



Research article

Partial symmetric regularized alternating direction method of multipliers for non-convex split feasibility problems

Yue Zhao¹, Meixia Li^{2,*}, Xiaowei Pan² and Jingjing Tan²

¹ College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China

² School of Mathematics and Statistics, Weifang University, Weifang 261061, China

* **Correspondence:** Email: limeixia001@163.com.

Abstract: In this paper, we propose a kind of partial symmetric regularized alternating direction method of multipliers for solving non-convex split feasibility problems. This algorithm adds an update to the Lagrangian multiplier term during the iteration process. Existing results about ADMM cannot be widely applied because the problem under consideration does not satisfy the assumptions usually claimed. The global convergence of the algorithm is proved under appropriate assumptions. And when the augmented Lagrangian function satisfies the KL property, the strong convergence of the algorithm is obtained. Furthermore, when the correlation function has a special structure, the sublinear and linear convergence rates of the algorithm are ensured. Finally, numerical experiments are conducted to show that the algorithm is effective.

Keywords: non-convex split feasibility problems; alternating direction method of multipliers; KL property; convergence rate

Mathematics Subject Classification: 47J20, 47J25, 47J30

1. Introduction

The split feasibility problem (SFP) [1] is formulated as

$$\text{to find } x \in C \text{ such that } Ax \in Q, \tag{1.1}$$

where C and Q are both non-emepty closed convex sets.

The SFP was first introduced by Censor and Elfving [1], which was used in modeling various inverse problems arising from signal processing [2], medical image reconstruction [3], and further studied by many researchers [4, 5]. Additionally, a variety of iterative algorithms have been employed to tackle the convex SFP [6–9].

When at least one of C and Q in Eq (1.1) is a non-convex set, it is called a non-convex SFP. The application of non-convex SFP is becoming increasingly widespread in various fields. Li and Pong [10] proposed the Douglas-Rachford algorithm for this type of problem. Moudafi introduced a semi-alternating algorithm in [11] to solve non-convex split equation problems and proved its global convergence under mild conditions. Chen et al. [12] developed S-Subgradient projection methods utilizing S-Subdifferential functions. Furthermore, Chen et al. [13] proposed an S-Subgradient projection algorithm incorporating an inertial technique for solving non-convex SFP, and proved the weak convergence of the algorithm in these two papers. Additionally, Guo proposed an inexact averaged projection algorithm for non-convex multiple sets SFP [14], and established the convergence of the proposed algorithm through the Kurdyka Lojasiewicz inequality.

Currently, there is a limited number of references to discuss this kind of problem. For solving non-convex SFP, Gibali et al. [15] presented four different optimization formulations and corresponding iterative algorithms. One of the formulations is as follows:

$$\begin{aligned} & \min d_C^2(x) + \delta_Q(y) \\ & \text{s.t. } Ax = y, \end{aligned} \quad (1.2)$$

where C is a closed convex set and Q is a closed non-convex set. This optimization problem is the sum of two non-negative functions, and when the optimal value of (1.2) is zero, then the achieved optimization solution is the solution of non-convex SFP.

By applying the augmented Lagrangian method, Gibali [15] processed the proximal alternating directions method (ADMM), linearized proximal ADMM, and weighted proximal ADMM to solve (1.2), which are the continuation of classic ADMM [16]. The iterative form of the weighted proximal ADMM (WPADMM) is the following:

$$\begin{cases} y^{k+1} \in \arg \min \left\{ \delta_Q(y) + \langle \lambda^k, Ax^k - y \rangle + \frac{\beta}{2} \|Ax^k - y\|^2 \right\}, \\ x^{k+1} \in \arg \min \left\{ d_C^2(x) + \langle \lambda^k, Ax - y^{k+1} \rangle + \frac{\beta}{2} \|Ax - y^{k+1}\|^2 + \frac{1}{2} \|x - x^k\|_G^2 \right\}, \\ \lambda^{k+1} = \lambda^k + \beta(Ax^{k+1} - y^{k+1}). \end{cases} \quad (1.3)$$

As the original problem can be reformulated as a separable nonconvex optimization problem (1.2), there are many advanced accelerated ADMM-type methods for solving the reformulation, such as [17] and [18]. Moreover, the last inexact one had been shown to be convergent with linear convergence rate and performed surprisingly efficient than some existing ADMM-type methods. In addition, some primal-dual splitting methods can also be used to solve it, such as [19] and [20]. However, the proposed methods involved in the last two references lack theoretical assurance since they focused on the convex case. And, due to the focus of [21] and [22] on convex problems, it is necessary to study the convergence of partially symmetric regularized ADMM in non-convex cases.

Furthermore, due to the slow convergence speed of existing ADMM, based on the above analysis and problem (1.2), we use the weighted proximal ADMM algorithm and add updates of Lagrangian multipliers to demonstrate global convergence.

In this article, the models for the functions and non-convex split feasibility problems are different, and the results of ADMM cannot be directly applied. And, we have added updates to the Lagrangian multiplier term to improve the convergence speed of the algorithm.

The structure of the paper is as follows. The basic concepts, definitions, and related results are described in the second part. The third section presents the algorithm and its proof. The fourth part

proves the convergence rate of the algorithm. The fifth section provides the corresponding numerical experiment, which verifies the feasibility and effectiveness of the algorithm.

2. Preliminaries

In this article, \mathbb{R}^n represents an n -dimensional Euclidean space, $\mathbb{R}^{m \times n}$ denotes $m \times n$ real matrices, \mathbb{N} denotes the set of natural number, and $\|\cdot\|$ represents the Euclidean norm. For any $x, y \in \mathbb{R}^n$, $\langle x, y \rangle = x^T y$. The G -norm $\|x\|_G = \sqrt{x^T G x}$ of x , where $G \succeq (>) 0$ is a symmetric positive semidefinite matrix. $\lambda_{\min}(G)$ and $\lambda_{\max}(G)$ represent the minimum and maximum eigenvalues of the symmetric matrix G , respectively. Then, $\lambda_{\min}(G)\|y\|^2 \leq \|y\|_G^2 \leq \lambda_{\max}(G)\|y\|^2, \forall y \in \mathbb{R}^n$.

To prove the partially symmetric regularized alternating direction multiplier algorithm for non-convex split feasibility problems, two functions play a crucial role which are the indicator function and the distance function. Mathematically speaking, given a non-empty closed set Q in the n -dimensional Euclidean space \mathbb{R}^n , the indicator function is defined as

$$\delta_Q(y) = \begin{cases} 0, & \text{if } y \in Q, \\ +\infty, & \text{otherwise.} \end{cases}$$

When Q is a closed non-convex set, the indicator function is proper, lower semi-continuous. The definition of proper function can be seen in reference [23].

The distance function of the set C , represented by $d_C : \mathbb{R}^n \rightarrow \mathbb{R}$, is given by

$$d_C(x) = \min\{\|u - x\| : u \in C\}.$$

In this article, we also write $d_C(x)$ as $d(x, C)$.

When C is a closed convex set, the square of the distance function is smooth. Specifically, if $C = \emptyset$, then $d_C(x) = +\infty$. Let $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. Define

$$\text{dom } g = \{y \in \mathbb{R}^n : g(y) < +\infty\}.$$

Given a non-empty closed subset C of Euclidean space, the orthogonal projection on set C is the operator $P_C(x) : \mathbb{R}^n \rightarrow C$, defined as

$$P_C(x) = \arg \min\{\|u - x\| : u \in C\}.$$

Definition 1. [24] If the function $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfies $g(y_0) \leq \liminf_{y \rightarrow y_0} g(y)$ at y_0 , then function g is said to be lower semi-continuous at y_0 . If g is lower semi-continuous at every point, then g is called a lower semi-continuous function.

Definition 2. [24] Let the function $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper lower semi-continuous.

