Linear and Nonlinear Analysis
Yokohama Publishers
ISSN 2188-8167

Volume 6, Number 1, 2020, 35–48

# EFFICIENCY OF INEXACT FIXED POINT QUASICONVEX SUBGRADIENT METHOD

KAZUHIRO HISHINUMA AND HIDEAKI IIDUKA

Abstract. This paper considers the efficiency of the fixed point quasiconvex subgradient method for solving constrained quasiconvex optimization problems. In particular, it focuses on the case when errors and noise appear in the computation. This paper presents two convergence rate analyses: one in terms of the objective functional value; the other in terms of the degree of approximation to the fixed point set. Numerical examples show that the effect of the computational inexactness on the obtained solution can be controlled in accordance with the convergence rate analysis.

## 1. Introduction

This paper considers the efficiency of the fixed point quasiconvex subgradient method [5] in the case where errors and noise appear in the computation. The fixed point quasiconvex subgradient method solves optimization problems whose objective functional is quasiconvex and whose constraint set is expressed as the fixed point set of some nonexpansive mapping. The fixed point quasiconvex subgradient method solves optimization problems whose objective functional is quasiconvex and whose constraint set is expressed as the fixed point set of some nonexpansive mapping. The class of optimization problems that the fixed point quasiconvex subgradient method can deal with includes many important instances appearing in various applied sciences [5, 6, 7, 8].

Reference [6] discusses the efficiency of the inexact quasiconvex subgradient method and provides a detailed analysis of the rate of convergence and numerical examples showing the effect of inexactnesses such as computational errors and noise on the quasiconvex subgradient method. However, the object of consideration in [6] is the quasiconvex subgradient method [6, 9], which assumes the constraint set is expressed as a metric projection that directly projects a given point onto the nearest point in the constraint set. Inspired by the idea of the Krasnosel'skiĭ-Mann algorithm [10, 11], which generates the sequence converging to a fixed point of a given nonexpansive mapping, Reference [5] provides a modification called the fixed point quasiconvex subgradient method. The fixed point quasiconvex subgradient method can be applied to optimization problems whose constraint set is not limited to one onto which the metric projection can be easily computed; instead, its constraint set can be expressed as the fixed point set of some nonexpansive mapping. However, the discussion in [5] does not consider cases in which computational errors

2010 *Mathematics Subject Classification.* 65K05, 90C26, 90C32.
*Key words and phrases.* Quasiconvex optimization, convergence rate analysis, inexact fixed point quasiconvex subgradient method.

and noise exist. Reference [4] performs a convergence analysis of the fixed point quasiconvex subgradient method including the effect of inexactness of its computation. However, to the best of the author's knowledge, there are no studies on the rate of convergence of (or experiments on) the fixed point quasiconvex subgradient method when computational errors and noise exist.

In this paper, we discuss the rate of convergence of the fixed point quasiconvex subgradient method including the effect of the inexactness in its computation. Two analyses are presented: the rate of convergence of the sequence generated by the fixed point quasiconvex subgradient method in terms of the objective functional value and the degree of approximation to the fixed point set. These analyses are extensions of the analyses presented in [5], which does not consider the effect of computational inexactness. Furthermore, this paper presents numerical examples including computational errors and noise. Through these discussions, this paper unravels the effect of the computational inexactness on the efficiency of the fixed point quasiconvex subgradient method.

This paper is organized as follows. Section 2 gives the mathematical preliminaries and defines the inexact fixed point quasiconvex subgradient method. Section 3 discusses the convergence rate analysis of the inexact fixed point quasiconvex subgradient method. Section 4 numerically compares the behaviors of the proposed method and the existing ones. Section 5 concludes this paper.

## 2. Mathematical preliminaries

Let $H$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ and its induced norm $\|\cdot\| : H \to \mathbb{R}$. $\mathbb{N}$ is the set of natural numbers without zero, and $\mathbb{R}$ is the set of real numbers. A functional $f : H \to \mathbb{R}$ is called *quasiconvex* if $f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\}$ for every $x, y \in H$ and $\alpha \in [0,1]$ [1, Definition 5.1], [3, Definition (4.4)]. The effective domain of a functional $f : H \to \mathbb{R}$ is defined as $\mathrm{dom}(f) := \{x \in H : f(x) < \infty\}$.

Here, let $f : H \to \mathbb{R}$ be a quasiconvex, continuous functional, and let $X \subset H$ be a nonempty, closed, convex set. Then, the main problem of this paper is to

$$\text{minimize } f(x) \text{ subject to } x \in X. \tag{2.1}$$

We define the *set of minima* and the *minimum value* of Problem (2.1) by $X^\star := \mathrm{argmin}_{x \in X} f(x)$ and $f_\star := \inf_{x \in X} f(x)$, respectively.

Let us define other terms and notations that will be used in the later discussion. $\mathbf{B} := \{x \in H : \|x\| \leq 1\}$ is the unit ball in this Hilbert space, and $\mathbf{S} := \{x \in H : \|x\| = 1\}$ is the unit sphere in that space. Id is the identity mapping of $H$ onto itself, and the closure of a set $C \subset H$ is denoted by $\mathrm{cl}\, C$.

The metric projection onto a closed, convex set $C \subset H$, denoted by $P_C$, and is defined as $P_C(x) \in C$ such that $\|x - P_C(x)\| = \inf_{y \in C} \|x - y\|$ for any $x \in H$. For any $\alpha \in \mathbb{R}$, the $\alpha$-slice of a functional $f : H \to \mathbb{R}$ is defined as $\mathrm{lev}_{<\alpha} f := \{x \in H : f(x) < \alpha\}$. A mapping $T : H \to H$ is said to be nonexpansive if $\|T(x) - T(y)\| \leq \|x - y\|$ for any $x, y \in H$, and it is said to be firmly nonexpansive if $\|T(x) - T(y)\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$ for any $x, y \in H$. Obviously, a firmly nonexpansive mapping is also a nonexpansive mapping [2, Subchapter 4.1]. The properties of these nonexpansivities are described in detail in [2, Chapter 4], [12, Chapter 6]. The fixed point set of a mapping $T : H \to H$ is defined as $\mathrm{Fix}(T) := \{x \in H : T(x) = x\}$.

