Applied Analysis and Optimization
Volume 9, Number 3, 2025, 263–278
Yokohama Publishers
ISSN 2432-1664 Online Journal
© Copyright 2025

TWO-STEP INERTIAL TSENG'S TYPE ALGORITHM FOR SOLVING NONCONVEX AND NONSMOOTH SEPARABLE OPTIMIZATION PROBLEMS

XINGMIN JIA AND JING ZHAO

Abstract. In this paper, we study an algorithm for solving a class of nonconvex and nonsmooth separable optimization problems. Based on Tseng's type algorithm, we propose a new iterative algorithm with two-step inertial extrapolation technique. By constructing appropriate benefit function, with the help of Kurdyka-Łojasiewicz property we establish the convergence of the whole sequence generated by proposed algorithm. Finally, we apply the algorithm to signal recovery and show the effectiveness of proposed algorithm.

1. Introduction

Many practical problems, such as image processing, support vector machines, compressed sensing, low rank matrix recovery, standard phase retrieval [12] can be transformed into the following separable optimization problems:

$$\min_{x\in\mathbb{R}^m} f(x)+h(x), \tag{1.1}$$

where $f:\mathbb{R}^m\to(-\infty,+\infty], h:\mathbb{R}^m\to\mathbb{R}$ are the proper, lower semicontinuous functions. In the full convex setting, there are a plenty of proximal algorithms for solving problem (1.1) effectively; see, [13, 16, 18–21, 23] The algorithms mentioned before both are splitting algorithms. Splitting algorithms have the advantage that the functions are evaluated in the iterative scheme separately. More precisely, a forward step means an evaluation of the smooth part through the gradient, while a backward step can evaluate the nonsmooth counterpart via its proximal operator.

Since nonconvex functions usually approximate the original problem better than convex functions, a large number of problems require to solve nonconvex minimization problems. In recent years, many scholars have paid attention to study nonconvex optimization problems, and some effective and stable algorithms have been proposed. Many algorithms have been proved to possess convergence properties also in the nonconvex case, see [1–3, 6, 8–11]. In this paper, we study optimization problem (1.1) in full nonconvex setting. Let $f:\mathbb{R}^m\to(-\infty,+\infty]$ be a proper, lower semicontinuous function, $h:\mathbb{R}^m\to\mathbb{R}$ be a *Fréchet* differentiable function such that ∇h is $L-Lipschitz$ continuous with $L>0$ and $f+h$ is bounded from below. We note that the above-mentioned functions f and h don't have the requirements of convexity. In this setting, the problem (1.1) is a nonconvex and nonsmooth

2010 *Mathematics Subject Classification.* 49M37, 49J45, 65K05.
Key words and phrases. Kurdyka-Łojasiewicz property, nonconvex and nonsmooth optimization, inertial extrapolation technique, Tseng's type algorithm.

optimization problem. Gao et al. [12] proposed an accelerated mirror descent algorithm for problem (1.1) with a constraint condition and get convergence properties by KŁproperty of the benefit function. Attouch et al. [4] proposed the following forward-backward algorithm for solving problem (1.1):

$$\begin{aligned} x_{n+1} &\in \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x)+\langle\nabla h(x_n),x\rangle+\frac{1}{2\alpha_n}\|x-x_n\|^2\right] \\ &= \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x)+\frac{1}{2\alpha_n}\|x-x_n+\alpha_n\nabla h(x_n)\|^2\right]. \end{aligned} \tag{1.2}$$

The convergence is obtained provided an appropriate regularization of the objective function satisfies the KŁproperty.

Inertial algorithm originates from the heavy ball method in physics dynamical systems. By utilizing the information from the first few steps during the iteration process, the computational complexity can remain basically unchanged and the numerical performance of the algorithm can be improved. Therefore, this acceleration method is widely used in nonconvex and nonsmooth optimization problems, see [14, 15]. Tseng's proximal algorithm fully utilizes the information of smooth functions. Bot et al. [7] proposed the following inertial version of the Tseng's type algorithm for solving nonconvex and nonsmooth minimization problem (1.1):

$$\begin{cases} z_n = \nabla h(x_n) + \frac{\beta_n}{\lambda_n}(x_{n-1}-x_n), \\ p_n \in \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x)+\langle z_n,x\rangle+\frac{1}{2\lambda_n}\|x-x_n\|^2\right], \\ x_{n+1} = p_n + \lambda_n(\nabla h(x_n)-\nabla h(p_n)). \end{cases} \tag{1.3}$$

They proved that the generated sequence globally converges to critical point of the objective function under the condition of the KŁproperty.

In this paper, we study an algorithm for solving nonconvex and nonsmooth separable optimization problem (1.1). Based on Tseng's type algorithm, we propose a new iterative algorithm with two-step inertial extrapolation and get convergence results in the full nonconvex setting. The methods used to prove the convergence of the numerical scheme rely on the same three key elements as those used in other algorithms for nonconvex optimization problems involving KŁ functions. More precisely, we demonstrate a sufficient decrease property for the iterates, establish the existence of a subgradient lower bound for the iterates gap, and ultimately, leverage certain analytical properties of the objective function to achieve convergence.

In Section 2, we review some concepts and important lemmas. We propose the two-step inertial Tseng type algorithm in Section 3. We analyze the convergence of the proposed algorithm in Section 4. Finally, in Ssction 5, the preliminary numerical example on signal recovery is provided to illustrate the behavior of the proposed algorithm.

2. Preliminaries

Let us recall some notions and results which are needed in the following. Let $\mathbb{N} := \{0, 1, 2, \dots\}$ be the set of nonnegative integers. For $m \geq 1$, the Euclidean scalar product and the induced norm on $\mathbb{R}^m$ are denoted by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$, respectively.

The domain of the function $f : \mathbb{R}^m \to (-\infty, +\infty]$ is defined by $\text{dom}f = \{x \in \mathbb{R}^m : f(x) < +\infty\}$. We say that f is proper if $\text{dom}f \neq \emptyset$.

Definition 2.1. The function $f : \mathbb{R}^m \to (-\infty, +\infty]$ is lower semicontinuous at the point $x_0 \in \text{dom}f$, if $f(x_0) \leq \liminf_{x \to x_0} f(x)$. f is lower semicontinuous function if it is lower semicontinuous at each point in the domain of the function.

Definition 2.2. Let $f : \mathbb{R}^m \to (-\infty, +\infty]$ be a proper lower semicontinuous function.

(i) If $x \in \text{dom}f$, the Fréchet subdifferential of f at x is defined by

$$\hat{\partial} f(x) = \left\{ x^* \in \mathbb{R}^m : \lim_{y \neq x} \inf_{y \to x} \frac{f(y) - f(x) - \langle x^*, y - x \rangle}{\|y - x\|} \geq 0 \right\},$$

for $x \notin \text{dom}f$ we set $\hat{\partial} f(x) = \emptyset$.