(1) The *Fréchet* subdifferential of g at $y \in \text{dom } g$ is defined as

$$\hat{\partial}g(y) = \{y^* \in \mathbb{R}^n : \liminf_{\substack{x \neq y \\ x \rightarrow y}} \frac{g(x) - g(y) - \langle y^*, x - y \rangle}{\|x - y\|} \geq 0\}.$$

(2) The limit subdifferential of g at $y \in \text{dom } g$ is defined as

$$\partial g(y) = \{y^* \in \mathbb{R}^n : \exists y_k \rightarrow y, g(y_k) \rightarrow g(y), \hat{y}_k \in \hat{\partial}g(y), \hat{y}_k \rightarrow y^*\}.$$

By definition, the closure of ∂g can be obtained, which means that $\{(y^k, u^k)\}$ is a sequence in $\mathbb{R}^n \times \mathbb{R}^n$ that satisfies the condition, for any $k \in N$,

$$(y^k, u^k) \in \text{Graph } \partial g = \{(y, u) \in \mathbb{R}^n \times \mathbb{R}^n : u \in \partial g(y)\},$$

and if (y^k, u^k) converges to (y, u) and $g(y^k)$ converges to $g(y)$, then $(y^k, u^k) \in \text{Graph } \partial g$.

Based on the first-order optimality condition of nonsmooth functions, the critical point set of g is defined as

$$\text{crit } g = \{y \in \mathbb{R}^n : 0 \in \partial g(y)\}.$$

Remrk 1. According to the definition of subdifferentials, we can obtain the following properties (see reference [24]):

- (1) $\forall y \in \mathbb{R}^n, \hat{\partial}g(y) \subseteq \partial g(y)$ and $\hat{\partial}g(y)$ is a closed convex set, $\partial g(y)$ is a closed set.
- (2) If $y_k^* \subseteq \hat{\partial}g(y)$ and $\lim_{k \rightarrow \infty} (y_k, y_k^*) = (y, y^*)$, then $y^* \in \partial g(y)$, i.e., $\partial g(y)$ is closed.
- (3) If $y \in \mathbb{R}^n$ is the minimum point of g , then $0 \in \partial g(y)$; if $0 \in \partial g(y)$, then y is the critical point of the function g .
- (4) If $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is proper lower semi-continuous and $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is continuously differentiable, then for any $y \in \text{dom } g$, we have $\partial(f(x) + g(y)) = \nabla f(x) + \partial g(y)$.

Definition 3. [25] (Kurdyka-Lojasiewicz property) The function $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a proper lower semi-continuous function, and g is said to have KL property in $y^* \in \text{dom } \partial g$ if there exists $\eta \in (0, +\infty]$, a certain neighborhood U of y^* , and a continuous concave function $\varphi : [0, \eta) \rightarrow \mathbb{R}^+$, such that

- (1) $\varphi(0) = 0$;
- (2) φ is continuously differentiable on $(0, \eta)$ and φ is also continuous at 0;
- (3) $\forall s \in (0, \eta), \varphi'(s) > 0$;
- (4) $\forall y \in U \cap [g(y^*) < g < g(y^*) + \eta]$, all KL inequalities hold, that is

$$\varphi'(g(y) - g(y^*))d(0, \partial g(y)) \geq 1.$$

Lemma 1. [26] (Consistent KL property) Let ϕ_n be the set of functions which satisfies the KL property. Assuming Ω is a compact set, and the function $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a proper lower semi-continuous function. If the function g is a constant on Ω and satisfies the KL property at every point in Ω , then there exist $\varepsilon > 0, \eta > 0, \varphi \in \phi_n$ such that for any $\bar{y} \in \Omega$ and all y belonging to the intersection

$$\{y \in \mathbb{R}^n : d(y, \Omega) < \varepsilon\} \cap [g(\bar{y}) < g < g(\bar{y}) + \eta],$$

the following inequality holds true:

$$\varphi'(g(y) - g(\bar{y}))d(0, \partial g(y)) \geq 1.$$

In some practical applications, many functions satisfy the KL property, such as semi-algebraic functions, real analytic functions, subanalytic functions, and strongly convex functions, etc., as seen in reference [27].

Lemma 2. [28] If $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a continuous differentiable function and $\nabla \psi$ is Lipschitz continuous, then there exists a Lipschitz constant $L_\psi > 0$, such that for any $x, y \in \mathbb{R}^n$, there is

$$|\psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle| \leq \frac{L_\psi}{2} \|x - y\|^2.$$

Definition 4. (Semi-algebraic sets and functions) (1) A subset S of \mathbb{R}^n is a real semi-algebraic set, if there exists a finite number of real polynomial functions $g_{ij}, h_{ij} : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$S = \bigcup_{j=1}^p \bigcap_{i=1}^q \{u \in \mathbb{R}^n : g_{ij}(u) = 0 \text{ and } h_{ij}(u) < 0\}.$$

(2) A function $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ is called semi-algebraic if its graph

$$\{(u, t) \in \mathbb{R}^{n+1} : f(u) = t\}$$

is a semi-algebraic subset of \mathbb{R}^{n+1} .

3. Non-convex split feasibility problems

In this section, we study the solving method for problem (1.2), where C is a non-empty closed convex set of \mathbb{R}^n and Q is a non-empty closed non-convex set of \mathbb{R}^m . According to the structure of C , the following conclusions hold.

(1) The function $d_C^2(\cdot)$ is continuously differentiable.

(2) The gradient of $d_C^2(\cdot)$ is Lipschitz continuous, i.e., there is a constant L such that $\|\nabla d_C^2(y) - \nabla d_C^2(x)\| \leq L\|y - x\|$.

Now, we give the algorithm.

Algorithm 1. Partial symmetric regularized alternating direction method of multipliers for non-convex split feasibility problems (PSRADMM)

Given three initial points x^0, y^0 , and λ^0 , $G \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix. β, γ , and s are constants. Iteratively update the sequence $\{\omega^k = (x^k, y^k, \lambda^k)\}$ as follows:

$$\left\{ \begin{array}{l} y^{k+1} \in \arg \min \left\{ \delta_Q(y) - \langle \lambda^k, Ax^k - y \rangle + \frac{\beta}{2} \|Ax^k - y\|^2 + \frac{1}{2} \|y - y^k\|_G^2 \right\}, \end{array} \right. \quad (3.1a)$$

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \gamma\beta(Ax^k - y^{k+1}), \quad (3.1b)$$

$$\left\{ \begin{array}{l} x^{k+1} = \arg \min \left\{ d_C^2(x) - \langle \lambda^{k+\frac{1}{2}}, Ax - y^{k+1} \rangle + \frac{\beta}{2} \|Ax - y^{k+1}\|^2 \right\}, \end{array} \right. \quad (3.1c)$$

$$\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} - y^{k+1}). \quad (3.1d)$$

Based on the optimality conditions for each subproblem in Algorithm 1, we have

$$\left\{ \begin{array}{l} 0 \in \partial \delta_Q(y^{k+1}) + \lambda^k + \beta(Ax^k - y^{k+1}) + G(y^{k+1} - y^k), \end{array} \right. \quad (3.2a)$$

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \gamma\beta(Ax^k - y^{k+1}), \quad (3.2b)$$

$$\left\{ \begin{array}{l} 0 = \nabla d_C^2(x^{k+1}) - A^T \lambda^{k+\frac{1}{2}} + \beta A^T (Ax^{k+1} - y^{k+1}), \end{array} \right. \quad (3.2c)$$

$$\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} - y^{k+1}). \quad (3.2d)$$

Assumption 1. Some assumptions and conditions regarding Algorithm 1 are provided below.

(1) $\gamma + s > 0$ and $\gamma(1 - 2s) + s - 6(1 - s)^2 > 0$.

(2) $\beta > \beta_0 = \frac{(\gamma+s) + \sqrt{(\gamma+s)^2 + 24[\gamma(1-2s)+s-6(1-s)^2]}}{2[\gamma(1-2s)+s-6(1-s)^2]}$.

(3) $\delta = \min\{\delta_1, \delta_2\}$, where

$$\delta_1 = -\frac{4(L + \|A\|^2(1-s)\beta)^2 + s\gamma\beta\|A\|^4}{\|A\|^2(\gamma+s)\beta} - \frac{L - \beta\|A\|^2}{2} > 0,$$

$$\delta_2 = -\frac{4\beta(1-s)^2}{s+\gamma} + \frac{1}{2}\lambda_{\min}(G) > 0.$$

(4) C, Q are both semi-algebraic sets.

