Jian et al. *Journal of Inequalities and Applications* (2019) 2019:193
https://doi.org/10.1186/s13660-019-2145-0

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

A regularized alternating direction method of multipliers for a class of nonconvex problems

Jin Bao Jian[1,2], Ye Zhang[1] and Mian Tao Chao[1*]

*Correspondence:
chaomiantao@126.com
[1]College of Mathematics and Information Science, Guangxi University, Nanning, China
Full list of author information is available at the end of the article

Abstract

In this paper, we propose a regularized alternating direction method of multipliers (RADMM) for a class of nonconvex optimization problems. The algorithm does not require the regular term to be strictly convex. Firstly, we prove the global convergence of the algorithm. Secondly, under the condition that the augmented Lagrangian function satisfies the Kurdyka–Łojasiewicz property, the strong convergence of the algorithm is established. Finally, some preliminary numerical results are reported to support the efficiency of the proposed algorithm.

Keywords: Nonconvex optimization problems; Alternating direction method of multipliers; Kurdyka–Łojasiewicz property; Convergence

1 Introduction

In this paper, we consider the following nonconvex optimization problem

$$\min \quad f(x) + g(Ax), \tag{1}$$

where $f : R^n \to R \cup \{+\infty\}$ is a proper lower semicontinuous function, $g : R^m \to R$ is a continuous differentiable function with ∇g Lipschitz continuous and modulus $L > 0$, while $A \in R^{m\times n}$ is a given matrix. When the functions f and g are convex, the problem (1) can be transformed into the split feasibility problem [1, 2]. Problem (1) is equivalent to the following constraint optimization problem:

$$\begin{aligned} \min \quad & f(x) + g(y), \\ \text{s.t. } & Ax - y = 0. \end{aligned} \tag{2}$$

The augmented Lagrangian function of (2) is defined as follows:

$$\mathcal{L}_\beta(x,y,\lambda) = f(x) + g(y) - \langle \lambda, Ax - y\rangle + \frac{\beta}{2}\|Ax - y\|^2, \tag{3}$$

where $\lambda \in R^m$ is the Lagrangian parameter and $\beta > 0$ is the penalty parameter.

The alternating direction method of multipliers (ADMM) was first proposed by Gabay and Mercier in 1970s, which is an effective algorithm for solving the two-block convex problems [3]. The iterative scheme of the classic ADMM for problem (2) is as follows:

$$\begin{cases} x^{k+1} \in \arg\min\{\mathcal{L}_\beta(x, y^k, \lambda^k)\}, \\ y^{k+1} \in \arg\min\{\mathcal{L}_\beta(x^{k+1}, y, \lambda^k)\}, \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} - y^{k+1}). \end{cases} \tag{4}$$

If f, g are convex functions, then the convergence of ADMM is well-understood and there are some recent convergence rate analysis results [4–8]. However, when the objective function is nonconvex, ADMM does not necessarily converge. Recently, some scholars have proposed various improved ADMM for nonconvex problems, and analyzed their convergence [9–15]. In particular, Guo et al. [16, 17] analyzed the strong convergence of classical ADMM for the nonconvex optimization problem of (2). Wang et al. [12, 14] studied the convergence of the Bregman ADMM for the nonconvex optimization problems, where they need the augmented Lagrangian function with respect to x or the Bregman distance in the x-subproblem to be strongly convex.

The first formula of (4) has the following structure:

$$\min\left\{f(x) + g(y^k) - \langle\lambda^k, Ax - y^k\rangle + \frac{\beta}{2}\|Ax - y^k\|^2\right\}. \tag{5}$$

When A is not the identity matrix, the above problem may not be easy. Regularization is a popular technique to simplify the optimization problems [12, 14, 18]. For example, the regular term $\frac{1}{2}\|x - x^k\|_G^2$ could be added to the above problem (5), where G is a symmetric semidefinite matrix. Specifically, when $G = \alpha I - \beta A^\top A$, problem (5) is converted into the following form:

$$\min\left\{f(x) + \frac{\alpha}{2}\|x - b^k\|^2\right\}, \tag{6}$$

with a certain known $b^k \in R^n$. Since the first formula of (4) has the form of (6) with $\alpha = 1$, this paper considers the following regularized ADMM (in short, RADMM) for problem (2):

$$\begin{cases} x^{k+1} \in \arg\min\{\mathcal{L}_\beta(x, y^k, \lambda^k) + \frac{1}{2}\|x - x^k\|_G^2\}, \\ y^{k+1} \in \arg\min\{\mathcal{L}_\beta(x^{k+1}, y, \lambda^k)\}, \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} - y^{k+1}), \end{cases} \tag{7}$$

where G is a symmetric semidefinite matrix, $\|x\|_G^2 := x^\top Gx$.

The framework of this paper is as follows. In Sect. 2, we present some preliminary materials that will be used in this paper. In Sect. 3, we prove the convergence of algorithm (7). In Sect. 4, we report some numerical results. In Sect. 5, we draw some conclusions.

2 Preliminaries

For a vector $x = (x_1, x_2, \dots, x_n)^\top \in R^n$, we let $\|x\| = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$, $\|x\|_1 = \sum_{i=1}^n |x_i|$, and $\|x\|_{\frac{1}{2}} = (\sum_{i=1}^n |x_i|^{\frac{1}{2}})^2$. Also $G \succeq 0$ ($\succ 0$) denotes that G is a positive semidefinite (positive definite) matrix. For a subset $S \subseteq R^n$ and a point $x \in R^n$, if S is nonempty, let $d(x,S) = \inf_{y\in S} \|y-x\|$. When $S = \emptyset$, we set $d(x,S) = +\infty$ for all x. A function $f : R^n \to (-\infty, +\infty]$ is said to be proper, if there exists at least one $x \in R^n$ such that $f(x) < +\infty$. The effective domain of f is defined through $\operatorname{dom} f = \{x \in R^n | f(x) < +\infty\}$.

Definition 2.1 ([19]) The function $f : R^n \to R \cup \{+\infty\}$ is lower semicontinuous at $\bar{x}$, if $f(\bar{x}) \leq \liminf_{x\to\bar{x}} f(x)$. If f is lower semicontinuous at every point $x \in R^n$, then f is called a lower semicontinuous function.

Definition 2.2 ([19]) Let $f : R^n \to R \cup \{+\infty\}$ be a proper lower semicontinuous function.