The distance of a vector $x \in H$ from a set $Z \subset H$ is defined as $\mathrm{dist}(x, Z) := \inf_{z \in Z} \|x - z\|$ [6, Subsection 2.1]. A functional $f : H \to \mathbb{R}$ is said to satisfy the *Hölder condition* of order $p > 0$ with modulus $\mu > 0$ on $H$ if $f(x) - f_\star \leq \mu(\mathrm{dist}(x, X^\star))^p$ holds for all $x \in H$ [6, Assumption 2]. For given a point $x \in H$ and for a nonnegative real $\epsilon \geq 0$, we call the set $\bar{\partial}^\star_\epsilon f(x) := \{g \in H : \langle g, y - x\rangle \leq 0 \ (y \in \mathrm{lev}_{<f(x)-\epsilon} f)\}$ the $\epsilon$-subdifferential of the quasiconvex functional $f$ at a point $x \in H$ [6, Definition 2.4]. We also call any of its elements a *subgradient*.

In this paper, we consider the fixed point quasiconvex subgradient method incorporating three kinds of computational inexactness, as shown in Algorithm 1. The

**Algorithm 1** Inexact fixed point quasiconvex subgradient method [4, Algorithm 1]

**Require:**
$f$: $H \to \mathbb{R}$, $T$: $H \to H$.
$\{v_k\} \subset (0, \infty)$, $\{\alpha_k\} \subset (0, 1]$. ▷ Hyperparameters
$\{\epsilon_k\} \subset [0, \infty)$, $\{r^f_k\} \subset H$, $\{r^T_k\} \subset H$. ▷ Inexactness

**Ensure:**
This algorithm generates a sequence $\{x_k\} \subset H$.

```
1: x_1 ∈ H.
2: for k = 1, 2, ... do
3:     g_k ∈ ∂̄*_{ε_k} f(x_k) ∩ S.
4:     g̃_k := g_k + r^f_k, T̃_k := T + r^T_k.
5:     x_{k+1} := α_k x_k + (1 − α_k) T̃_k(x_k − v_k g̃_k).
6: end for
```

difference from the original fixed point quasiconvex subgradient method [5, Algorithm 1] is the appearance of the sequences $\{\epsilon_k\}$, $\{r^f_k\}$, and $\{r^T_k\}$. The sequences $\{\epsilon_k\}$ and $\{r^f_k\}$ are from [6], and they express the computational errors and noises respectively. In addition to these sequences, the other sequence $\{r^T_k\}$ we consider expresses the noise appearing in the computation of the nonexpansive mapping.

The following assumption and propositions will be used in the later discussion.

**Assumption 2.1** ([4, Assumption 2.1])**.** We suppose that

(A1) the effective domain $\mathrm{dom}(f) := \{x \in H : f(x) < \infty\}$ coincides with the whole space $H$;
(A2) there exists some firmly nonexpansive mapping $T : H \to H$ whose fixed point set $\mathrm{Fix}(T)$ coincides with the constraint set $X$;
(A3) the constraint set $X$ is nonempty, and there exists at least one minimum, i.e. $X^\star \neq \emptyset$;
(A4) the generated sequence $\{x_k\}$ is bounded [6, Assumption 1];

(A5) the functional $f$ satisfies the Hölder condition of order $p > 0$ with modulus $\mu > 0$ on $H$ [6, Assumption 2];

(A6) the sequence $\{\alpha_k\} \subset (0,1]$ satisfies $0 < \liminf_{k\to\infty} \alpha_k \leq \limsup_{k\to\infty} \alpha_k < 1$ [5, Assumption 3.1];

(A7) there exist some $R_f, R_T, \epsilon \geq 0$ such that $\left\|r_k^f\right\| \leq R_f$ for all $k \in \mathbb{N}$, $\limsup_{k\to\infty} \left\|r_k^T\right\| = R_T$, and $\limsup_{k\to\infty} \epsilon_k = \epsilon$ [6, Assumption 3];

(A8) the sequence $\{v_k\} \subset (0,\infty)$ converges to some nonnegative real $v \in [0,\infty)$, $\sum_{k=1}^{\infty} v_k = \infty$, and there exists a nonnegative real $c \geq 0$ such that $\left\|r_k^T\right\| \leq c v_k$ for all $k \in \mathbb{N}$.

**Proposition 2.2** ([9, Lemma 6.(b)])**.** *If $\bar{x} + \bar{r}\mathbf{B} \subset \mathrm{cl}\left(\mathrm{lev}_{<f(x_k)-\epsilon_k} f\right)$ for some $\bar{x} \in H$ and $\bar{r} \geq 0$, then $\langle g_k, x_k - \bar{x}\rangle \geq \bar{r}$ holds.*

**Proposition 2.3** ([4, Lemma 2.3])**.** *Let $\{x_k\}$ be the sequence generated by Algorithm 1, and suppose that Assumption 2.1 holds. If $f(x_k) > f_\star + \mu \bar{r}^p + \epsilon_k$ holds for some $\bar{r} \geq 0$, then $\langle g_k, x_k - x^\star\rangle \geq \bar{r}$ for all $x^\star \in X^\star$.*

**Proposition 2.4** ([2, Corollary 2.15])**.** *Let $x, y \in H$, and let $\alpha \in \mathbb{R}$. Then,*

$$\|\alpha x + (1-\alpha)y\|^2 = \alpha \|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2$$

*holds.*

## 3. Convergence rate analysis

Before performing the convergence rate analyses, let us show the fundamental inequalities for evaluating the objective functional value and the degree of approximation to the fixed point set. We can use the lemma in [4, Lemma 3.1] for the first purpose. However, we must reexamine the degree of approximation to the fixed point set here, because the lemma shown in [4, Lemma 3.2] does not contain the term $v_k^2$, which is used in the convergence rate analysis.

