J. Nonlinear Funct. Anal. 2021 (2021), Article ID 41 https://doi.org/10.23952/jnfa.2021.41

Journal of Nonlinear Functional Analysis
Available online at http://jnfa.mathres.org

STRONG CONVERGENCE OF A RELAXED INERTIAL THREE-OPERATOR SPLITTING ALGORITHM FOR THE MINIMIZATION PROBLEM OF THE SUM OF THREE OR MORE FUNCTIONS

YANNI GUO*, WEIXIA WANG

College of Science, Civil Aviation University of China, Tianjin 300300, China

Abstract. This paper investigates a relaxed inertial three-operator splitting algorithm for solving the convex optimal problems of the sum of three functions in a real Hilbert space. The corresponding summable perturbation algorithm is also studied. The algorithm is then applied to the general minimization problem of the sum of finite convex functions. Strong convergence of the algorithms are obtained. These algorithms improve the existing results, and the feasibility of them are illustrated by two numerical examples.

Keywords. Convex minimization problem; Fixed point; Inexact algorithm; Relaxed inertial splitting algorithm.

1. INTRODUCTION

We consider the minimization problem of the sum of one convex differentiable function and two convex functions in a real Hilbert space H, and the following more general problem:

$$\min_{x\in H}[\sum_{i=1}^{K} f_i(x)+\sum_{j=1}^{N} g_j(x)], \tag{1.1}$$

where f_i, g_j, $i=1,2,\cdots,K$, $j=1,2,\cdots,N$ are proper, lower semicontinuous, and convex functions, f_i is differentiable Lipschitz continuous with Lipschitz constants L_i for each i, K, N are some finite positive integers. This kind of problems has wide applications in signal processing, control, and machine learning [1, 2, 3, 4]. To solve (1.1) at $K=1$ and $N=2$, Raguet et al. [5] proposed a generalized forward-backward splitting algorithm. Davis and Yin [6] proposed the following three-operator splitting algorithm: for $n=0,1,2\cdots$,

$$y_n = prox_{\gamma g_2} z_n,$$

*Correspoinding author.
E-mail addresses: ynguo@amss.ac.cn (Y. Guo), 1073916711@qq.com (W. Wang).
Received October 3, 2021; Accepted November 12, 2021.

$$x_n = prox_{\gamma g_1}[2y_n - z_n - \gamma\nabla f(y_n)],$$
$$z_{n+1} = z_n + \lambda_n(x_n - y_n).$$

They proved that $prox_{\gamma g_2} z_n$ converges weakly to a critical point of the problem. Subsequently, Yurtsever et al. [7] proposed a stochastic three-operator splitting algorithm. Zong et al. [8] studied an inexact three-operator splitting algorithm. Zhang and Chen [9] proposed a parameterized three-operator splitting algorithm. Cui et al. [10] studied the following inertial three-operator splitting algorithm: for $n = 1, 2\cdots$,

$$\begin{aligned}
\theta_n &= z_n + \xi_n(z_n - z_{n-1}),\\
y_n &= prox_{\gamma_n g_2}\theta_n,\\
x_n &= prox_{\gamma_n g_1}[2y_n - \theta_n - \gamma_n\nabla f(y_n)],\\
z_{n+1} &= (1-\lambda_n)\theta_n + \lambda_n(\theta_n - y_n + x_n).
\end{aligned} \tag{1.2}$$

They proved the weak convergence of algorithm (1.2) based on the inertial method, which is also called the heavy ball method. The method was first introduced by Polyak [11] for speed-up the classical gradient algorithm. In 1983, Nesterov [12] modified the heavy ball method and proposed accelerated gradient algorithms. This modification makes the inertial method more and more popular. For recent results, we refer to [13, 14, 15, 16, 17, 18] and the references therein.

Another technique for acceleration is the relaxation method, which was introduced by Richardson [19] for solving linear systems. In 1992, Eckstein and Bertsekas [20] proposed a relaxed proximal point algorithm to accelerate the standard proximal point algorithm, and pointed out that over-relaxation may indeed speed up the convergence. For more recent results and applications, we refer to [21, 22, 23] and the references therein.

Due to the advantages of relaxation techniques and inertial effects in solving monotone inclusion and convex optimization problems, Alvarez [24] proposed the following relaxed and inertial hybrid proximal point algorithm to solve the inclusion problem: find $x \in H$ such that $0 \in A(x)$ by coupling this two acceleration strategies in an iteration:

$$x_{n+1} = [(1-\rho_n)I + \rho_n J_{\lambda_n A}](x_n + \alpha_n(x_n - x_{n-1})) \tag{1.3}$$

where A is a maximal monotone operator, $J_{\lambda_n A} = (I + \lambda_n A)^{-1}$ is the resolvent of A, $\lambda_n > 0$, and the relaxed factor $\rho_n \in (0,2)$. They established the weak convergence under some assumptions. In 2008, Maingé [25] improved (1.3) by introducing the algorithm:

$$y_n = z_n + \xi_n(z_n - z_{n-1}),$$
$$z_{n+1} = (1-\alpha_n)y_n + \alpha_n T_n(y_n)$$

for finding a common fixed point of a countably infinite family of nonexpansive operators $\{T_n\}$ and discussed the weak convergence of it. Very recently, Attouch and Cabot, in [26], studied algorithm (1.3) under different parameters conditions, and in [27] specialized it in the nonsmooth convex minimization problems to obtain the convergence rates of the algorithm:

$$y_n = x_n + \alpha_n(x_n - x_{n-1}),$$
$$x_{n+1} = (1-\rho_n)y_n + \rho_n prox_{\lambda_n A} y_n.$$

Furthermore, Attouch and Cabot [28] studied the weak convergence properties of a class of relaxed inertial forward-backward algorithms:

$$\begin{aligned} y_n &= x_n + \alpha_n(x_n - x_{n-1}), \\ x_{n+1} &= (1-\rho_n)y_n + \rho_n J_{\mu_n A}(y_n - \mu_n B(y_n)) \end{aligned} \tag{1.4}$$

for the resolution of the structured monotone inclusion:

$$\text{find } x \in H \text{ such that } 0 \in Ax + Bx, \tag{1.5}$$

which is a natural extension of [24]. Besides, Alves et al. [17] derived the alternating direction method of multipliers with inertia and over-relaxation for solving the convex optimization problem $\min_{x\in\mathbb{R}^d}(f(x)+g(x))$. The authors [29, 30] addressed a relaxed inertial Tseng's type method for solving (1.5). Iutzeler and Hendrickx [21] proposed generic acceleration schemes based on relaxation and inertia.

Motivated by the above results, we propose a relaxed inertial three-operator splitting algorithm for solving the minimization problem of the sum of three convex functions and analyze the strong convergence, which further improves (1.4). We also study the corresponding perturbation algorithm as well as the convergence property. In order to solve problem (1.1) via the proposed algorithm, we transform (1.1) into a form in a product space to make it formally the minimization problem of the sum of three functions. As a result, the strong convergence follows.

The remainder of this paper is organized as follows. In Section 2, we briefly review some definitions and lemmas. In Section 3, we propose the relaxed inertial three-operator splitting algorithm, and prove the strong convergence. We also explore the associated summable perturbation algorithm. In Section 4, we generalize the relaxed inertial three-operator splitting algorithm to solve problem (1.1). In the last section, Section 5, we present two numerical examples to illustrate the performance of the proposed algorithms.

2. Preliminaries

Let H be a real Hilbert space endowed with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. Let $\Gamma_0(H)$ be the set of the proper, lower semicontinuous, and convex functions. Let $\{z_n\}$ be a sequence in H, and let z be an element of H. $z_n \to z$ denotes $\{z_n\}$ converges strongly to z, and $z_n \rightharpoonup z$ denotes $\{z_n\}$ converges weakly to z. If a subsequence $\{z_{n_j}\}$ of $\{z_n\}$ converges weakly to z, we call z a weak cluster point of $\{z_n\}$. The set of all weak cluster points of $\{z_n\}$ is denoted by $\omega_w(\{z_n\})$.

Let $T : H \to H$ be a nonlinear operator. Denote the fixed point set of T by $FixT$, that is, $FixT := \{z \in H : Tz = z\}$.

Definition 2.1. (i) T is said to be nonexpansive if

$$\|Tx - Ty\| \le \|x-y\|, \quad \forall x,y \in H.$$

(ii) T is said to be firmly nonexpansive if

$$\|Tx-Ty\|^2 \le \|x-y\|^2 - \|(I-T)x-(I-T)y\|^2, \quad \forall x,y \in H.$$

(iii) T is said to be L-Lipschitz continuous with $L \ge 0$ if

$$\|Tx-Ty\| \le L\|x-y\|, \quad \forall x,y \in H.$$

Especially, T is said to be a contractive mapping if $0 \leq L < 1$.

(iv) T is said to be α-averaged if

$$T = (1-\alpha)I + \alpha S$$

where $\alpha \in (0,1)$ and $S : H \to H$ is a nonexpansive mapping.

(v) T is said to be v-inverse strongly monotone (v-ism) with $v > 0$ if

$$\langle Tx - Ty, x - y\rangle \geq v\|Tx - Ty\|^2, \quad \forall x, y \in H.$$

Definition 2.2. [31] Let $g \in \Gamma_0(H)$. The proximal operator of g is defined by

$$prox_g(x) := \arg\min_{y\in H}\left\{\frac{\|y-x\|^2}{2} + g(y)\right\}, \quad \forall x \in H.$$

The proximal operator of g of order $\alpha > 0$ is defined as the proximal operator of αg, that is,

$$prox_{\alpha g}(x) := \arg\min_{y\in H}\left\{\frac{\|y-x\|^2}{2\alpha} + g(y)\right\}, \quad \forall x \in H.$$

The following lemmas 2.3 and 2.4 give us the property of proximity operators.

