Liu et al. *Journal of Inequalities and Applications* (2024) 2024:117
https://doi.org/10.1186/s13660-024-03197-z

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

A modified inertial proximal alternating direction method of multipliers with dual-relaxed term for structured nonconvex and nonsmooth problem

Yang Liu[1], Long Wang[2] and Yazheng Dang[2*]

*Correspondence: jgdyz@163.com
[2]School of Management, University of Shanghai for Science and Technology, Jungong Road 516, Yangpu District, Shanghai, PRC 200093, China
Full list of author information is available at the end of the article

Abstract

In this research, we introduce a novel optimization algorithm termed the dual-relaxed inertial alternating direction method of multipliers (DR-IADM), tailored for handling nonconvex and nonsmooth problems. These problems are characterized by an objective function that is a composite of three elements: a smooth composite function combined with a linear operator, a nonsmooth function, and a mixed function of two variables. To facilitate the iterative process, we adopt a straightforward parameter selection approach, integrate inertial components within each subproblem, and introduce two relaxed terms to refine the dual variable update step. Within a set of reasonable assumptions, we establish the boundedness of the sequence generated by our DR-IADM algorithm. Furthermore, leveraging the Kurdyka–Łojasiewicz (KŁ) property, we demonstrate the global convergence of the proposed method. To validate the practicality and efficacy of our algorithm, we present numerical experiments that corroborate its performance. In summary, our contribution lies in proposing DR-IADM for a specific class of optimization problems, proving its convergence properties, and supporting the theoretical claims with numerical evidence.

Keywords: Inertial effect; Dual-relaxed-term; ADMM; Kurdyka–Łojasiewicz property

1 Introduction

The present paper deals with the following nonconvex and nonsmooth problem as in [5]:

$$\min_{(x,y)\in R^m\times R^q}\left\{F(Ax)+G(y)+H(x,y)\right\}, \tag{1.1}$$

where $F: R^p \to R$ is a continuously Lipschitz differentiable function, $G: R^q \to R\cup\{+\infty\}$ is a proper and lower semicontinuous function, $H: R^m \times R^q \to R$ is a Frechet differentiable function with Lipschitz continuous gradient, and $A: R^m \to R^p$ is a linear operator. Many application problems can be modeled as (1.1), e.g., in compressed sensing [2, 9], matrix factorization [4], sparse approximations of signals and images [16, 19], and so on.

Obviously, when $m = p$ and A is the identity operator, (1.1) can be written as

$$\min_{(x,y)\in R^m\times R^q}\left\{F(x)+G(y)+H(x,y)\right\}. \tag{1.2}$$

A general method for addressing problem (1.2) is the alternating minimization method, as mentioned in the literature [3, 17, 22]. In the context of nonconvex and nonsmooth problems, Bolte et al. investigated a proximal alternating linearized minimization (PALM) algorithm in [4]. Following this, Driggs et al. introduced a generic stochastic variant of the PALM algorithm in [10], which allows for various variance-reduced gradient approximations. The PALM algorithm is essentially a blockwise implementation of the well-known proximal forward–backward algorithm, as referenced in [8, 13],

$$\min_{x\in R^m}\left\{F(Ax)+H(x)\right\}, \tag{1.3}$$

where $H : R^m \to R$ is Frechet differentiable and possesses a Lipschitz continuous gradient. In the convex case, the alternating direction method of multipliers (ADMM) [1, 3] and linearized ADMMi [11, 18, 23, 24] have proven to be highly effective in solving problem (1.3). Following that, Bot et al. [6] introduced a proximal linearized ADMM algorithm, and Liu et al., as seen in [14], presented a two-block linearized ADMM and a multi-block parallel linearized ADMM for the nonconvex case.

For the problem denoted by equation (1.1), Bot [5] converted it into a three-block nonseparable problem by introducing an additional variable:

$$\begin{aligned}&\min_{(x,y,z)\in R^m\times R^q\times R^p} F(z)+G(y)+H(x,y)\\ &\text{such that}\quad Ax=z.\end{aligned} \tag{1.4}$$

The augmented Lagrangian function $L_\beta : R^m \times R^q \times R^p \times R^p \to R \cup \{+\infty\}$ associated with problem (1.4) reads

$$L_\beta(x,y,z,u)=F(z)+G(y)+H(x,y)+\langle u,Ax-z\rangle+\frac{\beta}{2}\|Ax-z\|^2,\quad \beta>0, \tag{1.5}$$

where u is the Lagrangian multiplier and β is the penalty parameter. Bot gave a proximal minimization algorithm (PMA) to solve it in [5], which takes the following iterative form:

$$\begin{aligned}
y^{k+1} &\in \arg\min_{y\in \mathrm{R}^q}\left\{G(y)+\left\langle \nabla_y H\left(x^k,y^k\right),y\right\rangle+\frac{\mu}{2}\left\|y-y^k\right\|^2\right\},\\
z^{k+1} &\in \arg\min_{z\in \mathrm{R}^p}\left\{F(z)+\left\langle u^k,Ax^k-z\right\rangle+\frac{\beta}{2}\left\|Ax^k-z\right\|^2\right\rangle\right\},\\
x^{k+1} &:= x^k-\tau^{-1}\left(\nabla_x H\left(x^k,y^{k+1}\right)+A^T u^k+\beta A^T\left(Ax^k-z^{k+1}\right)\right),\\
u^{k+1} &:= u^k+\sigma\beta\left(Ax^{k+1}-z^{k+1}\right),
\end{aligned}$$

where $\tau > 0$, $0 < \sigma < 1$. In [5], sufficient conditions are established to ensure that the sequence generated is bounded, and it is demonstrated that the global convergence is achieved in accordance with the Kurdyka–Łojasiewicz inequality.

Recently, numerous scholars have integrated the inertial effect with ADMM for various nonconvex problems to enhance convergence [15, 25]. For example, Le et al. [12] introduced an inertial Alternating Direction Method of Multipliers (iADMM) tailored for tackling a category of nonconvex, nonsmooth multi-block composite optimization challenges characterized by linear constraints. In [23], an inertial proximal partially symmetric ADMM was introduced by Wang for tackling linearly constrained multi-block nonconvex separable optimization problems. This method involves updating the Lagrange multiplier not once but twice and incorporates distinct relaxation factors [20, 21] at every iteration. For the problem (1.3), Chao et al. [7] combined the inertial technique with ADMM and employed the KŁ assumption to achieve global convergence in the nonconvex setting.

Motivated by the aforementioned algorithms, we are poised to present a novel approach in this document. This approach is a dual-relaxed variant of the inertial proximal alternating direction method of multipliers, tailored for addressing the challenges posed by nonconvex and nonsmooth problems, specifically referring to problem (1.1). The key contributions of this paper are delineated as follows:

(1) In contrast to the approach described in [5], our algorithm integrates the fundamental concepts of the ADMM with an inertial component applied uniformly across all subproblems, rather than selectively to certain subproblems. This strategic implementation significantly enhances convergence rates.

(2) In contrast to the conventional ADMM or its variants, our proposed algorithm introduces two relaxation terms (instead of merely one) during the dual variable update phase, which consequently establishes a novel iterative dynamic for the dual ascent procedure.

(3) We provide straightforward sufficient conditions for the boundedness of the sequence generated by our algorithm. Unlike other studies, there is no need to assume that the sequence generated by the algorithm is bounded a priori.

The structure of the paper is as follows. In Sect. 2, we compile a collection of useful definitions and findings that will serve as a foundation for our convergence analysis. In Sect. 3, we introduces a novel weak inertial proximal minimization algorithm and delves into its convergence properties. A numerical experiment aimed at validating the efficacy of our proposed algorithm is conducted in the fourth section. The paper concludes with a summary of key points in the fifth section.

2 Notation and preliminaries

In the following, R^n stands for the n-dimensional Euclidean space,

$$\langle x, y\rangle = x^T y = \sum_{i=1}^{n} x_i y_i, \qquad \|x\| = \sqrt{\langle x, x\rangle},$$

where T stands for the transpose operation. For a set $S \subset R^n$ and a point $x \in R^n$, let $\operatorname{dist}(x, S) = \inf_{y\in S} \|y - x\|^2$. If $S = \emptyset$, we set $\operatorname{dist}(x, S) = +\infty$ for $\forall x \in R^n$.

Definition 2.1 (Lipschitz differentiability) Function $f(x)$ is said to be L_f Lipschitz differentiable if for all x, y we have

$$\left\|\nabla f(x) - \nabla f(y)\right\| \le L_f \|x - y\|.$$

Lemma 2.1 ([17] (Descent lemma)) *Let* $f : R^n \to \mathrm{R}$ *be Frechet differentiable such that its gradient is Lipschitz continuous with constant* $\ell > 0$. *Then, for* $\forall x, y \in R^n$ *and* $z = \{(1-t)x + ty : t \in [0,1]\} \in [x,y]$, *it holds that*

$$f(y) \le f(x) + \left\langle \nabla f(z), y - x \right\rangle + \frac{\ell}{2}\|y - x\|^2.$$

Lemma 2.2 ([23]) *Suppose the sequence of real numbers* $\{a_k\}_{k\ge 0}$ *is bounded from below,* $\{b_k\}_{k\ge 0}$ *is a sequence of real nonnegative numbers, and for* $\forall k \ge 0$,

$$a_{k+1} + b_k \le a_k.$$

Then the following statements are valid:

(*i*) *The sequence* $\{a_k\}_{k\ge 0}$ *is monotonically decreasing and convergent.*

(*ii*) *The sequence* $\{b_k\}_{k\ge 0}$ *is summable, namely* $\sum_{k\ge 0} b_k < \infty$.