(i) The Fréchet subdifferential, or regular subdifferential, of f at $x \in \operatorname{dom} f$, denoted by $\hat{\partial} f(x)$, is the set of all elements $u \in R^n$ which satisfy

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

when $x \notin \operatorname{dom} f$, let $\hat{\partial} f(x) = \emptyset$;

(ii) The limiting-subdifferential, or simply the subdifferential, of f at $x \in \operatorname{dom} f$, denoted by $\partial f(x)$, is defined as

$$\partial f(x) = \left\{ u \in R^n \mid \exists x^k \to x, f\left(x^k\right) \to f(x), u^k \in \hat{\partial} f\left(x^k\right) \to u, k \to \infty \right\}.$$

Proposition 2.1 ([20]) *Let $f : R^n \to R \cup \{+\infty\}$ be a proper lower semicontinuous function, then*

(i) *$\hat{\partial} f(x) \subseteq \partial f(x)$ for each $x \in R^n$, where the first set is closed and convex while the second one is only closed;*

(ii) *Let $u^k \in \partial f(x^k)$ and $\lim_{k\to\infty}(x^k, u^k) = (x, u)$, then $u \in \partial f(x)$;*

(iii) *A necessary condition for $x \in R^n$ to be a minimizer of f is*

$$0 \in \partial f(x); \tag{8}$$

(iv) *If $g : R^n \to R$ is continuously differentiable, then $\partial(f+g)(x) = \partial f(x) + \nabla g(x)$ for any $x \in \operatorname{dom} f$.*

A point that satisfies (8) is called a critical point or a stationary point. The set of critical points of f is denoted by $\operatorname{crit} f$.

Lemma 2.1 ([21]) *Suppose that $H(x,y) = f(x) + g(y)$, where $f : R^n \to R \cup \{+\infty\}$ and $g : R^m \to R$ are proper lower semicontinuous functions, then*

$$\partial H(x,y) = \partial_x H(x,y) \times \partial_y H(x,y) = \partial f(x) \times \partial g(y),$$

for all $(x,y) \in \operatorname{dom} H = \operatorname{dom} f \times \operatorname{dom} g$.

The following lemma is very important for the convergence analysis.

Lemma 2.2 ([22]) *Let function $h : R^n \to R$ be continuously differentiable and its gradient ∇h be Lipschitz continuous with modulus $L > 0$, then*

$$\left|h(y) - h(x) - \langle \nabla h(x), y - x\rangle\right| \le \frac{L}{2}\|y - x\|^2, \quad \textit{for all } x, y \in R^n.$$

Definition 2.3 We say that (x^*, y^*, λ^*) is a critical point of the augmented Lagrangian function $\mathcal{L}_\beta(\cdot)$ of (2) if it satisfies

$$\begin{cases} A^\top \lambda^* \in \partial f(x^*), \\ \nabla g(y^*) = \lambda^*, \\ Ax^* - y^* = 0. \end{cases} \tag{9}$$

Obviously, (2) is equivalent to $0 \in \partial \mathcal{L}_\beta(x^*, y^*, \lambda^*)$.

Definition 2.4 ([21] (Kurdyka–Łojasiewicz property)) Let $f : R^n \to R \cup \{+\infty\}$ be a proper lower semicontinuous function. If there exist $\eta \in (0, +\infty]$, a neighborhood U of $\hat{x}$, and a concave function $\varphi : [0, \eta) \to R_+$ satisfying the following conditions:

(i) $\varphi(0) = 0$;
(ii) φ is continuously differentiable on $(0, \eta)$ and continuous at 0;
(iii) $\varphi'(s) > 0$ for all $s \in (0, \eta)$;
(iv) $\varphi'(f(x) - f(\hat{x}))d(0, \partial f(x)) \ge 1$, for all $x \in U \cap [f(\hat{x}) < f(x) < f(\hat{x}) + \eta]$,

then f is said to have the Kurdyka–Łojasiewicz (KL) property at $\hat{x}$.

Lemma 2.3 ([23] (Uniform KL property)) *Let Φ_η be the set of concave functions which satisfy* (i), (ii) *and* (iii) *in Definition* 2.4. *Suppose that $f : R^n \to R \cup \{+\infty\}$ is a proper lower semicontinuous function and Ω is a compact set. If $f(x) \equiv a$ for all $x \in \Omega$ and f satisfies the KL property at each point of Ω, then there exist $\varepsilon > 0$, $\eta > 0$, and $\varphi \in \Phi_\eta$ such that*

$$\varphi'\bigl(f(x) - a\bigr)d\bigl(0, \partial f(x)\bigr) \ge 1,$$

for all $x \in \{x \in R^n | d(x, \Omega) < \varepsilon\} \cap [x : a < f(x) < a + \eta]$.

3 Convergence analysis

In this section, we prove the convergence of algorithm (7). Throughout this section, we assume that the sequence $\{z^k := (x^k, y^k, \lambda^k)\}$ is generated by RADMM (7). Firstly, the global convergence of the algorithm is established by the monotonically nonincreasing sequence $\{\mathcal{L}_\beta(z^k)\}$. Secondly, the strong convergence of the algorithm is proved under the condition that $\mathcal{L}_\beta(\cdot)$ satisfies the KL property. From optimality conditions of each subproblem in (7), we have

$$\begin{cases} 0 \in \partial f(x^{k+1}) - A^\top \lambda^k + \beta A^\top (Ax^{k+1} - y^k) + G(x^{k+1} - x^k), \\ 0 = \nabla g(y^{k+1}) + \lambda^k - \beta(Ax^{k+1} - y^{k+1}), \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} - y^{k+1}). \end{cases} \tag{10}$$

That is,

$$\begin{cases} A^\top\lambda^{k+1}-\beta A^\top(y^{k+1}-y^k)-G(x^{k+1}-x^k)\in\partial f(x^{k+1}), \\ -\lambda^{k+1}=\nabla g(y^{k+1}), \\ \lambda^{k+1}=\lambda^k-\beta(Ax^{k+1}-y^{k+1}). \end{cases} \tag{11}$$

We need the following basic assumptions on problem (2).

Assumption 3.1

(i) $f:R^n\to R\cup\{+\infty\}$ is proper lower semicontinuous;

(ii) $g:R^m\to R$ is continuously differentiable, with $\|\nabla g(u)-\nabla g(v)\|\le L\|u-v\|$, $\forall u,v\in R^n$;

(iii) $\beta>2L$ and $\delta:=\frac{\beta-L}{2}-\frac{L^2}{\beta}>0$;

(iv) $G\succeq 0$ and $G+A^\top A\succ 0$.

The following lemma implies that sequence $\{\mathcal{L}_\beta(z^k)\}$ is monotonically nonincreasing.