Lemma 2.3. [32] *Let $g \in \Gamma_0(H)$ and $\alpha > 0, \mu > 0$. Then*

(i) *$prox_{\alpha g}(x) = (I + \alpha\partial g)^{-1}(x)$;*

(ii) *$prox_{\alpha g}(x) = prox_{\mu g}(\frac{\mu}{\alpha}x + (1 - \frac{\mu}{\alpha})prox_{\alpha g}x)$.*

Lemma 2.4. [33] *Let $g \in \Gamma_0(H)$ and $\alpha > 0$. The proximity operator $prox_{\alpha g}$ is nonexpansive, that is,*

$$\|prox_{\alpha g}(x) - prox_{\alpha g}(y)\| \leq \|x - y\|, \quad \forall x, y \in H.$$

Lemma 2.5. [34] *Let $T : H \to H$ be a nonexpansive mapping with $FixT \neq \emptyset$. If $\{x_n\}$ is a sequence in H converging weakly to x, and if $\{(I-T)x_n\}$ converges strongly to y, then $(I - T)x = y$.*

Lemma 2.6. [35] *Assume that $\{a_n\}$ is a sequence of nonnegative real numbers satisfying*

$$a_{n+1} \leq (1-\bar{\lambda}_n)a_n + \bar{\lambda}_n\beta_n + U_n, \quad n \geq 0,$$

where $\{\bar{\lambda}_n\}$, $\{\beta_n\}$ and $\{U_n\}$ satisfy

(i) *$\bar{\lambda}_n \in [0,1]$, $\sum_{n=0}^{\infty}\bar{\lambda}_n = \infty$;*

(ii) *$\limsup_{n\to\infty}\beta_n \leq 0$;*

(iii) *$U_n \geq 0$ $(n \geq 0)$, $\sum_{n=0}^{\infty}U_n < \infty$.*

Then $\lim_{n\to\infty}a_n = 0$.

Lemma 2.7. [6] *Suppose that T_1, $T_2 : H \to H$ are firmly nonexpansive, and $C : H \to H$ is β-ism with $\beta > 0$. Let $\gamma \in (0, 2\beta)$. Then*

$$T_\gamma = I - T_2 + T_1 \circ (2T_2 - I - \gamma C \circ T_2)$$

is α-averaged with coefficient $\alpha = \frac{2\beta}{4\beta-\gamma} < 1$. In particular, it holds

$$\|T_\gamma z - T_\gamma w\|^2 \leq \|z-w\|^2 - \frac{1-\alpha}{\alpha}\|(I-T_\gamma)z - (I-T_\gamma)w\|^2, \forall z, w \in H.$$

Lemma 2.8. [6] *Let $f, g_1, g_2 \in \Gamma_0(H)$, $x \in H$, and $\gamma > 0$. Assume that f has Lipschitz continuous gradient ∇f, and the solution set of the problem:*

$$\min_{x\in H} f(x) + g_1(x) + g_2(x) \tag{2.1}$$

is nonempty. Then x is a solution of problem (2.1) if and only if there exists $z \in H$ satisfying $x = prox_{\gamma g_2} z$ such that z solves the fixed point equation:

$$z = z - prox_{\gamma g_2} z + prox_{\gamma g_1}(2prox_{\gamma g_2} - I - \gamma\nabla f(prox_{\gamma g_2}))z.$$

Denote by $T_\gamma = I - prox_{\gamma g_2} + prox_{\gamma g_1}(2prox_{\gamma g_2} - I - \gamma\nabla f(prox_{\gamma g_2}))$. Then

$$FixT_\gamma = \{x + \gamma a | a \in \partial g_2(x) \cap (\nabla f(x) + \partial g_1(x))\}.$$

In view of [36], we have the following result.

Lemma 2.9. *Let $f, g_1, g_2 \in \Gamma_0(H)$. Assume that f has Lipschitz continuous gradient ∇f with Lipschitz constant $L > 0$, and $0 < \gamma < \frac{2}{L}$. Set $T_\gamma = I - prox_{\gamma g_2} + prox_{\gamma g_1}(2prox_{\gamma g_2} - I - \gamma\nabla f(prox_{\gamma g_2}))$, and $k = \frac{4-\gamma L}{2}$. Then $kT_\gamma + (1-k)I$ is nonexpansive, that is,*

$$\|(kT_\gamma + (1-k)I)x - (kT_\gamma + (1-k)I)y\| \le \|x - y\|, \ \forall x, y \in H.$$

3. Iterative Algorithms and Their Convergence Analysis

In this section, we present the relaxed inertial three-operator splitting algorithm (RITSA), and discuss the strong convergence of it in subsection 3.1. Then we study the corresponding perturbation algorithm and the convergence property in subsection 3.2. To begin with, let us denote by S the solution set of problem (2.1) or (1.1) and always assume that S is nonempty.

3.1. **Relaxed inertial three-operator splitting algorithm.**

The proposed algorithm is as follows.

Let $z_0,\ z_{-1} \in H$ be two arbitrary initials. For $n = 0, 1, 2, \cdots$, the iterative sequence $\{z_{n+1}\}$ is generated by

$$\begin{aligned}
\theta_n &= z_n + \xi_n(z_n - z_{n-1}),\\
y_n &= prox_{\gamma_n g_2}\theta_n,\\
x_n &= prox_{\gamma_n g_1}[2y_n - \theta_n - \gamma_n D(y_n)\nabla f(y_n)],\\
z_{n+1} &= \lambda_n h(\theta_n) + t_n\theta_n + \alpha_n(\theta_n - y_n + x_n),
\end{aligned} \tag{3.1}$$

where $\gamma_n > 0$, $\{\xi_n\} \subset (0,1)$, $\{\lambda_n\} \subset (0,1)$, $\{t_n\} \subset (-1,1)$, $\{\alpha_n\} \subset (0,2)$ such that $\lambda_n + t_n + \alpha_n = 1$ for all $n \ge 0$, $h : H \to H$ is a ρ-contraction mapping for some $\rho \in [0,1)$, and $D(x) : H \to H$ is a linear operator for each $x \in H$ satisfying

$$\sum_{n=0}^{\infty}\|\delta(y_n)\| := \sum_{n=0}^{\infty}\|\nabla f(y_n) - D\nabla f(y_n)\| := \sum_{n=0}^{\infty}\|\nabla f(y_n) - D(y_n)\nabla f(y_n)\| < \infty.$$

Theorem 3.1. *Suppose that the following conditions hold:*

(i) There exists $\gamma > 0$ such that $\sum_{n=0}^{\infty}|\gamma_n - \gamma| < \infty$, where $0 < \gamma_n \le \overline{\gamma} := \sup_n \gamma_n < \frac{2}{L}$;
(ii) $\lim_{n\to\infty}\lambda_n = 0$, $\sum_{n=0}^{\infty}\lambda_n = \infty$;
(iii) $\sum_{n=0}^{\infty}\xi_n\|z_n - z_{n-1}\| < \infty$;
(iv) $\lim_{n\to\infty}\frac{\lambda_n}{\alpha_n} = 0$, $\sup_n\frac{\alpha_n}{1-\lambda_n} < \frac{4-\overline{\gamma}L}{2}$.

Then the sequence $\{z_n\}$ *generated by algorithm (3.1) converges strongly to a point* $z^* \in FixT_\gamma$*, where* $T_\gamma = I - prox_{\gamma g_2} + prox_{\gamma g_1}(2prox_{\gamma g_2} - I - \gamma\nabla f(prox_{\gamma g_2}))$*. Hence,* $x^* = prox_{\gamma g_2} z^*$ *is a solution of problem (2.1).*

Proof. The proof is split into two parts. In the first part, we verify the weak cluster points of some subsequence of $\{z_n\}$ are the fixed points of T_γ, where $T_\gamma = I - prox_{\gamma g_2} + prox_{\gamma g_1}(2prox_{\gamma g_2} - I - \gamma\nabla f(prox_{\gamma g_2}))$. In the second part, we prove that $\{z_n\}$ converges strongly to some z^* in $Fix(T_\gamma)$. From Lemma 2.8, we have that $x^* = prox_{\gamma g_2} z^*$ is a solution of problem (2.1).

To complete the first part, it is necessary to prove that $\{z_n\}$ is a bounded sequence in H. To this end, we denote by

$$\begin{aligned} D_n &:= I - prox_{\gamma_n g_2} + prox_{\gamma_n g_1}(2prox_{\gamma_n g_2} - I - \gamma_n D\nabla f(prox_{\gamma_n g_2})), \\ T_n &:= I - prox_{\gamma_n g_2} + prox_{\gamma_n g_1}(2prox_{\gamma_n g_2} - I - \gamma_n \nabla f(prox_{\gamma_n g_2})), \\ E_n &:= D_n - T_n, \end{aligned} \tag{3.2}$$

and observe that

$$\begin{aligned} \|E_n\theta_n\| &= \|D_n\theta_n - T_n\theta_n\| \\ &= \|\theta_n - prox_{\gamma_n g_2}\theta_n + prox_{\gamma_n g_1}(2prox_{\gamma_n g_2} - I - \gamma_n D\nabla f(prox_{\gamma_n g_2}))\theta_n \\ &\quad - \theta_n + prox_{\gamma_n g_2}\theta_n - prox_{\gamma_n g_1}(2prox_{\gamma_n g_2} - I - \gamma_n\nabla f(prox_{\gamma_n g_2}))\theta_n\| \\ &= \|prox_{\gamma_n g_1}(2prox_{\gamma_n g_2} - I - \gamma_n D\nabla f(prox_{\gamma_n g_2}))\theta_n \\ &\quad - prox_{\gamma_n g_1}(2prox_{\gamma_n g_2} - I - \gamma_n\nabla f(prox_{\gamma_n g_2}))\theta_n\| \\ &\le \gamma_n\|D\nabla f(prox_{\gamma_n g_2}\theta_n) - \nabla f(prox_{\gamma_n g_2}\theta_n)\|, \end{aligned}$$

which yields

$$\begin{aligned} \sum_{n=0}^{\infty}\|E_n\theta_n\| &= \sum_{n=0}^{\infty}\|D_n\theta_n - T_n\theta_n\| \\ &\le \sum_{n=0}^{\infty}\gamma_n\|D\nabla f(prox_{\gamma_n g_2}\theta_n) - \nabla f(prox_{\gamma_n g_2}\theta_n)\| \\ &< \infty. \end{aligned} \tag{3.3}$$