Lemma 2.3 ([23]) *Let* $\{a_k\}_{k\in N}$ *and* $\{b_k\}_{k\in N}$ *be nonnegative real sequences such that* $\sum_{k\in N} b_k < \infty$ *and* $a_{k+1} \le a \cdot a_k + b \cdot a_{k-1} + b_k$, *for* $\forall k \ge 1$, *where* $a \in R$, $b \ge 0$ *and* $a + b < 1$. *Then* $\sum_{k\in N} a_k < \infty$.

We proceed to introduce a function that exhibits the Kurdyka–Łojasiewicz property. This particular class of functions will be integral in establishing the convergence outcomes for our recommended algorithm.

Definition 2.2 ([2]) Let $\eta \in (0, +\infty]$. We use Φ_η to denote the set of all concave and continuous functions $\varphi : [0, \eta) \to [0, +\infty)$. A function φ belonging to the set Φ_η for $\eta \in (0, +\infty]$ is called a desingularization function if it satisfies the following conditions:

(i) $\varphi(0) = 0$.

(ii) φ is continuously differentiable on $(0, \eta)$ and continuous at 0.

(iii) $\varphi'(s) > 0$ for any $s \in (0, \eta)$.

Definition 2.3 ([2] (Kurdyka–Łojasiewicz property)) Let $f : R^n \to R \cup \{+\infty\}$ be proper and lower semicontinuous. The function f is said to have the Kundyka–Łojasiewicz (KŁ) property at a point $\hat{v} \in \operatorname{dom} \partial f := \{v \in R^n : \partial f(v) \neq \emptyset\}$ if there exists $\eta \in (0, +\infty]$, a neighborhood V of $\hat{v}$, and a function $\varphi \in f_\eta$ such that

$$\varphi'\big(f(v) - f(\hat{v})\big) \cdot \operatorname{dist}\big(\mathbf{0}, \partial f(v)\big) \ge 1,$$

for any

$$v \in V \cap \big\{v \in R^n : f(\hat{v}) < f(v) < f(\hat{v}) + \eta\big\}.$$

If f satisfies the KŁproperty at each point of dom ∂f, then f is called a KŁ function. Next, we recall the following result which is called the uniformized KŁproperty.

Lemma 2.4 ([2] (Uniformized KŁproperty)) *Let* Ω *be a compact set and* $f : R^n \to R \cup \{+\infty\}$ *be a proper and lower semicontinuous function. Assume that* f *is constant on* Ω *and satisfies* *the KŁ property at each point of* Ω. *Then there exist* $\varepsilon > 0$, $\eta > 0$, *and* $\varphi \in f_\eta$ *such that*

$$\varphi'\big(f(v) - f(\hat{v})\big) \cdot \operatorname{dist}\big(\mathbf{0}, \partial f(v)\big) \geq 1,$$

for any $\hat{v} \in \Omega$ *and every element* v *in the intersection*

$$\big\{v \in R^n : \operatorname{dist}(v, \Omega) < \varepsilon\big\} \cap \big\{v \in R^d : f(\hat{v}) < f(v) < f(\hat{v}) + \eta\big\}.$$

Definition 2.4 ([2] (Subdifferentials)) Let $f : R^n \to (-\infty, +\infty]$ be a proper and lower semicontinuous function. Suppose

$$\liminf_{y \neq x} \frac{f(y) - f(x) - \langle u, y - x\rangle}{\|y - x\|} \geq 0.$$

When $x \notin \operatorname{dom} f$, we set $\widehat{\partial} f(x) = \emptyset$.

(ii) The limiting-subdifferential, or simply the subdifferential, of f at $x \in R^n$, written $\partial f(x)$, is defined through the following closure process $\partial f(x) := \{u \in R^n : \exists x^k \to x, f(x^k) \to f(x)$ and $u^k \in \widehat{\partial} f(x^k) \to u$ as $k \to \infty\}$.

3 Algorithm and its convergence

In this section, we put forward a synchronized approach for solving the optimization problem (1.1) through an inertial proximal minimization algorithm with dual relaxation and subsequently examine its convergence properties.

Algorithm 3.1 Let α, β, $\tau > 0$, $0 < \theta < 1$. For the starting points $(x^0, y^0, z^0) = (x^1, y^1, z^1) \in R^m \times R^q \times R^p$ and $u^1 \in R^p$. The sequence $\{(x^k, y^k, z^k, u^k)\}_{k \geq 0}$ for $\forall k \geq 1$ is generated by:

$$y^{k+1} = \arg\min\big\{G(y) + H\big(x^k, y\big) + \tau \big\|y - z_y^k\big\|^2\big\}, \tag{3.1a}$$

$$z^{k+1} = \arg\min\left\{F(z) + < u^k, Ax^k - z > + \frac{\beta}{2}\big\|Ax^k - z\big\|^2 + \tau\big\|z - z_z^k\big\|^2\right\}, \tag{3.1b}$$

$$\begin{aligned} x^{k+1} = \arg\min\bigg\{ & H\big(x, y^{k+1}\big) + < u^k, Ax - z^{k+1} > + \frac{\beta}{2}\big\|Ax - z^{k+1}\big\|^2 \\ & + \frac{\alpha}{2}\big\|x - x^k\big\|^2 + \tau\big\|x - z_x^k\big\|^2\bigg\}, \end{aligned} \tag{3.1c}$$

$$u^{k+1} = u^k - \beta\big(z^{k+1} - Ax^{k+1}\big) - 2\tau\big(z^{k+1} - z_z^k\big), \tag{3.1d}$$

where

$$\begin{aligned} z_x^k &= x^k - \theta\big(x^k - x^{k-1}\big), \\ z_y^k &= y^k - \theta\big(y^k - y^{k-1}\big), \\ z_z^k &= z^k - \theta\big(z^k - z^{k-1}\big). \end{aligned}$$

Remark 3.1 Inertial terms $\tau\|\cdot\|^2$ are added into the y-, z-, and x- subproblem, respectively, and there exist two relaxed terms $\beta(z^{k+1} - Ax^{k+1})$ and $2\tau(z^{k+1} - z_z^k)$ in the dual update step in (3.1d). Hence we call our algorithm as dual-relaxed inertial proximal ADMM.

We will analyze the convergence of Algorithm 3.1 under the following assumptions:

Assumption A (i) The function F is Lipschitz differentiable, i.e.,

$$\|\nabla F(z) - \nabla F(z')\|^2 \le l_F^2 \|z - z'\|^2.$$

(ii) The function L_β is bounded from below and there exists a constant $\underline{L}$ such that

$$\underline{L} := \inf_{(x,y,\lambda)\in R^m\times R^q\times R^p\times R^p} \{L_\beta(x,y,z,u)\} > -\infty.$$

(iii) For any fixed $y \in R^q$, there exists $\ell_1(y) \ge 0$ such that

$$\|\nabla_y H(x,y) - \nabla_y H(x',y)\| \le \ell_1(y)\|x - x'\|, \quad \forall x, x' \in R^m.$$

Furthermore, there exists $\ell_{1,+} > 0$ such that $\sup_{y\in R^q} \ell_1(y) \le \ell_{1,+}$.

(iv) The parameters satisfy

$$0 < \theta < 0.5, \qquad \tau > 0, \qquad \beta > \frac{3(1+\tau)l_F^2 + 2\tau + 2\tau\theta^2}{(1-2\theta)\tau},$$
$$\alpha > -2(1-2\theta)\tau + 12\beta(1+\tau)\|A\|^2.$$

(v) Let $T := R^m \times R^q \times R^p \times R^p$. The set $\{\omega \in T : L_\beta(\omega) \le L_\beta(\omega^1)\}$ is bounded.

Lemma 3.1 *By the definitions of z_x^k, z_y^k, z_z^k, it holds that*

$$\|x^k - z_x^k\|^2 - \|x^{k+1} - z_x^k\|^2 \le -(1-\theta)\|x^k - x^{k+1}\|^2 + \theta\|x^k - x^{k-1}\|^2, \tag{3.2a}$$
$$\|y^k - z_y^k\|^2 - \|y^{k+1} - z_y^k\|^2 \le -(1-\theta)\|y^k - y^{k+1}\|^2 + \theta\|y^k - y^{k-1}\|^2 \quad \textit{and} \tag{3.2b}$$
$$\|z^k - z_z^k\|^2 - \|z^{k+1} - z_z^k\|^2 \le -(1-\theta)\|z^k - z^{k+1}\|^2 + \theta\|z^k - z^{k-1}\|^2. \tag{3.2c}$$

Proof By the definition of $\|x^k - z_x^k\|^2$, we have

$$\begin{aligned}\|x^k - z_x^k\|^2 - \|x^{k+1} - z_x^k\|^2 &\le \theta^2\|x^k - x^{k-1}\|^2 - \|x^k - x^{k+1} - \theta(x^k - x^{k-1})\|^2 \\ &= -\|x^k - x^{k+1}\|^2 + 2\theta\langle x^k - x^{k+1}, x^k - x^{k-1}\rangle \\ &\le -\|x^k - x^{k+1}\|^2 + \theta\|x^k - x^{k+1}\|^2 + \theta\|x^k - x^{k-1}\|^2 \\ &= -(1-\theta)\|x^k - x^{k+1}\|^2 + \theta\|x^k - x^{k-1}\|^2.\end{aligned}$$

Similarly, we get

$$\|y^k - z_y^k\|^2 - \|y^{k+1} - z_y^k\|^2 \le -(1-\theta)\|y^k - y^{k+1}\|^2 + \theta\|y^k - y^{k-1}\|^2 \quad \text{and}$$
$$\|z^k - z_z^k\|^2 - \|z^{k+1} - z_z^k\|^2 \le -(1-\theta)\|z^k - z^{k+1}\|^2 + \theta\|z^k - z^{k-1}\|^2.$$

The proof is completed. □

The following Lemmas 3.2 and 3.3 provide the descent properties of the key function defined in (3.11) and are important for the convergence.