**Lemma 3.1** ([4, Lemma 3.1])**.** *Let $\{x_k\}$ be the sequence generated by Algorithm 1, and let Assumption 2.1 hold. Suppose that $x^\star \in X^\star$ and $M_1 := \|x^\star\| + \sup_{k\in\mathbb{N}} \|x_k\| < \infty$. Then,*

$$\begin{aligned}\|x_{k+1} - x^\star\|^2 \leq \|x_k - x^\star\|^2 - (1-\alpha_k)\Bigg(2v_k\Bigg(&\langle g_k, x_k - x^\star\rangle - R_f M_1 - \frac{1}{2}v_k(R_f+1)^2 \\ &- \left\|r_k^T\right\|(R_f+1)\Bigg) - \left\|r_k^T\right\|\left(\left\|r_k^T\right\| + 2M_1\right)\Bigg).\end{aligned}$$

*holds for all $k \in \mathbb{N}$.*

**Lemma 3.2.** *Let $\{x_k\}$ be the sequence generated by Algorithm 1, and let Assumptions (A1–7) hold. Suppose that $x^\star \in X^\star$, and assume that the sequence $\{v_k\}$ is bounded from above. Then, a constant $M_2 \geq 0$ exists such that*

$$\begin{aligned}\|x_{k+1} - x^\star\|^2 \leq \|x_k - x^\star\|^2 &- \alpha_k(1-\alpha_k)\left\|x_k - \tilde{T}_k(x_k - v_k\tilde{g}_k)\right\|^2 \\ &+ (1-\alpha_k)v_k^2(1+R_f)^2 + (1-\alpha_k)\left(2v_k R_f M_1 + \left\|r_k^T\right\|\left(2M_2 + \left\|r_k^T\right\|\right)\right).\end{aligned}$$

*for all $k \in \mathbb{N}$ which satisfies $f_\star < f(x_k) - \epsilon_k$.*

*Proof.* Fix $x^\star \in X^\star$ and $k \in \mathbb{N}$ arbitrarily, and suppose that $f_\star < f(x_k) - \epsilon_k$. By using Proposition 2.4, we obtain

$$\begin{aligned} \|x_{k+1} - x^\star\|^2 &= \left\|\alpha_k(x_k - x^\star) + (1-\alpha_k)(\tilde{T}_k(x_k - v_k\tilde{g}_k) - x^\star)\right\|^2 \\ &= \alpha_k \|x_k - x^\star\|^2 + (1-\alpha_k)\left\|\tilde{T}_k(x_k - v_k\tilde{g}_k) - x^\star\right\|^2 \\ &\quad - \alpha_k(1-\alpha_k)\left\|x_k - \tilde{T}_k(x_k - v_k\tilde{g}_k)\right\|^2. \end{aligned} \tag{3.1}$$

Let us further evaluate the term $\|\tilde{T}_k(x_k - v_k\tilde{g}_k) - x^\star\|^2$ appearing on the right side of the above inequality. Since $\tilde{T}_k$ is defined as $\tilde{T}_k := T + r_k^T$, we expand the term $\|\tilde{T}_k(x_k - v_k\tilde{g}_k) - x^\star\|^2$ into

$$\begin{aligned} \left\|\tilde{T}(x_k - v_k\tilde{g}_k) - x^\star\right\|^2 &= \left\|T(x_k - v_k\tilde{g}_k) - x^\star + r_k^T\right\|^2 \\ &= \|T(x_k - v_k\tilde{g}_k) - x^\star\|^2 \\ &\quad + 2\left\langle r_k^T, T(x_k - v_k\tilde{g}_k) - x^\star\right\rangle + \left\|r_k^T\right\|^2. \end{aligned} \tag{3.2}$$

Now let us furthermore proceed to evaluate the term $\|T(x_k - v_k\tilde{g}_k) - x^\star\|^2$. The nonexpansivity of the mapping $T$ and the Cauchy-Schwarz inequality lead to

$$\begin{aligned} \|T(x_k - v_k\tilde{g}_k) - x^\star\|^2 &\le \|x_k - x^\star - v_k\tilde{g}_k\|^2 \\ &\le \|x_k - x^\star\|^2 + v_k^2\|\tilde{g}_k\|^2 \\ &\quad - 2v_k\langle g_k, x_k - x^\star\rangle + 2v_k\left\|r_k^f\right\|\|x_k - x^\star\| \\ &\le \|x_k - x^\star\|^2 + v_k^2(1+R_f)^2 \\ &\quad - 2v_k\langle g_k, x_k - x^\star\rangle + 2v_k R_f\|x_k - x^\star\|. \end{aligned}$$

Here, the definition of the constant $M_1$ in Lemma 3.1 guarantees $\|x_k - x^\star\| \le M_1$; thus, the above inequality can be simplified as

$$\begin{aligned} \|T(x_k - v_k\tilde{g}_k) - x^\star\|^2 \le \|x_k - x^\star\|^2 &+ v_k^2(1+R_f)^2 \\ &- 2v_k\langle g_k, x_k - x^\star\rangle + 2v_k R_f M_1. \end{aligned}$$

The assumption of this lemma states that $f(x^\star) = f_\star < f(x_k) - \epsilon_k$, i.e., $x^\star \in \mathrm{lev}_{<f(x_k - \epsilon_k)} f$. Hence, the definition of $g_k \in \bar{\partial}^\star_{\epsilon_k} f(x_k)$ ensures that $0 \le \langle g_k, x_k - x^\star\rangle$, and we have

$$\|T(x_k - v_k\tilde{g}_k) - x^\star\|^2 \le \|x_k - x^\star\|^2 + v_k^2(1+R_f)^2 + 2v_k R_f M_1.$$

Hence, together with inequality (3.2), we have

$$\begin{aligned} \left\|\tilde{T}(x_k - v_k\tilde{g}_k) - x^\star\right\|^2 &\le \|x_k - x^\star\|^2 + v_k^2(1+R_f)^2 + 2v_k R_f M_1 \\ &\quad + 2\left\langle r_k^T, T(x_k - v_k\tilde{g}_k) - x^\star\right\rangle + \left\|r_k^T\right\|^2. \end{aligned}$$

Here, Assumption (A4) implies that there exists a constant $M \ge 0$ such that $\|x_k\| \le M$ for all $k \in \mathbb{N}$, and the sequence $\{\|x^\star - T(x_k - v_k\tilde{g}_k)\|\}$ is also bounded, because

