

Research Article

Ying Zhao, Heng-you Lan*, and Hai-yang Xu

Convergence of Peaceman-Rachford splitting method with Bregman distance for three-block nonconvex nonseparable optimization

<https://doi.org/10.1515/dema-2024-0036>

received July 27, 2023; accepted May 17, 2024

Abstract: It is of strong theoretical significance and application prospects to explore three-block nonconvex optimization with nonseparable structure, which are often modeled for many problems in machine learning, statistics, and image and signal processing. In this article, by combining the Bregman distance and Peaceman-Rachford splitting method, we propose a novel three-block Bregman Peaceman-Rachford splitting method (3-BPRSM). Under a general assumption, global convergence is presented via optimality conditions. Furthermore, we prove strong convergence when the augmented Lagrange function satisfies Kurdyka-Łojasiewicz property. In addition, if the association function possessing the Kurdyka-Łojasiewicz property exhibits a distinctive structure, then linear and sublinear convergence rate of 3-BPRSM can be guaranteed. Finally, a preliminary numerical experiment demonstrates the effectiveness.

Keywords: convergence analysis, Peaceman-Rachford splitting method with Bregman distance, three-block nonconvex nonseparable optimization, Kurdyka-Łojasiewicz property, optimality condition

MSC 2020: 90C26, 65K05, 49K35, 41A25

1 Introduction

As we all know, optimization plays a crucial role in solving various complex models across diverse fields. One notable example is the application of heuristic computing with sequential quadratic programming to solve a nonlinear hepatitis B virus model, which is proposed by Umar et al. [1]. Optimization techniques are employed to fine-tune the parameters of the model, which can enable a more accurate representation of the complex dynamics of hepatitis B virus transmission. The iterative nature of sequential quadratic programming allows researchers to minimize the model's nonlinearity and to enhance its predictive capabilities. This approach not only contributes to a deeper understanding of biological mechanisms underlying hepatitis B virus propagation, but also highlights a pivotal role of optimization to refine and optimize mathematical models in the field of virology. The integration of heuristic computing with optimization techniques proves instrumental in advancing research and providing valuable insights into the dynamics of infectious diseases. Other models solved by the optimization process can be found in [2–8] and the references therein.

Nonconvex optimization is indispensable in practical engineering applications and offers versatile solutions to complex problems across various domains. For instance, Bian and Chen [9] addressed image

* **Corresponding author: Heng-you Lan**, College of Mathematics and Statistics, Sichuan University of Science and Engineering, South Sichuan Center for Applied Mathematics, Zigong 643000, China, e-mail: hengyoulan@163.com

Ying Zhao: College of Mathematics and Statistics, Sichuan University of Science and Engineering, Zigong 643000, China, e-mail: z_zhaoying@163.com

Hai-yang Xu: College of Mathematics and Statistics, Sichuan University of Science and Engineering, Zigong 643000, China, e-mail: haiyangxu96@163.com

processing in Examples 5.1 and 5.2, Xu et al. applied to a typical compressed sensing problem, i.e., the sparse signal recovery problem [10, (52), p. 1020] for solving the brand new feature of $l_{\frac{1}{2}}$ regularization, Lin et al. tested the performance of linearized alternating direction method with parallel splitting and adaptive penalty on sparse representation and low-rank recovery problems (2), (3), and (5) on page 289 of [11]. Be with meanwhile, Ames and Hong extended sparse principal component analysis and estimated the solution of any penalized eigenproblem of the form (4.1) [12, p. 741] with nonconcave objective function in general. Further, Lin et al. demonstrated utility of the alternating direction method of multipliers (ADMM) algorithm through identified sparsity structures, among mass-spring system, network with 100 unstable nodes and block sparsity: A bio-chemical reaction example (see, respectively, pages 2429 and 2430 in [13]). Thus, it can be seen that studying nonconvex optimization holds both strong theoretical significance and promising application prospects. Especially, for giving a partial answer, the open problem on convergence analysis of ADMM for solving large-scale nonconvex separable optimization, and Guo et al. [14] considered the following nonconvex optimization for two-block problem:

$$\begin{aligned} \min \quad & \xi(x) + \zeta(y) \\ \text{s.t.} \quad & Mx + y = d, \end{aligned} \quad (1.1)$$

which can be caused based on acquiring the sparsest solution of a linear system [15, p. 14]. Moreover, the formulation as (1.1), finding a minimizer of some composite objective functions in the fields of signal and image processing and machine learning, was investigated by Wang et al. [15], where $M \in \mathbb{R}^{m \times n}$ is a given matrix, $d \in \mathbb{R}^m$ is the observed data, $\xi: \mathbb{R}^n \rightarrow \mathbb{R}$ is usually a (quadratic, or logistic) loss function, and $\zeta: \mathbb{R}^m \rightarrow \mathbb{R}$ is often a regularizer such as the l_1 norm or $l_{\frac{1}{2}}$ quasi-norm.

However, when one considers the case that location parameters of a linear regression model are different for each groups, that is, it can be regarded that location parameters express individual effects of groups, one of the composite objective functions in the form of (1.1) is often nonseparable. For example, Ohishia et al. [16] presented that the parameter vector μ is estimated by minimizing the penalized residual sum of squares, which includes nonseparable objective function $\|y - X\beta - R\mu\|_2^2$, where β is a vector of regression coefficients. Indeed, nonseparable function exists objectively in image and signal processing including fused Lasso [16], group Lasso [17], total variation regularization [18], statistics, and machine learning such as large-scale problems [19]. Because of impersonally existing nonseparable structure, Liu et al. [20] challenged convergence analysis of a splitting method for a constrained nonconvex nonseparable optimization as follows:

$$\begin{aligned} \min \quad & \xi(x) + \zeta(y) + \ell(x, y) \\ \text{s.t.} \quad & Mx + y = d, \end{aligned} \quad (1.2)$$