Lemma 3.2 *Suppose that Assumption A holds, while L_β is defined as (1.5). Then,*

$$\begin{aligned}
&L_\beta(x^{k+1},y^{k+1},z^{k+1},u^{k+1}) + \left(3\beta(1+\tau)\|A\|^2+\theta\tau\right)\left\|x^{k+1}-x^k\right\|^2 + \theta\tau\left\|y^{k+1}-y^k\right\|^2 \\
&\quad + \left(\frac{2\tau\theta^2}{\beta}+\theta\tau\right)\left\|z^{k+1}-z^k\right\|^2 + C_1\left\|x^{k+1}-x^k\right\|^2 + C_2\left\|y^{k+1}-y^k\right\|^2 + C_3\left\|z^{k+1}-z^k\right\|^2 \\
&\le L_\beta(x^k,y^k,z^k,u^k) + \left(3\beta(1+\tau)\|A\|^2+\theta\tau\right)\left\|x^k-x^{k-1}\right\|^2 \\
&\quad + \theta\tau\left\|y^k-y^{k-1}\right\|^2 + \left(\frac{2\tau\theta^2}{\beta}+\theta\tau\right)\left\|z^k-z^{k-1}\right\|^2,
\end{aligned}$$

where

$$\begin{aligned}
C_1 &= \frac{\alpha}{2} + (1-\theta)\tau - 6\beta(1+\tau)\|A\|^2 - \theta\tau, \\
C_2 &= (1-\theta)\tau - \theta\tau, \\
C_3 &= (1-\theta)\tau - \frac{(1+\tau)3l_F^2}{\beta} - \frac{2\tau}{\beta} - \frac{2\tau\theta^2}{\beta} - \theta\tau.
\end{aligned}$$

Proof From (3.1a), (3.1b), and (3.1c), we have

$$G(y^{k+1}) + H(x^k,y^{k+1}) + \tau\left\|y^{k+1}-z_y^k\right\|^2 \le G(y^k) + H(x^k,y^k) + \tau\left\|y^k-z_y^k\right\|^2, \tag{3.3}$$

$$\begin{aligned}
&F(z^{k+1}) + \langle u^k, Ax^k - z^{k+1}\rangle + \frac{\beta}{2}\left\|Ax^k-z^{k+1}\right\|^2 + \tau\left\|z^{k+1}-z_z^k\right\|^2 \\
&\le F(z^k) + \langle u^k, Ax^k - z^k\rangle + \frac{\beta}{2}\left\|Ax^k-z^k\right\|^2 + \tau\left\|z^k-z_z^k\right\|^2,
\end{aligned} \tag{3.4}$$

and

$$\begin{aligned}
&H(x^{k+1},y^{k+1}) + \langle u^k, Ax^{k+1}-z^{k+1}\rangle + \frac{\beta}{2}\left\|Ax^{k+1}-z^{k+1}\right\|^2 \\
&\quad + \frac{\alpha}{2}\left\|x^{k+1}-x^k\right\|^2 + \tau\left\|x^{k+1}-z_x^k\right\|^2 \\
&\le H(x^k,y^{k+1}) + \langle u^k, Ax^k-z^{k+1}\rangle + \frac{\beta}{2}\left\|Ax^k-z^{k+1}\right\|^2 + \tau\left\|x^k-z_x^k\right\|^2,
\end{aligned} \tag{3.5}$$

respectively. Adding (3.3), (3.4) and (3.5) yields

$$\begin{aligned}
&F(z^{k+1}) + G(y^{k+1}) + H(x^{k+1},y^{k+1}) + \langle u^k, Ax^{k+1}-z^{k+1}\rangle \\
&\quad + \frac{\beta}{2}\left\|Ax^{k+1}-z^{k+1}\right\|^2 + \tau\left\|x^{k+1}-z_x^k\right\|^2 + \tau\left\|y^{k+1}-z_y^k\right\|^2 \\
&\quad + \frac{\alpha}{2}\left\|x^{k+1}-x^k\right\|^2 + \tau\left\|z^{k+1}-z_z^k\right\|^2 \\
&\le F(z^k) + G(y^k) + H(x^k,y^k) + \langle u^k, Ax^k-z^k\rangle \\
&\quad + \frac{\beta}{2}\left\|Ax^k-z^k\right\|^2 + \tau\left\|x^k-z_x^k\right\|^2 + \tau\left\|y^k-z_y^k\right\|^2 + \tau\left\|z^k-z_z^k\right\|^2.
\end{aligned}$$

By the definition of L_β, we have

$$\begin{aligned} &L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^k\right)+\tau\left\|x^{k+1}-z_x^k\right\|^2+\tau\left\|y^{k+1}-z_y^k\right\|^2\\ &\quad+\tau\left\|z^{k+1}-z_z^k\right\|^2+\frac{\alpha}{2}\left\|x^{k+1}-x^k\right\|^2\\ &\quad\le L_\beta\left(x^k,y^k,z^k,u^k\right)+\tau\left\|x^k-z_x^k\right\|^2+\tau\left\|y^k-z_y^k\right\|^2+\tau\left\|z^k-z_z^k\right\|^2. \end{aligned}$$

Then,

$$\begin{aligned} &L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^k\right)\\ &\quad\le L_\beta\left(x^k,y^k,z^k,u^k\right)+\tau\left\|x^k-z_x^k\right\|^2-\tau\left\|x^{k+1}-z_x^k\right\|^2-\frac{\alpha}{2}\left\|x^{k+1}-x^k\right\|^2\\ &\qquad+\tau\left\|y^k-z_y^k\right\|^2-\tau\left\|y^{k+1}-z_y^k\right\|^2+\tau\left\|z^k-z_z^k\right\|^2-\tau\left\|z^{k+1}-z_z^k\right\|^2\\ &\quad\le L_\beta\left(x^k,y^k,z^k,u^k\right)+\left(-(1-\theta)\tau-\frac{\alpha}{2}\right)\left\|x^{k+1}-x^k\right\|^2+\theta\tau\left\|x^k-x^{k-1}\right\|^2\\ &\qquad+\left(-(1-\theta)\tau\right)\left\|y^{k+1}-y^k\right\|^2+\theta\tau\left\|y^k-y^{k-1}\right\|^2\\ &\qquad+\left(-(1-\theta)\tau\right)\left\|z^{k+1}-z^k\right\|^2+\theta\tau\left\|z^k-z^{k-1}\right\|^2. \end{aligned} \tag{3.6}$$

The optimality condition for (3.1b) implies

$$\nabla F\left(z^{k+1}\right)-u^k+\beta\left(z^{k+1}-Ax^k\right)+2\tau\left(z^{k+1}-z_z^k\right)=0. \tag{3.7}$$

Combing (3.7) and (3.1d), we obtain

$$u^{k+1}=\nabla F\left(z^{k+1}\right)+\beta\left(Ax^{k+1}-Ax^k\right). \tag{3.8}$$

Hence,

$$\begin{aligned} \left\|u^{k+1}-u^k\right\|^2&=\left\|\nabla F\left(z^{k+1}\right)-\nabla F\left(z^k\right)+\beta\left(Ax^{k+1}-Ax^k\right)-\beta\left(Ax^k-Ax^{k-1}\right)\right\|^2\\ &\le 3l_F^2\left\|z^{k+1}-z^k\right\|^2+3\beta^2\|A\|^2\left\|x^{k+1}-x^k\right\|^2+3\beta^2\|A\|^2\left\|x^k-x^{k-1}\right\|^2. \end{aligned} \tag{3.9}$$

Inserting the u-updating rule (3.1d), we get

$$\begin{aligned} &L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^{k+1}\right)-L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^k\right)\\ &\quad=\left\langle u^{k+1}-u^k,Ax^{k+1}-z^{k+1}\right\rangle\\ &\quad=\frac{1}{\beta}\left\langle u^{k+1}-u^k,u^{k+1}-u^k+2\tau\left(z^{k+1}-z_z^k\right)\right\rangle\\ &\quad=\frac{1}{\beta}\left\|u^{k+1}-u^k\right\|^2+\frac{2\tau}{\beta}\left\langle u^{k+1}-u^k,z^{k+1}-z_z^k\right\rangle\\ &\quad\le\frac{1}{\beta}\left\|u^{k+1}-u^k\right\|^2+\frac{\tau}{\beta}\left\|u^{k+1}-u^k\right\|^2+\frac{\tau}{\beta}\left\|z^{k+1}-z_z^k\right\|^2\\ &\quad\le\frac{(1+\tau)}{\beta}\left\|u^{k+1}-u^k\right\|^2+\frac{2\tau}{\beta}\left\|z^{k+1}-z^k\right\|^2+\frac{2\tau\theta^2}{\beta}\left\|z^k-z^{k-1}\right\|^2. \end{aligned} \tag{3.10}$$

