, **16** (1979), 964–979.

[30] I. Lobel, A. Ozdaglar, and D. Feijer: Distributed multi-agent optimization with state-dependent communication. *Mathematical Programming*, **129** (2011), 255–284.

[31] S. Low and D.E. Lapsley: Optimization flow control–I: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, **7** (1999), 861–874.

[32] P. Maillé and L. Toka: Managing a peer-to-peer data storage system in a selfish society. *IEEE Journal on Selected Areas in Communications*, **26** (2008), 1295–1301.

[33] P.E. Maingé: The viscosity approximation process for quasi-nonexpansive mappings in Hilbert spaces. *Computers and Mathematics with Applications*, **59** (2010), 74–79.

[34] J. Mo and J. Walrand: Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, **8** (2000), 556–567.

[35] A. Nedić, A. Olshevsky, A. Ozdaglar, and J.N. Tsitsiklis: On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, **54** (2009), 2506–2517.

[36] A. Nedić and A. Ozdaglar: Distributed subgradient methods for multi agent optimization. *IEEE Transactions on Automatic Control*, **54** (2009), 48–61.

[37] A. Nedić and A. Ozdaglar: Cooperative distributed multi-agent optimization. In D.P. Palomar and Y.C. Eldar (eds.): *Convex Optimization in Signal Processing and Communications* (Cambridge University Press, Cambridge, 2010), 340–386.

[38] J.C. Pesquet and N. Pustelnik: A parallel inertial proximal optimization method. *Pacific Journal of Optimization*, **8** (2012), 273–306.

[39] R.T. Rockafellar: *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).

[40] S. Sharma and D. Teneketzis: An externalities-based decentralized optimal power allocation algorithm for wireless networks. *IEEE/ACM Transactions on Networking*, **17** (2009), 1819–1831.

[41] K. Slavakis and I. Yamada: Robust wideband beamforming by the hybrid steepest descent method. *IEEE Transactions on Signal Processing*, **55** (2007), 4511–4522.

[42] R. Srikant: *The Mathematics of Internet Congestion Control* (Birkhauser, Boston, MA, 2004).

[43] V.V. Vasin and A.L. Ageev: *Ill-posed problems with a priori information* (V.S.P. Intl Science, Utrecht, 1995).

[44] I. Yamada: 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, and S. Reich (eds.): *Inherently Parallel Algorithms for Feasibility and Optimization and Their Applications* (Elsevier, New York, 2001), 473–504.

[45] I. Yamada and N. Ogura: Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings. *Numerical Functional Analysis and Optimization*, **25** (2004), 619–655.

[46] I. Yamada, M. Yukawa, and M. Yamagishi: Minimizing the Moreau envelope of nonsmooth convex functions over the fixed point set of certain quasi-nonexpansive mappings. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz (eds.): *Fixed-Point Algorithms for Inverse Problems in Science and Engineering* (Springer, New York, 2011), 345–390.

[47] S. Zenios and Y. Censor: *Parallel Optimization: Theory, Algorithms, and Applications* (Oxford Univ Pr on Demand, New York, 1998).

Hideaki Iiduka
Department of Computer Science, Meiji University
1-1-1 Higashimita, Tama-ku, Kawasaki-shi
Kanagawa 214-8571, Japan
E-mail: `iiduka@cs.meiji.ac.jp`