(ii) The limiting subdifferential is defined at a point $x \in \text{dom}f$ by

$$\partial f(x) = \{x^* \in \mathbb{R}^m : \exists\, x_k \to x,\ f(x_k) \to f(x),\ \hat{x}_k \in \hat{\partial} f(x_k),\ \hat{x}_k \to x^*\},$$

while for $x \notin \text{dom}f$, one takes $\partial f(x) = \emptyset$.

Notice that in case f is convex, these notions coincide with the convex subdifferential, which means that $\hat{\partial} f(x) = \partial f(x) = \{x^* \in \mathbb{R}^m : f(y) \geq f(x) + \langle x^*, y - x \rangle, \forall y \in \mathbb{R}^m\}$ for all $x \in \text{dom}f$.

Remark 2.3. We give some remarks on subdifferential.

(i) For each $x \in \mathbb{R}^m, \hat{\partial} f(x) \subseteq \partial f(x)$, where the first set is convex and closed while the second one is closed.
(ii) Let $x_k^* \in \partial f(x_k)$ and $\lim_{k \to \infty}(x_k, x_k^*) = (x, x^*)$, then $x^* \in \partial f(x)$, which means that $\partial f(x)$ is a closed set.
(iii) If $x \in \mathbb{R}^m$ is a minimum point of f, then $0 \in \partial f(x)$. A point x is called the critical point of f if $0 \in \partial f(x)$, the set of critical points is denoted $\text{crit}f$.
(iv) Let $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function and $g : \mathbb{R}^m \to \mathbb{R}$ be continuously differential, then $\partial(f + g)(x) = \partial f(x) + \nabla g(x)$, for all $x \in \text{dom}f$.

We turn now our attention to functions satisfying the Kurdyka-Łojasiewicz property. This class of functions will play a crucial role in the convergence results of the proposed algorithm.

Definition 2.4 (Kurdyka-Łojasiewicz property)**.** Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper lower semicontinuous function and $\bar{x} \in \text{dom}f$. Denote $[\eta_1 < f < \eta_2] := \{x \in \mathbb{R}^n : \eta_1 < f(x) < \eta_2\}$. If there exists $\eta \in (0, +\infty)$, a neighborhood U of $\bar{x}$ and a continuous concave function $\varphi : [0, \eta) \to [0, +\infty)$ such that

(i) $\varphi(0) = 0$,
(ii) φ is continuously differentiable on $(0, \eta)$ with $\varphi'(s) > 0\, (\forall s \in (0, \eta))$,
(iii) for all $x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$, the following Kurdyka-Łojasiewicz inequality holds,

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) \geq 1.$$

Then the function f is said to have the Kurdyka-Łojasiewicz (KŁ) property at $\bar{x} \in \text{dom} f$.

We call f is a KŁ function, if f satisfies the KŁ property at each point of $\text{dom}\partial f$. Denote Φ_η the set of functions φ which satisfy $(i)(ii)$ in Definition 3. Then, we recall the uniformized KŁ property established in [17] as follows, which is important for further analysis.

Lemma 2.5 (Uniformity KŁ property [17]). *Let Ω be a compact set and let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper and lower semicontinuous function. Assume that f is constant on Ω, and f satisfies the KŁ property at each point of Ω. Then there exist $\epsilon > 0, \eta > 0, \varphi \in \Phi_\eta$, such that for all $\bar{x} \in \Omega$ and for all $x \in \{x \in \mathbb{R}^n : dist(x, \Omega) < \epsilon\} \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$, the following inequality holds*

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) \geq 1.$$

In practical applications, many functions satisfy KŁ properties [6], such as semi algebraic functions, strongly convex functions, real analytic functions, sub-analytic functions, etc.

In the following, we present two convergence results that will play a crucial role in the proof of the results provided in section 4. The first result has frequently been used in the literature in the context of Fejér monotonicity techniques for establishing convergence results of classical algorithms in convex optimization problems, or more broadly, for monotone inclusion problems (see [5]). The second result is likely already well-known; however, we include some details of its proof to ensure completeness.

Lemma 2.6. *Let $\{a_n\}$ and $\{b_n\}$ be real sequences such that $b_n \geq 0$ for all $n \geq 1$, $\{a_n\}$ is bounded below and $a_{n+1} + b_n \leq a_n$ for all $n \geq 1$. Then, $\{a_n\}$ is a monotonically nonincreasing convergent sequence and $\sum_{n\geq 1} b_n < +\infty$.*

Lemma 2.7. *Let $\{\xi_n\}$ and $\{\varepsilon_n\}$ be sequences in $[0, +\infty)$ such that $\sum_{n\geq 0} \varepsilon_n < +\infty$ and $\xi_{n+1} \leq a\xi_n + b\xi_{n-1} + c\xi_{n-2} + \varepsilon_n$ for all $n \geq 2$, where $a \in \mathbb{R}, b, c \geq 0, a+b+c < 1$. Then $\sum_{n\geq 0} \xi_n < +\infty$.*

Proof. Fix a positive integer $k \geq 2$. Summing up the inequality from the hypotheses for $n = 2, \ldots, k$, we obtain

$$\sum_{n=2}^{k} \xi_{n+1} \leq \sum_{n=2}^{k} (a\xi_n + b\xi_{n-1} + c\xi_{n-2} + \varepsilon_n).$$

We can simply transfer items to obtain

$$\begin{aligned}(1 - a - b - c) \sum_{n=0}^{k} \xi_n \leq \xi_2 - \xi_{k+1} + (1 - a)\xi_1 + (1 - a)\xi_0 - b(\xi_0 + \xi_k) \\ - c(\xi_{k-1} + \xi_k) + \sum_{n=2}^{k} \varepsilon_n.\end{aligned}$$

Moreover, we get

$$(1-a-b-c)\sum_{n=0}^{k}\xi_n \leq \xi_2 + (1-a)\xi_1 + (1-a)\xi_0 - b\xi_0 + \sum_{n=2}^{k}\varepsilon_n.$$

Since $a \in \mathbb{R}, b, c \geq 0, a+b+c<1$, $\sum_{n\geq 2}\varepsilon_n < +\infty$, and let $k \to \infty$, the conclusion follows. □

Lemma 2.8 (Descend Lemma)**.** *Let $h : \mathbb{R}^m \to \mathbb{R}$ be a Fréchet differentiable function and such that ∇h is $L-$ Lipschitz continuous with $L > 0$. Then,*

$$h(y) \leq h(x) + \langle \nabla h(x), y - x\rangle + \frac{L}{2}\|y-x\|^2, \quad \forall (x,y) \in \mathbb{R}^m \times \mathbb{R}^m. \tag{2.1}$$

3. Two-step inertial Tseng's type algorithm

In this section, we propose a two-step inertial Tseng's type algorithm for solving nonconvex and nonsmooth separable optimization problems (1.1). Before this, we give some assumptions and use the assumptions throughout the rest of the paper.