From (3.9), (3.10), and (3.6), we have

$$\begin{aligned}
&L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^{k+1}\right)\\
&\le L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^k\right)+\frac{1+\tau}{\beta}\left(3l_F^2\left\|z^{k+1}-z^k\right\|^2+3\beta^2\|A\|^2\left\|x^{k+1}-x^k\right\|^2\right.\\
&\quad\left.+3\beta^2\|A\|^2\left\|x^k-x^{k-1}\right\|^2\right)+\frac{2\tau}{\beta}\left\|z^{k+1}-z^k\right\|^2+\frac{2\tau\theta^2}{\beta}\left\|z^k-z^{k-1}\right\|^2\\
&=L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^k\right)+\left(\frac{(1+\tau)3l_F^2}{\beta}+\frac{2\tau}{\beta}\right)\left\|z^{k+1}-z^k\right\|^2\\
&\quad+3\beta(1+\tau)\|A\|^2\left\|x^{k+1}-x^k\right\|^2\\
&\quad+3\beta(1+\tau)\|A\|^2\left\|x^k-x^{k-1}\right\|^2+\frac{2\tau\theta^2}{\beta}\left\|z^k-z^{k-1}\right\|^2\\
&\le L_\beta\left(x^k,y^k,z^k,u^k\right)+\left(3\beta(1+\tau)\|A\|^2-(1-\theta)\tau-\frac{\alpha}{2}\right)\left\|x^{k+1}-x^k\right\|^2\\
&\quad+\left(3\beta(1+\tau)\|A\|^2+\theta\tau\right)\left\|x^k-x^{k-1}\right\|^2+\left(-(1-\theta)\tau\right)\left\|y^{k+1}-y^k\right\|^2\\
&\quad+\theta\tau\left\|y^k-y^{k-1}\right\|^2+\left(\frac{(1+\tau)3l_F^2}{\beta}+\frac{2\tau}{\beta}-(1-\theta)\tau\right)\left\|z^{k+1}-z^k\right\|^2\\
&\quad+\left(\frac{2\tau\theta^2}{\beta}+\theta\tau\right)\left\|z^k-z^{k-1}\right\|^2,
\end{aligned}$$

which can be written as

$$\begin{aligned}
&L_\beta\left(x^{k+1},y^{k+1},z^{k+1},u^{k+1}\right)+\left(3\beta(1+\tau)\|A\|^2+\theta\tau\right)\left\|x^{k+1}-x^k\right\|^2+\theta\tau\left\|y^{k+1}-y^k\right\|^2\\
&\quad+\left(\frac{2\tau\theta^2}{\beta}+\theta\tau\right)\left\|z^{k+1}-z^k\right\|^2+C_1\left\|x^{k+1}-x^k\right\|^2\\
&\quad+C_2\left\|y^{k+1}-y^k\right\|^2+C_3\left\|z^{k+1}-z^k\right\|^2\\
&\le L_\beta\left(x^k,y^k,z^k,u^k\right)+\left(3\beta(1+\tau)\|A\|^2+\theta\tau\right)\left\|x^k-x^{k-1}\right\|^2\\
&\quad+\theta\tau\left\|y^k-y^{k-1}\right\|^2+\left(\frac{2\tau\theta^2}{\beta}+\theta\tau\right)\left\|z^k-z^{k-1}\right\|^2,
\end{aligned}$$

where

$$\begin{aligned}
&C_1=\frac{\alpha}{2}+(1-\theta)\tau-6\beta(1+\tau)\|A\|^2-\theta\tau,\\
&C_2=(1-\theta)\tau-\theta\tau,\\
&C_3=(1-\theta)\tau-\frac{(1+\tau)3l_F^2}{\beta}-\frac{2\tau}{\beta}-\frac{2\tau\theta^2}{\beta}-\theta\tau.
\end{aligned}$$

The proof is completed. □

Remark 3.2 Obviously, Assumption A(iv) implies $C_1>0$, $C_2>0$, and $C_3>0$.

Based on Lemma 3.2, we define the following key function (the regularized augmented Lagrangian function)

$$\widehat{L}_\beta(\hat{\omega}^k) = L_\beta(x^k,y^k,z^k,u^k) + \eta_1\|x^k - x^{k-1}\|^2 + \eta_2\|y^k - y^{k-1}\|^2 + \eta_3\|z^k - z^{k-1}\|^2, \quad (3.11)$$

where $\eta_1 = 3\beta(1+\tau)\|A\|^2 + \theta\tau$, $\eta_2 = \theta\tau$, and $\eta_3 = \frac{2\tau\theta^2}{\beta} + \theta\tau$.

Let $\hat{\omega} = (x,y,z,u,x',y',z')$, $\hat{\omega}^k = (x^k,y^k,z^k,u^k,x^{k-1},y^{k-1},z^{k-1})$, $\omega^k = (x^k,y^k,z^k,u^k)$. Then the following lemma implies that the sequence $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\geq1}$ is decreasing. It is of great importance for the following convergence analysis.

Lemma 3.3 (Descent property) *Suppose that Assumption A holds. Let $\hat{L}_\beta(\hat{\omega}^k)$ be defined as in (3.11). Then we have $C_1, C_2, C_3 > 0$ such that*

$$\hat{L}_\beta(\hat{\omega}^{k+1}) + C_1\|x^{k+1} - x^k\|^2 + C_2\|y^{k+1} - y^k\|^2 + C_3\|z^{k+1} - z^k\|^2 \leq \hat{L}_\beta(\hat{\omega}^k). \quad (3.12)$$

Proof The result follows directly from Lemma 3.2 and Remark 3.2. The proof is completed. □

Theorem 3.1 (Boundedness) *Suppose that Assumption A holds. Suppose $\{\omega^k\}_{k\geq0}$ is a sequence generated by Algorithm 3.1, then the following statements are true:*

(*i*) *The sequence $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\geq1}$ is bounded from below and convergent.*

(*ii*) *One has*

$$x^{k+1} - x^k \to 0, \qquad y^{k+1} - y^k \to 0,$$
$$z^{k+1} - z^k \to 0, \quad \text{and} \quad u^{k+1} - u^k \to 0 \quad \text{as } k \to +\infty.$$

(*iii*) *The sequence $\{L_\beta(\omega^k)\}_{k\geq1}$ is convergent.*

(*iv*) *The sequence $\{(x^k,y^k,z^k,u^k)\}_{k\geq0}$ is bounded.*

Proof For $\eta_1 > 0$, $\eta_2 > 0$, $\eta_3 > 0$, one can obtain

$$L_\beta(x^k,y^k,z^k,u^k) \leq L_\beta(x^k,y^k,z^k,u^k) + \eta_1\|x^k - x^{k-1}\|^2 + \eta_2\|y^k - y^{k-1}\|^2 + \eta_3\|z^k - z^{k-1}\|^2,$$

that is,

$$L_\beta(\omega^k) \leq \hat{L}_\beta(\hat{\omega}^k). \quad (3.13)$$

From Assumption A(ii), we know that $\hat{L}_\beta(\hat{\omega}^k) \geq \underline{L}$, which implies that the sequence $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\geq1}$ is bounded from below. Combining (3.12) and Lemma 2.1, it is easy to get that the sequence $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\geq1}$ is convergent and also that

$$x^{k+1} - x^k \to 0, \qquad y^{k+1} - y^k \to 0, \qquad z^{k+1} - z^k \to 0 \quad \text{as } k \to \infty.$$

Then, according to (3.9), it follows that $u^{k+1} - u^k \to 0$ as $k \to \infty$. By the definition of $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\geq1}$, we obtain that $\{L_\beta(\omega^k)\}$ is convergent. From (3.12), we have that $\hat{L}_\beta(\hat{\omega}^k) \leq$

$\hat{L}_\beta(\hat{\omega}^1)$, for $\forall k > 0$. In addition, $\hat{L}_\beta(\hat{\omega}^1) = L_\beta(\omega^1)$ due to $x^0 = x^1$, $y^0 = y^1$, and $z^0 = z^1$. So, from (3.13), we get

$$L_\beta(\omega^k) \le L_\beta(\omega^1) \quad \forall k > 0.$$

Therefore, it follows that the sequence $\{(x^k, y^k, z^k, u^k)\}_{k\ge 0}$ generated by Algorithm 3.1 is bounded by Assumption A(v). The proof is completed. □

The next lemma provides upper estimates for the limiting subgradients of $\hat{L}_\beta(\hat{\omega}^k)$.