$\|x^\star - T(x_k - v_k\tilde{g}_k)\| \leq \|x^\star\| + \|x_k\| + v_k\|\tilde{g}_k\| \leq \|x^\star\| + M + (\sup_{j\in\mathbb{N}} v_j)(R_f+1) < \infty$ for all $k \in \mathbb{N}$. Let us define the constant $M_2$ as this upper bound. Using the Cauchy-Schwarz inequality, we can simplify the above inequality as follows:

$$\begin{aligned}\left\|\tilde{T}(x_k - v_k\tilde{g}_k) - x^\star\right\|^2 &\leq \|x_k - x^\star\|^2 + v_k^2(1+R_f)^2 + 2v_kR_fM_1 \\ &\quad + \left\|r_k^T\right\|\left(2\left\|T(x_k - v_k\tilde{g}_k) - x^\star\right\| + \left\|r_k^T\right\|\right) \\ &\leq \|x_k - x^\star\|^2 + v_k^2(1+R_f)^2 + 2v_kR_fM_1 \\ &\quad + \left\|r_k^T\right\|\left(2M_2 + \left\|r_k^T\right\|\right).\end{aligned}$$

Thus, together with inequality (3.1), we have

$$\begin{aligned}\|x_{k+1} - x^\star\|^2 \leq \|x_k - x^\star\|^2 &- \alpha_k(1-\alpha_k)\left\|x_k - \tilde{T}_k(x_k - v_k\tilde{g}_k)\right\|^2 \\ &+ (1-\alpha_k)v_k^2(1+R_f)^2 \\ &+ (1-\alpha_k)\left(2v_kR_fM_1 + \left\|r_k^T\right\|\left(2M_2 + \left\|r_k^T\right\|\right)\right).\end{aligned}$$

This completes the proof. □

Let us discuss the efficiency of Algorithm 1. First, we will examine the rate of convergence in terms of the objective functional value. For this, we introduce two notions: the sequence of best approximate points $\{x_k^\star\} \subset H$ and the inradius of the slice defined by these points $\{r_k\} \subset \mathbb{R} \cup \{-\infty, \infty\}$.

**Definition 3.3** ([6, Definition (5.3)], [5, Definition Appendix G.1])**.** Let $\{x_k\}$ be the sequence generated by Algorithm 1. Suppose that the minimum $x^\star \in X^\star$ is fixed in context. Then, we define

$$\begin{aligned}&\Gamma(k) := \max\left\{i \leq k : f(x_i) - \epsilon_i = \min_{j\leq k}\left(f(x_j) - \epsilon_j\right)\right\}, \\ &x_k^\star := x_{\Gamma(k)}, \\ &r_k := \sup\{r > 0 : x^\star + r\mathbf{B} \subset \operatorname{lev}_{<f(x_k^\star)-\epsilon_{\Gamma(k)}} f\}\end{aligned}$$

for each $k \in \mathbb{N}$.

The following theorem shows the rate of convergence of the sequence generated by Algorithm 1 in terms of the inradius $\{r_k\}$.

**Theorem 3.4.** *Let $\{x_k\}$ be the sequence generated by Algorithm 1, and let Assumption 2.1 hold. Then, a number $k_0 \in \mathbb{N}$ exists such that*

$$\begin{aligned}r_k \leq R_fM_1 &+ \left(\sup_{j\geq i}\left\|r_j^T\right\|\right)(R_f+1) + \frac{c}{2}\left(\left(\sup_{j\geq i}\left\|r_j^T\right\|\right) + 2M_1\right) \\ &+ \frac{\|x_i - x^\star\| + (R_f+1)^2\sum_{j=i}^k(1-\alpha_j)v_j^2}{2\sum_{j=i}^k(1-\alpha_j)v_j}\end{aligned}$$

*for any $x^\star \in X^\star$, $k \geq k_0$, and $i \in \{1, 2, \ldots, k\}$.*

*In particular, $k_0 = 1$, i.e., the above inequality holds for any $k \in \mathbb{N}$ when $\{\alpha_k\} \subset (0,1)$.*

*Proof.* Before proving the theorem, let us give the number $k_0$, which is the lower bound of usable numbers for $k$. In the following, we will use the property that $0 < 1 - \alpha_k$ for (at least) a fixed number $k \in \mathbb{N}$. If $\{\alpha_k\} \subset (0, 1)$, $0 < 1 - \alpha_k$ holds for all $k \in \mathbb{N}$. Hence, we can set $k_0 := 1$ in this case. In the opposite case, i.e., $\alpha_k \in \{0, 1\}$ for some $k \in \mathbb{N}$, Assumption (A6) guarantees a number $k_0 \in \mathbb{N}$ exists such that $0 < (1 - \limsup_{k\to\infty} \alpha_k)/2 < 1 - \alpha_k$ holds for all $k \geq k_0$. Hence, we use this beginning number $k_0$ in this case. In both cases, $0 < 1 - \alpha_k$ is guaranteed for all $k \geq k_0$.

Fix $x^\star \in X^\star$, $k \geq k_0$, and $i \in \{1, 2, \dots, k\}$ arbitrarily. If $r_k$ is nonpositive, the statement obviously holds. Therefore, let us consider the case where $r_k$ is positive. Fix $\delta \in (0, r_k)$ arbitrarily. Then, the definition of $r_k$ implies that

$$\begin{aligned} x^\star + \delta \mathbf{B} &\subset \mathrm{lev}_{<f(x_{\Gamma(k)})-\epsilon_{\Gamma(k)}} f \\ &\subset \mathrm{lev}_{<f(x_j)-\epsilon_j} f. \end{aligned}$$

for all $j \leq k$. Hence, from Lemma 2.2, we have

$$\delta \leq \langle g_j, x_j - x^\star \rangle$$

for all $j \leq k$. Together with Lemma 3.1, we have

$$\begin{aligned} \|x_{k+1} - x^\star\|^2 &\leq \|x_k - x^\star\|^2 \\ &\quad - (1-\alpha_k)\Bigg(2v_k\Bigg(\langle g_k, x_k - x^\star\rangle - R_f M_1 \\ &\quad - \frac{1}{2}v_k(R_f+1)^2 - \left\|r_k^T\right\|(R_f+1)\Bigg) - \left\|r_k^T\right\|\left(\left\|r_k^T\right\| + 2M_1\right)\Bigg) \\ &\leq \|x_i - x^\star\| - 2\delta \sum_{j=i}^{k}(1-\alpha_j)v_j \\ &\quad + \sum_{j=i}^{k}(1-\alpha_j)\Bigg(2v_j\left(R_f M_1 + \frac{1}{2}v_j(R_f+1)^2 + \left\|r_j^T\right\|(R_f+1)\right) \\ &\quad + \left\|r_j^T\right\|\left(\left\|r_j^T\right\| + 2M_1\right)\Bigg). \end{aligned}$$