Lemma 3.1

$$\mathcal{L}_\beta(z^{k+1})\le\mathcal{L}_\beta(z^k)-\delta\|y^k-y^{k+1}\|^2-\frac{1}{2}\|x^{k+1}-x^k\|_G^2. \tag{12}$$

Proof From the definition of the augmented Lagrangian function $\mathcal{L}_\beta(\cdot)$ and the third formula of (11), we have

$$\begin{aligned}\mathcal{L}_\beta(x^{k+1},y^{k+1},\lambda^{k+1})&=\mathcal{L}_\beta(x^{k+1},y^{k+1},\lambda^k)+\langle\lambda^k-\lambda^{k+1},Ax^{k+1}-y^{k+1}\rangle\\&=\mathcal{L}_\beta(x^{k+1},y^{k+1},\lambda^k)+\frac{1}{\beta}\|\lambda^k-\lambda^{k+1}\|^2\end{aligned} \tag{13}$$

and

$$\begin{aligned}&\mathcal{L}_\beta(x^{k+1},y^{k+1},\lambda^k)-\mathcal{L}_\beta(x^{k+1},y^k,\lambda^k)\\&\quad=g(y^{k+1})-g(y^k)-\langle\lambda^k,y^k-y^{k+1}\rangle-\frac{\beta}{2}\|Ax^{k+1}-y^k\|^2\\&\qquad+\frac{\beta}{2}\|Ax^{k+1}-y^{k+1}\|^2.\end{aligned} \tag{14}$$

From Assumption 3.1(ii), Lemma 2.2 and (11), we have

$$g(y^{k+1})-g(y^k)\le\langle-\lambda^{k+1},y^{k+1}-y^k\rangle+\frac{L}{2}\|y^k-y^{k+1}\|^2. \tag{15}$$

Inserting (15) into (14) yields

$$\begin{aligned}&\mathcal{L}_\beta(x^{k+1},y^{k+1},\lambda^k)-\mathcal{L}_\beta(x^{k+1},y^k,\lambda^k)\\&\quad\le\langle\lambda^k-\lambda^{k+1},y^{k+1}-y^k\rangle-\frac{\beta}{2}\|Ax^{k+1}-y^k\|^2\\&\qquad+\frac{\beta}{2}\|Ax^{k+1}-y^{k+1}\|^2+\frac{L}{2}\|y^k-y^{k+1}\|^2.\end{aligned} \tag{16}$$

Since $\lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} - y^{k+1})$, we have

$$Ax^{k+1} - y^{k+1} = \frac{1}{\beta}\left(\lambda^k - \lambda^{k+1}\right) \tag{17}$$

and

$$Ax^{k+1} - y^k = \frac{1}{\beta}\left(\lambda^k - \lambda^{k+1}\right) - \left(y^k - y^{k+1}\right).$$

It follows that

$$\begin{aligned} &\left\langle \lambda^k - \lambda^{k+1}, y^{k+1} - y^k\right\rangle - \frac{\beta}{2}\left\|Ax^{k+1} - y^k\right\|^2 \\ &\quad = -\frac{\beta}{2}\left\|y^{k+1} - y^k\right\|^2 - \frac{1}{2\beta}\left\|\lambda^k - \lambda^{k+1}\right\|^2. \end{aligned} \tag{18}$$

Combining (16), (17) and (18), we have

$$\mathcal{L}_\beta\left(x^{k+1}, y^{k+1}, \lambda^k\right) - \mathcal{L}_\beta\left(x^{k+1}, y^k, \lambda^k\right) \leq -\frac{\beta - L}{2}\left\|y^{k+1} - y^k\right\|^2. \tag{19}$$

From $-\lambda^{k+1} = \nabla g(y^{k+1})$ and Assumption 3.1(ii), we have

$$\frac{1}{\beta}\left\|\lambda^k - \lambda^{k+1}\right\|^2 \leq \frac{L^2}{\beta}\left\|y^{k+1} - y^k\right\|^2. \tag{20}$$

Adding (13), (19) and (20), one has

$$\mathcal{L}_\beta\left(x^{k+1}, y^{k+1}, \lambda^{k+1}\right) \leq \mathcal{L}_\beta\left(x^{k+1}, y^k, \lambda^k\right) - \left(\frac{\beta - L}{2} - \frac{L^2}{\beta}\right)\left\|y^{k+1} - y^k\right\|^2.$$

Since x^{x+1} is the optimal solution of the first subproblem of (7), one has

$$\mathcal{L}_\beta\left(x^{k+1}, y^k, \lambda^k\right) \leq \mathcal{L}_\beta\left(x^k, y^k, \lambda^k\right) - \frac{1}{2}\left\|x^{k+1} - x^k\right\|_G^2.$$

Thus

$$\begin{aligned} &\mathcal{L}_\beta\left(x^{k+1}, y^{k+1}, \lambda^{k+1}\right) \\ &\quad \leq \mathcal{L}_\beta\left(x^k, y^k, \lambda^k\right) - \left(\frac{\beta - L}{2} - \frac{L^2}{\beta}\right)\left\|y^k - y^{k+1}\right\|^2 - \frac{1}{2}\left\|x^{k+1} - x^k\right\|_G^2. \end{aligned}$$ □

Lemma 3.2 *If the sequence* $\{z^k\}$ *is bounded, then*

$$\sum_{k=0}^{+\infty}\left\|z^k - z^{k+1}\right\|^2 < +\infty.$$

Proof Since $\{z^k\}$ is bounded, $\{z^k\}$ has at least one cluster point. Let $z^* = (x^*, y^*, \lambda^*)$ be a cluster point of $\{z^k\}$ and let a subsequence $\{z^{k_j}\}$ converge to z^*. Since f is lower semicontinuous and g is continuously differentiable, then $\mathcal{L}_\beta(\cdot)$ is lower semicontinuous, and hence

$$\mathcal{L}_\beta(z^*) \le \liminf_{k_j \to +\infty} \mathcal{L}_\beta(z^{k_j}). \tag{21}$$

Thus $\{\mathcal{L}_\beta(z^{k_j})\}$ is bounded from below. Furthermore, by Lemma 3.1, sequence $\{\mathcal{L}_\beta(z^k)\}$ is nonincreasing, and so $\{\mathcal{L}_\beta(z^k)\}$ is convergent. Moreover, $\mathcal{L}_\beta(z^*) \le \mathcal{L}_\beta(z^k)$, for all k.

