

RESEARCH

Open Access



Strong convergence and bounded perturbation resilience of a modified proximal gradient algorithm

Yanni Guo* and Wei Cui

*Correspondence:
ynguo@amss.ac.cn
College of Science, Civil Aviation
University of China, Tianjin, China

Abstract

The proximal gradient algorithm is an appealing approach in finding solutions of non-smooth composite optimization problems, which may only has weak convergence in the infinite-dimensional setting. In this paper, we introduce a modified proximal gradient algorithm with outer perturbations in Hilbert space and prove that the algorithm converges strongly to a solution of the composite optimization problem. We also discuss the bounded perturbation resilience of the basic algorithm of this iterative scheme and illustrate it with an application.

Keywords: Strong convergence; Bounded perturbation resilience; Modified proximal gradient algorithm; Viscosity approximation; Convex minimization problem

1 Introduction

Let H be a real Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and an induced norm $\| \cdot \|$. Let $\Gamma_0(H)$ be the class of convex, lower semi-continuous, and proper functions from H to $(-\infty, +\infty]$. Consider the following non-smooth composite optimization problem:

$$\min_{x \in H} (f(x) + g(x)), \quad (1)$$

where $f, g \in \Gamma_0(H)$, f is differentiable and ∇f is L -Lipschitz continuous on H with $L > 0$. g may not be differentiable. If further, $f + g =: \Phi$ is coercive, that is,

$$\lim_{\|x\| \rightarrow +\infty} \Phi(x) = +\infty, \quad (2)$$

then Φ has a minimizer over H , that is, $S := \text{Argmin}(\Phi) \neq \emptyset$, see [1, page 159, Proposition 11.14]. Problem (1) has a typical scenario in linear inverse problems [2], it has applications in compressed sensing, machine learning, data recovering and so on (see [3–6] and the references therein).

Proximal gradient methods are among the methods used for solving problem (1), which allow to decouple the contribution of the functions f and g in a gradient descent step determined by f and in a proximal step induced by g [7, 8]. For the classical proximal gradient method, the initial value $x_0 \in H$ is given, and the iterative algorithm for generating

sequence $\{x_n\}$ is defined as follows:

$$x_{n+1} = \text{prox}_{\lambda g}(I - \lambda \nabla f)(x_n), \quad \forall n \geq 0, \tag{3}$$

where $\lambda > 0$ is the step size, $\text{prox}_{\lambda g}$ is a proximal operator (see Sect. 2). If $S \neq \emptyset$ and $0 < \lambda < \frac{2}{L}$, then any sequence generated by algorithm (3) converges weakly to an element of S [1, Corollary 27.9]. Xu [9] put forward the following slightly more general proximal gradient algorithm:

$$x_{n+1} = \text{prox}_{\lambda_n g}(I - \lambda_n \nabla f)(x_n) \tag{4}$$

for problem (1), where the weak convergence of the generated sequence $\{x_n\}$ was obtained. Besides, it was noted that no strong convergence is guaranteed if $\dim H = \infty$. In 2017, Guo, Cui and Guo [10] proposed the following proximal gradient algorithm with perturbations:

$$x_{n+1} = \text{prox}_{\lambda_n g}(I - \lambda_n D \nabla f + e)(x_n). \tag{5}$$

The generated sequence $\{x_n\}$ again converges weakly to a solution of (1).

On the other hand, it is well known that the viscosity approximation method proposed by Moudafi [11] generates a sequence $\{x_n\}$:

$$x_{n+1} = t_n h(x_n) + (1 - t_n) T x_n, \tag{6}$$

which converges strongly to a fixed point x^* of T for some contractive operator h . In 2004, Xu [12, Theorem 3.1] further proved that the above x^* is also the unique solution of the variational inequality:

$$\langle (I - h)x^*, x - x^* \rangle \geq 0, \quad \forall x \in \text{Fix}(T), \tag{7}$$

provided that $\{t_n\}$ satisfies certain conditions.

This paper is based on viscosity algorithm (6) and proximal gradient algorithm (4) to generate a sequence with perturbations, which converges strongly to a solution of problem (1). We also apply this algorithm to solve the linear inverse problem.

An objective of this paper considering the perturbation is the superiorization methodology introduced by [13]. The superiorization method may not find an optimal solution to the given objective function. It might try to find a point with a lower cost function value than other points by a rather simple algorithm, which is known as the basic algorithm (see [14–18] for more details). It is a heuristic method with less time consuming that makes it applicable to some important practical problems such as medical image recovery [19, 20], computed tomography [21], intensity-modulated radiation therapy [22] and the like. However, the superiorization method needs to investigate the basic iterative algorithm’s bounded perturbation resilience. Hence, there raises a new problem whether the basic algorithm is bounded perturbation resilient. Very recently, several articles focused on this topic [23–27]. So another task of this paper is to discuss the bounded perturbation resilience of the modified proximal gradient algorithm.

1.1 Results and discussion

In view of the facts that the sequence generated by (5) converges weakly to a solution of (1), the viscosity method can convert a weakly convergent sequence to a strongly convergent one, and that the applied widely superiorization method introduced by [13] is based on the bounded perturbation resilience of basic algorithms, we discuss the strong convergence problem of a modified proximal gradient algorithm with perturbations as well as the bounded perturbation resilience of the responding basic algorithm.

The structure of this paper is as follows. In Sect. 2, we introduce some definitions and lemmas that will be used to prove the main results in the subsequent sections. In Sect. 3, we present the modified proximal gradient algorithm with perturbations and prove that the generated sequence $\{x_n\}$ converges strongly to a solution of problem (1). We conclude this section with several corollaries. In Sect. 4, we introduce the definition of bounded perturbation resilience and certify the corresponding strong convergence result. In Sect. 5, we apply our algorithm to the linear inverse problem, and illustrate it with a specific numerical example. Finally, we give a conclusion in Sect. 6.

2 Preliminaries

Let $\{x_n\}$ be a sequence in Hilbert space H and $x \in H$. Let $T : H \rightarrow H$ be an operator (linear or nonlinear). We list some notations.

$x_n \rightarrow x$ means $\{x_n\}$ converges strongly to x .

$x_n \rightharpoonup x$ means $\{x_n\}$ converges weakly to x .

If there exists a subsequence $\{x_{n_j}\}$, which converges weakly to a point z , we will call z a weak cluster point of $\{x_n\}$. The set of all cluster points of $\{x_n\}$ is denoted by $\omega_w(x_n)$.

$\text{Fix}(T) := \{x \in H : Tx = x\}$.