Here, Assumption (A8) states that $\|r_j^T\| \leq cv_j$. Hence, we can simplify the above inequality as follows:

$$\begin{aligned} \|x_{k+1} - x^\star\|^2 \leq \|x_i - x^\star\| &- 2\delta \sum_{j=i}^{k}(1-\alpha_j)v_j + (R_f+1)^2 \sum_{j=i}^{k}(1-\alpha_j)v_j^2 \\ &+ 2\left(R_f M_1 + \left(\sup_{j\geq i}\left\|r_j^T\right\|\right)(R_f+1)\right. \\ &\left. + \frac{c}{2}\left(\left(\sup_{j\geq i}\left\|r_j^T\right\|\right) + 2M_1\right)\right)\sum_{j=i}^{k}(1-\alpha_j)v_j. \end{aligned}$$

Here, the discussion on the number $k_0$ at the beginning of this proof ensures $0 < (1-\alpha_k)v_k \leq \sum_{j=i}^{k}(1-\alpha_j)v_j$. Hence, the nonnegativity of the left side of the above inequality and positivity of the term $2\sum_{j=i}^{k}(1-\alpha_j)v_j$ leads to

$$\begin{aligned}\delta &\leq R_f M_1 + \left(\sup_{j\geq i}\left\|r_j^T\right\|\right)(R_f+1) + \frac{c}{2}\left(\left(\sup_{j\geq i}\left\|r_j^T\right\|\right) + 2M_1\right) \\ &\quad + \frac{\|x_i - x^\star\| + (R_f+1)^2\sum_{j=i}^{k}(1-\alpha_j)v_j^2}{2\sum_{j=i}^{k}(1-\alpha_j)v_j}.\end{aligned}$$

The arbitrariness of $\delta \in (0, r_k)$ implies that this theorem holds. This completes the proof. □

If there are no errors or noise, i.e., $\epsilon_k := 0$, $r_k^f := 0$, $r_k^T := 0$, and $c := 0$ for all $k \in \mathbb{N}$, Theorem 3.4 coincides with the existing theorem [5, Lemma Appendix G.1] for the exact fixed point quasiconvex subgradient method. This implies that this theorem is an extension of the existing result.

The following proposition shows the relationship between the inradius of the slice and the objective functional value.

**Proposition 3.5.** *Let $\{x_k\}$ be the sequence generated by Algorithm 1, and let Assumption 2.1 hold. Then,*

$$f(x_k^\star) - f_\star \leq \mu r_k^p + \epsilon_{\Gamma(k)}$$

*holds for any $k \in \mathbb{N}$ such that $f_\star < f(x_k^\star) - \epsilon_{\Gamma(k)}$.*

*Proof.* Fix $x^\star \in X^\star$ arbitrarily. The continuity of the functional $f$ and the assumption $f_\star < f(x_k^\star) - \epsilon_{\Gamma(k)}$ guarantees there exists an open ball of nonzero radius and center $x^\star$ contained in the slice $\operatorname{lev}_{<f(x_k^\star)-\epsilon_{\Gamma(k)}} f$. This implies that $r_k$ has a positive value. Since $r_k$ is the supremum of the set $\{r > 0 : x^\star + r\mathbf{B} \subset \operatorname{lev}_{<f(x_k^\star)-\epsilon_{\Gamma(k)}} f\}$,

$$x^\star + \left(r + \frac{1}{j}\right)\mathbf{B} \cap f^{-1}([f(x_k^\star) - \epsilon_{\Gamma(k)}, \infty)) \neq \emptyset$$

holds for all $j \in \mathbb{N}$. Here, we pick a point $u_j$ from this intersection for each $j \in \mathbb{N}$ and define them as the sequence $\{u_j\} \subset f^{-1}([f(x_k^\star) - \epsilon_{\Gamma(k)}, \infty))$. Then, we have

$$\|x^\star - u_j\| \leq r_k + \frac{1}{j}$$

for all $j \in \mathbb{N}$. Since Assumption (A5) guarantees that the functional $f$ satisfies the Hölder condition of order $p > 0$ with modulus $\mu > 0$, we have

$$\begin{aligned}f(x_k^\star) - f_\star &\leq f(u_j) - f_\star + \epsilon_{\Gamma(k)} \\ &\leq \mu\left\|u_j - x^\star\right\|^p + \epsilon_{\Gamma(k)} \\ &\leq \mu\left(r_k + \frac{1}{j}\right)^p + \epsilon_{\Gamma(k)}\end{aligned}$$

for all $j \in \mathbb{N}$. The arbitrariness of $j \in \mathbb{N}$ in the above inequality implies that the statement of this proposition holds. This completes the proof. $\square$

The following theorem gives the convergence rate in terms of the degree of approximation to the fixed point set with respect to the averaged norm. Similarly to Theorem 3.4, this theorem is an extension of the existing result [5, Theorem Appendix G.2].

**Theorem 3.6.** *Let $\{x_k\}$ be the sequence generated by Algorithm 1, and let Assumptions (A1–7) hold. Suppose that $f_\star < f(x_k) - \epsilon_k$ for all $k \in \mathbb{N}$ and $\sum_{k=1}^{\infty} v_k^2 < \infty$. Then, a constant $M \geq 0$ and a monotone increasing function $R : [0,\infty)^2 \to [0,\infty)$ exist such that $R(0,0) = 0$ and*

$$\frac{1}{k}\sum_{j=1}^{k} \|x_j - T(x_j)\|^2 \leq \frac{M}{k} + R\left(\sup_{j\in\mathbb{N}}\|r_j^T\|, R_f\right)$$