On the other hand, summing up of (12) for $k = 0, 1, 2, \ldots, p$, it follows that

$$\begin{aligned} \frac{1}{2}\sum_{k=0}^{p} \|x^{k+1} - x^k\|_G^2 + \delta \sum_{k=0}^{p} \|y^{k+1} - y^k\|^2 &\le \mathcal{L}_\beta(z^0) - \mathcal{L}_\beta(z^{p+1}) \\ &\le \mathcal{L}_\beta(z^0) - \mathcal{L}_\beta(z^*) \\ &< +\infty. \end{aligned}$$

Since $\delta > 0$, $G \succeq 0$, and p is chosen arbitrarily,

$$\sum_{k=0}^{+\infty} \|y^{k+1} - y^k\|^2 < +\infty, \qquad \sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|_G^2 < +\infty. \tag{22}$$

From (20) we have

$$\sum_{k=0}^{+\infty} \|\lambda^{k+1} - \lambda^k\|^2 < +\infty.$$

Next we prove $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$. From $\lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} - y^{k+1})$, we have

$$\lambda^{k+1} - \lambda^k = \lambda^k - \lambda^{k-1} + \beta(Ax^k - Ax^{k+1}) + \beta(y^{k+1} - y^k).$$

Then

$$\begin{aligned} &\|\beta(Ax^k - Ax^{k+1})\|^2 \\ &\quad = \|(\lambda^{k+1} - \lambda^k) - (\lambda^k - \lambda^{k-1}) - \beta(y^{k+1} - y^k)\|^2 \\ &\quad \le 3(\|\lambda^{k+1} - \lambda^k\|^2 + \|\lambda^k - \lambda^{k-1}\|^2 + \beta^2\|y^k - y^{k+1}\|^2). \end{aligned} \tag{23}$$

Therefore, we have $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|_{A^\top A}^2 < +\infty$. Taking into account (22), we have

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|_{(A^\top A + G)}^2 < +\infty.$$

Since $A^\top A + G \succ 0$ (see Assumption 3.1(iv)), one has $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$.

Therefore, $\sum_{k=0}^{+\infty} \|z^{k+1} - z^k\|^2 < +\infty$. □

Lemma 3.3 *Define*

$$\begin{cases} \varepsilon_x^{k+1} = \beta A^\top(y^k - y^{k+1}) + A^\top(\lambda^k - \lambda^{k+1}) - G(x^{k+1} - x^k), \\ \varepsilon_y^{k+1} = \lambda^{k+1} - \lambda^k, \\ \varepsilon_\lambda^{k+1} = \frac{1}{\beta}(\lambda^{k+1} - \lambda^k). \end{cases}$$

Then $(\varepsilon_x^{k+1}, \varepsilon_y^{k+1}, \varepsilon_\lambda^{k+1})^\top \in \partial\mathcal{L}_\beta(z^{k+1})$. *Furthermore, if* $A^\top A \succ 0$, *then there exists* $\tau > 0$ *such that*

$$d(0, \partial\mathcal{L}_\beta(z^{k+1})) \le \tau(\|y^{k+1} - y^k\| + \|y^k - y^{k-1}\|), \quad k \ge 1.$$

Proof By the definition of $\mathcal{L}_\beta(\cdot)$, one has

$$\begin{cases} \partial_x\mathcal{L}_\beta(z^{k+1}) = \partial f(x^{k+1}) - A^\top\lambda^{k+1} + \beta A^\top(Ax^{k+1} - y^{k+1}), \\ \partial_y\mathcal{L}_\beta(z^{k+1}) = \nabla g(y^{k+1}) + \lambda^{k+1} - \beta(Ax^{k+1} - y^{k+1}), \\ \partial_\lambda\mathcal{L}_\beta(z^{k+1}) = -(Ax^{k+1} - y^{k+1}). \end{cases} \tag{24}$$

Combining (24) and (11), we get

$$\begin{cases} \beta A^\top(y^k - y^{k+1}) + A^\top(\lambda^k - \lambda^{k+1}) - G(x^{k+1} - x^k) \in \partial_x\mathcal{L}_\beta(z^{k+1}), \\ \lambda^{k+1} - \lambda^k \in \partial_y\mathcal{L}_\beta(z^{k+1}), \\ \frac{1}{\beta}(\lambda^{k+1} - \lambda^k) \in \partial_\lambda\mathcal{L}_\beta(z^{k+1}). \end{cases} \tag{25}$$

From Lemma 2.1, one has $(\varepsilon_x^{k+1}, \varepsilon_y^{k+1}, \varepsilon_\lambda^{k+1})^\top \in \partial\mathcal{L}_\beta(z^{k+1})$.

On the other hand, it is easy to see that there exists $\tau_1 > 0$ such that

$$\|(\varepsilon_x^{k+1}, \varepsilon_y^{k+1}, \varepsilon_\lambda^{k+1})^\top\| \le \tau_1(\|y^{k+1} - y^k\| + \|\lambda^{k+1} - \lambda^k\| + \|x^{k+1} - x^k\|). \tag{26}$$

Due to $A^\top A \succ 0$ and (23), there exists $\tau_2 > 0$ such that

$$\|x^{k+1} - x^k\| \le \tau_2(\|y^{k+1} - y^k\| + \|\lambda^{k+1} - \lambda^k\| + \|\lambda^k - \lambda^{k-1}\|), \quad k \ge 1. \tag{27}$$

Since $(\varepsilon_x^{k+1}, \varepsilon_y^{k+1}, \varepsilon_\lambda^{k+1})^\top \in \partial\mathcal{L}_\beta(z^{k+1})$, from (26), (20) and (27), there exists $\tau > 0$ such that

$$d(0, \partial\mathcal{L}_\beta(z^{k+1})) \le \|(\varepsilon_x^{k+1}, \varepsilon_y^{k+1}, \varepsilon_\lambda^{k+1})^\top\| \le \tau(\|y^{k+1} - y^k\| + \|y^k - y^{k-1}\|), \quad k \ge 1.$$ □

Theorem 3.1 (Global convergence) *Let* Ω *denote the cluster point set of the sequence* $\{z^k\}$. *If* $\{z^k\}$ *is bounded, then*

(i) Ω *is a nonempty compact set, and* $d(z^k, \Omega) \to 0$, *as* $k \to +\infty$,

(ii) $\Omega \subseteq \operatorname{crit}\mathcal{L}_\beta$,

(iii) $\mathcal{L}_\beta(\cdot)$ *is constant on* Ω, *and* $\lim_{k\to+\infty}\mathcal{L}_\beta(z^k) = \mathcal{L}_\beta(z^*)$ *for all* $z^* \in \Omega$.

Proof (i) From the definitions of Ω and $d(z^k, \Omega)$, the claim follows trivially.