Furthermore, we rewrite algorithm (3.1) as

$$\begin{aligned} z_{n+1} &= \lambda_n h(\theta_n) + t_n\theta_n + \alpha_n(\theta_n - y_n + x_n) \\ &= \lambda_n h(\theta_n) + t_n\theta_n + \alpha_n D_n\theta_n \\ &= \lambda_n h(z_n) + (1-\lambda_n)u_n, \end{aligned}$$

where $u_n = \frac{t_n}{1-\lambda_n}\theta_n + \frac{\alpha_n}{1-\lambda_n}D_n\theta_n$. Letting $\rho_n = \frac{\alpha_n}{1-\lambda_n}$, we have

$$0 < \rho_n \le \sup_n \rho_n = \sup_n \frac{\alpha_n}{1-\lambda_n} < \frac{4-\bar{\gamma}L}{2} \le \frac{4-\gamma_n L}{2} < 2 \tag{3.4}$$

so that $u_n = (1-\rho_n)\theta_n + \rho_n D_n\theta_n$. Now, fix $x \in S$. Lemma 2.8 states that there exist v, $v_n \in H$ satisfying $v_n \in (I + \gamma_n\partial g_2)x$ and $v \in (I + \gamma\partial g_2)x$ such that $v_n = T_n v_n$, $v = Tv$ and $v_n - v =$

$(\gamma_n-\gamma)a,\ n=0,1,2,\cdots$. Utilizing Lemma 2.7, we have

$$\begin{aligned}\|T_n\theta_n-v_n\|^2 &\le \|\theta_n-v_n\|^2-\frac{2-\gamma_nL}{2}\|(I-T_n)\theta_n-(I-T_n)v_n\|^2\\ &=\|\theta_n-v_n\|^2-\frac{2-\gamma_nL}{2}\|\theta_n-T_n\theta_n\|^2,\end{aligned}$$

and

$$\begin{aligned}&\|(1-\rho_n)\theta_n+\rho_nT_n\theta_n-v_n\|^2\\ &\le(1-\rho_n)\|\theta_n-v_n\|^2-\rho_n(1-\rho_n)\|\theta_n-T_n\theta_n\|^2+\rho_n\Big[\|\theta_n-v_n\|^2-\frac{2-\gamma_nL}{2}\|\theta_n-T_n\theta_n\|^2\Big]\\ &=\|\theta_n-v_n\|^2-\rho_n\big(\frac{4-\gamma_nL}{2}-\rho_n\big)\|\theta_n-T_n\theta_n\|^2\\ &\le\|\theta_n-v_n\|^2.\end{aligned}$$

Consequently,

$$\begin{aligned}&\|z_{n+1}-v\|\\ &\le\lambda_n\|h(\theta_n)-v\|+(1-\lambda_n)\|u_n-v\|\\ &\le\lambda_n\|h(\theta_n)-h(v)\|+\lambda_n\|h(v)-v\|+(1-\lambda_n)\|u_n-v\|\\ &\le\lambda_n\rho\|\theta_n-v\|+\lambda_n\|h(v)-v\|+(1-\lambda_n)\|(1-\rho_n)\theta_n+\rho_nT_n\theta_n-v_n\|\\ &\quad+(1-\lambda_n)\rho_n\|D_n\theta_n-T_n\theta_n\|+(1-\lambda_n)\|v_n-v\|\\ &\le\lambda_n\rho\|\theta_n-v\|+\lambda_n\|h(v)-v\|+(1-\lambda_n)\|\theta_n-v_n\|\\ &\quad+(1-\lambda_n)\rho_n\|E_n\theta_n\|+(1-\lambda_n)\|v_n-v\|\\ &<(1-\lambda_n(1-\rho))\|\theta_n-v\|+\lambda_n\|h(v)-v\|\\ &\quad+(1-\lambda_n)[2\|E_n\theta_n\|+2\|v_n-v\|]\qquad(\text{since }0<\rho_n<2)\\ &\le(1-\lambda_n(1-\rho))\big[\|\theta_n-v\|+2\|E_n\theta_n\|+2\|v_n-v\|\big]+\lambda_n\|h(v)-v\|\\ &\le\max\big\{\|z_n-v\|+\xi_n\|z_n-z_{n-1}\|+2\|E_n\theta_n\|+2|\gamma_n-\gamma|\|a\|,\frac{\|h(v)-v\|}{1-\rho}\big\}.\end{aligned}$$

An induction argument shows that

$$\begin{aligned}&\|z_{n+1}-v\|\\ &<\max\{\|z_0-v\|+\sum_{i=0}^{n}(\xi_i\|z_i-z_{i-1}\|+2\|E_i\theta_i\|+2|\gamma_i-\gamma|\|a\|),\frac{\|h(v)-v\|}{1-\rho}\|\}\\ &\le\max\Big\{\|z_0-v\|+M_1,\frac{\|h(v)-v\|}{1-\rho}\Big\},\end{aligned}$$

where $M_1=\sum_{i=0}^{+\infty}(\xi_i\|z_i-z_{i-1}\|+2\|E_i\theta_i\|+2|\gamma_i-\gamma|\|a\|)<\infty$. This is due to (3.3) and the conditions (i) and (iii) of Theorem 3.1. Hence, $\{z_n\}$ is a bounded sequence. In addition, $\{\theta_n\}$, $\{h(z_n)\}$, and $\{prox_{\gamma g_2}z_n\}$ are also bounded sequences since $\theta_n=z_n+\xi_n(z_n-z_{n-1})$, $\{\xi_n\}\subset(0,1)$, h is a ρ-contraction, and $prox_{\gamma g_2}$ is a nonexpansive operator.

Next, we employ Lemma 2.5 to prove that there exists a subsequence $\{z_{n_j}\} \subset \{z_n\}$ such that $\omega_w(\{z_{n_j}\}) \subset FixT_\gamma$. To this end, let us prove $\lim_{j\to\infty} \|T_{n_j}\theta_{n_j} - \theta_{n_j}\| = 0$. We need to estimate

$$\begin{aligned}
\|z_{n+1} - v\|^2 &= \|\lambda_n(h(\theta_n) - h(v)) + (1-\lambda_n)(u_n - v) + \lambda_n(h(v) - v)\|^2 \\
&\le \|\lambda_n(h(\theta_n) - h(v)) + (1-\lambda_n)(u_n - v)\|^2 + 2\lambda_n\langle z_{n+1} - v, h(v) - v\rangle \\
&= \lambda_n\|h(\theta_n) - h(v)\|^2 + (1-\lambda_n)\|u_n - v\|^2 \\
&\quad - \lambda_n(1-\lambda_n)\|h(\theta_n) - h(v) - (u_n - v)\|^2 + 2\lambda_n\langle z_{n+1} - v, h(v) - v\rangle \\
&\le \lambda_n\rho^2\|\theta_n - v\|^2 + (1-\lambda_n)\|u_n - v\|^2 + 2\lambda_n\langle z_{n+1} - v, h(v) - v\rangle.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\|u_n - v\|^2 \\
&= \|\theta_n - v + \rho_n(T_n\theta_n - \theta_n) + \rho_n E_n\theta_n\|^2 \\
&\le \|\theta_n - v + \rho_n(T_n\theta_n - \theta_n)\|^2 + \rho_n^2\|E_n\theta_n\|^2 + 2\rho_n\|E_n\theta_n\|\|\theta_n - v + \rho_n(T_n\theta_n - \theta_n)\| \\
&\le \|\theta_n - v\|^2 + \rho_n^2\|T_n\theta_n - \theta_n\|^2 - 2\rho_n\langle\theta_n - v, \theta_n - T_n\theta_n\rangle + M_2\|E_n\theta_n\| \\
&= (1-\rho_n)\|\theta_n - v\|^2 - \rho_n(1-\rho_n)\|T_n\theta_n - \theta_n\|^2 + \rho_n\|T_n\theta_n - v\|^2 + M_2\|E_n\theta_n\|,
\end{aligned}$$

where $M_2 = \sup_n\left(\rho_n^2\|E_n\theta_n\| + 2\rho_n\|\theta_n - v + \rho_n(T_n\theta_n - \theta_n)\|\right)$, and

$$\begin{aligned}
&\|T_n\theta_n - v\|^2 \\
&\le \|T_n\theta_n - T_nv_n\|^2 + \|v_n - v\|^2 + 2\|\theta_n - v_n\|\|v_n - v\| \\
&\le \|\theta_n - v_n\|^2 - \frac{2-\gamma_n L}{2}\|T_n\theta_n - \theta_n\|^2 + \|v_n - v\|^2 + 2\Big(\|\theta_n - v\|) + \|v_n - v\|\Big)\|v_n - v\| \\
&\le \|\theta_n - v\|^2 + 2\|\theta_n - v\|\|v_n - v\| - \frac{2-\gamma_n L}{2}\|T_n\theta_n - \theta_n\|^2 + 4\|v_n - v\|^2 + 2\|\theta_n - v\|\|v_n - v\| \\
&= \|\theta_n - v\|^2 - \frac{2-\gamma_n L}{2}\|T_n\theta_n - \theta_n\|^2 + 4\|v_n - v\|^2 + 4\|\theta_n - v\|\|v_n - v\|.
\end{aligned}$$