The following definitions are needed in proving our main results.

Definition 2.1 Let $T, A : H \rightarrow H$ be operators.

(i) T is nonexpansive if

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

(ii) T is L -Lipschitz continuous with $L \geq 0$, if

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

We call T a contractive mapping if $0 \leq L < 1$.

(iii) T is α -averaged if

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

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

(iv) A is ν -inverse strongly monotone (ν -ism) with $\nu > 0$, if

$$\langle Ax - Ay, x - y \rangle \geq \nu \|Ax - Ay\|^2, \quad \forall x, y \in H.$$

Given $g \in \Gamma_0(H)$, [1, Proposition 12.15] ensures that $\frac{\|y-x\|^2}{2} + g(y)$ has exact one minimizer over H for each $x \in H$. So we have

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

$$\text{prox}_g(x) := \arg \min_{y \in H} \left\{ \frac{\|y - x\|^2}{2} + g(y) \right\}, \quad x \in H. \tag{8}$$

The proximal operator of g of order $\alpha > 0$ is defined as the proximal operator of αg . Moreover, it satisfies (see [3, Remark 12.24])

$$\text{prox}_{\alpha g}(x) := \arg \min_{y \in H} \left\{ \frac{\|y - x\|^2}{2\alpha} + g(y) \right\}, \quad x \in H. \tag{9}$$

The following lemmas (Lemma 2.3 and Lemma 2.4) describe the properties of the proximal operators.

Lemma 2.3 ([9, Lemma 3.1]) *Let $g \in \Gamma_0(H)$, and $\alpha > 0, \mu > 0$. Then*

$$\text{prox}_{\alpha g}(x) = \text{prox}_{\mu g} \left(\frac{\mu}{\alpha} x + \left(1 - \frac{\mu}{\alpha} \right) \text{prox}_{\alpha g} x \right).$$

Lemma 2.4 ([8, Lemma 2.4], [1, Remark 4.24]) *Let $g \in \Gamma_0(H)$, and $\alpha > 0$. Then the proximity operator $\text{prox}_{\alpha g}$ is $\frac{1}{2}$ -averaged. In particular, it is nonexpansive, that is,*

$$\| \text{prox}_{\alpha g}(x) - \text{prox}_{\alpha g}(y) \| \leq \| x - y \|, \quad \forall x, y \in H. \tag{10}$$

Lemma 2.5 ([9, Proposition 3.2]) *Let $f, g \in \Gamma_0(H)$, $z \in H$ and $\alpha > 0$. Assume that f is differentiable on H . Then z is a solution to (1) if and only if z solves the fixed point equation*

$$z = \text{prox}_{\alpha g}(I - \alpha \nabla f)z. \tag{11}$$

The following two lemmas play an important role in proving the strong convergence result.

Lemma 2.6 ([1, Theorem 4.17]) *Let $T : H \rightarrow H$ be a nonexpansive mapping with $\text{Fix}(T) \neq \emptyset$. 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 2.7 ([28, Lemma 2.5]) *Assume that $\{a_n\}$ is a sequence of nonnegative real numbers satisfying*

$$a_{n+1} \leq (1 - \gamma_n)a_n + \gamma_n \delta_n + \beta_n, \quad n \geq 0,$$

where $\{\gamma_n\}, \{\delta_n\}$ and $\{\beta_n\}$ satisfy the conditions:

- (i) $\{\gamma_n\} \subset [0, 1]$, $\sum_{n=0}^{\infty} \gamma_n = \infty$, or equivalently, $\prod_{n=0}^{\infty} (1 - \gamma_n) = 0$;
- (ii) $\limsup_{n \rightarrow \infty} \delta_n \leq 0$;
- (iii) $\beta_n \geq 0$ ($n \geq 0$), $\sum_{n=0}^{\infty} \beta_n < \infty$.

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

3 Convergence analysis

In this section, let H be a Hilbert space, $h : H \rightarrow H$ a ρ -contractive operator with $\rho \in (0, 1)$. $f, g \in \Gamma_0(H)$. f is differentiable, and ∇f is Lipschitz continuous with Lipschitz constant $L > 0$. Given $x_0 \in H$, we propose the following modified proximal gradient algorithm for solving (1):

$$x_{n+1} := t_n h(x_n) + (1 - t_n) \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n) + e(x_n), \quad n \geq 0, \tag{12}$$

where $\{t_n\}$ is a sequence in $[0, 1]$. $0 < a = \inf_n \alpha_n \leq \alpha_n < \frac{2}{L}$. $e : H \rightarrow H$ represents a perturbation operator and satisfies

$$\sum_{n=0}^{\infty} \|e(x_n)\| < +\infty. \tag{13}$$

We also introduce the following iterative scheme as a special case to (12):

$$x_{n+1} := t_n h(x_n) + (1 - t_n) \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f + e)(x_n), \quad n \geq 0. \tag{14}$$

We state the main strong convergence theorem.

Theorem 3.1 *Let S be the solution set of (1), and assume that $S \neq \emptyset$. Given $x_0 \in H$. Let $\{x_n\}$ be generated by (12). If (13) and the following conditions hold:*

- (i) $0 < a = \inf_n \alpha_n \leq \alpha_n < \frac{2}{L}$, $\sum_{n=0}^{\infty} |\alpha_{n+1} - \alpha_n| < \infty$;
- (ii) $\{t_n\} \subset (0, 1)$, $\lim_{n \rightarrow \infty} t_n = 0$;
- (iii) $\sum_{n=0}^{\infty} t_n = \infty$, $\sum_{n=0}^{\infty} |t_{n+1} - t_n| < \infty$.

Then $\{x_n\}$ converges strongly to a point $x^ \in S$, where x^* is the unique solution of the following variational inequality problem:*

$$\langle (I - h)x^*, x - x^* \rangle \geq 0, \quad \forall x \in S. \tag{15}$$

Proof We point out that $\text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)$ is nonexpansive for each n . Let us follow the proof of [29]. At first, that ∇f is L -Lipschitzian means that ∇f is $\frac{1}{L}$ -ism ([1, Theorem 18.15]). Consequently, $I - \alpha_n \nabla f$ is $\frac{\alpha_n L}{2}$ -averaged as $0 < \alpha_n < \frac{2}{L}$ ([1, Proposition 4.33]). Besides, $\text{prox}_{\alpha_n g}$ is $\frac{1}{2}$ -averaged by Lemma 2.4, the composite $\text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)$ is $\frac{\alpha_n L + 2}{4}$ -averaged ([29, Proposition 3.2]). Then it is nonexpansive ([1, Remark 4.24]).