Remark 2. The solution set of this inequality system (1) in Assumption 1 is represented as $(\gamma, s) \in D = D_1 \cup D_2$, where

$$D_1 = \left(\frac{6(1-s)^2 - s}{1-2s}, +\infty\right) \times \left(-\infty, \frac{1}{2}\right)$$

and

$$D_2 = \left(-s, \frac{6(1-s)^2 - s}{1-2s}\right) \times \left(\frac{3 - \sqrt{3}}{2}, \frac{3 + \sqrt{3}}{2}\right).$$

It can be seen that (γ, s) has a wide range of choices. Specifically, when $s \in (\frac{3}{4}, 1)$ occurs, $(\gamma = s, s) \in D_2$, which means that the parameters γ and s of PSRADMM can take the same value in this interval.

The augmented Lagrangian function for problem (1.2) is

$$L_\beta(\omega) = d_C^2(x) + \delta_Q(y) - \langle \lambda, Ax - y \rangle + \frac{\beta}{2}\|Ax - y\|^2,$$

where $\omega = (x, y, \lambda)$, $\lambda \in \mathbb{R}^m$ is a Lagrangian multiplier, and $\beta > 0$ is a penalty parameter.

Without loss of generality, we assume that $\{\omega^k = (x^k, y^k, \lambda^k)\}$ produced by Algorithm 1 is bounded. The reason to do this will be explained in Lemma 7.

Lemma 3. The $\{\omega^* = (x^*, y^*, \lambda^*)\}$ is the critical point of the augmented Lagrangian function $L_\beta(x, y, \lambda)$ for problem (1.2), i.e., $0 \in \partial L_\beta(\omega^*)$, if and only if

$$\begin{cases} -\lambda^* \in \partial\delta_Q(y^*), \\ \nabla d_C^2(x^*) = A^T \lambda^*, \\ Ax^* - y^* = 0. \end{cases}$$

Proof. Due to the simplicity of the proof, it is omitted here. \square

Similar to the approach in reference [29], Lemma 4 below indicates that the sequence $\{L_\beta(\omega^k)\}$ monotonically decreases. For ease of analysis, let $z = (x, y)$ and $z^k = (x^k, y^k)$.

Lemma 4. If Assumption 1 holds, then

$$L_\beta(\omega^{k+1}) \leq L_\beta(\omega^k) - \delta \|z^{k+1} - z^k\|^2. \quad (3.3)$$

Proof. First, based on the iterative formulas (3.2a), (3.2d), and (3.2b), one has

$$L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) = L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+\frac{1}{2}}) + s\beta\|Ax^{k+1} - y^{k+1}\|^2 \quad (3.4)$$

and

$$L_\beta(x^k, y^{k+1}, \lambda^{k+\frac{1}{2}}) = L_\beta(x^k, y^{k+1}, \lambda^k) + \gamma\beta\|Ax^k - y^{k+1}\|^2. \quad (3.5)$$

On the other hand, based on the definition of the augmented Lagrangian function, Eq (3.2c), the Lipschitz continuity of ∇d_C^2 , and Lemma 2, then

$$\begin{aligned} & L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+\frac{1}{2}}) - L_\beta(x^k, y^{k+1}, \lambda^{k+\frac{1}{2}}) \\ &= d_C^2(x^{k+1}) - d_C^2(x^k) - \frac{\beta}{2} \|A\|^2 \|x^{k+1} - x^k\|^2 - \langle \lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} - y^{k+1}), A(x^{k+1} - x^k) \rangle \\ &= d_C^2(x^{k+1}) - d_C^2(x^k) - \frac{\beta}{2} \|A\|^2 \|x^{k+1} - x^k\|^2 - \langle \nabla d_C^2(x^{k+1}), x^{k+1} - x^k \rangle \\ &\leq \frac{L - \beta \|A\|^2}{2} \|x^{k+1} - x^k\|^2. \end{aligned} \quad (3.6)$$

And, because y^{k+1} is the solution of (3.1a), then

$$L_\beta(x^k, y^{k+1}, \lambda^k) - L_\beta(x^k, y^k, \lambda^k) \leq -\frac{1}{2} \|y^{k+1} - y^k\|_G^2 \leq -\frac{1}{2} \lambda_{\min}(G) \|y^{k+1} - y^k\|^2. \quad (3.7)$$

Therefore, adding inequalities (3.4), (3.5), (3.6), and (3.7), we have

$$\begin{aligned} & L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) - L_\beta(x^k, y^k, \lambda^k) \\ &\leq -\frac{1}{2} \lambda_{\min}(G) \|y^{k+1} - y^k\|^2 + s\beta \|Ax^{k+1} - y^{k+1}\|^2 + \frac{L - \beta \|A\|^2}{2} \|x^{k+1} - x^k\|^2 \\ &\quad + \gamma\beta \|Ax^k - y^{k+1}\|^2. \end{aligned} \quad (3.8)$$

Also, according to (3.1b) and (3.1d), we have

$$Ax^k - y^{k+1} = \frac{1}{(\gamma + s)\beta} (\lambda^k - \lambda^{k+1}) - \frac{s}{\gamma + s} (Ax^{k+1} - Ax^k). \quad (3.9)$$

$$Ax^{k+1} - y^{k+1} = \frac{1}{(\gamma + s)\beta} (\lambda^k - \lambda^{k+1}) + \frac{\gamma}{\gamma + s} (Ax^{k+1} - Ax^k). \quad (3.10)$$

On the other hand, from (3.2c) and (3.2d), it can be concluded that

$$A^T \lambda^{k+1} = \nabla d_C^2(x^{k+1}) + \beta(1 - s)A^T (Ax^{k+1} - y^{k+1}). \quad (3.11)$$

Combining the Lipschitz continuity of $\nabla d_C^2(x)$, we obtain

$$\begin{aligned} & \|A^T\| \|\lambda^{k+1} - \lambda^k\| \\ &\leq (L + \|A\|^2(1 - s)\beta) \|x^{k+1} - x^k\| + \beta(1 - \gamma)\|A^T\| \|y^{k+1} - y^k\|. \end{aligned} \quad (3.12)$$

Furthermore, using the Cauchy inequality for (3.12), we have

$$\begin{aligned} & \|A\|^2 \|\lambda^{k+1} - \lambda^k\|^2 \\ &\leq 4(L + \|A\|^2(1 - s)\beta)^2 \|x^{k+1} - x^k\|^2 + 4\beta^2(1 - s)^2 \|A\|^2 \|y^{k+1} - y^k\|^2. \end{aligned} \quad (3.13)$$

Combining Eqs (3.9), (3.10), and inequality (3.13), one has

$$\begin{aligned} & \gamma\beta \|Ax^k - y^{k+1}\|^2 + s\beta \|Ax^{k+1} - y^{k+1}\|^2 \\ &= \frac{1}{(\gamma + s)\beta} \|\lambda^{k+1} - \lambda^k\|^2 + \frac{s\gamma\beta}{\gamma + s} \|Ax^{k+1} - Ax^k\|^2 \\ &\leq \frac{4(L + \|A\|^2(1 - s)\beta)^2 + s\gamma\beta \|A\|^4}{\|A\|^2(\gamma + s)\beta} \|x^{k+1} - x^k\|^2 + \frac{4\beta(1 - s)^2}{(\gamma + s)} \|y^{k+1} - y^k\|^2. \end{aligned} \quad (3.14)$$

Substituting inequality (3.14) into inequality (3.8), we have

$$\begin{aligned} & L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) - L_\beta(x^k, y^k, \lambda^k) \\ & \leq -\left(-\frac{4\beta(1-s)^2}{(s+\gamma)} + \frac{1}{2}\lambda_{\min}(G)\right)\|y^{k+1} - y^k\|^2 - \left(-\frac{4(L + \|A\|^2(1-s)\beta)^2 + s\gamma\beta\|A\|^4}{\|A\|^2(\gamma+s)\beta}\right. \\ & \quad \left. - \frac{L - \beta\|A\|^2}{2}\right)\|x^{k+1} - x^k\|^2. \end{aligned}$$

Therefore,

$$L_\beta(\omega^{k+1}) \leq L_\beta(\omega^k) - \delta_1\|x^{k+1} - x^k\|^2 - \delta_2\|y^{k+1} - y^k\|^2 \leq L_\beta(\omega^k) - \delta\|z^{k+1} - z^k\|^2.$$

□

Lemma 5. If Assumption 1 holds and the iteration sequence $\{\omega^k = (x^k, y^k, \lambda^k)\}$ generated by Algorithm 1 is bounded, then

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

Proof. Due to $\{\omega^k\}$ being bounded, $\{\omega^k\}$ has at least one aggregation point ω^* . Hence, there exists a subsequence $\{\omega^{k_j}\}$ such that $\lim_{k_j \rightarrow +\infty} \omega^{k_j} = \omega^*$. Because $\delta_Q(\cdot)$ is a lower semi-continuous function and $d_C^2(\cdot)$ is continuous differentiable, then $L_\beta(\cdot)$ is lower semi-continuous. So $L_\beta(\omega^*) \leq \liminf_{k_j \rightarrow +\infty} L_\beta(\omega^{k_j})$, i.e., $\{L_\beta(\omega^{k_j})\}$ has a lower bound.