Assumption 3.1. $f : \mathbb{R}^m \to (-\infty, +\infty]$ is a proper, lower semicontinuous function, $h : \mathbb{R}^m \to \mathbb{R}$ is a *Fréchet* differentiable function such that ∇h is $L-$ *Lipschitz* continuous with $L > 0$.

Algorithm 3.1. Let $x_0, x_1, x_2 \in \mathbb{R}^m$, $\overline{\alpha}, \underline{\alpha} > 0$, $a_1 > 0$, $a_2 > 0$, the sequences $\{a_{1,n}\}, \{a_{2,n}\}, \{\alpha_n\}$ fulfill

$$0 \leq a_{1,n} \leq a_1, 0 \leq a_{2,n} \leq a_2$$

and

$$0 < \underline{\alpha} \leq \alpha_n \leq \overline{\alpha},$$

consider the iterative scheme

$$\begin{cases} z_n = \nabla h(x_n) + a_{1,n}(x_{n-1} - x_n) + a_{2,n}(x_{n-2} - x_{n-1}), \\ p_n \in \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x) + \langle z_n, x\rangle + \frac{1}{2\alpha_n}\|x - x_n\|^2\right], \\ x_{n+1} = p_n + \alpha_n(\nabla h(x_n) - \nabla h(p_n)). \end{cases} \tag{3.1}$$

In Algorithm 3.1, we notice that

$$\begin{aligned} p_n &\in \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x) + \langle z_n, x\rangle + \frac{1}{2\alpha_n}\|x - x_n\|^2\right] \\ &= \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x) + \frac{1}{2\alpha_n}\|x - x_n + \alpha_n z_n\|^2\right]. \end{aligned} \tag{3.2}$$

Remark 3.2. We give special cases of Algorithm 3.1.

(i) Let $a_{1,n} \equiv a_{2,n} \equiv 0$, the algorithm (3.1) becomes

$$\begin{cases} p_n \in \underset{x\in\mathbb{R}^m}{\operatorname{argmin}}\left[f(x) + \frac{1}{2\alpha_n}\|x - x_n + \alpha_n \nabla h(x_n)\|^2\right], \\ x_{n+1} = p_n + \alpha_n(\nabla h(x_n) - \nabla h(p_n)), \end{cases}$$

which is the classical Tseng's type algorithm in nonconvex setting.

(ii) Let $a_{2,n} \equiv 0$, the algorithm (3.1) becomes

$$\begin{cases} z_n = \nabla h(x_n) + a_{1,n}(x_{n-1} - x_n), \\ p_n \in \underset{x \in \mathbb{R}^m}{\operatorname{argmin}} \left[f(x) + \langle z_n, x \rangle + \frac{1}{2\alpha_n} \|x - x_n\|^2 \right], \\ x_{n+1} = p_n + \alpha_n (\nabla h(x_n) - \nabla h(p_n)). \end{cases}$$

By variable transform $\alpha_n = \lambda_n$, $a_{1,n} = \frac{\beta_n}{\lambda_n}$, it becomes algorithm (1.3) proposed by Bot et al. [10].

4. Convergence analysis

In this section, we provide convergence analysis of the two-step inertial Tseng's type algorithm (3.1) proposed in this paper for solving nonsmooth and nonconvex problem (1.1).

Lemma 4.1. *Consider the sequences $\{z_n\}, \{p_n\}, \{x_n\}$ generated by Algorithm 3.1. Then for every $v, u, s > 0$, the following inequality holds:*

$$\begin{aligned} f(p_n) + h(p_n) + M_1 \|x_n - p_n\|^2 \le f(p_{n-1}) + h(p_{n-1}) + M_2 \|x_{n-1} - p_{n-1}\|^2 \\ + M_3 \|x_{n-2} - p_{n-2}\|^2, \end{aligned} \tag{4.1}$$

where

$$M_1 = \frac{1}{2\overline{\alpha}} - (s + L + ua_1 + va_2), \tag{4.2}$$

$$M_2 = (s + L + ua_1 + va_2)\overline{\alpha}^2 L^2 + \left(\frac{L^2}{2s} + \frac{1}{2\underline{\alpha}} \right) \overline{\alpha}^2 L^2 + \frac{a_1}{2u}(1 + \overline{\alpha} L)^2, \tag{4.3}$$

$$M_3 = \frac{a_2}{2v}(1 + \overline{\alpha} L)^2. \tag{4.4}$$

Proof. By Algorithm 3.1, fixing $n \ge 3$, we have

$$\begin{aligned} f(p_n) + \langle z_n, p_n \rangle + \frac{1}{2\alpha_n} \|p_n - x_n\|^2 \le f(p_{n-1}) + \langle z_n, p_{n-1} \rangle \\ + \frac{1}{2\alpha_n} \|p_{n-1} - x_n\|^2 \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} \langle z_n, p_{n-1} - p_n \rangle = \langle \nabla h(x_n), p_{n-1} - p_n \rangle + a_{1,n} \langle x_{n-1} - x_n, p_{n-1} - p_n \rangle \\ + a_{2,n} \langle x_{n-2} - x_{n-1}, p_{n-1} - p_n \rangle. \end{aligned} \tag{4.6}$$

Combining (4.5) and (4.6), we get

$$\begin{aligned} f(p_n) + \frac{1}{2\alpha_n} \|p_n - x_n\|^2 \le f(p_{n-1}) + \langle \nabla h(x_n), p_{n-1} - p_n \rangle \\ + a_{1,n} \langle x_{n-1} - x_n, p_{n-1} - p_n \rangle \\ + a_{2,n} \langle x_{n-2} - x_{n-1}, p_{n-1} - p_n \rangle \\ + \frac{1}{2\alpha_n} \|p_{n-1} - x_n\|^2. \end{aligned} \tag{4.7}$$

Moreover, according to Lemma 2.4, we have

$$h(p_n) \leq h(p_{n-1}) + \langle \nabla h(p_{n-1}), p_n - p_{n-1} \rangle + \frac{L}{2} \|p_n - p_{n-1}\|^2. \tag{4.8}$$

Combining (4.7) and (4.8), we get

$$\begin{aligned} f(p_n) + \frac{1}{2\alpha_n} \|p_n - x_n\|^2 + h(p_n) \leq\ & f(p_{n-1}) + \langle \nabla h(x_n), p_{n-1} - p_n \rangle \\ & + a_{1,n} \langle x_{n-1} - x_n, p_{n-1} - p_n \rangle \\ & + \frac{1}{2\alpha_n} \|p_{n-1} - x_n\|^2 \\ & + a_{2,n} \langle x_{n-2} - x_{n-1}, p_{n-1} - p_n \rangle \\ & + \frac{L}{2} \|p_n - p_{n-1}\|^2 \\ & + h(p_{n-1}) + \langle \nabla h(p_{n-1}), p_n - p_{n-1} \rangle. \end{aligned} \tag{4.9}$$