which reduces to (1.1) when the nonseparable function $\ell(x, y) \equiv 0$ for any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. As He and Yuan [21] pointed out “many applications such as sparse and/or low-rank optimization, compressive sensing, statistical learning, computer vision and large-scale distributed wireless network can be modeled or reformulated as a convex minimization model with linear constraints and a separable objective function in form of the sum of more than one function,” numerous problems in practical engineering applications can be effectively modeled as three-block nonconvex (nonseparable) optimization problem. For instance, He and Yuan successfully addressed the patch low-rank image decomposition model ([21, (7.2), p. 815]). In addition, the SCAD- l_2 model extensively explored by Zeng and Xie [22] serves as a notable example. Furthermore, Zhang et al. [23] applied pADMMz to resolve the challenge of precisely extracting background/foreground from a given video. So the models (1.1) or (1.2) may not fully describe the aforementioned practical application problems and additional variables need to be added to better represent the complexity of these application problems. Hence, it is necessary and valuable to explore the following three-block nonconvex constrained optimization with nonseparable coupled term:

$$\begin{aligned} \min \quad & f(x) + g(y) + h(z) + l(x, y, z) \\ \text{s.t.} \quad & Ax + y + z = b, \end{aligned} \quad (1.3)$$

where $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is proper and lower semicontinuous, $g: \mathbb{R}^m \rightarrow \mathbb{R}$, $h: \mathbb{R}^m \rightarrow \mathbb{R}$, and $l(x, y, z): \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ is continuously differentiable, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are the given matrix

and vector, respectively. If the nonseparable term $l(x, y, z) = 0$ for each $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, (1.3) is equivalent to a separable three-block optimization problem with linear constraints due to He and Yuan [21]. In fact, the form of (1.3) has many applications in various fields such as image processing [9,17,21], video processing [23], and background/foreground extraction [23,24].

On the other hand, to solve (1.1) or (1.2), many methods have been presented, such as ADMM [25], Douglas-Rachford splitting method [26], Peaceman-Rachford splitting method (PRSM) [27], and forward-backward splitting algorithm [28]. However, one can know that the biggest difference between PRSM and ADMM in the iterative process is to add the middle update item, which can also be seen as a symmetric version of ADMM. As Gabay [29] pointed out, although PRSM is less robust and requires more constraints than ADMM to guarantee convergence, PRSM converges faster than ADMM subject to convergence. Many scholars found that adding appropriate Bregman distance to ADMM can effectively simplify subproblem calculation or make the subproblem closed solution and further can improve numerical effect. Indeed, Li and Pong [30] proposed a variant of ADMM, that is, adding Bregman distance parameter to the second subproblem of the classical ADMM. More significant advances have been accomplished in the past few years upon ADMM or its variant version. For more detail, one can refer to [31,32] and the references therein.

In 2008, Wang et al. [33] indicated that as a natural extension of ADMM, multiblock ADMM is a widely used scheme and has also been found very useful in solving various nonconvex optimization problems. Meanwhile, considering the three-block nonconvex optimization problem with linear equality constraint, where the coefficient matrix of one variable is negative identity matrix, Yang et al. [24] studied the classical ADMM of Lagrange multiplier update step with relaxation factor. Furthermore, Chao et al. [34] declared that ADMM or its directly extend version may not converge, when the involved number of blocks is more than two, or there is a nonconvex function, or there is a nonseparable structure. Thus, a natural idea is to employ or extend PRSM to solve the problem (1.3). However, the work of Chen et al. [35] presented a counterexample that demonstrated the lack of convergence in the direct extension of the multiblock convex optimization problem. This discovery has aroused significant attention from the research community. Hence, PRSM or its direct extension may not convergent when the number of blocks exceeds two, or nonconvex functions are presented, or a nonseparable structure is involved. What is more, the study of PRSM is limited to two-block problem and a few three-block separable problems, and considering the three-block nonseparable problem has not been reported. Indeed, Bnouhachem and Rassias [36] proposed a Bregman proximal PRSM for solving a separable convex minimization model and investigated global convergence and convergence rate of the method in the ergodic sense under standard assumptions. Li et al. [37] applied strict contraction PRSM to a convex minimization problem with two linear constraints on nonseparable structure and proved convergence. In addition, Liu et al. [38] dealt with a special class of multiblock convex optimization problems, whose nonseparable structure is a quadratic function. Chao et al. [39] adopted the substitution method to consider the near block minimization method of multipliers, obtained the convergence of such multiblock optimization problems, and provided the corresponding iterative complexity results.

Inspired by the aforementioned work and challenged by investigation of the three-block nonseparable problems, the purpose and main contribution of this article is to propose and prove convergence of a three-block Bregman Peaceman-Rachford splitting method (3-BPRSM) for the nonconvex and nonseparable problem (1.3).

The novelty of this article can be summarized as follows:

- (N1) We extend PRSM from two-block to three-block, that is, the model (1.3) in this article is more general comparing with the existing work in [36–39], and prove its convergence and obtain convergence rate.
- (N2) Adding Bregman distance to the subproblem of $x \in \mathbb{R}^n$ allows the function $f(x)$ and the matrix A to obtain the descent properties of Lagrange function without any additional restrictions.
- (N3) The introduction of relaxation factor makes the iteration widely representative, including several types of ADMM and PRSM. Therefore, 3-BPRSM provides a relatively broad splitting algorithm framework for the three-block problem (1.3), which can uniformly analyze the theoretical properties of a large class of Bregman-type Peaceman-Rachford splitting algorithms.

The remainder of this article is organized as follows: some basic concepts and necessary foreshadowing for further analysis are presented in Section 2. We present 3-BPRSM and analyze its convergence for three-

block nonconvex and nonseparable problems in Section 3. In Section 4, a numerical example is implemented to illustrate and validate our major results. Finally, Section 5 concludes with some discussions of future work.