According to Lemma 4, $L_\beta(\omega^{k_j})$ monotonically decreases and then $L_\beta(\omega^{k_j})$ converges. Therefore, by summing up (3.3) from $k = 0$ to n , we can obtain that

$$\delta \sum_{k=0}^n \|z^{k+1} - z^k\|^2 \leq L_\beta(\omega^0) - L_\beta(\omega^{n+1}) \leq L_\beta(\omega^0) - L_\beta(\omega^*) < +\infty.$$

According to inequality (3.13), there are $\sum_{k=0}^{\infty} \|\lambda^{k+1} - \lambda^k\|^2 < +\infty$, so $\sum_{k=0}^{\infty} \|\omega^{k+1} - \omega^k\|^2 < +\infty$. □

Let Ω be the set of aggregation points of $\{\omega^k\}$ and based on the results of Lemma 5. We can establish the global convergence of Algorithm 1.

Theorem 1. (Global convergence) If Assumption 1 holds and $\{\omega^k\}$ is bounded, then the following conclusions hold.

(1) Ω is a non-empty compact set and

$$\lim_{k \rightarrow +\infty} d(\omega^k, \Omega) = 0.$$

(2) $\{L_\beta(\omega^k)\}$ is convergent, and for all $\omega^* \in \Omega$,

$$L_\beta(\omega^*) = \inf_{k \in N} L_\beta(\omega^k) = \lim_{k \rightarrow +\infty} L_\beta(\omega^k).$$

(3) $\Omega \subseteq \text{crit}L_\beta$.

Proof. (1) The conclusion can be drawn from the boundedness of $\{\omega^k\}$ and the definition of Ω .

(2) Let $\omega^* \in \Omega$. Therefore, there exists a subsequence $\{\omega^{k_j}\}$ of $\{\omega^k\}$ which converges to ω^* .

Also, according to the lower semi-continuity of $L_\beta(\cdot)$, i.e., $L_\beta(\omega^*) \leq \liminf_{k_j \rightarrow +\infty} L_\beta(\omega^{k_j+1})$, then

$\lim_{k_j \rightarrow +\infty} L_\beta(\omega^{k_j+1}) = L_\beta(\omega^*)$ and Lemma 4 states that $\{L_\beta(\omega^*)\}$ monotonically decreases. Therefore,

$$L_\beta(\omega^*) = \lim_{k \rightarrow +\infty} L_\beta(\omega^k) = \inf_{k \in \mathbb{N}} L_\beta(\omega^k), \forall \omega^* \in \Omega.$$

(3) Let $\omega^* = (x^*, y^*, \lambda^*) \in \Omega$. Therefore, there exists a subsequence $\{\omega^{k_j}\}$ of $\{\omega^k\}$ that converges to ω^* .

According to Lemma 5, $\lim_{k_j \rightarrow +\infty} \omega^{k_j+1} = \omega^*$. And, by (3.1b), $\{\lambda^{k_j+\frac{1}{2}}\}$ is convergent. Assume that

$\lim_{k_j \rightarrow +\infty} \lambda^{k_j+\frac{1}{2}} = \lambda^{**}$. Let $k = k_j$ in (3.1b) and (3.1d), and take the limit. Then, we have

$$\lambda^{**} = \lambda^* - \gamma\beta(Ax^* - y^*), \lambda^* = \lambda^{**} - s\beta(Ax^* - y^*),$$

which, combined with $(\gamma + s)\beta > 0$, it can be seen that $Ax^* - y^* = 0, \lambda^{**} = \lambda^*$.

As a result, x^* and y^* are the feasible points for problem (1.1). According to the following inequality (3.1a), we can obtain

$$\begin{aligned} & \delta_Q(y^{k_j+1}) + \langle \lambda^{k_j}, y^{k_j+1} \rangle + \frac{\beta}{2} \|Ax^{k_j} - y^{k_j+1}\|^2 + \frac{1}{2} \|y^{k_j+1} - y^{k_j}\|_G^2 \\ & \leq \delta_Q(y^*) + \langle \lambda^{k_j}, y^* \rangle + \frac{\beta}{2} \|Ax^{k_j} - y^*\|^2 + \frac{1}{2} \|y^* - y^{k_j}\|_G^2. \end{aligned}$$

Because $\lim_{k_j \rightarrow +\infty} \omega^{k_j} = \lim_{k_j \rightarrow +\infty} \omega^{k_j+1} = \omega^*$, we have $\lim_{k_j \rightarrow +\infty} \delta_Q(y^{k_j+1}) \leq \delta_Q(y^*)$. Also, due to the lower semi-continuity of $\delta_Q(y)$, one has $\lim_{k_j \rightarrow +\infty} \delta_Q(y^{k_j+1}) = \delta_Q(y^*)$. Furthermore, combining the closeness of $\partial\delta_Q$ and the continuity of ∇d_C^2 , and letting $k = k_j \rightarrow +\infty$ in (3.1), we have

$$\begin{cases} -\lambda^* \in \partial\delta_Q(y^*), \\ \nabla d_C^2(x^*) = A^T \lambda^*, \\ Ax^* - y^* = 0. \end{cases}$$

Therefore, according to Lemma 3, $\omega^* \in \text{crit}L_\beta$. Hence, $\Omega \subseteq \text{crit}L_\beta$. □

Lemma 6. If Assumption 1 holds, then there exists a constant $\tau > 0$, and, for all $k \geq 2$,

$$d(0, \partial L_\beta(\omega^{k+1})) \leq \tau(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\|).$$

Proof. Through the definition of $L_\beta(\cdot)$, it can be obtained that

$$\begin{cases} \partial_y L_\beta(\omega^{k+1}) = \partial\delta_Q(y^{k+1}) + \lambda^{k+1} - \beta(Ax^{k+1} - y^{k+1}), \\ \partial_x L_\beta(\omega^{k+1}) = \nabla d_C^2(x^{k+1}) - A^T \lambda^{k+1} + \beta A^T (Ax^{k+1} - y^{k+1}) \\ \partial_\lambda L_\beta(\omega^{k+1}) = -(Ax^{k+1} - y^{k+1}). \end{cases} \quad (3.15)$$

Let $\varepsilon^{k+1} = (\varepsilon_1^{k+1}, \varepsilon_2^{k+1}, \varepsilon_3^{k+1}) \in \partial L_\beta(\omega^{k+1})$. Combining (3.2) and (3.15), one has

$$\begin{cases} \varepsilon_1^{k+1} = \frac{\gamma s \beta}{\gamma + s} A^T A (x^{k+1} - x^k) - \frac{s A^T}{\gamma + s} (\lambda^{k+1} - \lambda^k), \\ \varepsilon_2^{k+1} = -(\lambda^k - \lambda^{k+1}) - \beta A (x^{k+1} - x^k) - G(y^{k+1} - y^k) \\ \varepsilon_3^{k+1} = \frac{1}{(\gamma + s)\beta} (\lambda^{k+1} - \lambda^k) - \frac{\gamma}{\gamma + s} A (x^{k+1} - x^k). \end{cases}$$

As a result, $\varepsilon_1^{k+1} \in \partial_x L_\beta(\omega^{k+1})$, $\varepsilon_2^{k+1} \in \partial_y L_\beta(\omega^{k+1})$, $\varepsilon_3^{k+1} \in \partial_\lambda L_\beta(\omega^{k+1})$, and $\varepsilon^{k+1} \in \partial L_\beta(\omega^{k+1})$. Hence, for all $k \geq 1$, there exists $\tau_0 > 0$ such that

$$d(0, \partial L_\beta(\omega^{k+1})) \leq \|\varepsilon^{k+1}\| \leq \tau_0(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\| + \|\lambda^{k+1} - \lambda^k\|). \quad (3.16)$$

According to the result of inequality (3.12), there exists a constant $\tau_1 > 0$ such that

$$\|\lambda^{k+1} - \lambda^k\| \leq \tau_1(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\|). \quad (3.17)$$

Combining (3.16) and (3.17), it can be concluded that there exists a constant $\tau_1 > 0$, and, for all $k \geq 2$,

$$d(0, \partial L_\beta(\omega^{k+1})) \leq \tau(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\|).$$

□

Now, we study the strong convergence of the Algorithm 1 by using Lemma 4, Lemma 6, and the relevant conclusions in Theorem 1.