Lemma 3.4 *Suppose that Assumption* A *holds. Denote* $v^k = (x^k, y^k, z^k)$. *Then there exists* $\zeta > 0$ *such that*

$$\operatorname{dist}\left(0, \partial \hat{L}_\beta\left(\hat{\omega}^{k+1}\right)\right) \le \zeta\left(\left\|v^{k+1} - v^k\right\| + \left\|v^k - v^{k-1}\right\|\right). \tag{3.14}$$

Proof Let $k \ge 1$ be fixed. Applying the calculus rules of the limiting subdifferential, we get

$$\begin{aligned} \partial_x \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = {} & \nabla_x H\left(x^{k+1}, y^{k+1}\right) + A^T u^{k+1} + \beta A^T\left(Ax^{k+1} - z^{k+1}\right) \\ & + 2\eta_1\left(x^{k+1} - x^k\right), \end{aligned} \tag{3.15a}$$

$$\partial_y \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = \partial G\left(y^{k+1}\right) + \nabla_y H\left(x^{k+1}, y^{k+1}\right) + 2\eta_2\left(y^{k+1} - y^k\right), \tag{3.15b}$$

$$\partial_z \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = \nabla F\left(z^{k+1}\right) - u^{k+1} - \beta\left(Ax^{k+1} - z^{k+1}\right) + 2\eta_3\left(z^{k+1} - z^k\right), \tag{3.15c}$$

$$\partial_u \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = Ax^{k+1} - z^{k+1} = \frac{1}{\beta}\left(u^{k+1} - u^k + 2\tau\left(z^{k+1} - z_z^k\right)\right), \tag{3.15d}$$

$$\partial_{x'} \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = -2\eta_1\left(x^{k+1} - x^k\right), \tag{3.15e}$$

$$\partial_{y'} \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = -2\eta_2\left(y^{k+1} - y^k\right), \tag{3.15f}$$

$$\partial_{z'} \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = -2\eta_3\left(z^{k+1} - z^k\right). \tag{3.15g}$$

By the optimality condition for (3.1c), we have

$$\nabla_x H\left(x^{k+1}, y^{k+1}\right) + A^T u^k + \beta A^T\left(Ax^{k+1} - z^{k+1}\right) + 2\tau\left(x^{k+1} - z_x^k\right) + \alpha\left(x^{k+1} - x^k\right) = 0.$$

Substituting it into (3.15a) leads to

$$\partial_x \hat{L}_\beta\left(\hat{\omega}^{k+1}\right) = A^T u^{k+1} - A^T u^k + 2\eta_1\left(x^{k+1} - x^k\right) - 2\tau\left(x^{k+1} - z_x^k\right) - \alpha\left(x^{k+1} - x^k\right).$$

By the optimality condition for (3.1a), we have

$$0 \in \partial G\left(y^{k+1}\right) + \nabla_y H\left(x^k, y^{k+1}\right) + 2\tau\left(y^{k+1} - z_y^k\right).$$

Substituting it into (3.15b) leads to

$$\partial_y \hat{L}_\beta\left(y^{k+1}\right) = \nabla_y H\left(x^{k+1}, y^{k+1}\right) - \nabla_y H\left(x^k, y^{k+1}\right) + (2\eta_2 - 2\tau)\left(y^{k+1} - y^k\right) + 2\tau\theta\left(y^k - y^{k-1}\right).$$

Substituting (3.7) into (3.15c) leads to

$$\partial_z \hat{L}_\beta(x^{k+1}) = u^k - u^{k+1} - \beta A(x^{k+1} - x^k) + (2\eta_3 - 2\tau)(z^{k+1} - z^k) + 2\tau\theta(z^k - z^{k-1}).$$

Let $D^k = (d_x^{k+1}, d_y^{k+1}, d_z^{k+1}, d_u^{k+1}, d_{x'}^{k+1}, d_{y'}^{k+1}, d_{z'}^{k+1})$, where

$$\begin{aligned}
d_x^{k+1} &= A^T u^{k+1} - A^T u^k + (2\eta_1 - \alpha - 2\tau)(x^{k+1} - x^k) - 2\tau(x^{k+1} - z_x^k),\\
d_y^{k+1} &= \nabla_y H(x^{k+1}, y^{k+1}) - \nabla_y H(x^k, y^{k+1}) + (2\eta_2 - 2\tau)(y^{k+1} - y^k) + 2\tau\theta(y^k - y^{k-1}),\\
d_z^{k+1} &= u^k - u^{k+1} - \beta A(x^{k+1} - x^k) + (2\eta_3 - 2\tau)(z^{k+1} - z^k) + 2\tau\theta(z^k - z^{k-1}),\\
d_u^{k+1} &= \frac{1}{\beta}\big(u^{k+1} - u^k + 2\tau(z^{k+1} - z^k) - 2\tau\theta(z^k - z^{k-1})\big),\\
d_{x'}^{k+1} &= -2\eta_1(x^{k+1} - x^k),\\
d_{y'}^{k+1} &= -2\eta_2(y^{k+1} - y^k),\\
d_{z'}^{k+1} &= -2\eta_3(z^{k+1} - z^k).
\end{aligned}$$

Then it follows that $D^{k+1} \in \partial \hat{L}_\beta(\hat{\omega}^{k+1})$ and $(d_x^{k+1}, d_y^{k+1}, d_z^{k+1}, d_u^{k+1}) \in \partial L_\beta(\omega^{k+1})$.
Thus $\operatorname{dist}^2(0, \partial \hat{L}_\beta(\omega^{k+1})) \le \|D^{k+1}\|^2$. By Assumption A(iii), we have

$$\|\nabla_y H(x^{k+1}, y^{k+1}) - \nabla_y H(x^k, y^{k+1})\| \le \ell_{1,+}\|x^{k+1} - x^k\|.$$

Then, there exists $\zeta_1 > 0$ such that

$$\begin{aligned}
\operatorname{dist}^2(0, \partial \hat{L}_\beta(\hat{\omega}^{k+1})) &\le \|D^{k+1}\|^2\\
&\le \zeta_1^2\big(\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2 + \|z^{k+1} - z^k\|^2\\
&\quad + \|u^{k+1} - u^k\|^2 + \|y^k - y^{k-1}\|^2 + \|x^k - x^{k-1}\|^2 + \|z^k - z^{k-1}\|^2\big).
\end{aligned}$$

Thus, by (3.9), there exists $\zeta > 0$ such that

$$\begin{aligned}
\operatorname{dist}^2(0, \partial \hat{L}_\beta(\hat{\omega}^{k+1})) &\le \zeta^2\big(\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2 + \|z^{k+1} - z^k\|^2\\
&\quad + \|x^k - x^{k-1}\|^2 + \|y^k - y^{k-1}\|^2 + \|z^k - z^{k-1}\|^2\big).
\end{aligned} \tag{3.16}$$

For $v^k = (x^k, y^k, z^k)$, it follows that

$$\|v^k - v^{k-1}\|^2 = \|x^k - x^{k-1}\|^2 + \|y^k - y^{k-1}\|^2 + \|z^k - z^{k-1}\|^2.$$

Combining with (3.16), the latter gives

$$\begin{aligned}
\operatorname{dist}(0, \partial \hat{L}_\beta(\hat{\omega}^{k+1})) &\le \sqrt{\zeta^2\big(\|v^{k+1} - v^k\|^2 + \|v^k - v^{k-1}\|^2\big)}\\
&\le \zeta\big(\|v^{k+1} - v^k\| + \|v^k - v^{k-1}\|\big).
\end{aligned}$$

The proof is completed. □

Now we prove that any cluster point of $\{(x^k, y^k, z^k, u^k)\}_{k\geq 0}$ is a KKT point of the optimization problem (1.1). Let Ω and $\hat{\Omega}$ denote the cluster point set of the sequences $\{\omega^k\}$ and $\{\hat{\omega}^k\}$, respectively.

Theorem 3.2 (Subsequence convergence) *Suppose that Assumption A holds. Then we have that*

(*i*) $\hat{\Omega}$ *is nonempty, compact, and connected.*

(*ii*) $\operatorname{dist}(\hat{\omega}^k, \hat{\Omega}) \to 0$ *as* $k \to \infty$.

(*iii*) *If* $\{(x^{k_j}, y^{k_j}, z^{k_j}, u^{k_j})\}_{j\geq 0}$ *is a subsequence of* $\{(x^k, y^k, z^k, u^k)\}_{k\geq 0}$ *that converges to* (x^*, y^*, z^*, u^*) *as* $k \to +\infty$ *and* $\hat{\omega} \in \hat{\Omega}$, *then*

$$\lim_{k\to+\infty} \hat{L}_\beta(\hat{\omega}^k) = \hat{L}_\beta(x^*, y^*, z^*, u^*, x^*, y^*, z^*) = \inf_k \hat{L}_\beta(\hat{\omega}^k). \tag{3.17}$$

(*iv*) $\hat{\Omega} \subset \operatorname{crit} \hat{L}_\beta(\hat{\omega})$.

(*v*) *The function* $\hat{L}_\beta$ *takes on* $\hat{\Omega}$ *the value*

$$\hat{L}_\beta^* = \lim_{k\to+\infty} \hat{L}_\beta(\hat{\omega}^k) = \lim_{k\to+\infty} \{F(z^k) + G(y^k) + H(x^k, y^k)\}.$$

Proof By the definition of Ω and $\hat{\Omega}$, (i) and (ii) are trivial.