(ii) Let $z^* = (x^*, y^*, \lambda^*) \in \Omega$, then there is a subsequence $\{z^{k_j}\}$ of $\{z^k\}$, such that $\lim_{k_j\to+\infty} z^{k_j} = z^*$. Since x^{k+1} is a minimizer of function $\mathcal{L}_\beta(x, y^k, \lambda^k) + \frac{1}{2}\|x - x^k\|_G^2$ for the variable x, one has

$$\mathcal{L}_\beta(x^{k+1}, y^k, \lambda^k) + \frac{1}{2}\|x^{k+1} - x^k\|_G^2 \le \mathcal{L}_\beta(x^*, y^k, \lambda^k) + \frac{1}{2}\|x^* - x^k\|_G^2,$$

that is,

$$\mathcal{L}_\beta(x^{k+1}, y^k, \lambda^k) \le \mathcal{L}_\beta(x^*, y^k, \lambda^k) + \frac{1}{2}\|x^* - x^k\|_G^2 - \frac{1}{2}\|x^{k+1} - x^k\|_G^2. \tag{28}$$

Lemma 3.1 implies that $\lim_{k\to\infty}\|x^{k+1}-x^k\|_G^2=0$. Since $\mathcal{L}_\beta(\cdot)$ is continuous with respect to y and λ, we have

$$\begin{aligned}\limsup_{k_j\to+\infty}\mathcal{L}_\beta\bigl(z^{k_j+1}\bigr)&=\limsup_{k_j\to+\infty}\mathcal{L}_\beta\bigl(x^{k_j+1},y^{k_j},\lambda^{k_j}\bigr)\\&\leq\limsup_{k_j\to+\infty}\mathcal{L}_\beta\bigl(x^*,y^k,\lambda^k\bigr)\\&=\mathcal{L}_\beta\bigl(z^*\bigr).\end{aligned}\tag{29}$$

On the other hand, since $\mathcal{L}(\cdot)$ is lower semicontinuous,

$$\liminf_{k_j\to+\infty}\mathcal{L}_\beta\bigl(z^{k_j+1}\bigr)\geq\mathcal{L}_\beta\bigl(z^*\bigr).\tag{30}$$

Combining (29) and (30), we get $\lim_{k_j\to+\infty}\mathcal{L}_\beta(z^{k_j})=\mathcal{L}_\beta(z^*)$. Then $\lim_{k_j\to+\infty}f(x^{k_j})=f(x^*)$. By taking the limit $k_j\to+\infty$ in (11), we have

$$\begin{cases}A^\top\lambda^*\in\partial f(x^*),\\\nabla g(y^*)=-\lambda^*,\\Ax^*-y^*=0.\end{cases}$$

that is, $z^*\in\operatorname{crit}\mathcal{L}_\beta$.

(iii) Let $z^*\in\Omega$. There exists $\{z^{k_j}\}$ such that $\lim_{k_j\to+\infty}z^{k_j}=z^*$. Combining $\lim_{k_j\to+\infty}\mathcal{L}_\beta(z^{k_j})=\mathcal{L}_\beta(z^*)$ and the fact that $\{\mathcal{L}_\beta(z^k)\}$ is monotonically nonincreasing, for all $z^*\in\Omega$, we have

$$\lim_{k\to+\infty}\mathcal{L}_\beta\bigl(z^k\bigr)=\mathcal{L}_\beta\bigl(z^*\bigr),$$

and so $\mathcal{L}_\beta(\cdot)$ is constant on Ω. □

Theorem 3.2 (Strong convergence) *If $\{z^k\}$ is bounded, $A^\top A\succ 0$, and $\mathcal{L}_\beta(z)$ satisfies the KL property at each point of Ω, then*

(i) $\sum_{k=0}^{+\infty}\|z^{k+1}-z^k\|<+\infty$,

(ii) *The sequence $\{z^k\}$ converges to a stationary point of $\mathcal{L}_\beta(\cdot)$.*

Proof (i) Let $z^*\in\Omega$. From Theorem 3.1, we have $\lim_{k\to+\infty}\mathcal{L}_\beta(z^k)=\mathcal{L}_\beta(z^*)$. We consider two cases:

(a) There exists an integer k_0, such that $\mathcal{L}_\beta(z^{k_0})=\mathcal{L}_\beta(z^*)$. From Lemma 3.1, we have

$$\begin{aligned}&\frac{1}{2}\bigl\|x^{k+1}-x^k\bigr\|_G^2+\delta\bigl\|y^{k+1}-y^k\bigr\|^2\\&\quad\leq\mathcal{L}_\beta\bigl(z^k\bigr)-\mathcal{L}_\beta\bigl(z^{k+1}\bigr)\leq\mathcal{L}_\beta\bigl(z^{k_0}\bigr)-\mathcal{L}_\beta\bigl(z^*\bigr)=0,\quad k\geq k_0.\end{aligned}$$

Then, one has $\|x^{k+1}-x^k\|_G^2=0$, $y^{k+1}=y^k$, $k\geq k_0$. From (20), one has $\lambda^{k+1}=\lambda^k$, $k>k_0$. Furthermore, from (23) and $A^\top A\succ 0$, we have $x^{k+1}=x^k$, $k>k_0+1$. Thus $z^{k+1}=z^k$, $k>k_0+1$. Therefore, the conclusions hold.

(b) Suppose that $\mathcal{L}_\beta(z^k) > \mathcal{L}_\beta(z^*)$, $k \geq 1$. From Theorem 3.1(i), it follows that for $\varepsilon > 0$, there exists $k_1 > 0$, such that $d(z^k, \Omega) < \varepsilon$, for all $k > k_1$. Since $\lim_{k\to+\infty} \mathcal{L}_\beta(z^k) = \mathcal{L}_\beta(z^*)$, for given $\eta > 0$, there exists $k_2 > 0$, such that $\mathcal{L}_\beta(z^k) < \mathcal{L}_\beta(z^*) + \eta$, for all $k > k_2$. Consequently, one has

$$d(z^k, \Omega) < \varepsilon, \mathcal{L}_\beta(z^*) < \mathcal{L}_\beta(z^k) < \mathcal{L}_\beta(z^*) + \eta, \quad \text{for all } k > \tilde{k} = \max\{k_1, k_2\}.$$

It follows from the KL property, that

$$\varphi'(\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^*))d(0, \partial\mathcal{L}_\beta(z^k)) \geq 1, \quad \text{for all } k > \tilde{k}. \tag{31}$$

By the concavity of φ and since $\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^{k+1}) = (\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^*)) - (\mathcal{L}_\beta(z^{k+1}) - \mathcal{L}_\beta(z^*))$, we have

$$\begin{aligned} &\varphi(\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^*)) - \varphi(\mathcal{L}_\beta(z^{k+1}) - \mathcal{L}_\beta(z^*)) \\ &\quad \geq \varphi'(\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^*))(\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^{k+1})). \end{aligned} \tag{32}$$

Let $\triangle_{p,q} = \varphi(\mathcal{L}_\beta(z^p) - \mathcal{L}_\beta(z^*)) - \varphi(\mathcal{L}_\beta(z^q) - \mathcal{L}_\beta(z^*))$. Combining $\varphi'(\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^*)) > 0$, (31) and (32), we have

$$\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^{k+1}) \leq \frac{\triangle_{k,k+1}}{\varphi'(\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^*))} \leq d(0, \partial\mathcal{L}_\beta(z^k))\, \triangle_{k,k+1}\,.$$

From Lemma 3.3, we obtain

$$\mathcal{L}_\beta(z^k) - \mathcal{L}_\beta(z^{k+1}) \leq \tau(\|y^k - y^{k-1}\| + \|y^{k-1} - y^{k-2}\|)\, \triangle_{k,k+1}\,.$$