Theorem 2. (Strong convergence) Assume that $\{\omega_k\}$ is bounded and $L_\beta(\cdot)$ satisfies the KL property on Ω . Then,

(1) $\lim_{k \rightarrow \infty} \|\omega^{k+1} - \omega^k\| = 0$.

(2) $\{\omega_k\}$ converges to the critical point of $L_\beta(\cdot)$.

Proof. (1) Theorem 1(2) means that $\lim_{k \rightarrow +\infty} L_\beta(\omega^k) = L_\beta(\omega^*)$ for each $\omega^* \in \Omega$. Now we divide into two cases for further analysis.

Case 1. There exists an integer k_0 such that $L_\beta(\omega^{k_0}) = L_\beta(\omega^*)$. By Lemma 4, for all $k \geq k_0$, we have

$$\delta \|z^{k+1} - z^k\|^2 \leq L_\beta(\omega^k) - L_\beta(\omega^{k+1}) \leq L_\beta(\omega^{k_0}) - L_\beta(\omega^*) = 0.$$

As a result, $x^{k+1} = x^k$, $y^{k+1} = y^k$. Also, from inequality (3.17), for all $k \geq k_0 + 2$, we have $\lambda^{k+1} = \lambda^k$ and $\omega^{k+1} = \omega^k$.

Case 2. Assume for all $k \geq 1$, $L_\beta(\omega^k) > L_\beta(\omega^*)$. Due to $d(\omega^k, \Omega) \rightarrow 0$, for any given $\varepsilon > 0$, there exists $k_1 > 0$, for $k > k_1$, such that $d(\omega^k, \Omega) < \varepsilon$. And, for any given $\eta > 0$, there exists $k_2 > 0$, for $k > \tilde{k} = \max\{k_1, k_2\}$, such that $L_\beta(\omega^k) < L_\beta(\omega^*) + \eta$.

Therefore,

$$d(\omega^k, \Omega^k) < \varepsilon, L_\beta(\omega^*) < L_\beta(\omega^k) < L_\beta(\omega^*) + \eta.$$

Because $\{\omega^k\}$ is bounded, according to Theorem 1, it is known that Ω is a non-empty compact set, and $L_\beta(\cdot)$ is a constant on Ω . Therefore, using Lemma 1, for all $k > \tilde{k}$,

$$\phi'(L_\beta(\omega^k) - L_\beta(\omega^*))d(0, \partial L_\beta(\omega^k)) \geq 1.$$

Because $\phi'(L_\beta(\omega^k) - L_\beta(\omega^*)) > 0$, therefore

$$\frac{1}{\phi'(L_\beta(\omega^k) - L_\beta(\omega^*))} \leq d(0, \partial L_\beta(\omega^k)). \quad (3.18)$$

Due to the convexity of the function ϕ , it can be concluded that

$$\phi(L_\beta(\omega^k) - L_\beta(\omega^*)) - \phi(L_\beta(\omega^{k+1}) - L_\beta(\omega^*)) \geq \phi'(L_\beta(\omega^k) - L_\beta(\omega^*))(L_\beta(\omega^k) - L_\beta(\omega^{k+1})).$$

Therefore,

$$0 \leq L_\beta(\omega^k) - L_\beta(\omega^{k+1}) \leq \frac{\phi(L_\beta(\omega^k) - L_\beta(\omega^*)) - \phi(L_\beta(\omega^{k+1}) - L_\beta(\omega^*))}{\phi'(L_\beta(\omega^k) - L_\beta(\omega^*))}. \quad (3.19)$$

Combining inequalities (3.18) and (3.19) with Lemma 6, one has

$$L_\beta(\omega^k) - L_\beta(\omega^{k+1}) \leq \tau(\|x^k - x^{k-1}\| + \|y^k - y^{k-1}\|)\Delta_{k,k+1}. \quad (3.20)$$

Note $\Delta_{p,q} = \phi(L_\beta(\omega^p) - L_\beta(\omega^*)) - \phi(L_\beta(\omega^q) - L_\beta(\omega^*))$.

For simplicity, let $\wedge_k = \|x^k - x^{k-1}\| + \|y^k - y^{k-1}\|$, and combining Lemma 4 and inequality (3.20), for all $k > \tilde{k}$, it can be derived that

$$\delta \|z^{k+1} - z^k\|^2 \leq \tau \wedge_k \Delta_{k,k+1}. \quad (3.21)$$

According to inequality (3.21), it can be concluded that

$$\|z^{k+1} - z^k\|^2 \leq \frac{\tau \wedge_k \Delta_{k,k+1}}{\delta}.$$

Furthermore, we have

$$5 \|z^{k+1} - z^k\| \leq 2 \wedge_k^{1/2} \times \left(\frac{5}{2} \sqrt{\frac{\tau}{\delta}} \Delta_{k,k+1}^{1/2}\right) \leq \wedge_k + \frac{25\tau}{4\delta} \Delta_{k,k+1}. \quad (3.22)$$

For inequality (3.22), summing up from $\tilde{k} + 1$ to m , it can be concluded that

$$5 \sum_{k=\tilde{k}+1}^m \|z^{k+1} - z^k\| \leq \sum_{k=\tilde{k}+1}^m \wedge_k + \frac{25\tau}{4\delta} \Delta_{k,k+1}.$$

Note that $\phi'(L(\omega^{m+1}) - L(\omega^*)) \geq 0$. According to the above inequality and by using the Cauchy inequality, at the same time letting $m \rightarrow +\infty$, we have

$$\begin{aligned} & \left(\frac{5\sqrt{2}}{2} - 1\right) \sum_{k=\tilde{k}+1}^m \|x^{k+1} - x^k\| + \left(\frac{5\sqrt{2}}{2} - 1\right) \|y^{k+1} - y^k\| \\ & \leq \|x^{\tilde{k}+1} - x^{\tilde{k}}\| + \|y^{\tilde{k}+1} - y^{\tilde{k}}\| + \frac{25\tau}{4\delta} \phi(L(\omega^{\tilde{k}}) - L(\omega^*)) < +\infty. \end{aligned}$$

Therefore, $\sum_{k=0}^{\infty} \|x^{k+1} - x^k\| < +\infty$ and $\sum_{k=0}^{\infty} \|y^{k+1} - y^k\| < +\infty$. By (3.17), it can be concluded that

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

Additionally, we have noticed that

$$\|\omega^{k+1} - \omega^k\| = (\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2 + \|\lambda^{k+1} - \lambda^k\|^2)^{\frac{1}{2}}$$

$$\leq \|x^{k+1} - x^k\| + \|y^{k+1} - y^k\| + \|\lambda^{k+1} - \lambda^k\|.$$

Hence, $\sum_{k=0}^{\infty} \|\omega^{k+1} - \omega^k\| < +\infty$.

In other words,

$$\lim_{k \rightarrow \infty} \|\omega^{k+1} - \omega^k\| = 0.$$

(2) From (1) we know that $\{\omega^k\}$ is a Cauchy sequence, and therefore it converges, and then from Theorem 1 (3), we know that $\{\omega^k\}$ converges to the critical point of the sequence $L_\beta(\cdot)$ as a whole. \square

Next, we wonder if the assumption about the boundedness of $\{\omega_k\}$ is reasonable, which will be ensured by the following lemma.

Lemma 7. Suppose that Assumption 1 and the following conditions (1)–(3) hold.

(1) A is a row full rank matrix.