It follows that

$$\begin{aligned}
\|u_n - v\|^2 &\le (1-\rho_n)\|\theta_n - v\|^2 - \rho_n(1-\rho_n)\|T_n\theta_n - \theta_n\|^2 + \rho_n\|T_n\theta_n - v\|^2 + M_2\|E_n\theta_n\| \\
&\le (1-\rho_n)\|\theta_n - v\|^2 - \rho_n(1-\rho_n)\|T_n\theta_n - \theta_n\|^2 + M_2\|E_n\theta_n\| \\
&\quad + \rho_n\Big[\|\theta_n - v\|^2 - \frac{2-\gamma_n L}{2}\|T_n\theta_n - \theta_n\|^2 + 4\|v_n - v\|^2 + 4\|\theta_n - v\|\|v_n - v\|\Big] \\
&= \|\theta_n - v\|^2 - \rho_n\left(\frac{4-\gamma_n L}{2} - \rho_n\right)\|T_n\theta_n - \theta_n\|^2 + M_2\|E_n\theta_n\| \\
&\quad + 4\rho_n\|v_n - v\|\Big[\|\theta_n - v\| + \|v_n - v\|\Big].
\end{aligned}$$

Consequently,

$$
\begin{aligned}
&\|z_{n+1}-v\|^2\\
&\le \lambda_n\rho^2\|\theta_n-v\|^2+(1-\lambda_n)\|u_n-v\|^2+2\lambda_n\langle z_{n+1}-v,h(v)-v\rangle\\
&\le \lambda_n\rho^2\|\theta_n-v\|^2+(1-\lambda_n)\left\{\|\theta_n-v\|^2-\rho_n\left(\frac{4-\gamma_nL}{2}-\rho_n\right)\|T_n\theta_n-\theta_n\|^2+M_2\|E_n\theta_n\|\right.\\
&\quad\left.+4\rho_n\|v_n-v\|\Big[\|\theta_n-v\|+\|v_n-v\|\Big]\right\}+2\lambda_n\langle z_{n+1}-v,h(v)-v\rangle\\
&\le (1-\lambda_n(1-\rho^2))\|\theta_n-v\|^2+2\lambda_n\langle z_{n+1}-v,h(v)-v\rangle-\alpha_n\left(\frac{4-\gamma_nL}{2}-\rho_n\right)\|T_n\theta_n-\theta_n\|^2\\
&\quad+(1-\lambda_n)\Big[M_2\|E_n\theta_n\|+4\rho_n\|v_n-v\|(\|\theta_n-v\|+\|v_n-v\|)\Big]\\
&\le (1-\lambda_n(1-\rho^2))\|z_n-v\|^2-\alpha_n\left(\frac{4-\gamma_nL}{2}-\rho_n\right)\|T_n\theta_n-\theta_n\|^2\\
&\quad+2\lambda_n\langle z_{n+1}-v,h(v)-v\rangle+(1-\lambda_n)\Big[M_2\|E_n\theta_n\|+4\rho_n\|v_n-v\|(\|\theta_n-v\|+\|v_n-v\|)\Big]\\
&\quad+(1-\lambda_n(1-\rho^2))[\xi_n^2\|z_n-z_{n-1}\|^2+2\xi_n\|z_n-z_{n-1}\|\|z_n-v\|]. \qquad (3.5)
\end{aligned}
$$

Take

$$\beta_n=\frac{2\langle z_{n+1}-v,h(v)-v\rangle}{1-\rho^2}-\frac{\alpha_n}{\lambda_n(1-\rho^2)}\left(\frac{4-\gamma_nL}{2}-\rho_n\right)\|T_n\theta_n-\theta_n\|^2. \qquad (3.6)$$

Then $\{\beta_n\}$ is a sequence with an upper bound. So there exists $\{\beta_{n_j}\}$, a subsequence of $\{\beta_n\}$, such that

$$
\begin{aligned}
+\infty &> \limsup_{n\to\infty}\beta_n\\
&=\lim_{j\to\infty}\beta_{n_j}\\
&=\lim_{j\to\infty}\Big[\frac{2}{1-\rho^2}\langle z_{n_j+1}-v,h(v)-v\rangle-\frac{\alpha_{n_j}}{\lambda_{n_j}(1-\rho^2)}(\frac{4-\gamma_{n_j}L}{2}-\rho_{n_j})\|T_{n_j}\theta_{n_j}-\theta_{n_j}\|^2\Big]\\
&=\frac{1}{1-\rho^2}\left(\lim_{j\to\infty}2\langle z_{n_j+1}-v,h(v)-v\rangle-\lim_{j\to\infty}\frac{\alpha_{n_j}}{\lambda_{n_j}}(\frac{4-\gamma_{n_j}L}{2}-\rho_{n_j})\|T_{n_j}\theta_{n_j}-\theta_{n_j}\|^2\right). \qquad (3.7)
\end{aligned}
$$

Here, without loss of generality, we assume that $\lim_{j\to\infty}\langle z_{n_j+1}-v,h(v)-v\rangle$ exists since $\langle z_{n+1}-v,h(v)-v\rangle$ is a bounded sequence. Then the second term in the parentheses of (3.7) implies

$$\lim_{j\to\infty}\|T_{n_j}\theta_{n_j}-\theta_{n_j}\|=0$$

owing to (3.4) and that $\lim_{j\to\infty}\frac{\alpha_{n_j}}{\lambda_{n_j}}=\infty$. To verify $\lim_{j\to\infty}\|T_\gamma z_{n_j}-z_{n_j}\|=0$, we set

$$S_{n_j}:=2prox_{\gamma_{n_j}g_2}-I-\gamma_{n_j}\nabla f(prox_{\gamma_{n_j}g_2}),S:=2prox_{\gamma g_2}-I-\gamma\nabla f(prox_{\gamma g_2}),$$

and observe

$$\begin{aligned}
&\|T_\gamma z_{n_j} - z_{n_j}\| \\
&\leq \|T_\gamma z_{n_j} - T_\gamma \theta_{n_j}\| + \|T_\gamma \theta_{n_j} - \theta_{n_j}\| + \|\theta_{n_j} - z_{n_j}\| \\
&\leq 2\|\theta_{n_j} - z_{n_j}\| + \|T_\gamma \theta_{n_j} - T_{n_j}\theta_{n_j} + T_{n_j}\theta_{n_j} - \theta_{n_j}\| \\
&\leq 2\xi_{n_j}\|z_{n_j} - z_{n_j-1}\| + \|T_\gamma \theta_{n_j} - T_{n_j}\theta_{n_j}\| + \|T_{n_j}\theta_{n_j} - \theta_{n_j}\|. \qquad (3.8)
\end{aligned}$$

By using Lemma 2.3 (ii) and Lemma 2.4, we conclude that

$$\begin{aligned}
&\|T_{n_j}\theta_{n_j} - T_\gamma \theta_{n_j}\| \\
&= \|\theta_{n_j} - prox_{\gamma_{n_j} g_2}\theta_{n_j} + prox_{\gamma_{n_j} g_1} S_{n_j}\theta_{n_j} - \theta_{n_j} + prox_{\gamma g_2}\theta_{n_j} - prox_{\gamma g_1} S\theta_{n_j}\| \\
&\leq \|prox_{\gamma g_2}\theta_{n_j} - prox_{\gamma_{n_j} g_2}\theta_{n_j}\| + \|prox_{\gamma g_1} S\theta_{n_j} - prox_{\gamma_{n_j} g_1} S_{n_j}\theta_{n_j}\| \\
&= \|prox_{\gamma_{n_j} g_2}(\frac{\gamma_{n_j}}{\gamma}\theta_{n_j} + (1 - \frac{\gamma_{n_j}}{\gamma}) prox_{\gamma g_2}\theta_{n_j}) - prox_{\gamma_{n_j} g_2}\theta_{n_j}\| \\
&\quad + \|prox_{\gamma_{n_j} g_1}(\frac{\gamma_{n_j}}{\gamma} S\theta_{n_j} + (1 - \frac{\gamma_{n_j}}{\gamma}) prox_{\gamma g_1} S\theta_{n_j}) - prox_{\gamma_{n_j} g_1} S_{n_j}\theta_{n_j}\| \\
&\leq \|(1 - \frac{\gamma_{n_j}}{\gamma})(prox_{\gamma g_2}\theta_{n_j} - \theta_{n_j})\| + \|\frac{\gamma_{n_j}}{\gamma} S\theta_{n_j} + (1 - \frac{\gamma_{n_j}}{\gamma}) prox_{\gamma g_1} S\theta_{n_j} - S_{n_j}\theta_{n_j}\| \\
&= \|(1 - \frac{\gamma_{n_j}}{\gamma})(prox_{\gamma g_2}\theta_{n_j} - \theta_{n_j})\| \\
&\quad + \|\frac{\gamma_{n_j}}{\gamma} S\theta_{n_j} - \frac{\gamma_{n_j}}{\gamma} S_{n_j}\theta_{n_j} + \frac{\gamma_{n_j}}{\gamma} S_{n_j}\theta_{n_j} - S_{n_j}\theta_{n_j} + (1 - \frac{\gamma_{n_j}}{\gamma}) prox_{\gamma g_1} S\theta_{n_j}\| \\
&\leq \left|1 - \frac{\gamma_{n_j}}{\gamma}\right| \Big(\|prox_{\gamma g_2}\theta_{n_j} - \theta_{n_j}\| + \|prox_{\gamma g_1} S\theta_{n_j} - S_{n_j}\theta_{n_j}\|\Big) + \frac{\gamma_{n_j}}{\gamma}\|S\theta_{n_j} - S_{n_j}\theta_{n_j}\|. \qquad (3.9)
\end{aligned}$$