2 Preliminaries

Denote $\|\cdot\|$ as Euclidean norm of a vector. Let $\text{dom}f = \{x \in \mathbb{R}^n : f(x) < +\infty\}$ be the domain of a function $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. For a symmetric semipositive definite matrix H , $\|x\|_H^2 = x^T H x$, where $H \succcurlyeq (>) 0$ indicates that H is a symmetric semi-positive (positive) definite matrix. $\lambda_{\min}(H)\|x\|^2 \leq x^T H x \leq \lambda_{\max}(H)\|x\|^2$ holds for all $x \in \mathbb{R}^n$, where $\lambda_{\min}(H)$ and $\lambda_{\max}(H)$ represent the minimum eigenvalue and the maximum eigenvalue of the symmetric matrix H , respectively.

Definition 2.1. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$. If for all $x, y \in \mathbb{R}^n$, there exists $L_f > 0$ such that $|f(x) - f(y)| \leq L_f \|x - y\|$, then f is called L_f -Lipschitz continuous.

Definition 2.2. Let $\mathbb{S} \subseteq \mathbb{R}^n$. The distance from point $x \in \mathbb{R}^n$ to \mathbb{S} is defined as $d(x, \mathbb{S}) = \inf_{y \in \mathbb{S}} \|y - x\|$. In particular, when $\mathbb{S} = \emptyset$, $d(x, \mathbb{S}) = +\infty$.

Definition 2.3. Let $\mathbb{S} \subset \mathbb{R}^n$ be a nonempty convex set, and f be a function defined on \mathbb{S} . If for any $x_1, x_2 \in \mathbb{S}$ and each $\alpha \in (0, 1)$, the following inequality holds:

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2),$$

then f is said to be convex on \mathbb{S} . In particular, if $x_1 \neq x_2$, we have

$$f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha)f(x_2),$$

then f is called strictly convex on \mathbb{S} .

Further, if for all $x_1, x_2 \in \mathbb{S}$ and any $\alpha \in (0, 1)$, there exists a constant $\sigma > 0$ such that

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) - \frac{1}{2}\sigma\alpha(1 - \alpha)\|x_1 - x_2\|^2,$$

then f is said to be strongly α -convex.

Definition 2.4. Let $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function.

(i) The Fréchet subdifferential of f at $x \in \text{dom}f$ is defined as follows:

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

And in particular, $\hat{\partial}f(x) = \emptyset$ when $x \notin \text{dom}f$.

(ii) The limiting subdifferential of f at $x \in \text{dom}f$ is decided by

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

Definition 2.5. [40] For differentiable convex function $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$, the corresponding Bregman distance is defined as follows:

$$\Delta_\phi(x, y) = \phi(x) - \phi(y) - \langle \nabla \phi(y), x - y \rangle, \quad \forall x, y \in \mathbb{R}^n.$$

In particular, when $\phi(x) = \|x\|^2$, the corresponding Bregman distance can be reduced to $\|x - y\|^2$, which is the classical Euclidean distance [41].

Definition 2.6. If f is continuously differentiable in the first order on a nonempty open convex set $\mathbb{S} \subset \mathbb{R}^n$, then the following equality holds:

$$f(y) = f(x) + \nabla f(x)^\top(y - x) + o(\|y - x\|), \quad \forall x, y \in \mathbb{S}.$$

Definition 2.7. A differentiable function f defined on a nonempty open convex set $\mathbb{S} \subset \mathbb{R}^n$, is

(i) convex if and only if

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \quad \forall x, y \in \mathbb{S},$$

or

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0, \quad \forall x, y \in \mathbb{S}.$$

(ii) strongly σ -convex if and only if

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\sigma}{2} \|y - x\|^2, \quad \forall x, y \in \mathbb{S}.$$

Definition 2.8. The subdifferential of a proper lower semicontinuous function f has many basic and important properties as follows:

- (i) $\partial f(x) \subseteq \partial f(x)$ holds for all $x \in \mathbb{R}^n$, and $\partial f(x)$ is a closed set if $\hat{\partial} f(x)$ is a closed convex set.
- (ii) $x^* \in \partial f(x)$ if $x_k^* \in \partial f(x_k)$ and $\lim_{k \rightarrow \infty} (x_k, x_k^*) = (x, x^*)$, so $\partial f(x)$ is closed.
- (iii) If $\hat{x} \in \mathbb{R}^n$ is the local minimum of f , then $0 \in \partial f(\hat{x})$. \hat{x} is the critical point of f when $0 \in \partial f(\hat{x})$, the set of critical points of f is denoted by $\text{crit } f$.
- (iv) If $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable, then $\partial(f + g)(x) = \partial f(x) + \nabla g(x)$ owns for all $x \in \text{dom } f$.

Let Φ_η be the set of all concave functions $\varphi : [0, \eta] \rightarrow [0, +\infty)$ that satisfy the following conditions:

- (i) $\varphi(0) = 0$; (ii) φ is continuous at 0 and continuous differentiable on $(0, \eta)$; (iii) $\varphi'(t) > 0$ for each $t \in (0, \eta)$.

Lemma 2.9. [42] (Kurdyka-Łojasiewicz property (KLP)) Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper lower semicontinuous function, $\bar{x} \in \text{dom}(\partial f) = \{x \in \mathbb{R}^n : \partial f(x) \neq \emptyset\}$. Let $[\eta_1 < f < \eta_2] = \{x \in \mathbb{R}^n : \eta_1 < f(x) < \eta_2\}$. If there exist $\eta \in (0, +\infty)$, a neighborhood U of \bar{x} and a concave function $\varphi \in \Phi_\eta$, then the following inequality owns for all $x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$:

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

where f is said to have **KLP** at \bar{x} . Meanwhile, φ is the associate function of f with **KLP**.