(2) The relaxation factor $s \in (\frac{3-\sqrt{3}}{2}, \frac{3+\sqrt{3}}{2})$.

(3) The function $\bar{f}(x^k) = d_C^2(x^k) - \frac{\|\nabla d_C^2(x^k)\|^2}{2\beta(2s-1)\|A\|^2}$ has a lower bound and is coercive, i.e.,

$$\inf \bar{f}(x^k) > -\infty \text{ and } \liminf_{\|x_k\| \rightarrow +\infty} \bar{f}(x^k) = +\infty.$$

Then, the sequence $\{\omega_k\}$ generated by Algorithm 1 is bounded.

Proof. According to $L_\beta(\omega^k)$ being monotonically decreasing, $L_\beta(\omega^k) \leq L_\beta(\omega^0)$, which, combined with (3.11), we have

$$\begin{aligned} L_\beta(\omega^0) &\geq L_\beta(\omega^k) \\ &= d_C^2(x^k) + \delta_Q(y^k) - \langle \lambda^k, Ax^k - y^k \rangle + \frac{\beta}{2} \|Ax^k - y^k\|^2 \\ &= d_C^2(x^k) + \delta_Q(y^k) - \langle (A^T)^+ \nabla d_C^2(x^k) + \beta(1-s), Ax^k - y^k \rangle \\ &= d_C^2(x^k) + \delta_Q(y^k) - \frac{\|\nabla d_C^2(x^k)\|^2}{2\beta(2s-1)\|A\|^2} + \frac{(2s-1)\beta}{2} \left\| Ax^k - y^k - \frac{(A^T)^+ \nabla d_C^2(x^k)}{(2s-1)\beta} \right\|^2, \end{aligned}$$

where $(A^T)^+$ represents the Morre-Penrose inverse of A^T .

Therefore, one has

$$\begin{cases} \bar{f}(x^k) + \delta_Q(y^k) + \frac{(2s-1)\beta}{2} \left\| Ax^k - y^k - \frac{(A^T)^+ \nabla d_C^2(x^k)}{(2s-1)\beta} \right\|^2 \leq L_\beta(\omega^0) < +\infty, \\ \inf_y \delta_Q(y) > -\infty, \quad \inf_{x^k} \bar{f}(x^k) > -\infty, \\ s \in \left(\frac{3-\sqrt{3}}{2}, \frac{3+\sqrt{3}}{2} \right), \quad (s > \frac{1}{2}). \end{cases}$$

It is not difficult to deduce that $Ax^k - y^k - \frac{(A^T)^+ \nabla d_C^2(x^k)}{(2s-1)\beta}$ is bounded, so $\nabla d_C^2(x^k)$ is bounded. Furthermore, combining (3.11), we can know that $\{\lambda^k\}$ is bounded. Therefore, the boundedness of $\{\omega^k\}$ has been proven. \square

4. Convergence rate analysis

In this section, we further discuss the convergence rate of PSRADMM under the premise of strong convergence. The main results are as follows.

Theorem 3. Assume that $\{\omega^k\}$ is a sequence generated by PSRADMM and converges to $\{\omega^*\}$. Let $L_\beta(\cdot)$ satisfy the KL property at ω^* and the correlation function $\varphi(t) = ct^{1-\theta}$, $\theta \in [0, 1)$, $c > 0$. Then, the following conclusions hold.

- (1) If $\theta = 0$, then sequence $\{\omega^k\}$ converges after finite iterations, i.e., there exists k such that $\omega^k = \omega^*$.
- (2) If $\theta \in (0, \frac{1}{2}]$, then there exists $\tau \in [0, 1)$, such that $\|\omega^k - \omega^*\| \leq O(\tau^k)$, i.e., PSRADMM is linearly convergent.
- (3) If $\theta \in (\frac{1}{2}, 1)$, then $\|\omega^k - \omega^*\| \leq O(k^{\frac{1-\theta}{1-2\theta}})$, i.e., PSRADMM is sublinearly convergent.

Proof. For the case that $\theta = 0$, we have $\varphi(t) = ct$, hence $\varphi'(t) = c$. If $\omega^k \neq \omega^*$ for all k , then for a sufficiently large k , it is known by the KL property, and $cd(0, \partial L_\beta(\omega^k)) \geq 1$, which contradicts (3.16).

Assume $\theta > 0$, and combine (3.22).

$$5 \sum_{k=\tilde{k}+1}^{\infty} \Lambda_{k+1} \leq \sum_{k=\tilde{k}+1}^{\infty} \Lambda_k + \frac{25\tau}{4\delta} \Delta_{k,k+1}, \forall k > \tilde{k}. \quad (4.1)$$

Note that as $L_\beta(\cdot)$ satisfies the KL property at ω^* , one has

$$\varphi'(L_\beta(\omega^k) - L_\beta(\omega^*))d(0, \partial L_\beta(\omega^k)) \geq 1, \forall k \geq \tilde{k},$$

and because $\varphi(t) = ct^{1-\theta}$, the above inequality is equivalent to

$$(L_\beta(\omega^k) - L_\beta(\omega^*))^\theta \leq c(1-\theta)d(0, \partial L_\beta(\omega^k)), \forall k \geq \tilde{k}. \quad (4.2)$$

Also, due to (3.16), we have

$$d(0, \partial L_\beta(\omega^k)) \leq \tau \Lambda_k, \quad (4.3)$$

and, combined with (4.2) and (4.3), there exists $\gamma > 0$ such that

$$\varphi(L_\beta(\omega^k) - L_\beta(\omega^*)) = c(L_\beta(\omega^k) - L_\beta(\omega^*))^{1-\theta} \leq \gamma \Lambda_k^{\frac{1-\theta}{\theta}}, \forall k \geq \tilde{k}.$$

By (4.1), it can be concluded that

$$5 \sum_{k=\tilde{k}+1}^{\infty} \Lambda_{k+1} \leq \sum_{k=\tilde{k}+1}^{\infty} \Lambda_k + \frac{25\tau}{4\delta} \Lambda_k^{\frac{1-\theta}{\theta}}, \forall k > \tilde{k}. \quad (4.4)$$

Based on inequality (4.4), further proof will be completed using the relevant conclusions of Attouch and Bolte [26].

- (1) If $\theta \in [0, \frac{1}{2})$, then, according to [26], there exist $c_1 > 0$ and $\tau \in [0, 1)$ such that

$$\|x^k - x^*\| \leq c_1 \tau^k, \text{ thus } \|x^k - x^*\| \leq O(\tau^k). \quad (4.5)$$

- (2) If $\theta \in (\frac{1}{2}, 1)$, then, according to [26], there exists $c_2 > 0$ such that

$$\|x^k - x^*\| \leq c_2 k^{\frac{1-\theta}{\theta}}, \text{ thus } \|x^k - x^*\| \leq O(k^{\frac{1-\theta}{1-2\theta}}). \quad (4.6)$$

Furthermore, from the Eq (3.11), Lemma 3, and the L -Lipshitz continuity of ∇d_C^2 , it can be inferred that

$$\begin{aligned} \|\lambda^k - \lambda^*\| &= \|(A^T)^+ [\nabla d_C^2(x^k) - \nabla d_C^2(x^*)] + \beta(1-s)(Ax^k - y^k - Ax^* + y^*)\| \\ &\leq O(\Lambda_k). \end{aligned}$$

Finally, according to (4.4)–(4.6), it can be concluded that conclusions (1)–(3) have been verified. \square

5. Numerical examples

In this section, we validated the feasibility and effectiveness of the PSRADMM compared to the WPADMM (1.3) by comparing the number of iterations and running time. The entire code is implemented in MATLAB R2018b. All numerical calculations were performed on the Lenovo Xiaoxin Air 13IWL, with a CPU clock of 1.6GHz and a memory of 8.00GB.