A similar calculation gives

$$\begin{aligned}
&\|S\theta_{n_j} - S_{n_j}\theta_{n_j}\| \\
&\leq \left|1 - \frac{\gamma_{n_j}}{\gamma}\right| (2 + \gamma_{n_j} L)\|prox_{\gamma g_2}\theta_{n_j} - \theta_{n_j}\| + \left|\gamma - \gamma_{n_j}\right| \|\nabla f(prox_{\gamma g_2})\theta_{n_j}\|. \qquad (3.10)
\end{aligned}$$

Substituting (3.9), (3.10) into (3.8) and noticing that $\lim_{n\to\infty}\gamma_n = \gamma$, $\{\|prox_{\gamma g_2}\theta_{n_j} - \theta_{n_j}\|\}$, $\{\|prox_{\gamma g_1} S\theta_{n_j} - S_{n_j}\theta_{n_j}\|\}$, and $\{\|\nabla f(prox_{\gamma g_2})\theta_{n_j}\|\}$ are bounded sequences, we deduce

$$\lim_{j\to\infty}\|T_\gamma z_{n_j} - z_{n_j}\| = 0. \qquad (3.11)$$

On the other hand, the boundedness of $\{z_n\}$ implies its subsequence $\{z_{n_j}\}$ is bounded. So there is a weak convergent subsequence of $\{z_{n_j}\}$. Without loss of generality, we use $\{z_{n_j}\}$ to represent any of its weak convergent subsequences, and assume that $\{z_{n_j}\}$ converges weakly to some point z^* as $j\to\infty$. The nonexpansiveness of T_γ together with (3.11) guarantees that $z^* \in Fix(T_\gamma)$, that is, $\omega_w(\{z_{n_j}\}) \subset FixT_\gamma$.

We now show that $\{z_n\}$ converges strongly to $z^* \in FixT_\gamma$ to complete the second part. In view of (3.5) and (3.6),

$$\begin{aligned}\|z_{n+1}-z^*\|^2 \leq\ & (1-\lambda_n(1-\rho^2))\|z_n-v\|^2+\lambda_n(1-\rho^2)\beta_n\\ &+(1-\lambda_n)\Big[M_2\|E_n\theta_n\|+4\rho_n\|v_n-v\|(\|\theta_n-v\|+\|v_n-v\|)\Big]\\ &+(1-\lambda_n(1-\rho^2))[\xi_n^2\|z_n-z_{n-1}\|^2+2\xi_n\|z_n-z_{n-1}\|\|z_n-v\|].\end{aligned}$$

Take

$$\begin{aligned}\bar{\lambda}_n &= \lambda_n(1-\rho^2)\in[0,1],\\ U_n &= (1-\lambda_n)\Big[M_2\|E_n\theta_n\|+4\rho_n\|v_n-v\|(\|\theta_n-v\|+\|v_n-v\|)\Big]\\ &\quad+(1-\lambda_n(1-\rho^2))[\xi_n^2\|z_n-z_{n-1}\|^2+2\xi_n\|z_n-z_{n-1}\|\|z_n-v\|]\geq 0.\end{aligned}$$

We verify that $\{\bar{\lambda}_n\}$, $\{\beta_n\}$, and $\{U_n\}$ meet the conditions of Lemma 2.6. Obviously, $\sum_{n=0}^{\infty}\bar{\lambda}_n=\infty$, $\sum_{n=0}^{\infty}U_n<\infty$. To explain $\limsup_{n\to\infty}\beta_n\leq 0$, let us certify $z_{n_j+1}\rightharpoonup z^*$ as $j\to\infty$. In fact, from $z_{n_j}\rightharpoonup z^*$ $(j\to\infty)$ and

$$\begin{aligned}&\|z_{n_j+1}-z_{n_j}\|\\ &=\|\lambda_{n_j}h(\theta_{n_j})+t_{n_j}\theta_{n_j}+\alpha_{n_j}D_{n_j}\theta_{n_j}-z_{n_j}\|\\ &=\|\lambda_{n_j}(h(\theta_{n_j})-z_{n_j})+t_{n_j}(\theta_{n_j}-z_{n_j})+\alpha_{n_j}(E_{n_j}\theta_{n_j}+T_{n_j}\theta_{n_j}-z_{n_j})\|\\ &\leq\lambda_{n_j}\|h(\theta_{n_j})-z_{n_j}\|+\|\theta_{n_j}-z_{n_j}\|+\alpha_{n_j}\left[\|E_{n_j}\theta_{n_j}\|+\|T_{n_j}\theta_{n_j}-\theta_{n_j}\|+\|\theta_{n_j}-z_{n_j}\|\right]\\ &\leq\lambda_{n_j}[\|h(\theta_{n_j})\|+\|z_{n_j}\|]+(1+\alpha_{n_j})\xi_{n_j}\|z_{n_j}-z_{n_j-1}\|+\alpha_{n_j}\left[\|E_{n_j}\theta_{n_j}\|+\|T_{n_j}\theta_{n_j}-\theta_{n_j}\|\right]\\ &\to 0,\ \text{as } j\to\infty,\end{aligned}$$

it is easy to see that $z_{n_j+1}\rightharpoonup z^*$ as $j\to\infty$. Hence, by choosing $v=z^*$ in (3.7), we have

$$\begin{aligned}&\limsup_{n\to\infty}\beta_n=\lim_{j\to\infty}\beta_{n_j}\\ &=\frac{1}{1-\rho^2}\left(\lim_{j\to\infty}2\langle z_{n_j+1}-z^*,h(z^*)-z^*\rangle-\lim_{j\to\infty}\frac{\alpha_{n_j}}{\lambda_{n_j}}\Big(\frac{4-\gamma_{n_j}L}{2}-\rho_{n_j}\Big)\|T_{n_j}\theta_{n_j}-\theta_{n_j}\|^2\right)\\ &\leq\frac{1}{1-\rho^2}\lim_{j\to\infty}2\langle z_{n_j+1}-z^*,h(z^*)-z^*\rangle=0.\end{aligned}$$

From Lemma 2.6, we have $\lim_{n\to\infty}\|z_n-z^*\|=0$. That is, $\{z_n\}$ converges strongly to $z^*\in FixT_\gamma$. In addition, Lemma 2.8 shows that $x^*=prox_{\gamma g_2}z^*$ solves problem (2.1). □

3.2. **Perturbation Algorithm.** In this subsection, we discuss an inexact version of (3.1) with summable perturbations, and prove the strong convergence under the same conditions as in Theorem 3.1. The sequence $\{\tilde{z}_n\}$ is generated by the following iteration.

Let $\tilde{z}_0$, $\tilde{z}_{-1}\in H$ be two arbitrary initials. For $n=0,1,2,\cdots$, define $\{\tilde{z}_{n+1}\}$ by

$$\begin{aligned}\tilde{\theta}_n &= \tilde{z}_n+\xi_n(\tilde{z}_n-\tilde{z}_{n-1}),\\ \tilde{y}_n &= prox_{\gamma_n g_2}(\tilde{\theta}_n+e_n),\\ \tilde{x}_n &= prox_{\gamma_n g_1}[2\tilde{y}_n-(\tilde{\theta}_n+e_n)-\gamma_n D(\tilde{y}_n)\nabla f(\tilde{y}_n)],\\ \tilde{z}_{n+1} &= \lambda_n h(\tilde{\theta}_n+e_n)+t_n(\tilde{\theta}_n+e_n)+\alpha_n(\tilde{\theta}_n+e_n-\tilde{y}_n+\tilde{x}_n),\end{aligned}\tag{3.12}$$

where h, D, and the parameters are defined as in (3.1). The perturbation terms e_n $(n=0,1,2,\cdots)$ satisfy $\sum_{n=0}^{\infty}\|e_n\|<\infty$.

Algorithm (3.12) is the general form of the bounded perturbation algorithm introduced in [37], which can be used to construct its superiorization version that has been found wide applications in practical problems, such as computed tomography, medical image recovery, convex feasibility problems, inverse problems of radiation therapy and so on. The convergence result of $\{\tilde{z}_n\}$ generated by (3.12) is as follows.