Lemma 2.10. [40] Let $\Delta_\phi(x, y)$ be the corresponding Bregman distance for differentiable convex function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$. Then

- (i) nonnegative: $\Delta_\phi(x, y) \geq 0$ and $\Delta_\phi(x, x) = 0$ hold for all $x, y \in \mathbb{R}^n$;
- (ii) convexity: $\Delta_\phi(x, y)$ is convex at x , not necessarily convex at y ;
- (iii) strong convexity: If ϕ is strongly σ -convex, then $\Delta_\phi(x, y) \geq \frac{\sigma}{2} \|x - y\|^2$ for all $x, y \in \mathbb{R}^n$.

Lemma 2.11. [42] Assume that $H(x, y, z) = p(x) + q(y) + h(z)$, here $p : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ and $q, h : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ are proper lower semicontinuous functions. Then the following equality holds for all $(x, y, z) \in \text{dom } H = \text{dom } p \times \text{dom } q \times \text{dom } h$:

$$\partial H(x, y, z) = \partial_x H(x, y, z) \times \partial_y H(x, y, z) \times \partial_z H(x, y, z).$$

Lemma 2.12. [43] (Uniform KLP) Let Ω be a compact set and $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper and lower semicontinuous. If f takes a fixed constant on Ω and satisfies **KLP** in Lemma 2.9 at any point in Ω , then there exist

$\zeta > 0, \eta > 0, \varphi \in \Phi_\eta$ such that for all $\bar{x} \in \Omega$ and $x \in \{x \in \mathbb{R}^n : d(x, \Omega) < \zeta\} \cap [f(\bar{x}) < f(x) < f(\bar{x}) + \eta]$, the following inequality holds:

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

Lemma 2.13. [44] Suppose that $p : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and differentiable, and gradient ∇p is L_p -Lipschitz continuous. Then

$$|p(y) - p(x) - \langle \nabla p(x), y - x \rangle| \leq \frac{L_p}{2} \|y - x\|^2, \quad \forall x, y \in \mathbb{R}^n.$$

Lemma 2.14.

- (i) The limit of $\{e^k\}$ exists when $\{e^k\}$ monotonically descends and has a lower bound.
- (ii) If $\{e^k\}$ is monotonic and there exists an infinite subsequence $e^{k_j} \rightarrow e$, then $\lim_{k \rightarrow \infty} e^{k_j} = e$.

3 Algorithm and convergence analysis

Aiming at the problem (1.3), combining with the Bregman distance and PRSM, the following 3-BPRSM is proposed, and Figure 1 is a flowchart of the proposed 3-BPRSM:

Algorithm 1. 3-BPRSM

Input: Initial point $x^0 \in \text{dom } f, y^0 \in \text{dom } g, z^0 \in \text{dom } h, (x^0, y^0, z^0) \in \text{dom } l$, matrix $b, \lambda^0, \beta, r, s$, and the number of iteration steps $k = 0$.

1: **for** $k = 0, 1, 2, 3, \dots$ **do**

2: $x^{k+1} \in \underset{x}{\operatorname{argmin}} \{ \mathcal{L}_\beta(x, y^k, z^k, \lambda^k) + \Delta_\phi(x, x^k) \}$ (3.1)

3: $\lambda^{k+\frac{1}{2}} = \lambda^k - r\beta(Ax^{k+1} + y^k + z^k - b)$ (3.2)

4: $y^{k+1} \in \underset{y}{\operatorname{argmin}} \{ \mathcal{L}_\beta(x^{k+1}, y, z^k, \lambda^{k+\frac{1}{2}}) \}$ (3.3)

5: $z^{k+1} \in \underset{z}{\operatorname{argmin}} \{ \mathcal{L}_\beta(x^{k+1}, y^{k+1}, z, \lambda^{k+\frac{1}{2}}) \}$ (3.4)

6: **if** $\|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|_2 \leq \sqrt{m}10^{-4}$, **then**

7: **break**

8: **else**

9: $\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b)$ (3.5)

10: **end if**

11: **return** $x^{k+1}, \lambda^{k+\frac{1}{2}}, y^{k+1}, z^{k+1}$

12: **end for**

The augmented Lagrange function for (1.3) is defined as follows:

$$\mathcal{L}_\beta(x, y, z, \lambda) = f(x) + g(y) + h(z) + l(x, y, z) - \langle \lambda, Ax + y + z - b \rangle + \frac{\beta}{2} \|Ax + y + z - b\|^2, \quad (3.6)$$

where $\lambda \in \mathbb{R}^m$ is the Lagrange multiplier and $\beta (>0)$ is the penalty parameter. Suppose that Δ_ϕ is the Bregman

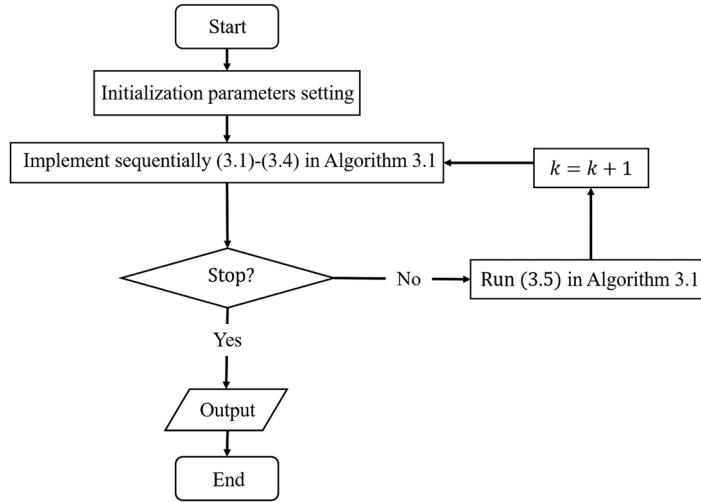


Figure 1: Flowchart of the proposed 3-BPRSM.