*holds for all $k \in \mathbb{N}$.*

*Proof.* Assumption (A6) guarantees the existence of a number $k_0 \in \mathbb{N}$ such that

$$0 < \frac{1}{2}\liminf_{j\to\infty} \alpha_j < \alpha_k < \frac{1}{2}\left(1 + \limsup_{j\to\infty} \alpha_j\right) < 1$$

for all $k \geq k_0$. For the later discussion, we define these lower and upper bounds as $\underline{\alpha} := \liminf_{k\to\infty} \alpha_k$ and $\bar{\alpha} := (1 + \limsup_{k\to\infty} \alpha_k)/2$, respectively. Fix $k \in \mathbb{N}$ arbitrarily, and let us consider the case of $k \geq k_0$. Using the convexity of $\|\cdot\|^2$, we have

$$\begin{aligned}\|x_j - T(x_j)\|^2 &= 4\left\|\frac{1}{2}\left(x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right) + \frac{1}{2}\left(\tilde{T}_j(x_j - v_j\tilde{g}_j) - T(x_j)\right)\right\|^2 \\ &\leq 2\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 + 2\left\|\tilde{T}_j(x_j - v_j\tilde{g}_j) - T(x_j)\right\|^2.\end{aligned} \tag{3.3}$$

for all $j \in \mathbb{N}$. Here, by using the Cauchy-Schwarz inequality, the term $\|\tilde{T}_j(x_j - v_j\tilde{g}_j) - T(x_j)\|^2$ appearing on the right side of the above inequality for $j \in \mathbb{N}$ can be expanded into

$$\begin{aligned}\left\|\tilde{T}_j(x_j - v_j\tilde{g}_j) - T(x_j)\right\|^2 &= \left\|T(x_j - v_j\tilde{g}_j) - T(x_j) + r_j^T\right\|^2 \\ &\leq \|T(x_j - v_j\tilde{g}_j) - T(x_j)\|^2 \\ &\quad + \left\|r_j^T\right\|\left(\|T(x_j - v_j\tilde{g}_j) - T(x_j)\| + \left\|r_j^T\right\|\right) \\ &\leq v_j^2\|\tilde{g}_j\|^2 + \left\|r_j^T\right\|\left(v_j\|\tilde{g}_j\| + \left\|r_j^T\right\|\right) \\ &\leq v_j^2(1+R_f)^2 + \left\|r_j^T\right\|\left(v_j(1+R_f) + \left\|r_j^T\right\|\right).\end{aligned}$$

Together with inequality (3.3), we have

$$\begin{aligned}\|x_j - T(x_j)\|^2 \leq{}& 2\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 + 2v_j^2(1+R_f)^2 \\ &+ 2\left\|r_j^T\right\|\left(v_j(1+R_f) + \left\|r_j^T\right\|\right)\end{aligned}$$

for $j \in \mathbb{N}$. Summing the above inequalities for $j \in \{1, 2, \ldots, k\}$ and dividing both sides by $k$, we obtain

$$\frac{1}{k}\sum_{j=1}^{k}\|x_j - T(x_j)\|^2 \leq \frac{2}{k}\sum_{j=1}^{k}\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 + \frac{2(1+R_f)^2}{k}\sum_{j=1}^{\infty}v_j^2 + 2\sup_{j\in\mathbb{N}}\left(\left\|r_j^T\right\|\left(v_j(1+R_f) + \left\|r_j^T\right\|\right)\right) \tag{3.4}$$

Next, let us evaluate the term $\sum_{j=k_0}^{k}\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\|^2$. From Lemma 3.2, we have

$$\begin{aligned}\|x_{j+1} - x^\star\|^2 \leq \|x_j - x^\star\|^2 - \alpha_j(1-\alpha_j)\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 + (1-\alpha_j)v_j^2(1+R_f)^2 \\ + (1-\alpha_j)\left(2v_jR_fM_1 + \left\|r_j^T\right\|\left(2M_2 + \left\|r_j^T\right\|\right)\right)\end{aligned}$$

for all $j \in \{k_0, k_0+1, \ldots, k\}$. Here, $\alpha_j \in (\underline{\alpha}, \bar{\alpha}) \subset (0,1)$ holds for all $j \in \{k_0, k_0 + 1, \ldots, k\}$. Hence, the above inequality implies that

$$\begin{aligned}\|x_{j+1} - x^\star\|^2 &\leq \|x_j - x^\star\|^2 - \underline{\alpha}(1-\bar{\alpha})\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 + (1+R_f)^2v_j^2 \\ &\quad + 2v_jR_fM_1 + \left\|r_j^T\right\|\left(2M_2 + \left\|r_j^T\right\|\right) \\ &\leq \|x_{k_0} - x^\star\|^2 - \underline{\alpha}(1-\bar{\alpha})\sum_{j=k_0}^{j}\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 \\ &\quad + (1+R_f)^2\sum_{j=k_0}^{j}v_j^2 \\ &\quad + (j - k_0 + 1)\sup_{j\geq k_0}\left(2v_jR_fM_1 + \left\|r_j^T\right\|\left(2M_2 + \left\|r_j^T\right\|\right)\right)\end{aligned}$$

for all $j \in \{k_0, k_0+1, \ldots, k\}$. Thus, we have

$$\begin{aligned}\sum_{j=k_0}^{k}\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2 &\leq \frac{\|x_{k_0} - x^\star\|^2}{\underline{\alpha}(1-\bar{\alpha})} + \frac{(1+R_f)^2}{\underline{\alpha}(1-\bar{\alpha})}\sum_{j=1}^{\infty}v_j^2 \\ &\quad + \frac{k}{\underline{\alpha}(1-\bar{\alpha})}\sup_{j\in\mathbb{N}}\left(2v_jR_fM_1 + \left\|r_j^T\right\|\left(2M_2 + \left\|r_j^T\right\|\right)\right).\end{aligned}$$

Together with inequality (3.4), we obtain

$$\begin{aligned}\frac{1}{k}\sum_{j=1}^{k}\|x_j - T(x_j)\|^2 &\leq \frac{2}{k}\left(\sum_{j=1}^{k_0-1}\left\|x_j - \tilde{T}_j(x_j - v_j\tilde{g}_j)\right\|^2\right. \\ &\quad \left. + \frac{\|x_{k_0} - x^\star\|^2}{\underline{\alpha}(1-\bar{\alpha})} + (1+R_f)^2\left(1 + \frac{1}{\underline{\alpha}(1-\bar{\alpha})}\right)\sum_{j=1}^{\infty}v_j^2\right) \\ &\quad + 2\sup_{j\in\mathbb{N}}\left(\left\|r_j^T\right\|\left(v_j(1+R_f) + \left\|r_j^T\right\|\right)\right) \\ &\quad + \frac{2}{\underline{\alpha}(1-\bar{\alpha})}\sup_{j\in\mathbb{N}}\left(2v_jR_fM_1 + \left\|r_j^T\right\|\left(2M_2 + \left\|r_j^T\right\|\right)\right).\end{aligned}$$