When C is a non-convex closed set and Q is a closed convex set, the above algorithm still holds true. For details, please see reference [29]. We consider a class of important problems that are often applied in practical fields such as image reconstruction and signal processing, that is, compressed sensing problems, which involves recovering a sparse signal x from a linear system $Ax = b$. Transform the problem into a non-convex SFP, that is, find a vector $x \in \mathbb{R}^n$ which satisfies

$$x \in C, Ax = b,$$

where $C = \{x \in \mathbb{R}^n : \|x\|_0 \leq 1\}$ (here, $\|x\|_0$ represents the number of non-zero elements in vector x), $b \in \mathbb{R}^m$, and the set $Q = \{b\}$. Therefore, this section starts from the perspective of solving non-convex SFP and solves the decompression sensing problem. We transform problem (1.2) into the following optimization problem, and take $G = uI - \beta A^T A$:

$$\begin{aligned} \min d_C^2(x) \\ \text{s.t. } Ax = b. \end{aligned} \tag{5.1}$$

Using WPADMM to solve (5.1), we have

$$\begin{cases} x^{k+1} = \frac{1}{1+\beta}(-\beta A^T b - A^T \lambda^k), \\ \lambda^{k+1} = \lambda^k - s\beta(Ax^{k+1} - b). \end{cases}$$

Using PSRADMM to solve (5.1), we have

$$\begin{cases} \lambda^{k+\frac{1}{2}} = \lambda^k - \gamma\beta(Ax^k - b), \\ x^{k+1} = \frac{1}{1+\beta}(-\beta A^T b - A^T \lambda^{k+\frac{1}{2}}), \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} - b). \end{cases}$$

In numerical experiments, A is a randomly generated $m \times n$ matrix. The initial point x^0 belongs to $\text{rand}(n, 1)$, and the initial point λ^0 is taken as the zero vector. In addition, the selection of correlation

coefficients in WPADMM and PSRADMM is consistent, which are $\beta = 12, s = r = \frac{10}{12}$. When $\|\omega^{k+1} - \omega^k\| < \varepsilon$, the program terminates. In this section, we set $\varepsilon = 10^{-5}$.

We solve problem (5.1) using WPADMM and PSRADMM, and selected six different dimensions of A for testing. Table 1 provides a comprehensive overview of the primary test outcomes pertaining to both algorithms, encompassing vital metrics such as iteration number and runtime. Figures 1 and 2 show the main numerical test results of the two algorithms, where iter (k) and Time/s represent the number of iterations and CPU calculation time in seconds, respectively.

Upon thorough examination of Figures 1 and 2, we can see that the PSRADMM consistently outperforms the WPADMM in terms of both runtime efficiency and convergence behavior, as reflected by the iteration number. This observation suggests that the PSRADMM exhibits enhanced computational efficiency and convergence speed compared to the WPADMM.

Table 1. Comparison of runtime and iteration number between WPADMM and PSRADMM.

| m | n | WPADMM | | PSRADMM | |
|------|------|-------------|---------|-------------|---------|
| | | Iter(k) | Time(s) | Iter(k) | Time(s) |
| 500 | 600 | 98 | 0.22677 | 89 | 0.15130 |
| 600 | 700 | 96 | 0.32404 | 87 | 0.24033 |
| 700 | 800 | 94 | 0.42846 | 86 | 0.33881 |
| 800 | 900 | 93 | 0.55322 | 85 | 0.45505 |
| 900 | 1000 | 92 | 0.67306 | 84 | 0.56656 |
| 1000 | 1000 | 92 | 0.73150 | 83 | 0.62536 |

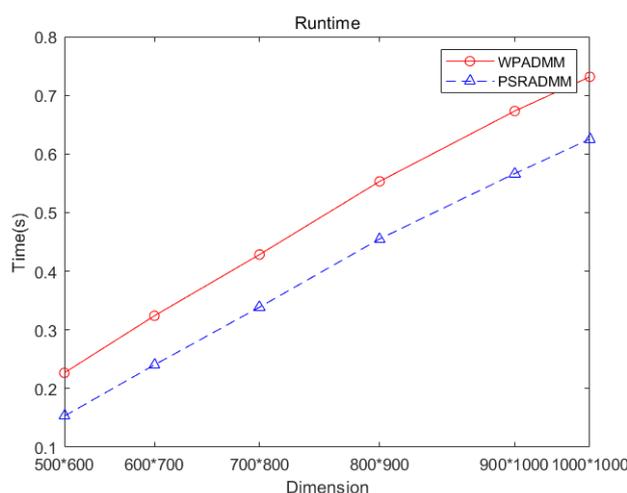


Figure 1. Comparison of runtime between WPADMM and PSRADMM.

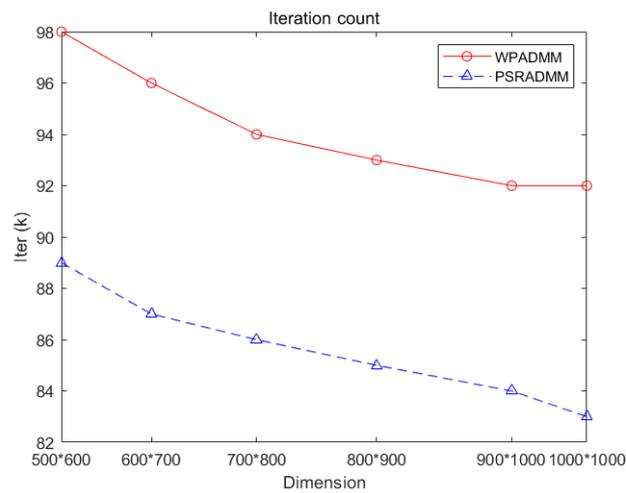


Figure 2. Comparison of iteration number between WPADMM and PSRADMM.

Inspired by Godwin et al. [8], we apply the algorithm to image processing for deblurring. First, we blur the original image using a Gaussian blur kernel with a size of 15×15 and a standard deviation with a size of 2 to obtain the blurred image (See Figures 3–6).



Figure 3. From left to right: original image of Cameraman, blurred image of Cameraman, WPADMM restored image of Cameraman, PSRADMM restored image of Cameraman.



Figure 4. From left to right: original image of Pepper, blurred image of Pepper, WPADMM restored image of Pepper, PSRADMM restored image of Pepper.

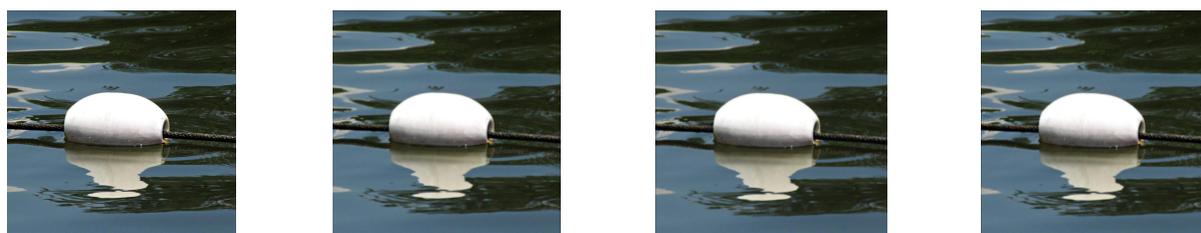


Figure 5. From left to right: original image of Boat, blurred image of Boat, WPADMM restored image of Boat, PSRADMM restored image of Boat.



Figure 6. From left to right: original image of Alex, blurred image of Alex, WPADMM restored image of Alex, PSRADMM restored image of Alex.

We can see from the Figures 3–6 that PSRADMM has a better deblurring effect than WPADMM. However, we still need to evaluate the effectiveness of deblurring based on two indicators: peak signal-to-noise ratio (PSNR:= $10\log_{10}\frac{255^2}{\sqrt{MSE}}$) and structural similarity index (SSIM).

From Table 2, we can see that the PSNR of PSRADMM is better than that of WPADMM restored images. Measure 1: A higher PSNR value indicates that the difference between the processed image and the original image is small, usually indicating higher image quality. The PSNR value of PSRADMM is higher. Measure 2: An SSIM value close to 1 indicates a higher similarity in brightness, contrast, and structure between the processed image and the original image, usually indicating better image quality. The SSIM value of the PSRADMM algorithm is closer to 1.

Table 2. The numerical results of the experiment.

| Index | Algorithm | Value |
|-------|-----------|--------|
| PSNR | PSRADMM | 43.22 |
| | WPADMM | 21.33 |
| SSIM | PSRADMM | 0.7656 |
| | WPADMM | 0.5274 |

Finally, we validate the feasibility and effectiveness of the algorithm through image enhancement. Figures 7 and 8 show the images displayed using image enhancement techniques using the WPADMM and PSRADMM algorithms, respectively.