Set $T_n = \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)$. For any $\bar{x} \in S$, we have

$$\|T_n x_n - \bar{x}\| \leq \|x_n - \bar{x}\|$$

by applying Lemma 2.5 and that T_n is nonexpansive. So,

$$\begin{aligned} & \|x_{n+1} - \bar{x}\| \\ &= \|t_n h(x_n) + (1 - t_n) T_n x_n + e(x_n) - \bar{x}\| \\ &\leq t_n \|h(x_n) - \bar{x}\| + (1 - t_n) \|T_n x_n - \bar{x}\| + \|e(x_n)\| \\ &\leq t_n \|h(x_n) - h(\bar{x})\| + t_n \|h(\bar{x}) - \bar{x}\| + (1 - t_n) \|T_n x_n - \bar{x}\| + \|e(x_n)\| \end{aligned}$$

$$\begin{aligned}
 &\leq t_n \rho \|x_n - \bar{x}\| + t_n \|h(\bar{x}) - \bar{x}\| + (1 - t_n) \|x_n - \bar{x}\| + \|e(x_n)\| \\
 &= (1 - t_n(1 - \rho)) \|x_n - \bar{x}\| + t_n(1 - \rho) \cdot \frac{\|h(\bar{x}) - \bar{x}\|}{1 - \rho} + \|e(x_n)\| \\
 &\leq \max \left\{ \|x_n - \bar{x}\|, \frac{\|h(\bar{x}) - \bar{x}\|}{1 - \rho} \right\} + \|e(x_n)\|. \tag{16}
 \end{aligned}$$

An induction argument shows that

$$\|x_{n+1} - \bar{x}\| \leq \max \left\{ \|x_0 - \bar{x}\|, \frac{\|h(\bar{x}) - \bar{x}\|}{1 - \rho} \right\} + \sum_{k=0}^{\infty} \|e(x_k)\|.$$

Hence $\{x_n\}$ is bounded as $\sum_{n=0}^{\infty} \|e(x_n)\| < \infty$. Consequently, we get the boundedness of $\{h(x_n)\}$ and $\{T_n x_n\}$.

We next prove that $\|x_{n+1} - x_n\| \rightarrow 0$ as $n \rightarrow \infty$. In fact,

$$\begin{aligned}
 &\|x_{n+1} - x_n\| \\
 &= \left\| t_n h(x_n) + (1 - t_n) T_n x_n + e(x_n) - t_{n-1} h(x_{n-1}) - (1 - t_{n-1}) T_{n-1} x_{n-1} - e(x_{n-1}) \right\| \\
 &\leq \left\| t_n [h(x_n) - h(x_{n-1})] + [t_n h(x_{n-1}) - t_{n-1} h(x_{n-1})] \right\| \\
 &\quad + \left\| (1 - t_n)(T_n x_n - T_n x_{n-1}) + (1 - t_n) T_n x_{n-1} - (1 - t_{n-1}) T_{n-1} x_{n-1} \right\| \\
 &\quad + \|e(x_n)\| + \|e(x_{n-1})\| \\
 &\leq t_n \rho \|x_n - x_{n-1}\| + |t_n - t_{n-1}| \|h(x_{n-1})\| \\
 &\quad + (1 - t_n) \|T_n x_n - T_n x_{n-1}\| + (1 - t_n) \|T_n x_{n-1} - T_{n-1} x_{n-1}\| \\
 &\quad + |t_n - t_{n-1}| \|T_{n-1} x_{n-1}\| + \|e(x_n)\| + \|e(x_{n-1})\| \\
 &\leq t_n \rho \|x_n - x_{n-1}\| + |t_n - t_{n-1}| \|h(x_{n-1})\| + (1 - t_n) \|x_n - x_{n-1}\| \\
 &\quad + (1 - t_n) \|T_n x_{n-1} - T_{n-1} x_{n-1}\| + |t_n - t_{n-1}| \|T_{n-1} x_{n-1}\| \\
 &\quad + \|e(x_n)\| + \|e(x_{n-1})\| \\
 &\leq (1 - t_n(1 - \rho)) \|x_n - x_{n-1}\| + |t_n - t_{n-1}| (\|h(x_{n-1})\| + \|T_{n-1} x_{n-1}\|) \\
 &\quad + (1 - t_n) \|T_n x_{n-1} - T_{n-1} x_{n-1}\| + \|e(x_n)\| + \|e(x_{n-1})\|. \tag{17}
 \end{aligned}$$

By applying Lemma 2.3 and Lemma 2.4, we compute

$$\begin{aligned}
 &\|T_n x_{n-1} - T_{n-1} x_{n-1}\| \\
 &= \left\| (\text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)) x_{n-1} - (\text{prox}_{\alpha_{n-1} g}(I - \alpha_{n-1} \nabla f)) x_{n-1} \right\| \\
 &= \left\| (\text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)) x_{n-1} - \text{prox}_{\alpha_n g} \left(\frac{\alpha_n}{\alpha_{n-1}} (I - \alpha_{n-1} \nabla f) x_{n-1} \right) \right. \\
 &\quad \left. + \left(1 - \frac{\alpha_n}{\alpha_{n-1}} \right) [\text{prox}_{\alpha_{n-1} g}(I - \alpha_{n-1} \nabla f)] x_{n-1} \right\| \\
 &\leq \left\| (I - \alpha_n \nabla f) x_{n-1} - \frac{\alpha_n}{\alpha_{n-1}} (I - \alpha_{n-1} \nabla f) x_{n-1} - \left(1 - \frac{\alpha_n}{\alpha_{n-1}} \right) T_{n-1} x_{n-1} \right\|
 \end{aligned}$$

$$\begin{aligned}
 &= \left| 1 - \frac{\alpha_n}{\alpha_{n-1}} \right| \|x_{n-1} - T_{n-1}x_{n-1}\| \\
 &\leq \frac{|\alpha_n - \alpha_{n-1}|}{a} \|x_{n-1} - T_{n-1}x_{n-1}\|,
 \end{aligned} \tag{18}$$

where $a = \inf_n \alpha_n > 0$.