(iii) Let $\{\omega^{k_j}\}$ be a subsequence of $\{\omega^k\}$ such that $\omega^{k_j} \to \omega^*$, $j \to \infty$. Since $L_\beta(\cdot)$ is lower semicontinuous, we have

$$\lim_{j\to\infty} \inf L_\beta(\omega^{k_j}) \geq L_\beta(\omega^*). \tag{3.18}$$

On the other hand, the definition of x^{k+1} shows that

$$\begin{aligned} &F(z^{k+1}) + \langle u^k, Ax^k - z^{k+1}\rangle + \frac{\beta}{2}\|Ax^k - z^{k+1}\|^2 + \tau\|z^{k+1} - z_z^k\|^2 \\ &\quad\leq F(z^*) + \langle u^k, Ax^k - z^*\rangle + \frac{\beta}{2}\|Ax^k - z^*\|^2 + \tau\|z^* - z_z^k\|^2, \end{aligned}$$

from which we get

$$L_\beta(x^k, y^k, z^{k+1}, u^k) + \tau\|z^{k+1} - z_z^k\|^2 - \tau\|z^* - z_z^k\|^2 \leq L_\beta(x^k, y^k, z^*, u^k).$$

Replacing x^k, y^k, z^{k+1}, u^k by $x^{k_j}, y^{k_j}, z^{k_j+1}, u^{k_j}$, we get

$$L_\beta(x^{k_j}, y^{k_j}, z^{k_j+1}, u^{k_j}) + \tau\|z^{k_j+1} - z_z^{k_j}\|^2 - \tau\|z^* - z_z^{k_j}\|^2 \leq L_\beta(x^{k_j}, y^{k_j}, z^*, u^{k_j}).$$

Combining with Theorem 3.1(ii), it follows that

$$\|\omega^{k+1} - \omega^k\| \to 0 \quad \text{as } k \to \infty,$$

and then we have

$$\|\omega^{k_j+1} - \omega^{k_j}\| \to 0 \quad \text{and} \quad \|\omega^{k_j} - \omega^*\| \to 0 \quad \text{as } j \to \infty,$$

which implies that

$$\lim_{j\to\infty}\sup L_\beta\left(x^{k_j},y^{k_j},z^{k_j+1},u^{k_j}\right)\le L_\beta\left(\omega^*\right).$$

From $z^{k+1}-z^k\to 0$ as $k\to\infty$, it is easy to get

$$\lim_{j\to\infty}\sup L_\beta\left(x^{k_j},y^{k_j},z^{k_j+1},u^{k_j}\right)=\lim_{j\to\infty}\sup L_\beta\left(\omega^{k_j}\right).$$

Then, we have

$$\lim_{j\to\infty}\sup L_\beta\left(\omega^{k_j}\right)\le L_\beta\left(\omega^*\right). \tag{3.19}$$

Therefore, from (3.18) and (3.19), it follows that

$$\lim_{j\to+\infty} L_\beta\left(\omega^{k_j}\right)=L_\beta\left(\omega^*\right).$$

By the definition of $\hat{L}_\beta(\hat{\omega}^k)$ and $\|\omega^k-\omega^{k-1}\|\to 0$ as $k\to\infty$, and since the sequence $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\ge 1}$ is convergent, so we have

$$\lim_{k\to+\infty}\hat{L}_\beta\left(\hat{\omega}^k\right)=\hat{L}_\beta\left(x^*,y^*,z^*,u^*,x^*,y^*,z^*\right)=\inf_k\hat{L}_\beta\left(\hat{\omega}^k\right).$$

(iv) For the sequence D^k defined in Lemma 3.4, for any $j\ge 1$, we have $D^{k_j}\in\partial\hat{L}_\beta(\hat{\omega}^{k_j})$. Then it also holds that

$$D^{k_j}\to 0\quad \text{as } j\to\infty,$$

and thus

$$\hat{\omega}^{k_j}\to\hat{\omega}^*\quad\text{and}\quad\hat{L}_\beta\left(\hat{\omega}^{k_j}\right)\to\hat{L}_\beta\left(\hat{\omega}^*\right)\quad\text{as } j\to\infty.$$

The closedness criterion of the limiting subdifferential guarantees that $0\in\partial\hat{L}_\beta(\hat{\omega}^{k_j})$, or, in other words, $\hat{\omega}^*\in\operatorname{crit}(\hat{L}_\beta)$.

(v) Due to Theorem 3.1(ii) and the fact that $\{u_n\}_{n\ge 0}$ is bounded, the sequences $\{\hat{L}_\beta(\hat{\omega}^k)\}_{k\ge 0}$ and $\{F(z^k)+G(y^k)+H(x^k,y^k)\}_{k\ge 0}$ have the same limit:

$$\hat{L}_\beta^*=\lim_{k\to+\infty}\hat{L}_\beta\left(\hat{\omega}^k\right)=\lim_{k\to+\infty}\left\{F\left(z^k\right)+G\left(y^k\right)+H\left(x^k,y^k\right)\right\}.$$

The conclusion follows by taking into consideration the statements (iii) and (iv). The proof is completed. □

Theorem 3.3 (Strong convergence) *Let $v^k=(x^k,y^k,z^k)$. Assume that $\hat{L}_\beta(\hat{\omega}^k)$ is a KŁ function and Assumption* A *is satisfied. Then we have*

(*i*) *The sequence $\{\omega^k\}$ has finite length, namely, $\sum_{k=1}^\infty\|\omega^{k+1}-\omega^k\|<\infty$.*

(*ii*) *The sequence $\{\omega^k\}$ converges to a critical point of $L_\beta(\omega^*)$.*

Proof (i) From the proof of Theorem 3.2, it follows that $\lim_{k\to+\infty}\hat{L}_\beta(\hat{\omega}^k)=\hat{L}_\beta(\hat{\omega}^*)$. We consider two cases.

Case 1. There exists an integer $k_0>0$ such that $\hat{L}_\beta(\hat{\omega}^{k_0})=\hat{L}_\beta(\hat{\omega}^*)$.

Since $\{\hat{L}_\beta(\hat{\omega}^k)\}$ is decreasing, we know that for all $k>k_0$,

$$\eta_1\|x^k-x^{k-1}\|^2+\eta_2\|y^k-y^{k-1}\|^2+\eta_3\|z^k-z^{k-1}\|^2\le\hat{L}_\beta(\hat{\omega}^*)-\hat{L}_\beta(\hat{\omega}^*)=0,$$

which implies that $x^{k+1}=x^k$, $y^{k+1}=y^k$, $z^{k+1}=z^k$ for $\forall k>k_0$. Then, from (3.6), we get $u^{k+1}=u^k$, for $\forall k>k_0+1$. Thus, $\omega^{k+1}=\omega^k$, the result is obtained.

Case 2. One has $\hat{L}_\beta(\hat{\omega}^k)>\hat{L}_\beta(\hat{\omega}^*)$ for $\forall k>0$.

Since $\operatorname{dist}(\hat{\omega}^k,\hat{\Omega})\to 0$, for $\forall\varepsilon_1>0$ there exists $k_1>0$ such that, for $\forall k>k_1$, $\operatorname{dist}(\hat{\omega}^k,\hat{\Omega})<\varepsilon_1$. Due to $\lim_{k\to+\infty}\hat{L}_\beta(\hat{\omega}^k)=\hat{L}_\beta(\hat{\omega}^*)$, for $\forall\varepsilon_2>0$ there exists $k_2>0$ such that $\hat{L}_\beta(\hat{\omega}^k)<\hat{L}_\beta(\hat{\omega}^*)+\varepsilon_2$, for $\forall k>k_2$. Therefore, for $\forall\varepsilon_1,\varepsilon_2>0$, when $k>\widetilde{k}=\max\{k_1,k_2\}$, we have $\operatorname{dist}(\hat{\omega}^k,\hat{\Omega})<\varepsilon_1$, $\hat{L}_\beta(\hat{\omega}^*)<\hat{L}_\beta(\hat{\omega}^k)<\hat{L}_\beta(\hat{\omega}^*)+\varepsilon_2$. Since $\{\omega^k\}$ is bounded, by Theorem 3.2, we know that $\hat{\Omega}$ is a nonempty compact set and $\hat{L}_\beta(\cdot)$ is constant on $\hat{\Omega}$. Applying Lemma 2.4, we deduce that, for $\forall k>\tilde{k}$,

$$\varphi'\big(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*)\big)\operatorname{dist}\big(0,\partial\hat{L}_\beta(\hat{\omega}^k)\big)\ge 1.$$

Since $\varphi'(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*))>0$, then

$$\frac{1}{\varphi'(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*))}\le\operatorname{dist}\big(0,\partial\hat{L}_\beta(\hat{\omega}^k)\big).$$

Making use of the concavity of φ, we get that

$$\begin{aligned}&\varphi\big(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*)\big)-\varphi\big(\hat{L}_\beta(\hat{\omega}^{k+1})-\hat{L}_\beta(\hat{\omega}^*)\big)\\&\quad\ge\varphi'\big(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*)\big)\big(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^{k+1})\big).\end{aligned}$$

Combining with the KŁ property, it follows that

$$\begin{aligned}&\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^{k+1})\\&\quad\le\frac{\varphi(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*))-\varphi(\hat{L}_\beta(\hat{\omega}^{k+1})-\hat{L}_\beta(\hat{\omega}^*))}{\varphi'(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*))}\\&\quad\le\operatorname{dist}\big(0,\partial\hat{L}_\beta(\hat{\omega}^k)\big)\big(\varphi\big(\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^*)\big)-\varphi\big(\hat{L}_\beta(\hat{\omega}^{k+1})-\hat{L}_\beta(\hat{\omega}^*)\big)\big).\end{aligned}\tag{3.20}$$

By Lemma 3.4, we get

$$\operatorname{dist}\big(0,\partial\hat{L}_\beta(\hat{\omega}^k)\big)\le\zeta\big(\|v^k-v^{k-1}\|+\|v^{k-1}-v^{k-2}\|\big).\tag{3.21}$$