Theorem 3.2. *Assume that the conditions in Theorem 3.1 are satisfied. Then the sequence* $\{\tilde{z}_n\}$ *generated by algorithm (3.12) converges strongly to a point* $z^* \in FixT_\gamma$*, and* $x^* = prox_{\gamma g_2} z^*$ *solves problem (2.1).*

Proof. Let D_n and T_n be defined as (3.2). Put

$$\tilde{e}_n = \lambda_n\Big(h(\tilde{\theta}_n+e_n)-h(\tilde{\theta}_n)\Big)+t_ne_n+\alpha_n\Big(D_n(\tilde{\theta}_n+e_n)-D_n\tilde{\theta}_n\Big).$$

We write (3.12) in the following form:

$$\begin{aligned}\tilde{z}_{n+1} &= \lambda_n h(\tilde{\theta}_n+e_n)+t_n(\tilde{\theta}_n+e_n)+\alpha_n(\tilde{\theta}_n+e_n-\tilde{y}_n+\tilde{x}_n),\\ &= \lambda_n h(\tilde{\theta}_n+e_n)+t_n(\tilde{\theta}_n+e_n)+\alpha_n D_n(\tilde{\theta}_n+e_n)\\ &= \lambda_n h(\tilde{\theta}_n)+t_n\tilde{\theta}_n+\alpha_n D_n\tilde{\theta}_n+\tilde{e}_n.\end{aligned}$$

In order to estimate $\|\tilde{e}_n\|$, we put $\rho_n=\frac{\alpha_n}{1-\lambda_n}$ (see (3.4)), $k_n=\frac{4-\gamma_n L}{2}$, and write nonexpansive operator $k_nT_n+(1-k_n)I=R_n$ (see Lemma 2.9). Then $0<\frac{\rho_n}{k_n}<1$, and

$$\rho_n[T_n(\tilde{\theta}_n+e_n)-T_n\tilde{\theta}_n]+(1-\rho_n)e_n=\frac{\rho_n}{k_n}[R_n(\tilde{\theta}_n+e_n)-R_n\tilde{\theta}_n]+(1-\frac{\rho_n}{k_n})e_n.$$

Consequently,

$$\begin{aligned}\|\tilde{e}_n\| &\le \lambda_n\|h(\tilde{\theta}_n+e_n)-h(\tilde{\theta}_n)\|+\|t_ne_n+\alpha_n\Big(D_n(\tilde{\theta}_n+e_n)-T_n(\tilde{\theta}_n+e_n)\\ &\quad +T_n(\tilde{\theta}_n+e_n)-T_n\tilde{\theta}_n+T_n\tilde{\theta}_n-D_n\tilde{\theta}_n\Big)\|\\ &\le \lambda_n\rho\|e_n\|+(1-\lambda_n)\|(1-\rho_n)e_n+\rho_n(T_n(\tilde{\theta}_n+e_n)-T_n\tilde{\theta}_n)\|\\ &\quad +\alpha_n\|D_n(\tilde{\theta}_n+e_n)-T_n(\tilde{\theta}_n+e_n)\|+\alpha_n\|T_n\tilde{\theta}_n-D_n\tilde{\theta}_n\|\\ &\le \lambda_n\rho\|e_n\|+(1-\lambda_n)\left\|\frac{\rho_n}{k_n}[R_n(\tilde{\theta}_n+e_n)-R_n\tilde{\theta}_n]+(1-\frac{\rho_n}{k_n})e_n\right\|\\ &\quad +\alpha_n\|\delta(\tilde{y}_n)\|+\alpha_n\|\delta(prox_{\gamma_n g_2}\tilde{\theta}_n)\|\\ &\le \lambda_n\rho\|e_n\|+(1-\lambda_n)\frac{\rho_n}{k_n}\|R_n(\tilde{\theta}_n+e_n)-R_n\tilde{\theta}_n\|+(1-\frac{\rho_n}{k_n})\|e_n\|\\ &\quad +\alpha_n\|\delta(\tilde{y}_n)\|+\alpha_n\|\delta(prox_{\gamma_n g_2}\tilde{\theta}_n)\|\\ &\le \lambda_n\rho\|e_n\|+(1-\lambda_n)\left[\frac{\rho_n}{k_n}\|e_n\|+(1-\frac{\rho_n}{k_n})\|e_n\|\right]\\ &\quad +\alpha_n\Big(\|\delta(\tilde{y}_n)\|+\|\delta(prox_{\gamma_n g_2}\tilde{\theta}_n)\|\Big)\\ &= (1-\lambda_n(1-\rho))\|e_n\|+\alpha_n\Big(\|\delta(\tilde{y}_n)\|+\|\delta(prox_{\gamma_n g_2}\tilde{\theta}_n)\|\Big),\end{aligned}$$

which implies $\sum_{n=0}^{\infty}\|\tilde{e}_n\|<\infty$.

Now, let $\{z_n\}$ be generated by algorithm (3.1). Then $\{z_n\}$ converges strongly to some fixed point z^* of T_γ by Theorem 3.1. By using

$$\begin{aligned}
&\|\tilde{z}_{n+1}-z_{n+1}\| \\
&\le \lambda_n\|h(\tilde{\theta}_n)-h(\theta_n)\|+\|t_n(\tilde{\theta}_n-\theta_n)+\alpha_n(D_n\tilde{\theta}_n-D_n\theta_n)\|+\|\tilde{e}_n\| \\
&\le \lambda_n\rho\|\tilde{\theta}_n-\theta_n\|+\|t_n(\tilde{\theta}_n-\theta_n)+\alpha_n(D_n\tilde{\theta}_n-T_n\tilde{\theta}_n+T_n\tilde{\theta}_n-T_n\theta_n+T_n\theta_n-D_n\theta_n)\|+\|\tilde{e}_n\| \\
&\le \lambda_n\rho\|\tilde{\theta}_n-\theta_n\|+\|\tilde{e}_n\|+(1-\lambda_n)\|(1-\rho_n)(\tilde{\theta}_n-\theta_n)+\rho_n(T_n\tilde{\theta}_n-T_n\theta_n)\| \\
&\quad +\alpha_n\|D_n\tilde{\theta}_n-T_n\tilde{\theta}_n\|+\alpha_n\|T_n\theta_n-D_n\theta_n\| \\
&= \lambda_n\rho\|\tilde{\theta}_n-\theta_n\|+\|\tilde{e}_n\|+(1-\lambda_n)\|(1-\frac{\rho_n}{k_n})(\tilde{\theta}_n-\theta_n)+\frac{\rho_n}{k_n}(R_n\tilde{\theta}_n-R_n\theta_n)\| \\
&\quad +\alpha_n\|\delta(\tilde{y}_n)\|+\alpha_n\|\delta(y_n)\| \\
&\le (1-\lambda_n(1-\rho))\|\tilde{\theta}_n-\theta_n\|+\alpha_n\|\delta(\tilde{y}_n)\|+\alpha_n\|\delta(y_n)\|+\|\tilde{e}_n\| \\
&\le (1-\lambda_n(1-\rho))\|\tilde{z}_n-z_n\|+\xi_n(\|\tilde{z}_n-\tilde{z}_{n-1}\|+\|z_n-z_{n-1}\|)+\alpha_n(\|\delta(\tilde{y}_n)\|+\|\delta(y_n)\|)+\|\tilde{e}_n\|
\end{aligned}$$

and taking $\bar{\lambda}_n=\lambda_n(1-\rho)$, $\beta_n=0$ and

$$U_n=\xi_n(\|\tilde{z}_n-\tilde{z}_{n-1}\|+\|z_n-z_{n-1}\|)+\alpha_n(\|\delta(\tilde{y}_n)\|+\|\delta(y_n)\|)+\|\tilde{e}_n\|$$

in Lemma 2.6, we obtain

$$\lim_{n\to\infty}\|\tilde{z}_{n+1}-z_{n+1}\|=0,$$

which guarantees that $\{\tilde{z}_n\}$ converges strongly to z^*. □

4. The Generalized Three-Operator Splitting Algorithm

In this section, we consider problem (1.1):

$$\min_{x\in H}\sum_{i=1}^{K}f_i(x)+\sum_{j=1}^{N}g_j(x),$$

which can be converted into the form:

$$\begin{aligned}
&\min_{x_i\in H}\sum_{i=1}^{N}[f_i(x_i)+g_i(x_i)] \\
&\text{subject to}\quad x_1=x_2=\cdots=x_N,
\end{aligned}\tag{4.1}$$

where $f_i(x)$ is a function in $\Gamma_0(H)$ with L_i-Lipschitz continuous gradient ∇f_i, $g_i(x)$ is a function in $\Gamma_0(H)$ that may not be differentiable for each $i=1,2,\cdots,N$, and N is an arbitrary positive integer.

In order to apply algorithm (3.1) to problem (4.1), we further turn (4.1) into a problem in the product space $H^N=H\times H\times\cdots\times H$. H^N becomes a Hilbert space with inner product $\langle\mathbf{x},\mathbf{y}\rangle_{H^N}=\sum_{i=1}^{N}\langle x_i,y_i\rangle$, $\forall\mathbf{x}=(x_1,x_2,\cdots,x_N),\mathbf{y}=(y_1,y_2,\cdots,y_N)\in H^N$, and induced norm $\|\mathbf{x}\|_{H^N}=(\sum_{i=1}^{N}\|x_i\|^2)^{\frac{1}{2}}$. Set

$$F(\mathbf{x}):=\sum_{i=1}^{N}f_i(x_i),\quad G(\mathbf{x}):=\sum_{i=1}^{N}g_i(x_i),\ \forall\mathbf{x}\in H^N.$$

Denote by $\mathbf{C}=\{\mathbf{x}=(x_1,\cdots,x_N)\in H^N | x_1=x_2=\cdots=x_N\}$. We transform (4.1) into

$$\min_{\mathbf{x}\in H^N}[F(\mathbf{x})+G(\mathbf{x})+\chi_{\mathbf{C}}(\mathbf{x})], \tag{4.2}$$

where

$$\chi_{\mathbf{C}}(\mathbf{x})=\begin{cases} 0, & \mathbf{x}\in\mathbf{C}, \\ +\infty, & \text{otherwise} \end{cases}$$

is the indicator function of set $\mathbf{C}$.