Figure 7. Image enhancement of WPADMM.



Figure 8. Image enhancement of PSRADMM.

From Table 3, we can see that the image enhancement effect of PSRADMM is better than that of the WPADMM by comparing PSNR and SSIM.

Table 3. The numerical results of the experiment.

| Index | Algorithm | Value |
|-------|-----------|--------|
| PSNR | PSRADMM | 40.52 |
| | WPADMM | 28.42 |
| SSIM | PSRADMM | 0.9093 |
| | WPADMM | 0.6574 |

6. Conclusions

In this article, we propose a partially symmetric regularized alternating direction multiplier method to solve non-convex split feasibility problems. A conclusion on global convergence is drawn. When the augmented Lagrangian function satisfies the KL property, the algorithm has strong convergence. In addition, this article provides the convergence rate of the algorithm. Finally, in the compressed sensing problem, numerical experiments have shown that our proposed algorithm outperforms existing weighted proximal symmetric regularized alternating direction multiplier methods for solving non-convex SFP. And, we applied the PSRADMM to image processing with good experimental results.

Author contributions

Yue Zhao: Methodology; Meixia Li: Supervision; Xiaowei Pan: Software, Visualization; Jingjing Tan: Writing—original draft, Writing—review & editing. All authors have read and approved the final manuscript.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This project is supported by Natural Science Foundation of China (Grant No.12071351, 11571120) and Shandong Provincial Natural Science Foundation (Grant No.ZR2020MA027, ZR2022MA034).

Conflict of interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projections in a product space, *Numer. Algor.*, **8** (1994), 221–239. <https://doi.org/10.1007/BF02142692>
2. C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, *Inverse Problems*, **18** (2002), 441–453. <https://doi.org/10.1088/0266-5611/18/2/310>
3. A. Moudafi, A. Gibali, $I_1 - I_2$ regularization of split feasibility problems, *Numer. Algor.*, **78** (2018), 739–757. <https://doi.org/10.1007/s11075-017-0398-6>
4. B. Qu, C. Y. Wang, N. H. Xiu, Analysis on Newton projection method for the split feasibility problem, *Comput. Optim. Appl.*, **67** (2017), 175–199. <https://doi.org/10.1007/s10589-016-9884-3>
5. B. Tan, X. L. Qin, J.-C. Yao, Strong convergence of self-adaptive inertial algorithms for solvingsplit variational inclusion problems with applications, *J. Sci. Comput.*, **87** (2021), 20. <https://doi.org/10.1007/s10915-021-01428-9>
6. A.-S. O.-E. Owolabi, A. Taiwo, L. O. Jolaoso, O. T. Mewomo, A self-adaptive hybrid inertial algorithm for split feasibility problems in Banach spaces, *Int. J. Nonlinear Anal. Appl.*, **13** (2022), 791–812. <https://doi.org/10.22075/ijnaa.2021.23402.2532>
7. B. Tan, X. L. Qin, X. F. Wang, Alternated inertial algorithms for split feasibility problems, *Numer. Algor.*, **95** (2024), 773–812. <https://doi.org/10.1007/s11075-023-01589-8>
8. E. C. Godwin, C. Izuchukwu, O. T. Mewomo, Image restorations using a modified relaxed inertial technique for generalized split feasibility problems, *Math. Method. Appl. Sci.*, **46** (2023), 5521–5544. <https://doi.org/10.1002/mma.8849>
9. H. Yu, F. H. Wang, A new relaxed method for the split feasibility problem in Hilbert spaces, *Optimization*, **73** (2024), 1417–1432. <https://doi.org/10.1080/02331934.2022.2158036>
10. G. Y. Li, T. K. Pong, Douglas–Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems, *Math. Program.*, **159** (2014), 371–401. <https://doi.org/10.1007/s10107-015-0963-5>
11. A. Moudafi, A semi-alternating algorithm for solving nonconvex split equality problems, *Numer. Funct. Anal. Optim.*, **42** (2021), 1747–1756. <https://doi.org/10.1080/01630563.2021.1933529>

12. J. Z. Chen, M. Postolache, Y. H. Yao, S-subgradient projection methods with S-subdifferential functions for nonconvex split feasibility problems, *Symmetry*, **11** (2019), 1517. <https://doi.org/10.3390/sym11121517>
13. J. Z. Chen, K. Chen, T. C. Yin, S-subgradient projection algorithm with inertial technique for non-convex split feasibility problems, *U.P.B. Sci. Bull. Series A*, **84** (2022), 3–12.
14. K. Guo, C. Zhu, Inexact averaged projection algorithm for nonconvex multiple-set split feasibility problems, *Journal of Mathematical Research with Applications*, **40** (2020), 534–542. <https://doi.org/10.3770/j.issn:2095-2651.2020.05.009>
15. A. Gibali, S. Sabach, S. Voldman, Non-convex split feasibility problems: models, algorithms and theory, *Open Journal of Mathematical Optimization*, **1** (2020), 1. <https://doi.org/10.5802/ojmo.1>
16. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends Mach. Learn.*, **3** (2011), 1–122. <https://doi.org/10.1561/22000000016>
17. J. C. Bai, Y. Chen, X. Yu, H. C. Zhang, Generalized asymmetric forward-backward-adjoint algorithms for convex-concave saddle-point problem, *J. Sci. Comput.*, **102** (2025), 80. <https://doi.org/10.1007/s10915-025-02802-7>
18. J. C. Bai, K. Guo, J. L. Liang, Y. Jing, H. C. Su, Accelerated symmetric ADMM and its applications in large-scale signal processing, *J. Comput. Math.*, **42** (2024), 1605–1626. <https://doi.org/10.4208/jcm.2305-m2021-0107>
19. J. C. Bai, L. Y. Jia, Z. Peng, A new insight on augmented lagrangian method with applications in machine learning, *J. Sci. Comput.*, **99** (2024), 53. <https://doi.org/10.1007/s10915-024-02518-0>
20. J. C. Bai, M. Zhang, H. C. Zhang, An inexact ADMM for separable nonconvex and nonsmooth optimization, *Comput. Optim. Appl.*, in press. <https://doi.org/10.1007/s10589-024-00643-y>
21. B. S. He, F. Ma, X. M. Yuan, Convergence study on the symmetric version of ADMM with larger step sizes, *SIAM J. Imaging Sci.*, **9** (2016), 1467–1501. <https://doi.org/10.1137/15M1044448>
22. J. C. Bai, J. C. Li, F. M. Xu, H. C. Zhang, Generalized symmetric ADMM for separable convex optimization, *Comput. Optim. Appl.*, **70** (2018), 129–170. <https://doi.org/10.1007/s10589-017-9971-0>
23. R. T. Rockafellar, *Convex analysis*, Princeton: Princeton University Press, 1970. <https://doi.org/10.1515/9781400873173>
24. R. T. Rockafellar, R. J. B. Wets, *Variational analysis*, Berlin, Heidelberg: Springer, 1998. <https://doi.org/10.1007/978-3-642-02431-3>
25. J. Bolte, A. Daniilidis, A. Lewis, The KŁojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, *SIAM J. Optimiz.*, **17** (2007), 1205–1223. <https://doi.org/10.1137/050644641>
26. H. Attouch, J. Bolte, B. F. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods, *Math. Program.*, **137** (2013), 91–129. <https://doi.org/10.1007/s10107-011-0484-9>

-
27. F. H. Wang, W. F. Cao, Z. B. Xu, Convergence of multi-block Bregman ADMM for nonconvex composite problems, *Sci. China Inf. Sci.*, **61** (2018), 122101. <https://doi.org/10.1007/s11432-017-9367-6>
28. Y. Nesterov, *Introductory lectures on convex optimization: a basic course*, 1 Eds., New York: Springer, 2004. <https://doi.org/10.1007/978-1-4419-8853-9>
29. J. B. Jian, P. J. Liu, X. Z. Jiang, A partially symmetric regularized alternating direction method of multipliers for nonconvex multi-block optimization, (Chinese), *Acta Mathematica Sinica, Chinese Series*, **64** (2021), 1005–1026. <https://doi.org/10.12386/A2021sxxb0084>



AIMS Press

© 2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)