Substituting (18) into (17), we obtain

$$\begin{aligned}
 &\|x_{n+1} - x_n\| \\
 &\leq (1 - t_n(1 - \rho)) \|x_n - x_{n-1}\| + |t_n - t_{n-1}| (\|h(x_{n-1})\| + \|T_{n-1}x_{n-1}\|) \\
 &\quad + (1 - t_n) \frac{|\alpha_n - \alpha_{n-1}|}{a} \|x_{n-1} - T_{n-1}x_{n-1}\| + \|e(x_n)\| + \|e(x_{n-1})\| \\
 &\leq (1 - t_n(1 - \rho)) \|x_n - x_{n-1}\| + \frac{|\alpha_n - \alpha_{n-1}|}{a} (\|x_{n-1}\| + \|T_{n-1}x_{n-1}\|) \\
 &\quad + t_n(1 - \rho) \cdot \frac{(-1) \cdot |\alpha_n - \alpha_{n-1}|}{a(1 - \rho)} (\|x_{n-1}\| + \|T_{n-1}x_{n-1}\|) \\
 &\quad + |t_n - t_{n-1}| (\|h(x_{n-1})\| + \|T_{n-1}x_{n-1}\|) + \|e(x_n)\| + \|e(x_{n-1})\| \\
 &\leq (1 - t_n(1 - \rho)) \|x_n - x_{n-1}\| \\
 &\quad + t_n(1 - \rho) \cdot \frac{(-1) \cdot |\alpha_n - \alpha_{n-1}|}{a(1 - \rho)} (\|x_{n-1}\| + \|T_{n-1}x_{n-1}\|) \\
 &\quad + M_1 (|t_n - t_{n-1}| + |\alpha_n - \alpha_{n-1}|) + \|e(x_n)\| + \|e(x_{n-1})\|,
 \end{aligned} \tag{19}$$

where $M_1 := \sup_{n \in \mathbb{N}} \{ \|h(x_{n-1})\| + \|T_{n-1}x_{n-1}\|, \frac{\|x_{n-1}\| + \|T_{n-1}x_{n-1}\|}{a} \}$ is well defined since $\{x_n\}$, $\{h(x_n)\}$ and $\{T_n x_n\}$ are bounded.

By taking $\gamma_n = t_n(1 - \rho)$, $\delta_n = \frac{(-1) \cdot |\alpha_n - \alpha_{n-1}|}{a(1 - \rho)} (\|x_{n-1}\| + \|T_{n-1}x_{n-1}\|)$ and $\beta_n = M_1 (|t_n - t_{n-1}| + |\alpha_n - \alpha_{n-1}|) + \|e(x_n)\| + \|e(x_{n-1})\|$ in (19), we get

$$\|x_{n+1} - x_n\| \rightarrow 0, \quad \text{as } n \rightarrow \infty \tag{20}$$

according to Lemma 2.7 and (i)–(iii) in Theorem 3.1.

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_j}\}$ such that $x_{n_j} \rightharpoonup z$ as $j \rightarrow \infty$. In the sequel, we shall verify that $z \in S$. To this end, assume that $\alpha_{n_j} \rightarrow \alpha$ ($j \rightarrow \infty$), and set $T := \text{prox}_{\alpha g}(I - \alpha \nabla f)$. We compute

$$\begin{aligned}
 &\|x_{n_j} - \text{prox}_{\alpha g}(I - \alpha \nabla f)x_{n_j}\| \\
 &\leq \|x_{n_j} - x_{n_j+1}\| + \|x_{n_j+1} - \text{prox}_{\alpha g}(I - \alpha \nabla f)x_{n_j}\| \\
 &= \|x_{n_j} - x_{n_j+1}\| + \|t_{n_j}h(x_{n_j}) + (1 - t_{n_j})T_{n_j}x_{n_j} + e(x_{n_j}) - Tx_{n_j}\| \\
 &\leq \|x_{n_j} - x_{n_j+1}\| + t_{n_j} \|h(x_{n_j}) - Tx_{n_j}\| + (1 - t_{n_j}) \|T_{n_j}x_{n_j} - Tx_{n_j}\| \\
 &\quad + \|e(x_{n_j})\|.
 \end{aligned} \tag{21}$$

By using Lemma 2.3, we get

$$\begin{aligned}
 &\|T_{n_j}x_{n_j} - Tx_{n_j}\| \\
 &= \|\text{prox}_{\alpha_{n_j} g}(I - \alpha_{n_j} \nabla f)x_{n_j} - \text{prox}_{\alpha g}(I - \alpha \nabla f)x_{n_j}\|
 \end{aligned}$$

$$\begin{aligned}
 &= \left\| \operatorname{prox}_{\alpha g} \left[\frac{\alpha}{\alpha_{n_j}} (I - \alpha_{n_j} \nabla f) x_{n_j} + \left(1 - \frac{\alpha}{\alpha_{n_j}} \right) \operatorname{prox}_{\alpha_{n_j} g} (I - \alpha_{n_j} \nabla f) x_{n_j} \right] \right. \\
 &\quad \left. - \operatorname{prox}_{\alpha g} (I - \alpha \nabla f) x_{n_j} \right\| \\
 &\leq \left\| \frac{\alpha}{\alpha_{n_j}} (I - \alpha_{n_j} \nabla f) x_{n_j} + \left(1 - \frac{\alpha}{\alpha_{n_j}} \right) \operatorname{prox}_{\alpha_{n_j} g} (I - \alpha_{n_j} \nabla f) x_{n_j} \right. \\
 &\quad \left. - (I - \alpha \nabla f) x_{n_j} \right\| \\
 &= \left\| \left(\frac{\alpha}{\alpha_{n_j}} - 1 \right) x_{n_j} + \left(1 - \frac{\alpha}{\alpha_{n_j}} \right) \operatorname{prox}_{\alpha_{n_j} g} (I - \alpha_{n_j} \nabla f) x_{n_j} \right\| \\
 &= \left| 1 - \frac{\alpha}{\alpha_{n_j}} \right| \|x_{n_j} - T_{n_j} x_{n_j}\| \\
 &\leq \left| 1 - \frac{\alpha}{\alpha_{n_j}} \right| \cdot (\|x_{n_j}\| + \|T_{n_j} x_{n_j}\|). \tag{22}
 \end{aligned}$$

Thus, we have

$$\|T_{n_j} x_{n_j} - Tx_{n_j}\| \rightarrow 0 \quad \text{as } j \rightarrow \infty \tag{23}$$

in view of that $\{\|x_{n_j}\|\}, \{\|T_{n_j} x_{n_j}\|\}$ are bounded for j , and $\lim_{j \rightarrow \infty} \alpha_{n_j} = \alpha$.

We combine (20), (21) and (23) to have

$$\|x_{n_j} - \operatorname{prox}_{\alpha g} (I - \alpha \nabla f) x_{n_j}\| \rightarrow 0 \quad \text{as } j \rightarrow \infty, \tag{24}$$

which implies that $z \in S$ owing to Lemma 2.6, and hence $\omega_w(x_n) \subset S$.