By Algorithm 3.1, we have the following inequality

$$\|x_n - p_{n-1}\| \leq \alpha_{n-1} L \|x_{n-1} - p_{n-1}\|, \tag{4.10}$$

which implies

$$\|p_n - p_{n-1}\|^2 \leq 2\|p_n - x_n\|^2 + 2\alpha_{n-1}^2 L^2 \|x_{n-1} - p_{n-1}\|^2. \tag{4.11}$$

According to (4.10), we get the following inequality

$$\|x_{n-1} - x_{n-2}\| \leq (1 + \alpha_{n-2} L) \|x_{n-2} - p_{n-2}\| \tag{4.12}$$

and

$$\|x_n - x_{n-1}\| \leq (1 + \alpha_{n-1} L) \|x_{n-1} - p_{n-1}\|. \tag{4.13}$$

Taking arbitrarily $v, u, s > 0$, by Young's inequality, we have

$$\langle \nabla h(p_{n-1}) - \nabla h(x_n), p_n - p_{n-1} \rangle \leq \frac{s}{2} \|p_n - p_{n-1}\|^2 + \frac{L^2}{2s} \|x_n - p_{n-1}\|^2, \tag{4.14}$$

$$\langle x_{n-1} - x_n, p_{n-1} - p_n \rangle \leq \frac{1}{2u} \|x_n - x_{n-1}\|^2 + \frac{u}{2} \|p_n - p_{n-1}\|^2 \tag{4.15}$$

and

$$\langle x_{n-2} - x_{n-1}, p_{n-1} - p_n \rangle \leq \frac{1}{2v} \|x_{n-1} - x_{n-2}\|^2 + \frac{v}{2} \|p_{n-1} - p_n\|^2. \tag{4.16}$$

Combining (4.9), (4.12)-(4.16), we can simply transfer items to obtain

$$\begin{aligned} f(p_n) + h(p_n) + \frac{1}{2\alpha_n} \|x_n - p_n\|^2 \leq\ & f(p_{n-1}) + h(p_{n-1}) \\ & + \frac{s + L + u a_{1,n} + v a_{2,n}}{2} \|p_{n-1} - p_n\|^2 \\ & + \frac{a_{1,n}}{2u} \left(1 + \alpha_{n-1} L\right)^2 \|x_{n-1} - p_{n-1}\|^2 \\ & + \frac{a_{2,n}}{2v} \left(1 + \alpha_{n-2} L\right)^2 \|x_{n-2} - p_{n-2}\|^2 \\ & + \left(\frac{L^2}{2s} + \frac{1}{2\alpha_n} \right) \|x_n - p_{n-1}\|^2. \end{aligned} \tag{4.17}$$

Combining (4.10), (4.11) and (4.17), we get

$$f(p_n) + h(p_n) + M_{1,n}\|x_n - p_n\|^2 \le f(p_{n-1}) + h(p_{n-1}) + M_{2,n}\|x_{n-1} - p_{n-1}\|^2 + M_{3,n}\|x_{n-2} - p_{n-2}\|^2,$$

where

$$M_{1,n} = \frac{1}{2\alpha_n} - (s + L + ua_{1,n} + va_{2,n}),$$

$$M_{2,n} = (s + L + ua_{1,n} + va_{2,n})\alpha_{n-1}^2 L^2 + \left(\frac{L^2}{2s} + \frac{1}{2\alpha_n}\right)\alpha_{n-1}^2 L^2 + \frac{a_{1,n}}{2u}(1 + \alpha_{n-1}L)^2,$$

$$M_{3,n} = \frac{a_{2,n}}{2v}(1 + \alpha_{n-2}L)^2.$$

Finally, by using the bounds given for the sequences of real numbers involved, we easily derive that $M_{1,n} > M_1, M_2 > M_{2,n}, M_3 > M_{3,n}$ and the conclusion follows. $\square$

We introduce the function $H : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$,

$$H(b, c, d, e) = f(x_1) + h(x_1) + (M_2 + M_3)\|c - b\|^2 + M_3\|e - d\|^2. \tag{4.18}$$

Define a sequence $\{\kappa^n\} = \{(p_n, x_n, p_{n-1}, x_{n-1})\}$, and use those throughout the rest of the paper.

According to Lemma 4.1, we can obtain the monotonicity of the sequence $\{H(\kappa^n)\}$, if $M_1 > M_2 + M_3$. Next, it will be proven that selecting appropriate parameters can satisfy the monotonicity of the sequence $\{H(\kappa^n)\}$.

Lemma 4.2. *For arbitrary $v, u, s > 0$, chose $\underline{\alpha} > 0, a_1 > 0$ and $a_2 > 0$ such that*

$$2(s + L + ua_1 + va_2)\underline{\alpha} + 2\left(s + L + ua_1 + va_2 + \frac{L^2}{2s} + \frac{1}{2\underline{\alpha}}\right)\underline{\alpha}^3 L^2 + \frac{a_1}{u}(1 + \underline{\alpha}L)^2\underline{\alpha} + \frac{a_2}{v}(1 + \underline{\alpha}L)^2\underline{\alpha} < 1. \tag{4.19}$$

Then, there exists $\overline{\alpha} > 0$ such that $\overline{\alpha} > \underline{\alpha}$ and the constants introduced in Lemma 4.1 *fulfill $M_1 > M_2 + M_3$.*

Proof. There exists $\rho > 0$ such that

$$2(s + L + ua_1 + va_2)(\underline{\alpha} + \rho) + 2\left(s + L + ua_1 + va_2 + \frac{L^2}{2s} + \frac{1}{2\underline{\alpha}}\right)(\underline{\alpha} + \rho)^3 L^2 + \frac{a_1}{u}(1 + (\underline{\alpha} + \rho)L)^2(\underline{\alpha} + \rho) + \frac{a_2}{v}(1 + (\underline{\alpha} + \rho)L)^2(\underline{\alpha} + \rho) < 1.$$

We define $\overline{\alpha} := \underline{\alpha} + \rho$ and $M_1 > M_2 + M_3$ follows straight forwardly from the above inequality. $\square$

We give now a decrease property which will be useful in the following.