distance with respect to the differentiable convex function ϕ and $\omega = (x, y, z, \lambda)$, according to the augmented Lagrange function (3.6), one can define as follows:

$$\begin{cases} \partial_x \mathcal{L}_\beta(\omega) = \partial f(x) + \nabla_x l(x, y, z) - A^\top \lambda + \beta A^\top (Ax + y + z - b), \\ \partial_y \mathcal{L}_\beta(\omega) = \nabla g(y) + \nabla_y l(x, y, z) - \lambda + \beta (Ax + y + z - b), \\ \partial_z \mathcal{L}_\beta(\omega) = \nabla h(z) + \nabla_z l(x, y, z) - \lambda + \beta (Ax + y + z - b), \\ \partial_\lambda \mathcal{L}_\beta(\omega) = -(Ax + y + z - b). \end{cases} \quad (3.7)$$

From (3.7), we can obtain the following result.

Lemma 3.1. *Let $(x^*, y^*, z^*, \lambda^*)$ be the stable point of $\mathcal{L}_\beta(x, y, z, \lambda)$. Then $0 \in \partial \mathcal{L}_\beta(x^*, y^*, z^*, \lambda^*)$ if and only if the following equalities hold by Lemma 2.11:*

$$\begin{aligned} A^\top \lambda^* - \nabla_x l(x^*, y^*, z^*) &\in \partial f(x^*), & \lambda^* - \nabla_y l(x^*, y^*, z^*) &= \nabla g(y^*), \\ \lambda^* - \nabla_z l(x^*, y^*, z^*) &= \nabla h(z^*), & Ax^* + y^* + z^* - b &= 0. \end{aligned}$$

The augmented Lagrange function $\mathcal{L}_\beta(x, y, z, \lambda)$ has the following properties for all $\theta \in \mathbb{R}$ and $(x, y, z, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m$:

$$\mathcal{L}_\beta(x, y, z, \lambda - \theta(Ax + y + z - b)) = \mathcal{L}_\beta(x, y, z, \lambda) + \theta \|Ax + y + z - b\|^2. \quad (3.8)$$

To simplify the analysis, we introduce symbols:

$$v = (x, y, z), v^* = (x^*, y^*, z^*), v^k = (x^k, y^k, z^k), \omega = (v, \lambda), \omega^* = (v^*, \lambda^*), \omega^k = (v^k, \lambda^k).$$

To analyze the convergence, some basic assumptions of the problem (1.3) are given.

Assumption 3.2.

- (i) $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is the normal lower semicontinuous, $g: \mathbb{R}^m \rightarrow \mathbb{R}$, $h: \mathbb{R}^m \rightarrow \mathbb{R}$ and $l(x, y, z): \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ are differentiable, where gradient ∇g , ∇h , and ∇l are Lipschitz continuous with constants L_g , L_h , and L_l , respectively.
- (ii) ϕ is strongly σ -convex differentiable function, and gradient $\nabla \phi$ is L_ϕ -Lipschitz continuous.

Remark 3.3. According to the optimality conditions of 3-BPRSM, one can obtain

$$\begin{cases} 0 \in \partial f(x^{k+1}) + \nabla_x l(x^{k+1}, y^k, z^k) - A^\top \lambda^k + \beta A^\top (Ax^{k+1} + y^k + z^k - b) + \nabla \phi(x^{k+1}) - \nabla \phi(x^k), \\ 0 = \nabla g(y^{k+1}) + \nabla_y l(x^{k+1}, y^{k+1}, z^k) - \lambda^{k+\frac{1}{2}} + \beta (Ax^{k+1} + y^{k+1} + z^k - b), \\ 0 = \nabla h(z^{k+1}) + \nabla_z l(v^{k+1}) - \lambda^{k+\frac{1}{2}} + \beta (Ax^{k+1} + y^{k+1} + z^{k+1} - b). \end{cases} \quad (3.9)$$

To analyze the monotonicity of $\{\mathcal{L}_\beta(\omega^k)\}$, let

$$\begin{cases} \delta_1 := \delta_1(r, s, \beta) = \frac{\sigma}{2} - \frac{6[L_l^2 + \beta^2(1-s)^2\lambda_{\max}(A^\top A)]}{(r+s)\beta}, \\ \delta_2 := \delta_2(r, s, \beta) = \frac{\beta - L_g - L_l}{2} - \frac{sr\beta}{r+s} - \frac{6[(L_l)^2 + \beta^2(1-s)^2]}{(r+s)\beta} \\ \quad = \frac{[r(1-2s) + s - 12(1-s)^2]\beta^2 - (r+s)(L_g + L_l)\beta - 12(L_l)^2}{2(r+s)\beta}, \\ \delta_3 := \delta_3(r, s, \beta) = \frac{\beta - L_h - L_l}{2} - \frac{sr\beta}{r+s} - \frac{6[(L_l + L_h)^2 + \beta^2(1-s)^2]}{(r+s)\beta} \\ \quad = \frac{[r(1-2s) + s - 12(1-s)^2]\beta^2 - (r+s)(L_h + L_l)\beta - 12(L_l + L_h)^2}{2(r+s)\beta}, \\ \delta := \delta(r, s, \beta) = \min\{\delta_1, \delta_2, \delta_3\}. \end{cases} \quad (3.10)$$

Lemma 3.4. If Assumption 3.2 holds, r and s in 3-BPRSM satisfy $r + s > 0$, then

$$\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^k) \leq -\delta \|v^{k+1} - v^k\|^2, \quad \forall k \geq 0, \quad (3.11)$$

where the definition of δ is shown in (3.10).