Finally, we prove that

$$\lim_{n \rightarrow \infty} \|x_n - x^*\| = 0.$$

We have, by utilizing Lemma 2.4,

$$\begin{aligned}
 &\|x_{n+1} - x^*\|^2 \\
 &= \|t_n h(x_n) + (1 - t_n) T_n x_n + e(x_n) - x^*\|^2 \\
 &= \|t_n (h(x_n) - x^*) + (1 - t_n) (T_n x_n - x^*) + e(x_n)\|^2 \\
 &\leq \|t_n (h(x_n) - x^*) + (1 - t_n) (T_n x_n - x^*)\|^2 + 2 \langle x_{n+1} - x^*, e(x_n) \rangle \\
 &\leq \|t_n (h(x_n) - x^*) + (1 - t_n) (T_n x_n - x^*)\|^2 + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 &= \|t_n (h(x_n) - h(x^*)) + (1 - t_n) (T_n x_n - x^*) + t_n (h(x^*) - x^*)\|^2 \\
 &\quad + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 &\leq \|t_n (h(x_n) - h(x^*)) + (1 - t_n) (T_n x_n - x^*)\|^2 \\
 &\quad + 2 \langle t_n (h(x^*) - x^*), x_{n+1} - e(x_n) - x^* \rangle + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 &\leq t_n \|h(x_n) - h(x^*)\|^2 + (1 - t_n) \|T_n x_n - x^*\|^2
 \end{aligned}$$

$$\begin{aligned}
 & + 2t_n \langle h(x^*) - x^*, x_{n+1} - e(x_n) - x^* \rangle + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 \leq & t_n \rho^2 \|x_n - x^*\|^2 + (1 - t_n) \|x_n - x^*\|^2 \\
 & + 2t_n \langle h(x^*) - x^*, x_{n+1} - e(x_n) - x^* \rangle + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 = & (1 - t_n(1 - \rho^2)) \|x_n - x^*\|^2 + 2t_n \langle h(x^*) - x^*, x_{n+1} - x^* \rangle \\
 & - 2t_n \langle h(x^*) - x^*, e(x_n) \rangle + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 \leq & (1 - t_n(1 - \rho^2)) \|x_n - x^*\|^2 + 2t_n \langle h(x^*) - x^*, x_{n+1} - x^* \rangle \\
 & + 2t_n \|h(x^*) - x^*\| \|e(x_n)\| + 2 \|x_{n+1} - x^*\| \|e(x_n)\| \\
 = & (1 - t_n(1 - \rho^2)) \|x_n - x^*\|^2 + 2t_n \langle h(x^*) - x^*, x_{n+1} - x^* \rangle \\
 & + 2(t_n \|h(x^*) - x^*\| + \|x_{n+1} - x^*\|) \|e(x_n)\| \\
 \leq & (1 - t_n(1 - \rho^2)) \|x_n - x^*\|^2 + 2t_n \langle h(x^*) - x^*, x_{n+1} - x^* \rangle \\
 & + M_2 \|e(x_n)\|, \tag{25}
 \end{aligned}$$

where $M_2 = \sup_{n \in \mathbb{N}} \{2(t_n \|h(x^*) - x^*\| + \|x_{n+1} - x^*\|)\} < \infty$ in view of $t_n \in [0, 1]$ and $\{x_n\}$ being bounded.

In order to apply Lemma 2.7 to (25), we need to prove

$$\limsup_{n \rightarrow \infty} \langle h(x^*) - x^*, x_n - x^* \rangle \leq 0. \tag{26}$$

Select a suitable subsequence $\{x_{n_i}\}$ from $\{x_n\}$ such that

$$\limsup_{n \rightarrow \infty} \langle h(x^*) - x^*, x_n - x^* \rangle = \lim_{i \rightarrow \infty} \langle h(x^*) - x^*, x_{n_i} - x^* \rangle. \tag{27}$$

Since $\{x_{n_i}\}$ is bounded, it has a weakly convergent subsequence. Without loss of generality, we denote the weakly convergent subsequence by $\{x_{n_i}\}$ and assume that $x_{n_i} \rightharpoonup \hat{z}$. Then $\hat{z} \in S$, and

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \langle h(x^*) - x^*, x_n - x^* \rangle \\
 = & \lim_{i \rightarrow \infty} \langle h(x^*) - x^*, x_{n_i} - x^* \rangle \\
 = & \langle h(x^*) - x^*, \hat{z} - x^* \rangle \\
 \leq & 0. \tag{28}
 \end{aligned}$$

Take $\gamma_n = t_n(1 - \rho^2)$, $\delta_n = \frac{2}{1 - \rho^2} \langle h(x^*) - x^*, x_{n+1} - x^* \rangle$, $\beta_n = M_2 \|e(x_n)\|$. Then all conditions in Lemma 2.7 are satisfied. Thus

$$\|x_{n+1} - x^*\|^2 \leq (1 - \gamma_n) \|x_n - x^*\|^2 + \gamma_n \delta_n + \beta_n, \tag{29}$$

which implies that $x_n \rightarrow x^*$ as $n \rightarrow \infty$. □

With $\{x_n\}$ generated by (14), we obtain the following.

Theorem 3.2 *With the conditions in Theorem 3.1 hold, given $x_0 \in H$, the sequence $\{x_n\}$ generated by (14) converges strongly to a point $x^* \in S$.*

Proof We complete the proof by translating (14) into the form of (12). Indeed, we can rewrite x_{n+1} as

$$\begin{aligned} x_{n+1} &= t_n h(x_n) + (1 - t_n) \operatorname{prox}_{\alpha_n g}(I - \alpha_n \nabla f + e)(x_n) \\ &= t_n h(x_n) + (1 - t_n) \operatorname{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n) + \hat{e}(x_n), \end{aligned} \tag{30}$$

where

$$\hat{e}(x_n) = (1 - t_n) [\operatorname{prox}_{\alpha_n g}(I - \alpha_n \nabla f + e)(x_n) - \operatorname{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n)].$$

Obviously, $\|\hat{e}(x_n)\| \leq \|e(x_n)\|$ owing to Lemma 2.4 and $t_n \in [0, 1]$. Thus we have $\sum_{n=0}^\infty \|\hat{e}(x_n)\| < \infty$. Since (12) was shown to converge, this immediately implies that (14) converges strongly to a solution of (1). \square

If $e(x_n) \equiv 0, n \geq 0$, the exact form of the two modified proximal gradient algorithms follows.