Since the assumption of this theorem ensures the convergence of $\sum_{j=1}^{\infty} v_j^2$, the coefficient of $1/k$ is finite and constant with respect to $k$. Hence, we define the larger of this value and $\sum_{j=1}^{k_0-1} \|x_j - T(x_j)\|^2$ (for the case where $k < k_0$) as $M$. Furthermore, the remaining part of the right side of the above inequality is bounded from above by the value related to $\sup_{j\in\mathbb{N}} \|r_j^T\|$ and $R_f$. In particular, it becomes zero if $\sup_{j\in\mathbb{N}} \|r_j^T\| = R_f = 0$. Therefore, the desired inequality holds. This completes the proof. $\square$

## 4. Numerical examples

We examined the behavior of Algorithm 1 under different step-size rules, noise levels, and errors. We applied it to the following $N$-dimensional constrained quasiconvex optimization, called the Cobb-Douglas production efficiency problem [5, Problem 4.1], [6, Problem (6.1)].

**Problem 4.1** ([5, Problem 4.1])**.** Suppose that $H := \mathbb{R}^n$. Let $a_0, c_0 > 0$, and let $a, c \in (0,\infty)^n$ such that $\sum_{i=1}^n a_i = 1$. Furthermore, let $b_i \in [0,\infty)^n$, $\underline{p}_i \in [0,\infty)^n$, and $\overline{p}_i \in (0,\infty]^n$ for $i = 1, 2, \ldots, m$. Then, we would like to

$$\begin{aligned} &\text{minimize } f(x) := \begin{cases} \frac{-a_0 \prod_{j=1}^n x_j^{a_j}}{\langle c, x\rangle + c_0} & (x \in [0,\infty)^n), \\ 0 & (\text{otherwise}), \end{cases} \\ &\text{subject to } \underline{p}_i \leq \langle b_i, x\rangle \leq \overline{p}_i \quad (i = 1, 2, \ldots, m), \\ &\qquad\qquad x \in [0, M]^n, \end{aligned}$$

where $M > 0$.

Here, the objective functional is quasiconvex [6, Section 6], and the subgradient $g \in \bar{\partial}^\star_{\epsilon_k} f(x)$ for an arbitrarily given point $x \in H$ can be computed [9, Lemma 4]. Furthermore, we can construct a firmly nonexpansive mapping whose fixed point set coincides with the intersection of all constraint sets [5, Section 4]. Hence, we can use Algorithm 1 to solve this problem.

We used a Mac Pro (Late 2013) computer with a 3 GHz 8 Cores Intel Xeon E5 CPU and 32GB 1800MHz DDR3 ECC memory. The experimental code was written in Python 3.7 with NumPy 1.18.1. To construct a firmly nonexpansive mapping expressing the desired constraint set, we used the *fpmlib*[1] toolbox for Python 3.

The parameters of Problem 4.1were set as follows: $n := 100$; $m := 100$; $a_0, c_0 \in (0, 10]$, $\tilde{a} \in (0,1]^n$, $c \in (0,10]^n$ were chosen randomly; $a := \tilde{a}/\sum_{i=1}^n a_i$; $b_i \in [0,1)^n$, $\underline{p}_i \in [0, 25\,\|b_i\|)$, $\overline{p}_i \in (75\,\|b_i\|, 100\,\|b_i\|]$ were chosen randomly for each $i = 1, 2, \ldots, m$; and $M := 100$. These parameter settings are the same as those used in [5, Subsection 4.2]. (Refer to [5] for a detailed evaluation of the behavior of the exact fixed point quasiconvex subgradient method.)

We ran Algorithm 1 under the following conditions:

**C1:** $v_k := 10^{-3}$, $r_k^f := 0$ and $r_k^T := 0$ for all $k \in \mathbb{N}$;

**C2:** $v_k := 10^{-3}$, $\|r_k^f\| \leq 10^{-3}$ and $\|r_k^T\| \leq 10^{-3}$ for all $k \in \mathbb{N}$;

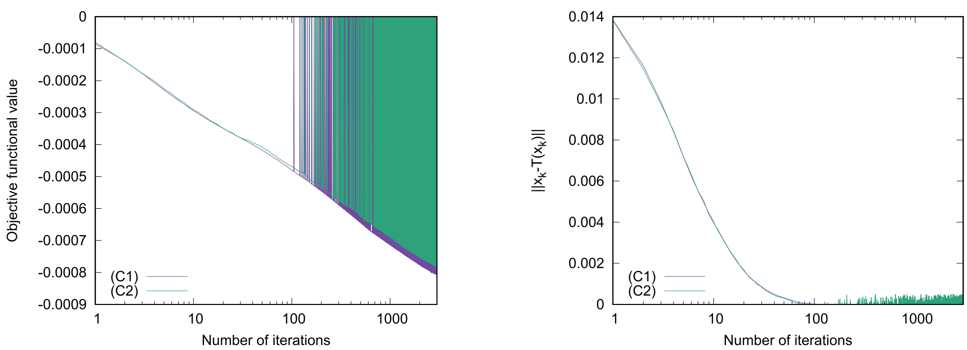
[1]fpmlib: `https://github.com/kazh98/fpmlib`

**D1:** $v_k := 10^{-3}/k$, $r_k^f := 0$ and $r_k^T := 0$ for all $k \in \mathbb{N}$;
**D2:** $v_k := 10^{-3}/k$, $\|r_k^f\| \leq 10^{-3}$ and $\|r_k^T\| \leq 10^{-3}$ for all $k \in \mathbb{N}$;
**D3:** $v_k := 10^{-3}/k$, $\|r_k^f\| \leq 10^{-3}$ and $\|r_k^T\| \leq 10^{-3}/k$ for all $k \in \mathbb{N}$.

Conditions C1 and D1 mean that no noise or error affects Algorithm 1, and the algorithm coincides with the existing exact fixed point quasiconvex subgradient method. Condition C2 expresses the case where Algorithm 1 runs with a constant step size and with noise added to it. Since the added noise is bounded, Assumptions (A7) and (A8) are satisfied. Hence, we can use Theorem 3.4 in this case. However, the constant sequence $\{v_k\}$ violates the assumption $\sum_{k=1}^{\infty} v_k^2 < \infty$ of Theorem 3.6 in this case. Condition D2 and D3 express the case where Algorithm 1 runs with a diminishing step-size and with added noise. Condition D2, unfortunately, violates the assumption of Theorem 3.6 because of its decreasing step size. In contrast to Condition D2, Condition D3 overcomes this violation by introducing a new assumption of decreasing added noise. In addition, we set the other parameters $\alpha_k := 10^{-16}$ and $\epsilon_k := 0$ for all $k \in \mathbb{N}$.