Lemma 4.3. *Suppose that $f + h$ is bounded from below and let the sequences $\{x_n\}$ and $\{p_n\}$ be generated by Algorithm* 3.1, *where $v, u, s, \overline{\alpha}, \underline{\alpha}, a_1, a_2$ are chosen as in Lemma* 4.2. *The constants M_1, M_2 and M_3 are chosen as in Lemma* 4.1. *Then, the following statements are true:*

(i) *the sequence $\{H(\kappa^n)\}$ is monotonically nonincreasing and convergent;*

(ii) $\sum_{n\geq 1}\|x_n - p_n\|^2 < +\infty$ *and* $\sum_{n\geq 1}\|x_{n+1} - x_n\|^2 < +\infty$*;*

(iii) *the sequence* $\{(f+h)(p_n)\}$ *is convergent.*

Proof. From Lemma 4.1, we deduce that, for every $n \geq 3$,

$$H(\kappa^n) + (M_1 - M_2 - M_3)\|x_n - p_n\|^2 \leq H(\kappa^{n-1}). \tag{4.20}$$

The conclusion follows from Lemma 4.2, Lemma 2.2 and relation (4.13). □

The following lemma provides an estimate for some elements in the limiting subdifferential.

Lemma 4.4. *Let the sequences* $\{x_n\}$ *and* $\{p_n\}$ *be generated by Algorithm* 3.1. *Then, for every* $n \geq 3$*,*

$$\xi_n \in \partial(f+h)(p_n), \tag{4.21}$$

where

$$\begin{aligned} \xi_n &= \nabla h(p_n) - \nabla h(x_n) + a_{1,n}\alpha_{n-1}\left(\nabla h(x_{n-1}) - \nabla h(p_{n-1})\right) \\ &\quad + a_{2,n}\alpha_{n-2}\left(\nabla h(x_{n-2}) - \nabla h(p_{n-2})\right) + a_{1,n}\left(p_{n-1} - x_{n-1}\right) \\ &\quad + a_{2,n}\left(p_{n-2} - x_{n-2}\right) - \frac{1}{\alpha_n}\left(p_n - x_n\right). \end{aligned} \tag{4.22}$$

Moreover,

$$\begin{aligned} \|\xi_n\| \leq \left(L + \frac{1}{\alpha_n}\right)\|p_n - x_n\| + (a_{1,n} + a_{1,n}\alpha_{n-1}L)\|p_{n-1} - x_{n-1}\| \\ + (a_{2,n} + a_{2,n}\alpha_{n-2}L)\|p_{n-2} - x_{n-2}\|. \end{aligned} \tag{4.23}$$

Proof. Take $n \geq 3$. It follows from (3.1) that

$$\begin{aligned} 0 &\in \partial f(p_n) + z_n + \frac{1}{\alpha_n}(p_n - x_n) \\ &= \partial\left(f+h\right)(p_n) - \nabla h(p_n) + \nabla h(x_n) \\ &\quad + a_{1,n}(x_{n-1} - x_n) + a_{2,n}(x_{n-2} - x_{n-1}) + \frac{1}{\alpha_n}(p_n - x_n). \end{aligned}$$

Relation (4.21) follows from the above identity, by using also that

$$\begin{aligned} x_{n-1} - x_n &= x_{n-1} - p_{n-1} + p_{n-1} - x_n \\ &= (x_{n-1} - p_{n-1}) - (\alpha_{n-1}\nabla h(x_{n-1}) - \alpha_{n-1}\nabla h(p_{n-1})) \end{aligned}$$

and

$$\begin{aligned} x_{n-2} - x_{n-1} &= x_{n-2} - p_{n-2} + p_{n-2} - x_{n-1} \\ &= (x_{n-2} - p_{n-2}) - (\alpha_{n-2}\nabla h(x_{n-2}) - \alpha_{n-2}\nabla h(p_{n-2})). \end{aligned}$$

The relation (4.21) follows from the definition of the sequence $\{\xi_n\}$. □

In the following, we use the notation $\omega(p_n)$ for the set of cluster points of the sequence $\{p_n\}$. Next, we will give some properties of this set (see [6]).

Lemma 4.5. *Suppose that the function* $f+h$ *is coercive(that is* $\lim_{\|x\|\to+\infty}(f+h)(x) = +\infty$*), and let the sequences* $\{x_n\}$ *and* $\{p_n\}$ *be generated in Algorithm* 3.1*, where* $v, u, s, \overline{\alpha}, \underline{\alpha}, a_1, a_2$ *are chosen as in Lemma* 4.2. *The constants* M_1, M_2 *and* M_3 *are chosen as in Lemma* 4.1. *Then, the following statements are true:*

(i) $\emptyset \neq \omega(p_n) \subseteq crit(f+h)$;
(ii) $\lim_{n\to+\infty} dist(p_n, \omega(p_n)) = 0$;
(iiii) $\omega(p_n)$ *is a nonempty, compact and connected set;*
(iv) $f+h$ *is finite and constant on* $\omega(p_n)$.

Proof. Since $f+h$ is a proper, lower semicontinuous and coercive function, it follows that $\inf_{x\in\mathbb{R}^m}[f(x)+h(x)]$ is finite and the infimum is attained (see [14]). Hence $f+h$ is bounded from below. According to Lemma 4.3, we have, for all $n \geq 2$,

$$\begin{aligned}(f+h)(p_n) &\leq (f+h)(p_n) + (M_2+M_3)\|x_n - p_n\|^2 + M_3\|x_{n-1} - p_{n-1}\|^2,\\ &\leq (f+h)(p_2) + (M_2+M_3)\|x_2 - p_2\|^2 + M_3\|x_1 - p_1\|^2.\end{aligned}$$

Since the function $f+h$ is coercive, its lower level sets are bounded and we conclude that $\{p_n\}$ is bounded. Hence $\omega(p_n) \neq \emptyset$, $\{x_n\}$ and $\{z_n\}$ are bounded. Take arbitrarily $p^* \in \omega(p_n)$. There exists a subsequence $\{p_{n_k}\}$ such that $p_{n_k} \to p^* (k \to +\infty)$. We show in the following that $\lim_{k\to+\infty} f(p_{n_k}) = f(p^*)$. Note that the lower semicontinuity of the function f ensures $\liminf_{k\to+\infty} f(p_{n_k}) \geq f(p^*)$. Moreover, from (3.1), we have that, for every $k \geq 2$,

$$f(p_{n_k}) + \langle z_{n_k}, p_{n_k}\rangle + \frac{1}{2\alpha_{n_k}}\|p_{n_k} - x_{n_k}\|^2 \leq f(p^*) + \langle z_{n_k}, p^*\rangle + \frac{1}{2\alpha_{n_k}}\|p^* - x_{n_k}\|^2.$$

By using Lemma 4.3, and by taking into consideration the bounds of the sequences involved, it follows $\limsup_{k\to\infty} f(p_{n_k}) \leq f(p^*)$, $\lim_{k\to\infty} f(p_{n_k}) = f(p^*)$. Further, using the definition of ξ_n in Lemma 4.4, we have $\xi_{n_k} \in \partial(f+h)(p_{n_k})$, for $k \geq 2$. By using (4.22), it follows from $p_{n_k} \to p^*$ that $\xi_{n_k} \to 0 (k \to \infty)$. Since we additionally have that $\lim_{k\to+\infty}(f+h)(p_{n_k}) = (f+h)(p^*)$, the closedness of the graph of the limiting subdifferential operator guarantees that $0 \in \partial(f+h)(p^*)$, thus $p^* \in \mathrm{crit}(f+h)$. Then (i) is proved.