From Lemma 3.2, we have

$$\begin{aligned}\hat{L}_\beta(\hat{\omega}^k)-\hat{L}_\beta(\hat{\omega}^{k+1})&\ge\eta_1\|x^k-x^{k-1}\|^2+\eta_2\|y^k-y^{k-1}\|^2+\eta_3\|z^k-z^{k-1}\|^2\\&\ge\eta\|v^{k+1}-v^k\|^2,\end{aligned}\tag{3.22}$$

where $\eta = \min\{\eta_1, \eta_2, \eta_3\}$. Putting (3.21) and (3.22) into (3.20), we obtain

$$\begin{aligned}\eta\|v^{k+1} - v^k\|^2 &\leq \sqrt{\zeta}\left(\|v^k - v^{k-1}\| + \|v^{k-1} - v^{k-2}\|\right) \\ &\quad \times \left(\varphi\left(\hat{L}_\beta(\hat{\omega}^k) - \hat{L}_\beta(\hat{\omega}^*)\right) - \varphi\left(\hat{L}_\beta(\hat{\omega}^{k+1}) - \hat{L}_\beta(\hat{\omega}^*)\right)\right). \end{aligned} \tag{3.23}$$

Set $b_k = \frac{\sqrt{\zeta}}{\eta}(\varphi(\hat{L}_\beta(\hat{\omega}^k) - \hat{L}_\beta(\hat{\omega}^*)) - \varphi(\hat{L}_\beta(\hat{\omega}^{k+1}) - \hat{L}_\beta(\hat{\omega}^*))) \geq 0$, $a_k = \|v^k - v^{k-1}\| \geq 0$. Then (3.23) can be equivalently rewritten as

$$a_{k+1}^2 \leq b_k(a_k + a_{k-1}). \tag{3.24}$$

Since $\varphi \geq 0$, we know that

$$\sum_{k=1}^{\infty} b_k \leq \frac{\sqrt{\zeta}}{\eta}\varphi\left(\hat{L}_\beta(\hat{\omega}^1) - \hat{L}_\beta(\hat{\omega}^*)\right),$$

hence $\sum_{k=1}^{\infty} b_k < \infty$. Note that from (3.24) we have

$$a_{k+1} \leq \sqrt{b_k(a_k + a_{k-1})} \leq \frac{1}{4}(a_k + a_{k-1}) + b_k.$$

So Lemma 2.4 gives that $\sum_{k=1}^{\infty} a_k < \infty$. Then,

$$\sum_{k=1}^{\infty}\|x^k - x^{k-1}\| < \infty, \qquad \sum_{k=1}^{\infty}\|y^k - y^{k-1}\| < \infty, \qquad \sum_{k=1}^{\infty}\|z^k - z^{k-1}\| < \infty.$$

Combining it with (3.8), we get

$$\sum_{k=1}^{\infty}\|u^k - u^{k-1}\| < \infty.$$

(ii) Statement (i) indicates that $\{\omega^k\}$ is a Cauchy sequence. So $\{\omega^k\}$ is convergent. Let $\omega^k \to \omega^*$, $k \to \infty$. According to Theorem 3.2(iv), it is clear that $\hat{\omega}^* \in \hat{\Omega} \subset \operatorname{crit}\hat{L}_\beta(\hat{\omega})$. Thus $\hat{\omega}^*$ is a critical point of $\hat{L}_\beta(\hat{\omega})$. Therefore, by the definition of $\hat{L}_\beta$, $\{\omega^k\}$ converges to a critical point of $L_\beta(\omega^*)$. The proof is completed. □

4 Numerical experiments

In this section, we illustrate two computational instances to contrast the efficacy of our methodology with the PMA technique detailed in [5]. The computational trials are executed on 64-bit MATLAB R2019b installed on a 64-bit computer equipped with an Intel(R) Core(TM) i7-6700HQ CPU operating at 2.6 GHz and possessing 32 GB of RAM.

Example 4.1 We consider the following optimization problem:

$$\min_{x,y} \frac{1}{2}\|Ax - b\|^2 + c_1\|y\|_{\frac{1}{2}}^{\frac{1}{2}} + \frac{c_2}{2}\|Bx - y\|^2,$$

which can be written as

$$\min_{x,y,z} \frac{1}{2}\|z-b\|^2 + c_1\|y\|_{\frac{1}{2}}^{\frac{1}{2}} + \frac{c_2}{2}\|Bx-y\|^2$$

$$\text{such that} \quad Ax = z.$$

Select random matrices $A = (a_{ij})_{p\times m}$ and $B = (b_{ij})_{q\times m}$, where $a_{ij}, b_{ij} \in (0,1)$. Let m, p, q be three positive integers with $m = q$. Take the initial points, $x^0 = x^{-1} = \text{zeros}(m,1)$, $y^0 = y^{-1} = \text{zeros}(q,1)$, $z^0 = z^{-1} = \text{zeros}(p,1)$, $u^0 = \text{zeros}(p,1)$ for Algorithm 3.1. The parameters are set as $l_F = 1$, $\tau = 10$, $\beta = 67$, $\alpha = 6.6\times 10^7$, $c_1 = c_2 = 1$. The initial points for PMA in [5] are also set as the previous x^0, y^0, z^0, u^0, and the parameter is taken as $\sigma = 0.1$. Define $\|Ax-z\|^2$ as the error, and select $\|Ax-z\|^2 < 10^{-4}$ as the stopping criterion. The results are presented in Table 1 for clarity and, to provide a clear evaluation of the algorithm's performance, we also depict the error curve. The respective outcomes are illustrated in Figs. 1 and 2. In the table, k denotes the number of iterations, s denotes the computing time.

Considering Table 1 and Fig. 1, we observe that the inclusion of an inertial factor positively impacts the convergence of Algorithm 3.1. Furthermore, a comparison between Table 1 and Fig. 2 suggests that our algorithm requires fewer iterations and achieves convergence at a faster rate compared to the PMA. In summary, empirical evidence indicates

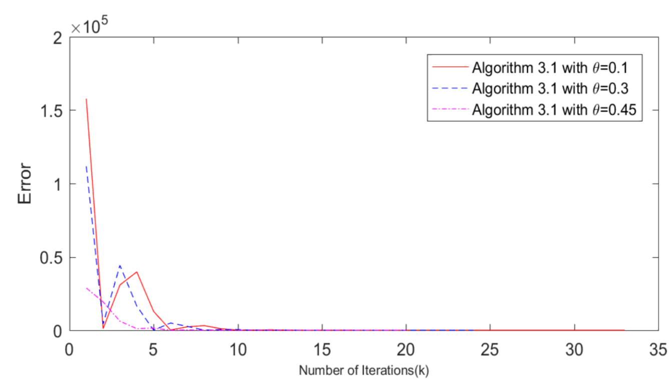

Figure 1 The performance of Algorithm 3.1 with $m = q = 100$ and $p = 300$ with different inertial factor

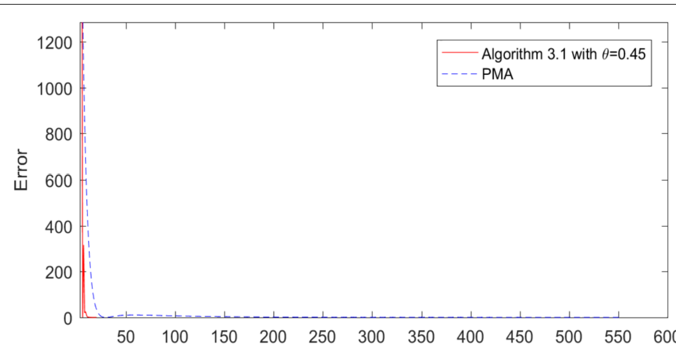

Figure 2 A comparison of Algorithm 3.1 and PMA with $m = q = 100$ and $p = 400$

Table 1 Numerical results of two algorithms for Example 4.1 under different inertial values and dimensions

	$m = q = 100; p = 200$	$m = q = 100; p = 300$	$m = q = 100; p = 500$
Algorithm 3.1 $\theta = 0.1$	$k = 31; s = 0.4375$	$k = 33; s = 0.5781$	$k = 34; s = 1.6719$
Algorithm 3.1 $\theta = 0.3$	$k = 24; s = 0.3906$	$k = 24; s = 0.5156$	$k = 33; s = 1.4844$
Algorithm 3.1 $\theta = 0.45$	$k = 20; s = 0.4531$	$k = 20; s = 0.4844$	$k = 21; s = 1.0156$
PMA	$k = 521; s = 1.1719$	$k = 528; s = 1.5625$	$k = 558; s = 4.6563$

that our algorithm, which incorporates an inertial approach, outperforms the PMA as reported in [5].