We evaluated the behavior of Algorithm 1 from two viewpoints: the objective functional value $f(x_k)$ and the degree of approximation to the fixed point set $\|x_k - T(x_k)\|$ in each iteration $k \in \mathbb{N}$. We gave $x_0 := 0$ as the initial point.

First, let us see the results for constant step size (Figure 1). The x-axes of

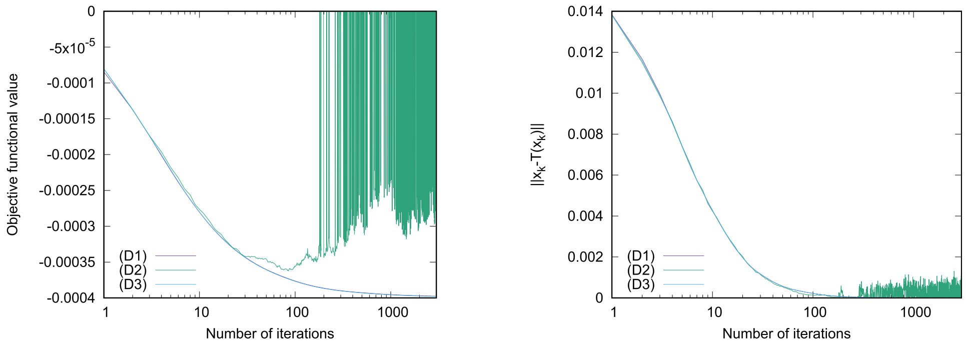



(A) Behavior of the objective functional (B) Behavior of $\|x_k - T(x_k)\|$ in each iteration $k$

FIGURE 1. Experimental results for a constant step size

these graphs show the number of iterations, while the y-axes show the measured values. Conditions C1 and C2 stably decrease both the objective functional value and the degree of approximation to the fixed point set until about 100 iterations. However, the values begin to oscillate after that. This behavior is reasonable because Theorem 3.4 states that the error term is proportional to the step size $v_k$ (and it does not converge to zero in the case of a constant step size).

Figure 2 shows the results in the case of a diminishing step size. Under Condition D3, both the objective functional value and the degree of approximation to the fixed point set decrease, similarly to the existing case (Condition D1). Under




(A) Behavior of the objective functional (B) Behavior of $\|x_k - T(x_k)\|$ in each iteration $k$

FIGURE 2. Experimental results for a diminishing step size

Condition D2, the objective functional value diverges and the degree of approximation to the fixed point set oscillates. These results imply that Assumption (A8) is important for obtaining stable convergence.

## 5. CONCLUSION

This paper discussed the efficiency of the fixed point quasiconvex subgradient method under the condition that errors and noise appear in the computation. The convergence rate analysis showed how much improvement can be expected relative to the number of iterations in terms of both the objective functional value and the degree of approximation to the fixed point set. The numerical examples showed that the proposed algorithm behaves in accordance with the convergence rate analysis, and it revealed the importance of the assumption used in the convergence rate analyses.

**Acknowledgments.** The authors would like to thank Professor Wataru Takahashi for giving us a chance to submit our paper to *Linear and Nonlinear Analysis*. This work was supported by the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Numbers JP17J09220, JP18K11184).

## REFERENCES

[1] D. Aussel. *New developments in quasiconvex optimization*, in: Fixed Point Theory, Variational Analysis, and Optimization, S. A. R. Al-Mezel, F. R. M. Al-Solamy, and Q. H. Ansari (eds), chapter 5, Chapman and Hall/CRC, 2014, pp. 139–169.

[2] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer International Publishing, second edition, 2017.

[3] N. Hadjisavvas. *Convexity, generalized convexity, and applications*, in: Fixed Point Theory, Variational Analysis, and Optimization, S. A. R. Al-Mezel, F. R. M. Al-Solamy, and Q. H. Ansari (eds), chapter 4, Chapman and Hall/CRC, 2014, pp. 139–169.

[4] K. Hishinuma and H. Iiduka, *Evaluation of fixed point quasiconvex subgradient method with computational inexactness*, Pure Appl.Func. Anal. (2020), accepted.

[5] K. Hishinuma and H. Iiduka, *Fixed point quasiconvex subgradient method*, European Journal of Operational Research **282** (2020), 428–437.

[6] Y. Hu, X. Yang, and C.-K. Sim, *Inexact subgradient methods for quasi-convex optimization problems*, European Journal of Operational Research **240** (2015), 315–327.

[7] Y. Hu, C. K. W. Yu, C. Li, and X. Yang, *Conditional subgradient methods for constrained quasi-convex optimization problems*, Journal of Nonlinear and Convex Analysis **17** (2016), 2143–2158.

[8] Y. Hu, C. K. W. Yu, and X. Yang, *Incremental quasi-subgradient methods for minimizing the sum of quasi-convex functions*, Journal of Global Optimization **75** (2019), 1003–1028.

[9] K. C. Kiwiel, *Convergence and efficiency of subgradient methods for quasiconvex minimization*, Mathematical Programming **90** (2001), 1–25.

[10] M. A. Krasnosel'skiĭ, *Two remarks on the method of successive approximations*, Uspekhi Matematicheskikh Nauk **10** (1955), 123–127.

[11] W. R. Mann, *Mean value methods in iteration*, Proceedings of the American Mathematical Society **4** (1953), 506–510.

[12] W. Takahashi, *Introduction to Nonlinear and Convex Analysis*, Yokohama Publishers, Inc., Yokohama, 2009.

*Manuscript received 5 February 2020*
*revised 7 March 2020*

K. Hishinuma
Digital Solutions Division 1st Group, NEC Solution Innovators, Ltd., 2-6-1 Kitamikata, Takatsu-ku, Kawasaki-shi, Kanagawa 213-8511, Japan
*E-mail address*: `kaz-hishinuma@nec.com`

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