Assume that the solution set $\mathbf{S}$ of problem (4.2) is nonempty. We have the generalized three-operator splitting algorithm from algorithm (3.1):

Let $\mathbf{z_0}=\{z_{1,0},z_{2,0},\cdots,z_{N,0}\}$ and $\mathbf{z_{-1}}=\{z_{1,-1},z_{2,-1},\cdots,z_{N,-1}\}\in H^N$ be two arbitrary initials. Define $\{\mathbf{z_n}\}$ by

$$\begin{aligned} &\theta_{\mathbf{n}}=\mathbf{z_n}+\xi_n(\mathbf{z_n}-\mathbf{z_{n-1}}),\\ &\mathbf{y_n}=prox_{\gamma_n\chi_{\mathbf{C}}}(\theta_{\mathbf{n}}),\\ &\mathbf{x_n}=prox_{\gamma_n\mathbf{G}}[(I-\gamma_n\mathbf{D}\nabla F)\mathbf{y_n}+\mathbf{y_n}-\theta_{\mathbf{n}}],\\ &\mathbf{z_{n+1}}=\lambda_n\mathbf{h}(\theta_{\mathbf{n}})+t_n\theta_{\mathbf{n}}+\alpha_n(\theta_{\mathbf{n}}-\mathbf{y_n}+\mathbf{x_n}). \end{aligned} \tag{4.3}$$

The corresponding component form is

$$\begin{aligned} &\theta_{i,n}=z_{i,n}+\xi_n(z_{i,n}-z_{i,n-1}),\\ &\bar{y}_n=\frac{1}{N}\sum_{i=1}^{N}\theta_{i,n},\\ &x_{i,n}=prox_{\gamma_n g_i}[(I-\gamma_n D_i\nabla f_i)\bar{y}_n+\bar{y}_n-\theta_{i,n}],\\ &z_{i,n+1}=\lambda_n h_i(\theta_{i,n})+t_n\theta_{i,n}+\alpha_n(x_{i,n}-\bar{y}_n+\theta_{i,n}). \end{aligned} \tag{4.4}$$

Here

$$\begin{aligned} &\theta_{\mathbf{n}}=(\theta_{i,n})_{i=1}^N,\ \mathbf{y_n}=(y_{i,n})_{i=1}^N,\ \mathbf{x_n}=(x_{i,n})_{i=1}^N,\\ &prox_{\gamma\chi_{\mathbf{C}}}(\mathbf{y})=P_{\mathbf{C}}(\mathbf{y})=\mathbf{j}(\frac{1}{N}\sum_{i=1}^{N}y_i),\\ &prox_{\gamma G}(\mathbf{y})=\big(prox_{\gamma g_1}y_1,prox_{\gamma g_2}y_2,\cdots,prox_{\gamma g_N}y_N\big),\\ &\nabla F(\mathbf{y})=(\nabla f_1(y_1),\nabla f_2(y_2),\cdots,\nabla f_N(y_N)),\ \forall\mathbf{y}=(y_1,y_2,\cdots,y_N)\in H^N. \end{aligned}$$

$\mathbf{D}(\mathbf{x}):H^N\to H^N$ defined by $\mathbf{D}(\mathbf{x})\mathbf{y}=(D_1(x_1)y_1,D_2(x_2)y_2,\cdots,D_N(x_N)y_N)$ satisfies

$$\sum_{i=1}^{N}\sum_{n=0}^{\infty}\|\nabla f_i(y_{i,n})-D_i(y_{i,n})\nabla f_i(y_{i,n})\|<\infty,$$

$\mathbf{h}(\mathbf{x})=(h_1(x_1),\cdots,h_N(x_N))$ is a ρ-contraction mapping for $\rho=\max\{\rho_1,\cdots,\rho_N\}\in[0,1)$, where ρ_i is the corresponding contraction constant for contraction mapping h_i, $i=1,2,\cdots,N$, $\mathbf{j}:H\to H^N$ is defined by $\mathbf{j}(x)=(x,x,\cdots,x)$, and $\{\lambda_n\}\subset(0,1)$, $\{t_n\}\subset(-1,1)$, and $\{\alpha_n\}\subset(0,2)$ such that $\lambda_n+t_n+\alpha_n=1$ for all $n\geq 0$.

Corollary 4.1. *Suppose the following conditions hold:*

(i) There exists some positive number γ such that $\sum_{n=0}^{\infty}|\gamma_n-\gamma|<\infty$, where $0<\gamma_n\leq\bar{\gamma}:=\sup\gamma_n<\frac{2}{L}$;

(ii) $\lim_{n\to\infty}\lambda_n = 0$ *and* $\sum_{n=0}^{\infty}\lambda_n = \infty$*;*
(iii) $\sum_{n=0}^{\infty}\xi_n\|\mathbf{z_n}-\mathbf{z_{n-1}}\|_{H^N} < \infty$*;*
(iv) $\lim_{n\to\infty}\frac{\lambda_n}{\alpha_n} = 0$, $\sup_n \frac{\alpha_n}{1-\lambda_n} < \frac{4-\bar{\gamma}L}{2}$.
Then the sequence $\{\mathbf{z_n}\}$ *generated by (4.3) converges strongly to a point* $\mathbf{z}^* \in Fix\mathbf{T}_\gamma$*, where* $\mathbf{T}_\gamma = I - P_{\mathbf{C}} + prox_{\gamma G}(2P_{\mathbf{C}} - I - \gamma\nabla F(P_{\mathbf{C}}))$*. In addition,* $x^* = \frac{1}{N}\sum_{i=1}^{N} z_i^*$ *solves problem (1.1).*

Similar to the proof of Theorem 3.1, we can obtain the strong convergence of the sequence $\{\mathbf{z_n}\}$ generated by (4.3). We omit the proof here.

5. Numerical Experiments

In this section, two numerical examples are presented to illustrate the validity of the proposed algorithms. Example 1 is a special case of problem (2.1), which can be solved by algorithms (1.2), (3.1), and (3.12), respectively. We compare the performances of these three algorithms in this example. Example 2 is the case of the general problem (1.1) to show the performance of algorithm (4.3).

Example 1. Let $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, and $\varphi(x) = -\ln(x_1x_2)$ for $x_1x_2 > 0$, $\varphi(x) = +\infty$ otherwise, where $x = (x_1, x_2)^T \in \mathbb{R}^2$.

We consider the following convex minimization problem:

$$\min_{x\in\mathbb{R}^2} \frac{1}{2}\|Ax-b\|^2 + \|x\|_1 + \varphi(x). \tag{5.1}$$

(5.1) is a case of the minimization problem (2.1) with $f(x) = \frac{1}{2}\|Ax-b\|^2$, $g_1(x) = \|x\|_1$, and $g_2(x) = \varphi(x)$. The gradient $\nabla f = A^T(Ax-b)$ satisfies

$$\|\nabla f(x) - \nabla f(y)\| \le \|A^TA\|\cdot\|x-y\|, \quad \forall x,y \in \mathbb{R}^2.$$

To show the validity of algorithms (3.1) and (3.12), and compare the performances of them with (1.2), we calculate the critical point of this problem:

$$x^* = (x_1, x_2)^T = (\frac{2+\sqrt{14}}{10}, \frac{2+\sqrt{14}}{10})^T.$$

Choose $\lambda_n = \frac{1}{10(n+1)}$, $\alpha_n = \frac{9n}{10(n+1)}$, $\gamma_n = \frac{1.99n}{L(n+1)}$, and

$$\xi_n = \begin{cases} \frac{1}{n^2\|x_n - x_{n-1}\|}, & \|x_n - x_{n-1}\| \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

the contraction operator $h(x_n) = 0.1x_n$, and the scaled operator $D(x_n) = diag\{1+\frac{1}{n^2}\}$. Take the perturbation terms $e_n = l_n \times v_n$ in algorithm (3.12), where $v_n = [\frac{n}{n+1}, \frac{n}{n+1}]^T$ and $l_n = 0.9^n$. Give the stopping criterion $\|x_n - x^*\| < \varepsilon$ and denote the number of iterations by "Iter", the minimum value of $f(x)+g_1(x)+g_2(x)$ by "Obj". We have Table 1 and Figure 1.

Table 1: Results for Example 1

Methods	$\|x_n - x^*\| < 10^{-6}$			
	Iter	Time(s)	x_n	Obj
Algorithm 1.2	1244	0.010317	$[0.574165, 0.574165]^T$	2.313011
Algorithm 3.1	378	0.003434	$[0.574166, 0.574166]^T$	2.313011
Algorithm 3.12	305	0.003065	$[0.574166, 0.574166]^T$	2.313011

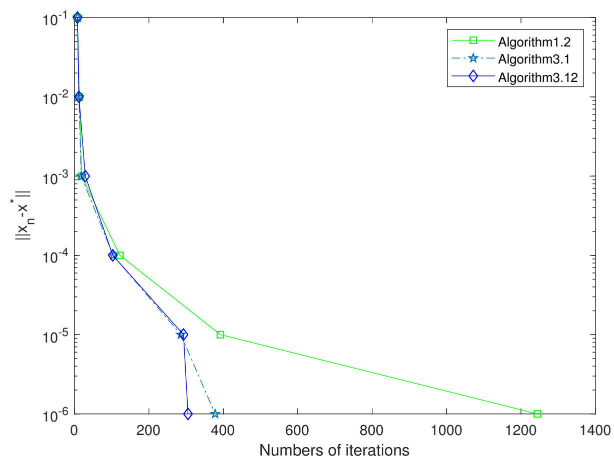

Figure 1: The numerical results of Algorithms (1.2), (3.1), and (3.12).

From the numerical results of Table 1 and Figure 1, we can see that the inertial algorithm (3.1) and the corresponding inertial bounded perturbation algorithm (3.12) improve the performance of algorithm (1.2) in running time and iteration steps under the same algorithm parameters. Especially, the bounded perturbation algorithm (3.12) achieves the same effect in less time and fewer iterative steps, which may indicate that the proposed algorithm in this paper provides a new choice when applying the superiorization method.