Proof. On the one hand, according to (3.6) and (3.4), we have

$$\begin{aligned} & \mathcal{L}_\beta(v^{k+1}, \lambda^{k+\frac{1}{2}}) - \mathcal{L}_\beta(x^{k+1}, y^{k+1}, z^k, \lambda^{k+\frac{1}{2}}) \\ &= h(z^{k+1}) - h(z^k) + l(v^{k+1}) - l(x^{k+1}, y^{k+1}, z^k) - \langle \lambda^{k+\frac{1}{2}}, z^{k+1} - z^k \rangle + \frac{\beta}{2} \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2 \\ & \quad - \frac{\beta}{2} \|Ax^{k+1} + y^{k+1} + z^k - b\|^2 \\ &= h(z^{k+1}) - h(z^k) + l(v^{k+1}) - l(x^{k+1}, y^{k+1}, z^k) - \langle \lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} + y^{k+1} - b), z^{k+1} - z^k \rangle - \frac{\beta}{2} \|z^{k+1} - z^k\|^2 \\ &= h(z^{k+1}) - h(z^k) + l(v^{k+1}) - l(x^{k+1}, y^{k+1}, z^k) - \langle \nabla h(z^{k+1}) + \nabla_z l(v^{k+1}), z^{k+1} - z^k \rangle - \frac{\beta}{2} \|z^{k+1} - z^k\|^2, \end{aligned}$$

where the second equality is obtained by $c^2 - d^2 = (c + d)(c - d)$. Take advantage of the Lipschitz continuity of ∇_h and ∇_l , and Lemma 2.13. Then we have

$$h(z^{k+1}) - h(z^k) - \langle \nabla h(z^{k+1}), z^{k+1} - z^k \rangle \leq \frac{L_h}{2} \|z^{k+1} - z^k\|^2$$

and

$$l(v^{k+1}) - l(x^{k+1}, y^{k+1}, z^k) - \langle \nabla_z l(v^{k+1}), z^{k+1} - z^k \rangle \leq \frac{L_l}{2} \|z^{k+1} - z^k\|^2.$$

Thus, one has

$$\mathcal{L}_\beta(v^{k+1}, \lambda^{k+\frac{1}{2}}) - \mathcal{L}_\beta(x^{k+1}, y^{k+1}, z^k, \lambda^{k+\frac{1}{2}}) \leq -\frac{\beta - L_h - L_l}{2} \|y^{k+1} - y^k\|^2. \quad (3.12)$$

Similarly, by (3.3) and (3.4), the following inequality can be deduced in the same way:

$$\mathcal{L}_\beta(x^{k+1}, y^{k+1}, z^k, \lambda^{k+\frac{1}{2}}) - \mathcal{L}_\beta(x^{k+1}, y^k, z^k, \lambda^{k+\frac{1}{2}}) \leq -\frac{\beta - L_g - L_l}{2} \|y^{k+1} - y^k\|^2. \quad (3.13)$$

Moreover, it is known from (3.2), (3.5), and (3.8) that

$$\mathcal{L}_\beta(x^{k+1}, y^k, z^k, \lambda^{k+\frac{1}{2}}) - \mathcal{L}_\beta(x^{k+1}, y^k, z^k, \lambda^k) = r\beta \|Ax^{k+1} + y^k + z^k - b\|^2, \quad (3.14)$$

$$\mathcal{L}_\beta(w^{k+1}) - \mathcal{L}_\beta(v^{k+1}, \lambda^{k+\frac{1}{2}}) = s\beta \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2. \quad (3.15)$$

In addition, since x^{k+1} is the optimal solution to (3.1), combining with Lemma 2.10, the following inequality holds:

$$\mathcal{L}_\beta(x^{k+1}, y^k, z^k, \lambda^k) - \mathcal{L}_\beta(w^k) \leq -\Delta_\phi(x^{k+1}, x^k) \leq -\frac{\sigma}{2} \|x^{k+1} - x^k\|^2. \quad (3.16)$$

Hence, according to (3.12)–(3.16), one can obtain

$$\begin{aligned} & \mathcal{L}_\beta(w^{k+1}) - \mathcal{L}_\beta(w^k) \\ & \stackrel{(3.15)}{=} \mathcal{L}_\beta(v^{k+1}, \lambda^{k+\frac{1}{2}}) + s\beta \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2 - \mathcal{L}_\beta(w^k) \\ & \stackrel{(3.13)}{\leq} \mathcal{L}_\beta(x^{k+1}, y^{k+1}, z^k, \lambda^{k+\frac{1}{2}}) - \frac{\beta - L_h - L_l}{2} \|y^{k+1} - y^k\|^2 + s\beta \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2 - \mathcal{L}_\beta(w^k) \\ & \stackrel{(3.12)}{\leq} \mathcal{L}_\beta(x^{k+1}, y^k, z^k, \lambda^{k+\frac{1}{2}}) - \frac{\beta - L_h - L_l}{2} \|z^{k+1} - z^k\|^2 - \frac{\beta - L_g - L_l}{2} \|y^{k+1} - y^k\|^2 + s\beta \|Ax^{k+1} + y^{k+1} \\ & \quad + z^{k+1} - b\|^2 - \mathcal{L}_\beta(w^k) \\ & \stackrel{(3.14)}{=} \mathcal{L}_\beta(x^{k+1}, y^k, z^k, \lambda^k) - \frac{\beta - L_h - L_l}{2} \|z^{k+1} - z^k\|^2 - \frac{\beta - L_g - L_l}{2} \|y^{k+1} - y^k\|^2 + r\beta \|Ax^{k+1} + y^k + z^k \\ & \quad - b\|^2 + s\beta \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2 - \mathcal{L}_\beta(w^k) \\ & \stackrel{(3.16)}{\leq} -\frac{\sigma}{2} \|x^{k+1} - x^k\|^2 - \frac{\beta - L_h - L_l}{2} \|z^{k+1} - z^k\|^2 - \frac{\beta - L_g - L_l}{2} \|y^{k+1} - y^k\|^2 \\ & \quad + r\beta \|Ax^{k+1} + y^k + z^k - b\|^2 + s\beta \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2. \end{aligned} \quad (3.17)$$