From Lemma 3.1 and the above inequality, we have

$$\begin{aligned} &\frac{1}{2}\|x^{k+1} - x^k\|_G^2 + \delta\|y^{k+1} - y^k\|^2 \\ &\quad \leq \tau(\|y^k - y^{k-1}\| + \|y^{k-1} - y^{k-2}\|)\triangle_{k,k+1}, \quad \text{for all } k > \tilde{k}. \end{aligned}$$

Thus

$$\|y^{k+1} - y^k\|^2 \leq \frac{\tau}{\delta}(\|y^k - y^{k-1}\| + \|y^{k-1} - y^{k-2}\|)\triangle_{k,k+1}, \quad \text{for all } k > \tilde{k}.$$

Furthermore,

$$\begin{aligned} &3\|y^{k+1} - y^k\| \\ &\quad \leq 2(\|y^k - y^{k-1}\| + \|y^{k-1} - y^{k-2}\|)^{\frac{1}{2}}\left(\frac{3}{2}\sqrt{\frac{\tau}{\delta}}\triangle_{k,k+1}^{\frac{1}{2}}\right), \quad \text{for all } k > \tilde{k}. \end{aligned}$$

Using the fact that $2ab \leq a^2 + b^2$, we obtain

$$\begin{aligned} &3\|y^{k+1} - y^k\| \\ &\quad \leq (\|y^k - y^{k-1}\| + \|y^{k-1} - y^{k-2}\|) + \frac{9\tau}{4\delta}\triangle_{k,k+1}, \quad \text{for all } k > \tilde{k}. \end{aligned} \tag{33}$$

Summing up the above inequality for $k = \tilde{k}+1,\dots,s$, yields

$$3\sum_{k=\tilde{k}+1}^{s}\|y^{k+1}-y^k\| \le \sum_{k=\tilde{k}+1}^{s}\left(\|y^k-y^{k-1}\|+\|y^{k-1}-y^{k-2}\|\right)+\frac{9\tau}{4\delta}\triangle_{\tilde{k}+1,s+1}.$$

Thus

$$\sum_{k=\tilde{k}+1}^{s}\|y^{k+1}-y^k\| \le \left(2\|y^{\tilde{k}+1}-y^{\tilde{k}}\|+\|y^{\tilde{k}}-y^{\tilde{k}-1}\|\right)+\frac{9\tau}{4\delta}\triangle_{\tilde{k}+1,s+1}.$$

Notice that $\varphi(\mathcal{L}_\beta(z^{s+1})-\mathcal{L}_\beta(z^*))>0$, so taking the limit $s\to+\infty$, we have

$$\begin{aligned}&\sum_{k=\tilde{k}+1}^{+\infty}\|y^{k+1}-y^k\|\\&\quad\le\left(2\|y^{\tilde{k}+1}-y^{\tilde{k}}\|+\|y^{\tilde{k}}-y^{\tilde{k}-1}\|\right)+\frac{9\tau}{4\delta}\varphi\left(\mathcal{L}_\beta(z^{\tilde{k}+1})-\mathcal{L}(z^*)\right).\end{aligned}\tag{34}$$

Thus

$$\sum_{k=\tilde{k}+1}^{+\infty}\|y^{k+1}-y^k\|<+\infty.$$

It follows from (20) that

$$\sum_{k=\tilde{k}+1}^{+\infty}\|\lambda^{k+1}-\lambda^k\|<+\infty.$$

From $A^\top A\succ 0$, (23) and the above two formulas, we obtain

$$\sum_{k=\tilde{k}+1}^{+\infty}\|x^{k+1}-x^k\|<+\infty.$$

Since

$$\begin{aligned}\|z^{k+1}-z^k\|&=\left(\|x^{k+1}-x^k\|^2+\|y^{k+1}-y^k\|^2+\|\lambda^{k+1}-\lambda^k\|^2\right)^{\frac{1}{2}}\\&\le\|x^{k+1}-x^k\|+\|y^{k+1}-y^k\|+\|\lambda^{k+1}-\lambda^k\|,\end{aligned}$$

we know

$$\sum_{k=\tilde{k}+1}^{m}\|z^{k+1}-z^k\|<+\infty.$$

(ii) From (i), we known that $\{z^k\}$ is a Cauchy sequence and so is convergent. Theorem 3.2(ii) follows immediately from Theorem 3.1(ii). □

In the above results, we have assumed the boundedness of the sequence $\{z^k\}$. Next, we present two sufficient conditions ensuring this requirement.

Lemma 3.4 *Suppose that* $A^\top A \succ 0$ *and*

$$\Gamma := \inf_{y \in R^m} \left\{ g(y) - \frac{1}{2L} \left\| \nabla g(y) \right\|^2 \right\} > -\infty.$$

If one of the following statements is true:

(i) f *is coercive, i.e.*, $\lim_{\|x\| \to +\infty} f(x) = +\infty$,

(ii) f *is bounded from below and* g *is coercive, i.e.*, $\inf_{x \in R^n} f(x) > -\infty$ *and* $\lim_{\|x\| \to +\infty} g(x) = +\infty$,

then $\{z^k\}$ *is bounded.*

Proof (i) Suppose that f is coercive. From Lemma 3.1, we know that $\mathcal{L}_\beta(z^k) \le \mathcal{L}_\beta(z^1) < +\infty$, for all $k \ge 1$. Combining with $\nabla g(x^k) = -\lambda^k$, one has

$$\begin{aligned}
\mathcal{L}_\beta(z^1) \ge{}& f(x^k) + g(y^k) - \langle \lambda^k, Ax^k - y^k \rangle + \frac{\beta}{2} \| Ax^k - y^k \|^2 \\
={}& f(x^k) + g(y^k) - \frac{1}{2\beta} \| \lambda^k \|^2 + \frac{\beta}{2} \left\| Ax^k - y^k - \frac{1}{\beta} \lambda^k \right\|^2 \\
={}& f(x^k) + \left(g(y^k) - \frac{1}{2L} \| \nabla g(y^k) \|^2 \right) + \left(\frac{1}{2L} - \frac{1}{2\beta} \right) \| \lambda^k \|^2 \\
&+ \frac{\beta}{2} \left\| Ax^k - y^k - \frac{1}{\beta} \lambda^k \right\|^2 \\
\ge{}& f(x^k) + \Gamma + \left(\frac{1}{2L} - \frac{1}{2\beta} \right) \| \lambda^k \|^2 + \frac{\beta}{2} \left\| Ax^k - y^k - \frac{1}{\beta} \lambda^k \right\|^2.
\end{aligned} \tag{35}$$

Since $\beta > 2L$ and f is coercive, it is easy to see that $\{x^k\}$, $\{\lambda^k\}$, and $\{\frac{\beta}{2} \|Ax^k - y^k - \frac{1}{\beta}\lambda^k\|^2\}$ are bounded. Furthermore, $\{y^k\}$ is bounded. Thus $\{z^k\}$ is bounded.