The proof of (ii) and (iii) can be done in the lines of [7]. Next we prove (iv) holds.

By Lemma 4.3, $\{(f+h)(p_n)\}$ is convergent. Let us denote its limit by $l \in \mathbb{R}$. Take arbitrarily $p^* \in \omega(p_n)$. There exists a subsequence $p_{n_k} \to p^* (k \to +\infty)$. As shown in (i) one has that $\lim_{k\to\infty}(f+h)(p_{n_k}) = (f+h)(p^*)$. On the other hand $\lim_{k\to\infty}(f+h)(p_{n_k}) = l$, hence $(f+h)(p^*) = l$. Thus, the restriction of $f+h$ to $\omega(p_n)$ equals l. □

Lemma 4.6. *Suppose that the function $f+h$ is coercive and let the sequences $\{x_n\}$ and $\{p_n\}$ be generated in Algorithm* 3.1, *where $v, u, s, \overline{\alpha}, \underline{\alpha}, a_1, a_2$ are chosen as in Lemma* 4.2. *The constants M_1, M_2 and M_3 are chosen as in Lemma* 4.1. *Then, the following statements are true:*

(i) $\emptyset \neq \omega(\kappa^n) \subseteq crit(H) = \{(x,x,x,x) \in \mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}^m : x \in crit(f+h)\}$;
(ii) $\lim_{n\to+\infty} dist(\kappa^n, \omega(\kappa^n)) = 0$;
(iii) $\omega(\kappa^n)$ *is a nonempty, compact and connected set;*
(iv) H *is finite and constant on* $\omega(\kappa^n)$.

Proof. The proof is similar to the one of Lemma 4.5, by noticing that

$$H(\kappa^n) + (M_1 - M_2 - M_3)\|x_n - p_n\|^2 \leq H(\kappa^{n-1}),$$

$$\zeta_n = (\xi_n + 2(M_2+M_3)(p_n - x_n), 2(M_2+M_3)(x_n - p_n), \tag{4.24}$$

$$2M_3(p_{n-1} - x_{n-1}), 2M_3(x_{n-1} - p_{n-1})) \in \partial H(\kappa^n), \tag{4.25}$$

where $n \geq 3$, $\{\xi_n\}$ is the sequence introduced in Lemma 4.4. □

We are now in position to prove the convergence of Algorithm 3.1 provided that H is a KŁ function.

Theorem 4.7. *Suppose that the function $f + h$ is coercive, the function H is a KŁ function. Let the sequences $\{x_n\}$ and $\{p_n\}$ be generated in Algorithm 3.1, where $v, u, s, \overline{\alpha}, \underline{\alpha}, a_1, a_2$ are chosen as in Lemma 4.2. The constants M_1, M_2 and M_3 are chosen as in Lemma 4.1. Then,*

(i) $\sum_{n\geq 1} \|x_n - p_n\| < +\infty$, $\sum_{n\geq 1} \|x_{n+1} - x_n\| < +\infty$;

(ii) *there exists $x \in crit(f + h)$ such that* $\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} p_n = x$.

Proof. According to Lemma 4.6, we can take an element $\kappa^* = (p^*, p^*, p^*, p^*) \in \omega(\kappa^n)$. In analogy to the proof of Lemma 4.5 one can easily show that $\lim_{n\to+\infty} H(\kappa^n) = H(\kappa^*)$. We consider two cases.

(i) There exists an integer $\overline{n} \in \mathbb{N}$ such that $H\left(\kappa^{\overline{n}}\right) = H(\kappa^*)$. The decrease property in (4.20) implies $\kappa^{n+1} = \kappa^n$ for any $n \geq \overline{n}$. It follows that $\kappa^n = \kappa^*$ for any $n \geq \overline{n}$ and the assertions hold trivially.

(ii) For all $n \geq 1$, we have $H(\kappa^n) > H(\kappa^*)$. Denote $\Omega := \omega(\kappa^n)$. Since H is a KŁ function, from Lemma 4.6 and Lemma 2.1, there exist $\epsilon, \eta > 0$ and $\varphi \in \Phi_\eta$ such that for all (x_1, y_1, x_2, y_2) in the intersection

$$\begin{aligned}&\{(x_1, y_1, x_2, y_2) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m : dist\left((x_1, y_1, x_2, y_2), \Omega\right) < \epsilon\} \\ &\quad \cap \{(x_1, y_1, x_2, y_2) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m : \\ &\qquad H(p^*, p^*, p^*, p^*) < H(x_1, y_1, x_2, y_2) < H(p^*, p^*, p^*, p^*) + \eta\},\end{aligned} \tag{4.26}$$

the following inequality holds:

$$\begin{aligned}&\varphi'(H(x_1, y_1, x_2, y_2) - H(p^*, p^*, p^*, p^*))dist((0,0,0,0), \partial H(x_1, y_1, x_2, y_2)) \\ &\geq 1.\end{aligned} \tag{4.27}$$

Take $n_1 > 0$ such that $H(\kappa^n) < H(\kappa^*) + \eta$ for $n \geq n_1$. Take $n_2 > 0$ such that $dist(\kappa^n, \Omega) < \epsilon$ for $n \geq n_2$. Let $N = max\{n_1, n_2\}$, thus κ^n is in the intersection (4.24) for all $n \geq N$. Then the following inequality holds:

$$\varphi'(H(\kappa^n) - H(\kappa^*))dist((0,0,0,0), \partial H(\kappa^n)) \geq 1. \tag{4.28}$$

Because of the concavity of φ, the following inequality holds:

$$\begin{aligned}&\varphi(H(\kappa^n) - H(\kappa^*)) - \varphi(H(\kappa^{n+1}) - H(\kappa^*)) \\ &\qquad \geq \varphi'(H(\kappa^n) - H(\kappa^*))(H(\kappa^n) - H(\kappa^{n+1})).\end{aligned} \tag{4.29}$$

From (4.27) and (4.24), for all $n \geq N$

$$\varphi'(H(\kappa^n) - H(\kappa^*)) \geq \frac{1}{\|\zeta_n - 0\|}, \tag{4.30}$$

where ζ_n is defined by (4.24). For the convenience of expression, let

$$\Delta_{n,n+1} := \varphi(H(\kappa^n) - H(\kappa^*)) - \varphi(H(\kappa^{n+1}) - H(\kappa^*)).$$