Example 4.2 In the second example, we consider the SCAD-l_2, which takes the form of

$$\min \sum_{i=1}^{n} f_k(|z_i|) + \|y\|^2 + \|Mx - My - c\|^2$$
$$\text{such that} \quad Ax = z,$$

where $A \in R^{m\times n}$, $c \in R^m$, and $f_k(|z_i|)$ is defined by

$$f_k(t) = \begin{cases} kt, & t \le k, \\ \frac{-t^2+2akt-k^2}{2(a+1)}, & k < t \le ak, \\ \frac{(a+1)k^2}{2}, & t > ak, \end{cases}$$

with $a > 2$ and $k > 0$ being the knots of a quadratic spline function. We select random $m \times n$ matrices $A, D \sim N(0, 1)$, and all columns are normalized. We select random sparse vectors in R^m with the density 0.01 as x^*, y^* and the vector $c = Mx^* - My^* + Q$ with the noise vector $Q \sim N(0, 10^{-3}I)$. For the sole purpose of showing the numerical efficiency, we fix the parameters $k = 3$, $a = 4$ as constants for $f_k(|z_i|)$. In addition, we set $l_F = 3$, $\tau = 10$, $\beta = 36.43$, $\alpha = 1.99 \times 10^4$ in Algorithm 3.1, and select $\sigma = 9.17 \times 10^{-4}$, $\beta = 1.67 \times 10^3$, $\tau = 4.33 \times 10^4$, $\mu = 276$ in PMA [5]. The initial points are selected as $x^0 = x^{-1} = \text{zeros}(m, 1)$, $y^0 = y^{-1} = \text{zeros}(q, 1)$, $z^0 = z^{-1} = \text{rand}(p, 1)$, $u = \text{ones}(p, 1)$ in Algorithm 3.1, and $x^0 = \text{zeros}(m, 1)$, $y^0 = \text{zeros}(q, 1)$, $z^0 = \text{rand}(p, 1)$, $u = \text{ones}(p, 1)$ in PMA [5]. The stopping criterion is taken as $\text{Error} = \|Ax - z\|^2 < 10^{-4}$.

Figures 3 and 4 show the results of evolution of the Error with respect to iterations when we run Algorithm 3.1 and PMA in [5]. Figure 4 shows that the Error of Algorithm 3.1 decreases faster than that of PMA. One can see that for larger values of θ, Algorithm 3.1 has a smaller error value in Table 2 and Fig. 3.

5 Conclusion

This paper presents a dual-relaxed inertial proximal minimization algorithm designed for addressing a specific category of structured nonconvex and nonsmooth optimization problems. The objective function in these problems is characterized by being the sum of a composite function, a nonsmooth function, and a mixed function. The algorithm introduced herein features an update mechanism for each subproblem that incorporates

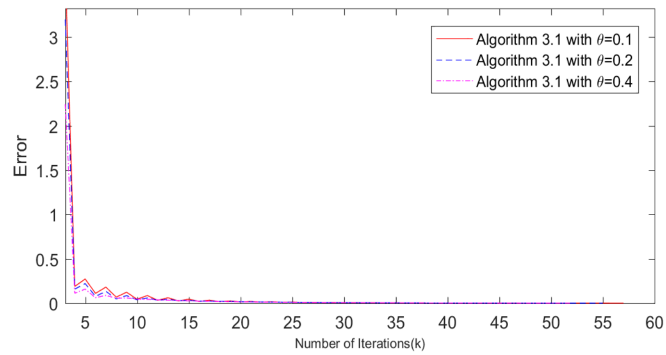

Figure 3 The performance of our algorithm with $m = q = 1200$, $p = 1100$

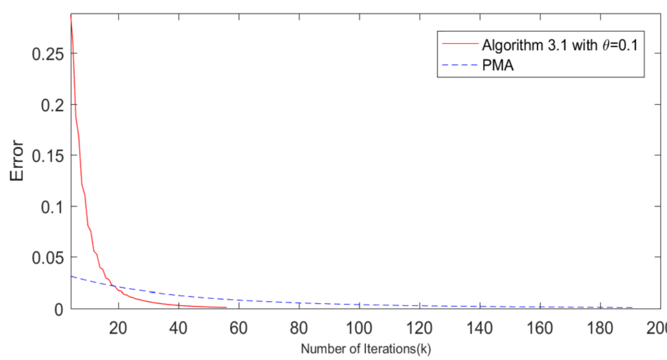

Figure 4 The performance of Algorithms 3.1 and PMA with $m = q = 1200$, $p = 1100$ and $\theta = 0.1$

Table 2 Numerical results of two algorithms for Example 4.2 under different inertial values and dimensions

	$m = q = 800; p = 1100$	$m = q = 900; p = 1100$	$m = q = 1200; p = 1100$
Algorithm 3.1 $\theta = 0.1$	$k = 53; s = 10.1563$	$k = 55; s = 13.6875$	$k = 57; s = 38.7969$
Algorithm 3.1 $\theta = 0.2$	$k = 51; s = 10.2031$	$k = 54; s = 13.9375$	$k = 55; s = 30.9219$
Algorithm 3.1 $\theta = 0.4$	$k = 48; s = 10.2188$	$k = 50; s = 12.6406$	$k = 52; s = 30.5313$
PMA	$k = 219; s = 7.7969$	$k = 204; s = 10.2031$	$k = 195; s = 13.4531$

inertial effects and employs two relaxed terms during the dual update phase. Additionally, the parameters within our algorithm are determined using a straightforward approach. Computational experiments demonstrate that our algorithm is both practical and effective.

Author contributions

Yang Liu wrote the main manuscript text and Yazheng Dang constructed the algorithm and proved the convergence. All authors reviewed the manuscript.

Funding

The National Natural Science Foundation of China (72071130; 71901145); National Key R and D Program Key Special Projects (2023YFC3306103, 2023YFC3306105); Shanghai Philosophy and Social Sciences Research Project (2023EFX011); Youth Fund for Humanities and Social Sciences of the Ministry of Education (20YJC820030); The Major Project of Chinese Society of Criminology (FZXXH2022A02).

Data Availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Ethics approval and consent to participate not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Information Science and Technology, East China University of Political Science and Law, Shanghai, PRC 200093, China. [2]School of Management, University of Shanghai for Science and Technology, Jungong Road 516, Yangpu District, Shanghai, PRC 200093, China.

Received: 7 April 2024 Accepted: 27 August 2024 Published online: 05 September 2024

References

1. Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka–Łojasiewicz inequality. Math. Oper. Res. **35**(2), 438–457 (2010)
2. Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. Math. Program. **137**, 91–129 (2013)
3. Auslender, A.: Méthodes numériques pour la décomposition et la minimisation de fonctions nondifferentiates. Numer. Math. **18**(3), 213–223 (1971)
4. Bolte, J., Sabach, S., Teboulle, M.: Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Math. Program. **146**, 459–494 (2014)
5. Bot, R.I., Csetnek, E.R., Nguyen, D.K.: A proximal minimization algorithm for structured nonconvex and nonsmooth problems. SIAM J. Optim. **29**(2), 1300–1328 (2019)
6. Bot, R.I., Nguyen, D.K.: The proximal alternating direction method of multipliers in the nonconvex setting: convergence analysis and rates. Math. Oper. Res. **45**(2), 682–712 (2020)
7. Chao, M.T., Zhang, Y., Jian, J.B.: An inertial proximal alternating direction method of multipliers for nonconvex optimization. Int. J. Comput. Math. **98**(6), 1199–1217 (2021)
8. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward–backward splitting. Multiscale Model. Simul. **4**(4), 1168–1200 (2005)
9. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
10. Driggs, D., Tang, J., Liang, J.: SPRING: a fast stochastic proximal alternating method for non-smooth non-convex optimization (2020). arXiv:2002.12266
11. Glowinski, R., Marrocco, A.: Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. ESAIM: Math. Model. Numer. Anal. **9**(2), 41–76 (1975)
12. Hien, L.T.K., Phan, D.N., Gillis, N.: Inertial alternating direction method of multipliers for non-convex non-smooth optimization. Comput. Optim. Appl. **83**, 247–285 (2022)
13. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(6), 964–979 (1979)
14. Liu, Q., Shen, X., Gu, Y.: Linearized ADMM for non-convex non-smooth optimization with convergence analysis. IEEE Access **7**, 76131–76144 (2017)
15. Miao, L., Tang, Y., Wang, C.: A parameterized three-operator splitting algorithm for non-convex minimization problems with applications. J. Nonlinear Var. Anal. **8**, 451–471 (2024)
16. Ochs, P., Brox, T., Pock, T.: iPiasco: inertial proximal algorithm for strongly convex optimization. J. Math. Imaging Vis. **53**(2), 171–181 (2015)
17. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, San Diego (1970)
18. Ouyang, Y., Chen, Y., Lan, G., Pasiliao, E.: An accelerated linearized alternating direction method of multipliers. SIAM J. Imaging Sci. **8**(1), 644–681 (2015)
19. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D, Nonlinear Phenom. **60**(1–4), 259–268 (1992)
20. Shi, S., Fu, Z., Wu, Q.: On the average operators, oscillatory integrals, singulars, singular integrals and their applications. J. Appl. Anal. Comput. **14**, 334–378 (2024)
21. Shi, S., Zhang, L.: Dual characterization of fractional capacity via solution of fractional p-Laplace equation. Math. Nachr. **293**, 2233–2247 (2020)
22. Tseng, P.: Convergence of a block coordinate descent method for nondifferentiable minimization. J. Optim. Theory Appl. **109**(3), 475–494 (2001)
23. Wang, X.Q., Shao, H., Liu, P.J., Wu, T.: An inertial proximal partially symmetric ADMM-based algorithm for linearly constrained multi-block nonconvex optimization problems with applications. J. Comput. Appl. Math. **420** (2022)
24. Yang, J., Yuan, X.: Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization. Math. Comput. **82**(281), 301–329 (2013)
25. Yu, Y., Yin, T.C.: Strong convergence theorems for a nonmonotone equilibrium problem and a quasi-variational inclusion problem. J. Nonlinear Convex Anal. **25**, 503–512 (2024)

Publisher's Note