Example 2. Let A be a non-zero $M \times J$ real matrix, b a non-zero $M \times 1$ real matrix, $C = \{x \in \mathbb{R}^J | \|x\| \leq 1\}$, and

$$\varphi(x) = \begin{cases} -\ln(x_1 x_2 \cdots x_J), & x_1 x_2 \cdots x_J > 0, \\ +\infty, & \text{otherwise}, \end{cases}$$

$x = (x_1, x_2, \cdots, x_J)^T$. We solve the minimization problem defined on the unit ball:

$$\min_{x \in C} \frac{1}{2}\|Ax - b\|^2 + \|x\|_1 + \varphi(x),$$

which can be converted into a minimization problem of the sum of four convex functions:

$$\min_{x \in \mathbb{R}^J} \frac{1}{2}\|Ax - b\|^2 + \|x\|_1 + \varphi(x) + \chi_C(x).$$

This is a scenario of (1.1) with $K = 1$, $N = 3$, $f(x) = \frac{1}{2}\|Ax-b\|^2$, $g_1(x) = \|x\|_1$, $g_2(x) = \varphi(x)$, and $g_3(x) = \chi_C(x)$. It is easy to see that $prox_{\gamma\chi_C}(x) = P_C(x), \forall x \in \mathbb{R}^J$.

Choose $\lambda_n = \frac{1}{10(n+1)}$,

$$\xi_n = \begin{cases} \frac{1}{n^2\|x_n - x_{n-1}\|}, & \|x_n - x_{n-1}\| \neq 0, \\ 0, & \text{otherwise}, \end{cases}$$

the contraction operator $h(x_n) = 0.1x_n$, and the scaled operator $D(x) = I$ in algorithm (4.4). We investigate the performances of (4.4) under the varying values of algorithm coefficients α_n and γ_n and obtain Figure 2 and Figure 3.

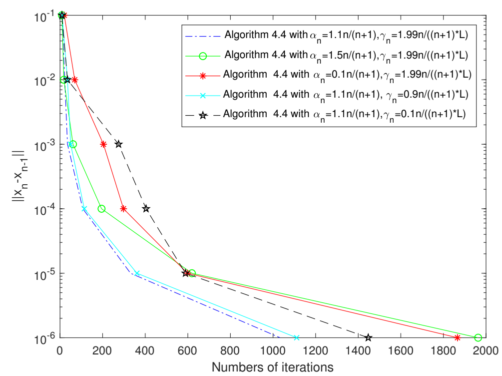

Figure 2: Numerical results of (4.4) for 25×25 matrix A

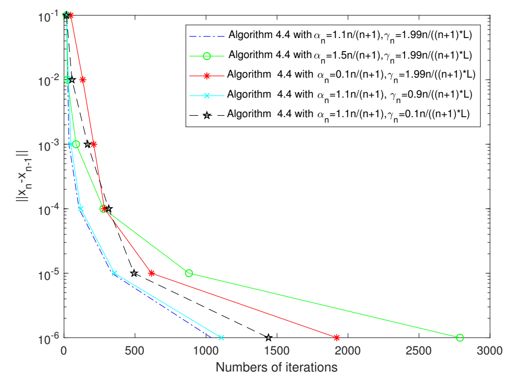

Figure 3: Numerical results of (4.4) for 256×256 matrix A

The numerical results in Figure 2 and Figure 3 show that algorithm (4.4) has similar performance under the stopping criterion $\|x_n - x_{n-1}\| \leq 10^{-6}$ when γ_n approaches to $2/L$, and the values of α_n are not near 1. Algorithm (4.4) has best performance when the relaxation parameters α_n approach to 1.1 whatever the proximal parameters γ_n approach to $2/L$ or $1/L$, which shows that the over-relaxation parameters can indeed reduce the iteration steps of the algorithm.

Acknowledgements

This paper was supported by the Fundamental Research Funds for the Central Universities under Grant No. 3122019185.

REFERENCES

[1] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward–backward splitting, Multiscale Model. Simul. 4 (2005), 1168-1200.

[2] P.L. Combettes, J.C. Pesquet, Proximal splitting methods in signal processing, In: H. Bauschke, R. Burachik, P. Combettes, V. Elser, D. Luke, H. Wolkowicz (eds), Fixed-Point Algorithms for Inverse Problems in Science and Engineering, vol. 49, pp. 185–221, Springer, 2011.

[3] B. O'Donoghue, G. Stathopoulos, S. Boyd, A splitting method for optimal control, IEEE. Trans. Control. Syst. Technol. 21 (2013), 2432-2442.

[4] E. Bonacker, A. Gibali, K.H. Küfer, Nesterov perturbations and projection methods applied to IMRT, J. Nonlinear Var. Anal. 4 (2020), 63-86.

[5] H. Raguet, J.Fadili, G.Peyré, A generalized forward–backward splitting, SIAM J. Imaging Sci. 6 (2013), 1199-1226.

[6] D. Davis, W. Yin, A three-operator splitting scheme and its optimization applications, Set-Valued Var. Anal. 25 (2017), 829-858.

[7] A. Yurtsever, B.C. Vu, V. Cevher, Stochastic three-composite convex minimization, 30th Conference on Neural Information Processing Systems (NIPS 2016), pp. 4329-4337, Barcelona, Spain, 2016.

[8] C.X. Zong, Y.C. Tang, Y.J. Cho, Convergence analysis of an inexact three-operator splitting algorithm, Symmetry, 10 (2018), 563.

[9] C.J. Zhang, J.J. Chen, A parameterized three-operator splitting algorithm and its expansion, J. Nonlinear Var. Anal. 5 (2021), 211-226.

[10] F. Cui, Y. Tang, Y. Yang, An inertial three-operator splitting algorithm with applications to image inpainting, Appl. Set-Valued Anal. Optim. 1 (2019), 113-134.

[11] B.T. Polyak, Some methods of speeding up the convergence of iterative methods, Zh. Vychisl. Mat. Mat. Fiz. 4 (1964), 1-17.

[12] Y. Nesterov, A method for solving the convex programming problem with convergence rate $O(1/k2)$, Dokl. Akad. Nauk SSSR. 269 (1983), 543-547.

[13] A. Moudafifi, M. Oliny, Convergence of a splitting inertial proximal method for monotone operators, J. Comput. Appl. Math. 155 (2003), 447-454.

[14] D.R. Sahu, J.C. Yao, M. Verma, K.K. Shukla, Convergence rate analysis of proximal gradient methods with applications to composite minimization problems, Optimization, 70 (2021), 75-100.

[15] H. Attouch, A. Cabot, Convergence of a relaxed inertial forward–backward algorithm for structured monotone inclusions, Appl. Math. Optim. 80 (2019), 547-598.

[16] S.Y. Cho, A convergence theorem for generalized mixed equilibrium problems and multivalued asymptotically nonexpansive mappings, J. Nonlinear Convex Anal. 21 (2020), 1017-1026.

[17] M. M. Alves, J. Eckstein, M. Geremia, et al., Relative-error inertial-relaxed inexact versions of Douglas-Rachford and ADMM splitting algorithms, Comput. Optim. Appl. 75 (2020), 389-422.

[18] V. Apidopoulos, J.F. Aujol, C. Dossal, Convergence rate of inertial forward–backward algorithm beyond Nesterov's rule, Math. Program. 180 (2020), 137-156.

[19] L.F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differiential equations, with an application to the stresses in a masonry dam, Philos. Trans. R. Soc. Lond. 201 (1911), 307-357.

[20] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program. 55 (1992), 293-318.

[21] F. Iutzeler, J.M. Hendrickx, A generic online acceleration scheme for optimization algorithms via relaxation and inertia, Optim. Method Softw. 34 (2019), 383-405.

[22] P. Cholamjiak, D.V. Hieu, Y.J. Cho, Relaxed forward–backward splitting methods for solving variational inclusions and applications, J. Sci. Comput. 85 (2021), 88.

[23] R. Borgard, S.N. Harding, H. Duba, et al., Accelerating the distributed Kaczmarz algorithm by strong overrelaxation, Linear Algebra Appl. 611 (2021), 334-355.

[24] F. Alvarez, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space, SIAM J. Optim. 14 (2004), 773-782.

[25] P. E. Maingé, Convergence theorems for inertial KM-type algorithms, J. Comput. Appl. Math. 219 (2008), 223-236.

[26] H. Attouch, A. Cabot, Convergence rate of a relaxed inertial proximal algorithm for convex minimization, Optimization, 69 (2020), 1281-1312.

[27] H. Attouch, A. Cabot, Convergence of a relaxed inertial proximal algorithm for maximally monotone operators, Math. Program. 184 (2020), 243-287.

[28] H. Attouch, A. Cabot, Convergence of a relaxed inertial forward-backward algorithm for structured monotone inclusions, Appl. Math. Optim. 80 (2019), 547-598.

[29] J. Abubakar, P. Kumam, A.H. Ibrahim, A. Padcharoen, Relaxed inertial Tseng's type method for solving the inclusion problem with application to image restoration, Mathematics 8 (2020), 818.

[30] H.A. Hammad, H.U. Rehman, M.D.L. Sen, Shrinking projection methods for accelerating relaxed inertial Tseng-type algorithm with applications, Math. Probl. Eng. 2020 (2020), 7487383.

[31] J.J. Moreau, Functions convexes duales et points proximaux dans un espace helbertien, C. R. Acad. Sci. Paris SER. A Math. 255 (1962), 2897-2899.
[32] G. Marino, H.K. Xu, Convergence of generalized proximal point algorithm, Commun. Pure Appl. Anal. 3 (2014), 791-808.
[33] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, Chamm 2011.
[34] K. Geobel, W.A. Kirk, Topics in Metric Fixed Point Theory, Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, (1990).
[35] H. K. Xu, Iterative algorithms for nonlinear operators, J. London Math. Soc. 66 (2002), 240-256.
[36] Y.N. Guo, X.Z. Zhao, Strong convergence of over-relaxed multi-parameter proximal scaled gradient algorithm and superiorization, Optimization 70 (2020), 461-481.
[37] Y. Censor, R. Davidi, G.T. Herman, Perturbation resilience and superiorization of iterative algorithms, Inverse Probl. 26 (2010), 65008.