(ii) Similar as with (i), we have

$$\begin{aligned}
\mathcal{L}_\beta(z^1) \ge{}& f(x^k) + g(y^k) - \frac{1}{2\beta} \| \lambda^k \|^2 + \frac{\beta}{2} \left\| Ax^k - y^k - \frac{1}{\beta} \lambda^k \right\|^2 \\
\ge{}& f(x^k) + \frac{1}{2} g(y^k) + \frac{1}{2} \Gamma + \frac{1}{4L} \| \nabla g(y^k) \|^2 - \frac{1}{2\beta} \| \lambda^k \|^2 \\
&+ \frac{\beta}{2} \left\| Ax^k - y^k - \frac{1}{\beta} \lambda^k \right\|^2 \\
\ge{}& f(x^k) + \frac{1}{2} g(y^k) + \frac{1}{2} \Gamma + \left(\frac{1}{4L} - \frac{1}{2\beta} \right) \| \lambda^k \|^2 + \frac{\beta}{2} \left\| Ax^k - y^k - \frac{1}{\beta} \lambda^k \right\|^2.
\end{aligned}$$

Notice that $\beta > 2L$, function f is bounded from below, g is coercive and Assumption 3.1(ii) holds, thus $\{y^k\}$, $\{\lambda^k\}$, and $\{\frac{\beta}{2} \|Ax^k - y^k - \frac{1}{\beta}\lambda^k\|^2\}$ are bounded. Since $A^\top A \succ 0$, $\{x^k\}$ is bounded. Thus $\{z^k\}$ is bounded. □

4 Numerical examples

In compressed sensing, one needs to find the sparsest solution of a linear system, which can be modeled as

$$\begin{aligned} \min \quad & \|x\|_0, \\ \text{s.t. } & Dx = b, \end{aligned} \tag{36}$$

where $D \in R^{m\times n}$ is the measuring matrix, $b \in R^m$ is observed data, $\|x\|_0$ denotes the number of nonzero elements of x, which is called the l_0 norm.

Problem (36) is NP-hard. In practical applications, one may relax the l_0 norm to the l_1 norm or $l_{\frac{1}{2}}$ norm, and consider their regularized versions, which lead to the following convex problem (37) and nonconvex problem (38):

$$\begin{aligned} \min \quad & \gamma \|x\|_1 + \|y\|^2, \\ \text{s.t. } & Dx - y = b, \end{aligned} \tag{37}$$

and

$$\begin{aligned} \min \quad & \gamma \|x\|_{\frac{1}{2}}^{\frac{1}{2}} + \|y\|^2, \\ \text{s.t. } & Dx - y = b. \end{aligned} \tag{38}$$

In this section, we will apply RADMM (7) to solve the above two problems. For simplicity, we set $b = 0$ throughout this section. Applying RADMM (7) to problem (37) with $G = \alpha I - \beta D^\top D$ yields

$$\begin{cases} x^{k+1} \in S(x^k - \frac{\beta}{\alpha} D^\top D x^k + \frac{1}{\alpha} D^\top (\beta y^k + \lambda^k)); \frac{\gamma}{2\alpha}), \\ y^{k+1} = \frac{1}{2+\beta}(\beta D x^{k+1} - \lambda^k), \\ \lambda^{k+1} = \lambda^k - \beta(D x^{k+1} - y^{k+1}), \end{cases} \tag{39}$$

where $S(\cdot;\mu) = \{s_\mu(x_1), s_\mu(x_2), \ldots, s_\mu(x_n)\}^\top$ is the soft shrinkage operator [24] defined as follows:

$$s_\mu(x_i) = \begin{cases} x_i + \frac{\mu}{2}, & \text{if } x_i \le -\frac{\mu}{2}, \\ 0, & \text{if } |x| < \frac{\mu}{2}, \\ x_i - \frac{\mu}{2}, & \text{if } x_i \ge \frac{\mu}{2}. \end{cases}$$

Applying RADMM (7) to problem (38) with $G = \alpha I - \beta D^\top D$ yields

$$\begin{cases} x^{k+1} \in H(x^k - \frac{\beta}{\alpha} D^\top D x^k + \frac{1}{\alpha} D^\top (\beta y^k + \lambda^k)); \frac{\gamma}{\alpha}), \\ y^{k+1} = \frac{1}{2+\beta}(\beta D x^{k+1} - \lambda^k), \\ \lambda^{k+1} = \lambda^k - \beta(D x^{k+1} - y^{k+1}), \end{cases} \tag{40}$$

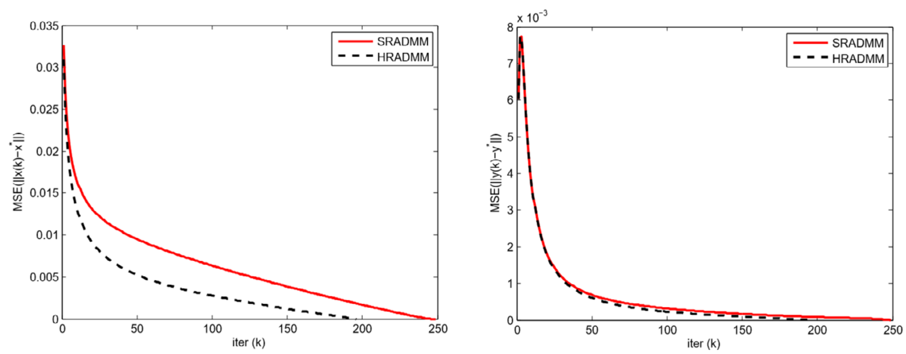

Figure 1 Comparison the performance of HRADMM and SRADMM

where $H(\cdot;\mu) = \{h_\mu(x_1), h_\mu(x_2), \ldots, h_\mu(x_n)\}^\top$ is the half-shrinkage operator [25] defined as follows:

$$h_\mu(x_i) = \begin{cases} \frac{2x_i}{3}(1 + \cos\frac{2}{3}(\pi - \varphi(|x_i|))), & |x_i| > \frac{\sqrt[3]{54}}{4}\mu^{\frac{2}{3}}, \\ 0, & \text{otherwise}, \end{cases}$$

with $\varphi(x) = \arccos(\frac{\mu}{8}(\frac{x_i}{3})^{-\frac{3}{2}})$.

For simplicity, we denote algorithms (39) and (40) by SRADMM and HRADMM, respectively. The selection of relevant parameters in numerical experiments is given below. We now conduct an experiment to verify convergence of the nonconvex RADMM, and reveal its advantages in sparsity-inducing and efficiency through comparing the performance of HRADMM and SRADMM. In the experiment, $m = 511$, $n = 512$, the matrix $D \in R^{511\times 512}$ is obtained by unitizing the matrix with randomly generated entries obeying the normal distribution $\mathcal{N}(0,1)$, the noise vector $\varepsilon \sim \mathcal{N}(0,1)$, the recovery vector $r = Dx^0 + \varepsilon$, and the regularization parameters are $\gamma = 0.0015$, $\beta = 0.8$, $\alpha = 2.5$.