Thus, from (4.27), (4.29) and (4.20), for all $n \geq N$, we have

$$\Delta_{n,n+1} \geq (M_1 - M_2 - M_3) \cdot \frac{\|x_{n+1} - p_{n+1}\|^2}{\|\zeta_n\|}, \tag{4.31}$$

moreover, accoding to Young's inequality, we have

$$\|x_{n+1} - p_{n+1}\| \leq \frac{\delta\|\zeta_n\|}{2} + \frac{\Delta_{n,n+1}}{2\delta(M_1 - M_2 - M_3)}, \tag{4.32}$$

where $\delta > 0$ satisfies

$$\begin{aligned}&\frac{\delta}{2}\sqrt{\left[3(L + \frac{1}{\underline{\alpha}} + 2(M_2 + M_3))^2 + 4(M_2 + M_3)^2\right]} + \\ &\frac{\delta}{2}\left(\sqrt{\left[3(a_1 + a_1\overline{\alpha}L)^2 + 8M_3^2\right]} + \sqrt{3}(a_2 + a_2\overline{\alpha}L)\right) < 1.\end{aligned}$$

From the definition of ζ_n in (4.24), we can obtain

$$\begin{aligned}\|\zeta_n\| \leq &\sqrt{\left[3(L + \frac{1}{\underline{\alpha}} + 2(M_2 + M_3))^2 + 4(M_2 + M_3)^2\right]}\|x_n - p_n\| \\ &+ \sqrt{\left[3(a_1 + a_1\overline{\alpha}L)^2 + 8M_3^2\right]}\|x_{n-1} - p_{n-1}\| \\ &+ \sqrt{3}\,(a_2 + a_2\overline{\alpha}L)\,\|x_{n-2} - p_{n-2}\|.\end{aligned}$$

Combining this inequality and (4.31), we get

$$\begin{aligned}\|x_{n+1} - p_{n+1}\| \leq a\|x_n - p_n\| + b\|x_{n-1} - p_{n-1}\| + c\|x_{n-2} - p_{n-2}\| \\ + \frac{\Delta_{n,n+1}}{2\delta(M_1 - M_2 - M_3)},\end{aligned} \tag{4.33}$$

where

$$\begin{aligned}a &:= \frac{\delta}{2}\sqrt{\left[3(L + \frac{1}{\underline{\alpha}} + 2(M_2 + M_3))^2 + 4(M_2 + M_3)^2\right]}, \\ b &:= \frac{\delta}{2}\sqrt{\left[3(a_1 + a_1\overline{\alpha}L)^2 + 8M_3^2\right]}, \\ c &:= \frac{\sqrt{3}\delta}{2}(a_2 + a_2\overline{\alpha}L).\end{aligned}$$

From the definition of $\Delta_{n,n+1}$, for all $k \geq 2$,

$$\sum_{n=2}^{k} \Delta_{n,n+1} \leq \varphi\left(H(\kappa^2) - H(\kappa^*)\right).$$

This indicates that

$$\sum_{n\geq 2} \frac{\Delta_{n,n+1}}{2\delta(M_1 - M_2 - M_3)} < +\infty.$$

From Lemma 2.3, we have $\sum_{n\geq 2}\|x_n - p_n\| < +\infty$. Combining this and (4.13) we get $\sum_{n\geq 2}\|x_{n+1} - x_n\| < +\infty$.

We easily derive that $\{x_n\}$ is a Cauchy sequence, hence it is convergent. Let $\lim_{n\to\infty} x_n = x$, then $\lim_{n\to\infty} p_n = x$ from $x_n - p_n \to 0 (n \to +\infty)$. We can get the conclusion (ii) by Lemma 4.5 (i). □

5. Numerical experiments

In all of the following experiments, the codes are written in MATLAB, using the Windows 11 operating system, an R7 5700X3D CPU @ 3GHz and RAM 32GB.

The sparse signal recovery problem has been studied extensively. Supposed x is an unknown vector in $\mathbb{R}^m$ (a signal), given a matrix $A \in \mathbb{R}^{n\times m}$ and an observation $b \in \mathbb{R}^n$, we plan to recover x from observation b such that x is of the sparsest structure(that is, x has the fewest nonzero components). In this section, the sparse signal recovery problem can be modulated by the following L_0-problem

$$\min_x \|x\|_0, s.t. Ax = b.$$

The L_0-problem in the form of $L_{1/2}$ regularization can be described as

$$\min_x \frac{1}{2}\|Ax - b\|_2^2 + \eta\|x\|_{1/2}^{1/2},$$

where $\eta > 0$ is used to balance regularization and data fitting. $\|x\|_{1/2}$ is the $L_{1/2}$ quasi-norm of $\mathbb{R}^m$, defined by $\|x\|_{1/2} = \left(\sum_{i=1}^m |x_i|^{1/2}\right)^2$. Now we illustrate how to implement Algorithm 3.1 for solving the above model. Let $f(x) = \eta\|x\|_{1/2}^{1/2}, h(x) = \frac{1}{2}\|Ax - b\|_2^2$, thus we can obtain the explicit expression of the subproblem based on Algorithm 3.1,

$$\begin{aligned}
z_n &= A^T(Ax_n - b) + a_{1,n}(x_{n-1} - x_n) + a_{2,n}(x_{n-2} - x_{n-1}),\\
p_n &\in \operatorname*{argmin}_{x\in\mathbb{R}^m}\left[f(x) + \langle z_n, x\rangle + \frac{1}{2\alpha_n}\|x - x_n\|^2\right]\\
&= \operatorname*{argmin}_{x\in\mathbb{R}^m}\left[f(x) + \frac{1}{2\alpha_n}\|x - x_n + \alpha_n z_n\|^2\right]\\
&= \mathcal{H}\left(x_n - \alpha_n z_n, \eta\alpha_n\right)
\end{aligned}$$

and

$$x_{n+1} = p_n + \alpha_n\left(A^T A(x_n - p_n)\right).$$

For all $\iota > 0$, $\mathcal{H}(\cdot, \iota)$ is called half shrinkage operator [22] defined as

$$\mathcal{H}(a, \iota) = \{h_\iota(a_1), h_\iota(a_2), \ldots, h_\iota(a_n)\}^T,$$

with

$$h_\iota(a_i) = \begin{cases} \frac{2a_i}{3}(1 + cos(\frac{2\pi}{3} - \frac{2}{3}\varphi_\iota(a_i))), & |a_i| > \frac{3}{4}\iota^{2/3},\\ 0, & otherwise,\end{cases}$$

and $\varphi_\iota(a_i) = \arccos\left(\frac{\iota}{8}(\frac{|a_i|}{3})^{-3/2}\right)$.