Corollary 3.3 *With the conditions in Theorem 3.1 holding, and given $x_0 \in H$, arbitrarily, any sequence $\{x_n\}$ defined by*

$$x_{n+1} := t_n h(x_n) + (1 - t_n) \operatorname{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n), \quad n \geq 0, \tag{31}$$

converges strongly to a point $x^ \in S$.*

We also get the following result of [12, Theorem 3.2] with $T = \operatorname{prox}_{\alpha g}(I - \alpha \nabla f)$ and $\alpha_n \equiv \alpha$.

Corollary 3.4 *With the conditions in Theorem 3.1 hold, given $x_0 \in H$, any sequence $\{x_n\}$ defined by*

$$x_{n+1} := t_n h(x_n) + (1 - t_n) \operatorname{prox}_{\alpha g}(I - \alpha \nabla f)(x_n), \quad n \geq 0 \tag{32}$$

converges strongly to a point $x^ \in S$.*

In addition, if h is some constant function, we have

Corollary 3.5 *Under the conditions given in Theorem 3.1, for any $x_0 \in H$, the sequence $\{x_n\}$ defined by*

$$x_{n+1} := t_n u + (1 - t_n) \operatorname{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n), \quad n \geq 0 \tag{33}$$

converges strongly to a point $x^ \in S$, where u is a point in H .*

4 Bounded perturbation resilience

The superiorization method can solve a broad class of nonlinear constrained optimal problems, which works by using the bounded perturbation resilience (BPR) of an original algorithm in order to steer the iterates of the algorithm towards to lower values of the objective function. In this paper, we investigate the BPR of the modified proximal gradient algorithm. The superiorization version of this scheme will be presented in the sequel paper.

Given a problem Φ . Assume that we have a basic algorithm operator $A : H \rightarrow H$, where H is a real Hilbert space. Then we have the following definition, which was originally given with a finite-dimensional Euclidean space [13].

Definition 4.1 ([14], Bounded perturbation resilience) An algorithmic operator A is said to be bounded perturbation resilient if the following condition holds: if the sequence $\{x_n\}$, generated by $x_{n+1} = Ax_n$ with $x_0 \in H$, converges to a solution of Φ , then any sequence $\{y_n\}$ generated by $y_{n+1} = A(y_n + \beta_n v_n)$ with any $y_0 \in H$, also converges to a solution of Φ , where the vector sequence $\{v_n\}_{n=0}^\infty$ is bounded, and the scalars $\{\beta_n\}_{n=0}^\infty$ are such that $\beta_n \geq 0$ for all $n \geq 0$, and $\sum_{n=0}^\infty \beta_n < \infty$.

If we treat the modified proximal gradient algorithm (31) as the basic algorithm A , the bounded perturbation of it is a sequence $\{x_n\}$ generated by

$$x_{n+1} = t_n h(x_n + \beta_n v_n) + (1 - t_n) \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n + \beta_n v_n). \tag{34}$$

We have the following result.

Theorem 4.2 *Let H be a real Hilbert space. Let $h : H \rightarrow H$ be a ρ -contractive operator, $\rho \in (0, 1)$ and $f, g \in \Gamma_0(H)$. Assume the solution set S of (1) is nonempty. Assume, in addition, that f is differentiable, ∇f is L -Lipschitz continuous on H . $\{\beta_n\}, \{v_n\}$ satisfy the conditions in Definition 4.1, $\{t_n\}$ and $\{\alpha_n\}$ satisfy the conditions in Theorem 3.1, respectively. Then any sequence $\{x_n\}$ generated by (34) converges strongly to a point x^* in S .*

Thus, the modified proximal gradient algorithm is bounded perturbation resilient.

Proof We rewrite (34) as

$$\begin{aligned} x_{n+1} &= t_n h(x_n + \beta_n v_n) + (1 - t_n) \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n + \beta_n v_n) \\ &= t_n h(x_n) + (1 - t_n) \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n) + \tilde{e}(x_n), \end{aligned} \tag{35}$$

where

$$\begin{aligned} \tilde{e}(x_n) &= t_n [h(x_n + \beta_n v_n) - h(x_n)] \\ &\quad + (1 - t_n) [\text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n + \beta_n v_n) - \text{prox}_{\alpha_n g}(I - \alpha_n \nabla f)(x_n)]. \end{aligned}$$

In view of Lemma 2.4 and the assumptions as regards h and f , we have

$$\begin{aligned} \|\tilde{e}(x_n)\| &\leq t_n \|\beta_n v_n\| + (1 - t_n) \|\beta_n v_n + \alpha_n [\nabla f(x_n) - \nabla f(x_n + \beta_n v_n)]\| \end{aligned}$$

$$\begin{aligned} &\leq t_n \|\beta_n v_n\| + (1 - t_n)(1 + \alpha_n L) \|\beta_n v_n\| \\ &= [1 + (1 - t_n)\alpha_n L] \|\beta_n v_n\|, \end{aligned} \tag{36}$$

which implies that $\sum_{n=0}^\infty \|\tilde{e}(x_n)\| < \infty$ owing to the conditions imposed on t_n, α_n, β_n and v_n . We then deduce the conclusion from Theorem 3.1. \square

5 An application and the numerical experiment

In this section, we apply Theorem 4.2 to linear inverse problem and show the numerical experiment.

5.1 Linear inverse problem

Let H be a real Hilbert space and $A : H \rightarrow H$ a bounded linear operator. Given $b \in H$. We consider the following linear inverse problem:

$$Ax = b + w, \quad x \in H, \tag{37}$$

which is used to estimate an unknown signal x from the noise measurement b in finite-dimensional space. w is an unknown noise vector. This problem can be solved via the regularized least-squares problem:

$$\min_{x \in H} \left\{ \frac{1}{2} \|Ax - b\|^2 + \gamma \|x\| \right\}, \tag{38}$$

where $\gamma > 0$ is a regularization parameter.

By applying algorithm (34) to (38), we obtain the following.