On the other hand, the equalities hold by the multipliers update formulas (3.2) and (3.5):

$$Ax^{k+1} + y^k + z^k - b = -\frac{1}{(r+s)\beta}(\lambda^{k+1} - \lambda^k) - \frac{s}{r+s}(y^{k+1} - y^k) - \frac{s}{r+s}(z^{k+1} - z^k), \quad (3.18)$$

$$Ax^{k+1} + y^{k+1} + z^{k+1} - b = -\frac{1}{(r+s)\beta}(\lambda^{k+1} - \lambda^k) + \frac{r}{r+s}(y^{k+1} - y^k) + \frac{r}{r+s}(z^{k+1} - z^k). \quad (3.19)$$

Then, from (3.5) and (3.9), we have

$$\lambda^{k+1} = \nabla h(z^{k+1}) + \nabla_z l(v^{k+1}) + \beta(1-s)(Ax^{k+1} + y^{k+1} + z^k - b). \quad (3.20)$$

Combining with the Lipschitz continuity of ∇_g and ∇_h , it is derived that

$$\begin{aligned} \|\lambda^{k+1} - \lambda^k\| &= \|\nabla h(z^{k+1}) - \nabla h(z^k) + \nabla_z l(v^{k+1}) - \nabla_z l(v^k) + \beta(1-s)[A(x^{k+1} - x^k) + (y^{k+1} - y^k) \\ & \quad + (z^{k+1} - z^k)]\| \\ &\leq L_l \|x^{k+1} - x^k\| + L_l \|y^{k+1} - y^k\| + (L_h + L_l) \|z^{k+1} - z^k\| \\ & \quad + \beta \|1-s\| (\|A(x^{k+1} - x^k)\| + \|y^{k+1} - y^k\| + \|z^{k+1} - z^k\|). \end{aligned} \quad (3.21)$$

Furthermore, according to the Cauchy inequality, the following inequality holds:

$$\begin{aligned} \frac{1}{6} \|\lambda^{k+1} - \lambda^k\|^2 &\leq \beta^2(1-s)^2 (\|A(x^{k+1} - x^k)\|^2 + \|y^{k+1} - y^k\|^2 + \|z^{k+1} - z^k\|^2) + L_l^2 \|x^{k+1} - x^k\|^2 \\ & \quad + L_l^2 \|y^{k+1} - y^k\|^2 + (L_h + L_l)^2 \|z^{k+1} - z^k\|^2. \end{aligned} \quad (3.22)$$

Thus, combining (3.18), (3.19), (3.22), and $r + s > 0$, one has

$$\begin{aligned}
 & r\beta \|Ax^{k+1} + y^k + z^k - b\|^2 + s\beta \|Ax^{k+1} + y^{k+1} + z^{k+1} - b\|^2 \\
 &= \frac{1}{(r+s)\beta} \|\lambda^{k+1} - \lambda^k\|^2 + \frac{rs\beta}{r+s} \|y^{k+1} - y^k\|^2 + \frac{rs\beta}{r+s} \|z^{k+1} - z^k\|^2 \\
 &\leq \frac{6[L_l^2 + \beta^2(1-s)^2\lambda_{\max}(A^T A)]}{(r+s)\beta} \|x^{k+1} - x^k\|^2 + \left(\frac{6[(L_l)^2 + \beta^2(1-s)^2]}{(r+s)\beta} + \frac{rs\beta}{r+s} \right) \|y^{k+1} - y^k\|^2 \\
 &\quad + \left(\frac{6[(L_h + L_l)^2 + \beta^2(1-s)^2]}{(r+s)\beta} + \frac{rs\beta}{r+s} \right) \|z^{k+1} - z^k\|^2. \tag{3.23}
 \end{aligned}$$

Finally, substituting (3.23) into (3.17), the following inequality saves:

$$\begin{aligned}
 & \mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^k) \\
 &\leq \left(-\frac{\sigma}{2} + \frac{6[L_l^2 + \beta^2(1-s)^2\lambda_{\max}(A^T A)]}{(r+s)\beta} \right) \|x^{k+1} - x^k\|^2 \\
 &\quad + \left(-\frac{\beta - L_g - L_l}{2} + \left(\frac{6[(L_l)^2 + \beta^2(1-s)^2]}{(r+s)\beta} + \frac{rs\beta}{r+s} \right) \right) \|y^{k+1} - y^k\|^2 \\
 &\quad + \left(\frac{6[(L_h + L_l)^2 + \beta^2(1-s)^2]}{(r+s)\beta} + \frac{rs\beta}{r+s} \right) \|z^{k+1} - z^k\|^2 \\
 &= -\delta_1(r, s, \beta) \|x^{k+1} - x^k\|^2 - \delta_2(r, s, \beta) \|y^{k+1} - y^k\|^2 - \delta_3(r, s, \beta) \|z^{k+1} - z^k\|^2 \\
 &\leq -\delta \|v^{k+1} - v^k\|^2.
 \end{aligned}$$

The proof is completed. \square

It follows that $\{\mathcal{L}_\beta(\omega^k)\}$ has sufficient descent property when $\delta > 0$, and this can be achieved by choosing a suitable value for (r, s, β, σ) . So some of the restrictions on the parameter (r, s, β, σ) are given below.