In this experiment, each entry of A is drawn from the standard normal distribution, and then all columns of A are normalized according to $L_{1/2}$ quasi-norm; we generate a random sparse vector as x; the vector without noise $b = Ax$; the noise vector $\omega \sim N(0, 10^{-3}I)$ and the vector $b = Ax + \omega$; the regularization parameter $\eta = 0.001\left\|A^T b\right\|_\infty$. In this experiment, we take $a_{1,n} \equiv 0.3, a_{2,n} \equiv 0.3, \alpha_n \equiv 0.2$ in

TABLE 1. the algorithms' performance while $n = 80, m = 500$

	$b = Ax$		$b = Ax + w$	
$n = 80, m = 500$	iteration no.	time(second)	iteration no.	time(second)
Alg (1.2)	2596	23.03	3126	26.01
Alg (1.3)	2010	20.59	2782	25.31
Alg 3.1	1742	16.67	2577	23.35

TABLE 2. the algorithms' performance while $n = 100, m = 600$

	$b = Ax$		$b = Ax + w$	
$n = 100, m = 600$	iteration no.	time(second)	iteration no.	time(second)
Alg (1.2)	2288	25.53	3148	35.03
Alg (1.3)	1863	23.32	2510	30.78
Alg 3.1	1264	19.07	1870	25.37

algorithm 3.1, $a_{1,n} \equiv 0.3, \alpha_n \equiv 0.2$ (i.e., $\lambda_n \equiv 0.2, \beta_n \equiv 0.06$) in algorithm (1.3) and $\alpha_n \equiv 0.2$ in algorithm (1.2). We take the origin point as the initial point for all algorithms and use

$$E_n = \|x_{n+1} - x_n\| < 10^{-4}$$

as the stopping criterion. In the following, we compare algorithm (1.2), algorithm (1.3) and Algorithm 3.1 for matrix A of different dimension with $(n, m) = (80, 500)$ and $(n, m) = (100, 600)$.

Table 1 and Table 2 report the number of iteration and CPU time for $(n, m) = (80, 500)$ and $(n, m) = (100, 600)$, respectively. Alg (1.2) represents algorithm (1.2), Alg (1.3) represents algorithm (1.3) and Alg 3.1 represents Algorithm 3.1. From Table 1 and Table 2, we can know that two inertial versions of Tseng's type algorithm are better than one inertial version of Tseng's type algorithm and forward backward algorithm.

6. Conclusions

In this paper, two-step inertial Tseng's type algorithm is proposed to solve nonconvex and nonsmooth separable optimization problems. We construct appropriate benefit function and obtain that the generated sequence is globally convergent to a critical point, under the assumptions that auxiliary function satisfies the Kurdyka-Łojasiewicz inequality and the parameters satisfy certain conditions. In numerical experiments, the solutions of subproblems have colsed form for solving the sparse signal recovery problem. Numerical results are reported to show the effectiveness of proposed algorithm.

References

[1] B. Azmi and M. Bernreuther, *On the forward-backward method with nonmonotone linesearch for infinite-dimensional nonsmooth nonconvex problems,* Comput. Optim. Appl. **91** (2025), 1263–1308.

[2] H. Attouch and J. Bolte, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features,* Math. Program. Ser. B **116** (2009), 5–16.

[3] H. Attouch, J. Bolte, P. Redont and A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka–Lojasiewicz inequality,* Math. Oper. Res. **35** (2010), 438–457.

[4] H. Attouch, J. Bolte and B.F. Svaiter, *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods,* Math. Program. Ser. **137** (2013), 91–129.

[5] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems,* SIAM J. Imaging Sci. **2** (2009), 183–202.

[6] J. Bolte, S. Sabach and M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems,* Math. Program. Ser. A **146** (2014), 459–494.

[7] R.I. Boţ and E.R. Csetnek, *An Inertial Tseng's Type Proximal Algorithm for Nonsmooth and Nonconvex Optimization Problems,* J. Optim. Theory Appl. **171** (2016), 600–616.

[8] Y. Chen, H. Liu, X. Ye and Q. Zhang, *Learnable descent algorithm for nonsmooth nonconvex image reconstruction,* SIAM J. Imag. Sci. **14** (2021), 1532–1564.

[9] E. Chouzenoux, J.C. Pesquet and A. Repetti, *Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function,* J. Optim. Theory Appl. **162** (2014), 107–132.

[10] Y. Chen, L. Liu and L. Zhang, *A learned proximal alternating minimization algorithm and its induced network for a class of two-block nonconvex and nonsmooth optimization,* J. Sci. Comput. **103** (2025): 56.

[11] P. Frankel, G. Garrigos and J. Peypouquet, *Splitting methods with variable metric for KŁ functions,* J. Optim. Theory Appl. **165** (2014), 874–900.

[12] X. Gao, Y. Jiang, X. Cai and W. Kai, *An accelerated mirror descent algorithm for constrained nonconvex problems,* Optim. **2025** (2025), 1–26.

[13] A. Moudafi, *Operator splitting schemes through a regularization approach,* Commun. Optim. Theory **2025** (2025): 2.

[14] P. Ochs, Y. Chen and T. Brox, *IPIANO : Inertial proximal algorithm for non-convex optimization,* SIAM J. Imaging Sci **7** (2014), 1388–1419.

[15] B.T. Polyak, *Some methods of speeding up the convergence of iteration methods,* USSR Comput. **4** (1964), 1–17.

[16] Y. Pei, Y. Chen and S. Song, *A novel accelerated algorithm for solving split variational inclusion problems and fixed point problems,* J. Nonlinear Funct. Anal. **2023** (2023): 19.

[17] R.T. Rockafellar and R. Wets, *Variational Analysis,* Springer Berlin (1998).

[18] H. Sadeghi, S. Banert and P. Giselsson, *Forward-backward splitting with deviations for monotone inclusions,* Appl. Set-Valued Anal. Optim. **6** (2024), 113–135.

[19] V.N. Tran, Y. Shehu, R. Xu and P.T. Vuong, *An inertial forward-backward-forward algorithm for solving non-convex mixed variational inequalities,* J. Appl. Numer. Optim. **5** (2023), 335–348.

[20] B. Tan and X. Qin, *On relaxed inertial projection and contraction algorithms for solving monotone inclusion problems,* Adv. Comput. Math. **50** (2024): 59.

[21] B. Tan and X. Qin, *An alternated inertial algorithm with weak and linear convergence for solving monotone inclusions,* Ann. Math. Sci. Appl. **8** (2023), 321–345.

[22] Z. Xu, X. Chang, F. Xu and H. Zhang, *$L_{1/2}$ Regularization: A Thresholding Representation Theory and a Fast Solver,* IEEE Trans. **23** (2012), 1013–1027.

[23] X. Zuo, S. Osher and W. Li, *Primal-dual damping algorithms for optimization*, Ann. Math. Sci. Appl. **9** (2024), 467–504.

Manuscript received June 1 2025
revised August 23 2025

X. Jia
College of Science, Civil Aviation University of China, Tianjin, China
E-mail address: jiaxingmin1224@163.com

J. Zhao
College of Science, Civil Aviation University of China, Tianjin, China
E-mail address: zhaojing200103@163.com