Theorem 5.1 *Let $h : H \rightarrow H$ be a ρ -contractive operator with $\rho \in (0, 1)$. Assume $A \neq 0$ and the solution set S of (38) is nonempty. Assume, in addition, $\{v_n\}$ is a bounded sequence in H , $\{\beta_n\} \subset (0, +\infty)$ such that $\sum_{n=0}^\infty \beta_n < \infty$. Given $x_0 \in H$, we define $\{x_n\}$ by the iterative scheme*

$$x_{n+1} = t_n h(x_n + \beta_n v_n) + (1 - t_n) \text{prox}_{\alpha_n \gamma \|\cdot\|} (I - \alpha_n A^* (A(x_n + \beta_n v_n) - b)) \tag{39}$$

(A^* is the adjoint of A), where

- (i) $0 < a \leq \alpha_n < \frac{2}{\|A\|^2}, \sum_{n=0}^\infty |\alpha_{n+1} - \alpha_n| < \infty$;
- (ii) $\{t_n\} \subset (0, 1), \lim_{n \rightarrow \infty} t_n = 0$;
- (iii) $\sum_{n=0}^\infty t_n = \infty, \sum_{n=0}^\infty |t_{n+1} - t_n| < \infty$.

Then $\{x_n\}$ converges strongly to a point $x^* \in S$, where x^* is the unique solution of the following variational inequality problem:

$$\langle (I - h)x^*, x - x^* \rangle, \quad \forall x \in S.$$

Proof Take $f(x) = \frac{1}{2} \|Ax - b\|^2, g(x) = \gamma \|x\|$. It is easy to see that $f, g \in \Gamma_0(H), \nabla f(x) = A^*(Ax - b)$, and

$$\begin{aligned} &\|\nabla f(x) - \nabla f(y)\| \\ &= \|A^*(Ax - b) - A^*(Ay - b)\| = \|A^*A(x - y)\| \leq \|A\|^2 \|x - y\|. \end{aligned} \tag{40}$$

So ∇f is Lipschitz continuous with $L = \|A\|^2$. In addition, g is subdifferentiable, and its subdifferential is

$$\partial \|\cdot\|(x) = \begin{cases} \{x/\|x\|\}, & \text{if } x \neq 0; \\ B(0; 1), & \text{if } x = 0. \end{cases} \tag{41}$$

So we can apply Theorem 4.2 to obtain this result. □

5.2 Numerical experiment

In this subsection, we apply the iterative scheme (34) to solve (38) with $H = \mathbb{R}^J$ to demonstrate the effectiveness of this algorithm. For finite-dimensional spaces, the least-squares problem (38) takes the form as follows:

$$\min \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \gamma \|x\|_1 : x \in \mathbb{R}^J \right\}, \tag{42}$$

where $A \in \mathbb{R}^{M \times J}$ is a matrix. The vector $b \in \mathbb{R}^M$. $f(x) = \frac{1}{2} \|Ax - b\|_2^2$ implies $\nabla f(x) = A^T(Ax - b)$ with $L = \|A^T A\|$, where A^T represents the transpose of A . $\text{prox}_{\alpha_n \gamma \|\cdot\|_1}(x_n) = (\text{prox}_{\alpha_n \gamma |\cdot|_1}(x_n^1), \dots, \text{prox}_{\alpha_n \gamma |\cdot|_1}(x_n^J))^T$, where $\text{prox}_{\alpha_n \gamma |\cdot|_1}(x_n^k) = \text{sgn}(x_n^k) \max\{|x_n^k| - \alpha_n \gamma, 0\}$, $k = 1, 2, \dots, J$. The bounded sequence $\{v_n\}$ and the summarizable nonnegative real sequence $\{\beta_n\}$ can be chosen as follows:

$$v_n = \begin{cases} -\frac{d_n}{\|d_n\|}, & \text{if } 0 \neq d_n \in \partial g(x_n), \\ 0, & \text{if } 0 \in \partial g(x_n). \end{cases} \tag{43}$$

$\beta_n = c^n$ for some $c \in (0, 1)$.

Throughout the experiments, $A \in \mathbb{R}^{M \times J}$ is a matrix whose entries are sampled independently from a Gaussian distribution of zero mean and unit variance. The vector $b \in \mathbb{R}^M$ is generated from a uniformly distribution in the interval $[-5, 5]$. The regularization parameter $\gamma = 0.05$. We choose $M = 50$ and $J = 200$. Given $t_n = \frac{1}{3n}$, $\alpha_n = \frac{n}{3\sqrt{L(n+1)}}$, we define the stopping criterion

$$Err := \|x_{n+1} - x_n\| < \varepsilon, \tag{44}$$

where ε is a given small positive constant. To see the behavior of algorithm (34), we plotted the evolutions of ‘Err’ defined by (44) with respect to the numbers of iterations in Fig. 1 for the initial point $x_0 = (0, 0, \dots, 0)^T \in \mathbb{R}^{200}$. The plots in Fig. 1 show that the proposed algorithm is reliable to solve (42). Besides, The iteration numbers (“Iter”), the computing time in seconds (“Time”), the error’s values (“Err”) and (“ $\|Ax_n - b\|$ ”) are reported in Table 1 when the stopping criterion $\varepsilon = 5 \times 10^{-5}$ is reached. We can see from Table 1 that the summarizable positive real sequence $\{\beta = c^n\}$ and the contractive constant ρ can have a large impact on the numerical performance. We also find that the sequence $\{x_n\}$ generated by algorithm (34) can get very close to the solution of the problem $Ax = b$.