The experimental results are shown in Fig. 1, where the restoration accuracy is measured by means of the mean squared error:

$$\mathrm{MSE}(\|x^* - x^k\|) = \frac{1}{n}\|x^* - x^k\|,$$
$$\mathrm{MSE}(\|y^* - y^k\|) = \frac{1}{n}\|y^* - y^k\|,$$

where $(x^*, y^*) = (0,0)$ is the optimal solution for the problems (37) and (38), respectively.

Programming is performed on Matlab R2014a, a computer running the program is configured as follows: Windows 7 system, Inter(R) Core(TM) i7-4790 CPU 3.60 GHz, 4 GB memory. Numerical results show that algorithm (7) is efficient and stable. As shown in Fig. 1, both sequences x^k and y^k were fairly near the true solution. i.e., the convergence is justified. It is readily seen that HRADMM converges faster than SRADMM.

5 Conclusion and outlook

In this paper, a regularized alternating direction method of multipliers is proposed for a class of nonconvex problems. Firstly, the global convergence of the algorithm is analyzed. Secondly, under the condition that the augmented Lagrangian function $\mathcal{L}_\beta(\cdot)$ satisfies the

Kurdyka–Łojasiewicz property, the strong convergence of the algorithm is analyzed. Finally, the effectiveness of the algorithm is verified by numerical experiments.

Funding
This research was supported by the National Natural Science Foundation of China (Nos. 11601095, 11771383) and by the Natural Science Foundation of Guangxi Province (Nos. 2016GXNSFBA380185, 2016GXNSFDA380019 and 2014GXNSFFA118001).

Availability of data and materials
The authors declare that all data and material in the paper are available and veritable.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally and significantly in writing this article. All authors wrote, read, and approved the final manuscript.

Authors' information
JinBao Jian, Ph.D., professor, E-mail: jianjb@gxu.edu.cn; Ye Zhang, E-mail: yezhang126@126.com; MianTao Chao, Ph.D., associate professor, E-mail: chaomiantao@126.com.

Author details
[1]College of Mathematics and Information Science, Guangxi University, Nanning, China. [2]College of Science, Guangxi University for Nationalities, Nanning, China.

Publisher's Note

Received: 11 December 2018 Accepted: 1 July 2019 Published online: 12 July 2019

References

1. Yao, Y.H., Liou, Y.C., Yao, J.C.: Iterative algorithms for the split variational inequality and fixed point problems under nonlinear transformations. J. Nonlinear Sci. Appl. **10**, 843–854 (2017)
2. Yao, Y.H., Yao, J.C., Liou, Y.C., Postolache, M.: Iterative algorithms for split common fixed points of demicontractive operators without prior knowledge of operator norms. Carpath. J. Math. **34**, 459–466 (2018)
3. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Comput. Math. Appl. **2**(1), 17–40 (1976)
4. He, B.S., Yuan, X.M.: On the $O(1/n)$ convergence rate of the Douglas–Rachford alternating direction method. SIAM J. Numer. Anal. **50**(2), 700–709 (2012)
5. Monteiro, R.D.C., Svaite, B.F.: Iteration-complexity of block-decomposition algorithms and the alternating direction method of multipliers. SIAM J. Optim. **23**(1), 475–507 (2013)
6. He, B.S., Yuan, X.M.: On non-ergodic convergence rate of Douglas–Rachford alternating direction method of multipliers. Numer. Math. **130**(3), 567–577 (2015)
7. Han, D.R., Sun, D.F., Zhang, L.W.: Linear rate convergence of the alternating direction method of multipliers for convex composite quadratic and semi-definite programming. IEEE Trans. Autom. Control **60**(3), 644–658 (2015)
8. Hong, M.Y.: A distributed, asynchronous, and incremental algorithm for nonconvex optimization: an ADMM approach. IEEE Trans. Control Netw. Syst. **5**(3), 935–945 (2018)
9. Li, G.Y., Pong, T.K.: Global convergence of splitting methods for nonconvex composite optimization. SIAM J. Optim. **25**(4), 2434–2460 (2015)
10. You, S., Peng, Q.Y.: A non-convex alternating direction method of multipliers heuristic for optimal power flow. In: IEEE International Conference on Smart Grid Communications, pp. 788–793. IEEE Press, New York (2014)
11. Hong, M.Y., Luo, Z.Q., Razaviyayn, M.: Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems. SIAM J. Optim. **26**(1), 3836–3840 (2014)
12. Wang, F.H., Xu, Z.B., Xu, H.K.: Convergence of Bregman alternating direction method with multipliers for nonconvex composite problems (2014). Preprint. Available at arXiv:1410.8625
13. Hong, M., Luo, Z.Q., Razaviyayn, M.: Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems. SIAM J. Optim. **26**(1), 337–364 (2016)
14. Wang, F.H., Cao, W.F., Xu, Z.B.: Convergence of multi-block Bregman ADMM for nonconvex composite problems. Sci. China Inf. Sci. **61**(12), 53–64 (2018)
15. Yang, L., Pong, T.K., Chen, X.J.: Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction. SIAM J. Imaging Sci. **10**(1), 74–110 (2017)
16. Guo, K., Han, D.R., Wu, T.T.: Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints. Int. J. Comput. Math. **94**(8), 1–18 (2016)
17. Guo, K., Han, D.R., Wang, Z.W., Wu, T.T.: Convergence of ADMM for multi-block nonconvex separable optimization models. Front. Math. China **12**(5), 1139–1162 (2017)
18. Zhao, X.P., Ng, K.F., Li, C., Yao, J.C.: Linear regularity and linear convergence of projection-based methods for solving convex feasibility problems. Appl. Math. Optim. **78**, 613–641 (2018)
19. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer, Berlin (1998)

20. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I: Basic Theory. Springer, Berlin (2006)
21. Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: 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)
22. Nesterov, Y.: Introduction Lectures on Convex Optimization: A Basic Course. Springer, Berlin (2013)
23. Frankel, P., Garrigos, G., Peypouquet, J.: Splitting methods with variable metric for Kurdyka–Lojasiewicz functions and general convergence rates. J. Optim. Theory Appl. **165**(3), 874–900 (2015)
24. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**(11), 1413–1457 (2003)
25. Xu, Z., Chang, X., Xu, F., Zhang, H.: $l_{\frac{1}{2}}$ regularization: a thresholding representation theory and a fast solver. IEEE Trans. Neural Netw. Learn. Syst. **23**(7), 1013–1027 (2012)