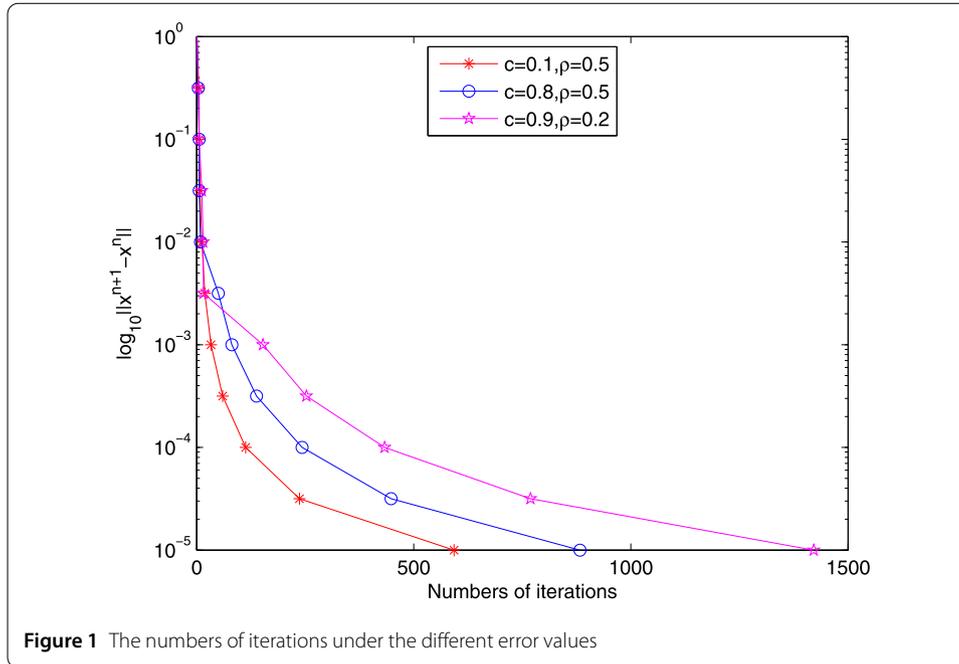


Figure 1 The numbers of iterations under the different error values

Table 1 Numerical results with different (c, ρ) and initial value x_0

(c, ρ)	$x_0 = (0, 0, \dots, 0)^T$			$x_0 = (1, 1, \dots, 1)^T$		
	Iter.	Time	$\ Ax_n - b\ $	Iter.	Time	$\ Ax_n - b\ $
(0.1, 0.5)	152	0.140	0.066	199	0.140	0.074
(0.8, 0.5)	440	0.608	0.220	260	0.172	0.070
(0.9, 0.2)	680	0.546	0.307	773	0.484	0.353

6 Conclusion

In this paper, we introduced a modified proximal gradient algorithm with perturbations in Hilbert space by making a convex combination of a proximal gradient operator and a contractive operator h . There exists a perturbation term in each iterative step (see (12)). We proved that the generated iterative sequence converges strongly to a solution of a non-smooth composite optimization problem. We also showed that the perturbation in computing the gradient of f in algorithm (14) actually can be seen as a special case of (12). Finally, as one of the main objectives of this paper, we verified that the exact modified algorithm is bounded perturbation resilient, a fact which, to some extent, extends the horizon of the recent developed superiorization methodology.

Acknowledgements

The authors would like to acknowledge the reviewers for the very valuable comments which helped to improve the presentation of this paper. The authors also would like to acknowledge the National Natural Science Foundation of China (No. 11402294, 61503385) for financial support.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally and significantly to this paper. All authors have read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 15 September 2017 Accepted: 20 April 2018 Published online: 02 May 2018

References

- Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011)
- Engl, H.W., Hanke, M., Neubauer, A.: *Regularization of Inverse Problems*. Kluwer Academic, Dordrecht (1996)
- Bach, F., Jenatton, R., Mairal, J., Obozinski, G.: Optimization with sparsity-inducing penalties. *Found. Trends Mach. Learn.* **4**, 1–106 (2012)
- Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **57**, 1413–1457 (2004)
- Tibshirani, R.: Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **73**, 273–282 (2011)
- Liu, Q.H., Liu, A.J.: Block SOR methods for the solution of indefinite least squares problems. *Calcolo* **51**, 367–379 (2014)
- Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**, 964–979 (1979)
- Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. *SIAM J. Multiscale Model. Simul.* **4**, 1168–1200 (2005)
- Xu, H.K.: Properties and iterative methods for the lasso and its variants. *Chin. Ann. Math., Ser. B* **35**, 501–518 (2014)
- Guo, Y.N., Cui, W., Guo, Y.S.: Perturbation resilience of proximal gradient algorithm for composite objectives. *J. Nonlinear Sci. Appl.* **10**, 5566–5575 (2017)
- Moudafi, A.: Viscosity approximation methods for fixed-points problems. *J. Math. Anal. Appl.* **241**, 46–55 (2000)
- Xu, H.K.: Viscosity approximation methods for nonexpansive mappings. *J. Math. Anal. Appl.* **298**, 279–291 (2004)
- Censor, Y., Davidi, R., Herman, G.T.: Perturbation resilience and superiorization of iterative algorithms. *Inverse Probl.* **26**, 065008 (2010)
- Censor, Y., Davidi, R., Herman, G.T., Schulte, R.W., Tetruashvili, L.: Projected subgradient minimization versus superiorization. *J. Optim. Theory Appl.* **160**, 730–747 (2014)
- Censor, Y., Zaslavski, A.J.: Strict Fejér monotonicity by superiorization of feasibility-seeking projection methods. *J. Optim. Theory Appl.* **165**, 172–187 (2015)
- Cegielski, A., Al-Musallam, F.: Superiorization with level control. *Inverse Probl.* **33**, 044009 (2017)
- Garduño, E., Herman, G.: Superiorization of the ML-EM algorithm. *IEEE Trans. Nucl. Sci.* **61**, 162–172 (2014)
- He, H., Xu, H.K.: Perturbation resilience and superiorization methodology of averaged mappings. *Inverse Probl.* **33**, 044007 (2017)
- Helou, E.S., Zibetti, M.V.W., Miqueles, E.X.: Superiorization of incremental optimization algorithms for statistical tomographic image reconstruction. *Inverse Probl.* **33**(4), 044010 (2017)
- Nikazad, T., Davidi, R., Herman, G.T.: Accelerated perturbation-resilient block-iterative projection methods with application to image reconstruction. *Inverse Probl.* **28**, 035005 (2012)
- Schrapp, M.J., Herman, G.T.: Data fusion in X-ray computed tomography using a superiorization approach. *Rev. Sci. Instrum.* **85**, 053701 (2014)
- Davidi, R., Censor, Y., Schulte, R.W., Geneser, S., Xing, L.: Feasibility-seeking and superiorization algorithm applied to inverse treatment planning in radiation therapy. *Contemp. Math.* **636**, 83–92 (2015)
- Censor, Y., Zaslavski, A.J.: Convergence and perturbation resilience of dynamic string averaging projection methods. *Comput. Optim. Appl.* **54**, 65–76 (2013)
- Dong, Q.L., Zhao, J., He, S.N.: Bounded perturbation resilience of the viscosity algorithm. *J. Inequal. Appl.* **2016**, 299 (2016)
- Jin, W., Censor, Y., Jiang, M.: Bounded perturbation resilience of projected scaled gradient methods. *Comput. Optim. Appl.* **63**, 365–392 (2016)
- Nikazad, T., Abbasi, M.: A unified treatment of some perturbed fixed point iterative methods with an infinite pool of operators. *Inverse Probl.* **33**, 044002 (2017)
- Xu, H.K.: Bounded perturbation resilience and superiorization techniques for the projected scaled gradient method. *Inverse Probl.* **33**, 044008 (2017)
- Xu, H.K.: Iterative algorithms for nonlinear operators. *J. Lond. Math. Soc.* **66**, 240–256 (2002)
- Xu, H.K.: Averaged mappings and the gradient-projection algorithm. *J. Optim. Theory Appl.* **150**, 360–378 (2011)

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com