Assumption 3.5. In 3-BPRSM, the strongly convex coefficient σ of function ϕ and parameter (r, s, β) meet the following conditions:

(i) The solution of $r + s > 0$ and $r(1 - 2s) + s - 12(1 - s)^2 > 0$ can be expressed as follows:

$$(r, s) \in D = \left\{ \left(\frac{12(1-s)^2 - s}{1-2s}, +\infty \right) \times \left(-\infty, \frac{1}{2} \right) \right\} \cup \left\{ \left(-s, \frac{12(1-s)^2 - s}{1-2s} \right) \times \left(\frac{6 - \sqrt{6}}{5}, \frac{6 + \sqrt{6}}{5} \right) \right\};$$

(ii) $\beta > \beta_0 := \max\{\varpi, \varrho\}$, where

$$\begin{aligned}
 \varpi &= \frac{(r+s)(L_g + L_h) + \sqrt{(r+s)^2(l_g + l_h)^2 + 48[r(1-2s) + s - 12(1-s)^2](L_l^2)}}{2[r(1-2s) + s - 12(1-s)^2]}, \\
 \varrho &= \frac{(r+s)(L_g + L_h) + \sqrt{(r+s)^2(l_g + l_h)^2 + 48[r(1-2s) + s - 12(1-s)^2]((L_h + L_l)^2)}}{2[r(1-2s) + s - 12(1-s)^2]};
 \end{aligned}$$

(iii) $\sigma > \frac{12[L_l^2 + \beta^2(1-s)^2\lambda_{\max}(A^T A)]}{(r+s)\beta}$.

In particular, $r = s$ holds when $s \in (6/7, 1)$. Therefore, the parameter (r, s) can take the same value in this range.

Lemma 3.6. If both Assumptions 3.2 and 3.5 hold, then δ_1 , δ_2 , and δ defined by (3.10) are positive. Thus, $\{\mathcal{L}_\beta(\omega^k)\}$ decreases monotonically.

Proof. $\delta_1 > 0$ and $\delta_2 > 0$ clearly hold from Assumption 3.2. Further, $\delta > 0$ hold by (3.10). Finally, $\{\mathcal{L}_\beta(\omega^k)\}$ decreases monotonically by (3.11). \square

Lemma 3.7. *If both Assumptions 3.2 and 3.5 own, and $\{\omega^k\}$ in 3-BPRSM is bounded, then $\sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\|^2 < +\infty$.*

Proof. By the boundedness of the sequence $\{\omega^k\}$, we can find a subsequence $\{\omega^{k_j}\}$ such that $\lim_{k_j \rightarrow +\infty} \omega^{k_j} = \omega^*$. Since the function $f(x)$ is lower semicontinuous, and functions $g(y)$, $h(z)$, and $l(x, y, z)$ are continuous, one can conclude that the Lagrange function $\mathcal{L}_\beta(\cdot)$ is also lower semicontinuous. Hence, we have $\mathcal{L}_\beta(\omega^*) \leq \liminf_{k_j \rightarrow +\infty} \mathcal{L}_\beta(\omega^{k_j}) \leq \mathcal{L}_\beta(\omega^0) < +\infty$. This implies that $\{\mathcal{L}_\beta(\omega^{k_j})\}$ is bounded from below. By combining the lower semicontinuity of $\{\mathcal{L}_\beta(\omega^{k_j})\}$, we know that the entire sequence $\{\mathcal{L}_\beta(\omega^{k_j})\}$ is convergent. So its entire sequence converges and $\lim_{k \rightarrow +\infty} \mathcal{L}_\beta(\omega^k) = \inf_k \mathcal{L}_\beta(\omega^k) \geq \mathcal{L}_\beta(\omega^*)$ holds since $\{\mathcal{L}_\beta(\omega^k)\}$ is monotonically decreasing and has convergent subsequences combining with Lemma 2.14. Therefore, the inequality $\mathcal{L}_\beta(\omega^k) \geq \mathcal{L}_\beta(\omega^*)$ holds for all k . In addition, by (3.11), one has

$$\delta \|v^{k+1} - v^k\|^2 \leq \mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^{k+1}), \quad \forall k \geq 0. \quad (3.24)$$

Summing up (3.24) from $k = 0$ to $k = q$, and the following inequality holds when $\mathcal{L}_\beta(\omega^0) < +\infty$:

$$\delta \sum_{k=0}^q \|v^{k+1} - v^k\|^2 \leq \mathcal{L}_\beta(\omega^0) - \mathcal{L}_\beta(\omega^{q+1}) \leq \mathcal{L}_\beta(\omega^0) - \mathcal{L}_\beta(\omega^*) < +\infty.$$

Then because $\delta > 0$, one immediately knows that $\sum_{k=0}^{+\infty} \|v^{k+1} - v^k\|^2 < +\infty$ holds. So we can know

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty, \quad \sum_{k=0}^{+\infty} \|y^{k+1} - y^k\|^2 < +\infty, \quad \sum_{k=0}^{+\infty} \|z^{k+1} - z^k\|^2 < +\infty.$$

Finally, by combining with (3.22), we obtain $\sum_{k=0}^{+\infty} \|\lambda^{k+1} - \lambda^k\|^2 < +\infty$, and so $\sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\|^2 < +\infty$. The proof is complete. \square

Theorem 3.8. (Global convergent) *Let Ω be the set of accumulation point of $\{\omega^k\}$. Assuming that Assumptions 3.2 and 3.5 hold, and $\{\omega^k\}$ generated by 3-BPRSM is bounded, then the following three results are true:*

- (i) Ω is a nonempty convex set, and $d(\omega^k, \Omega) \rightarrow 0, k \rightarrow +\infty$;
- (ii) $\Omega \subseteq \text{crit } \mathcal{L}_\beta$;
- (iii) *The entire sequence of $\{\mathcal{L}_\beta(\omega^k)\}$ is convergent, and $\mathcal{L}_\beta(\omega^*) = \lim_{k \rightarrow +\infty} \mathcal{L}_\beta(\omega^k) = \inf_k \mathcal{L}_\beta(\omega^k)$ holds for all $\omega^* \in \Omega$. Further, $\mathcal{L}_\beta(\cdot)$ is finite and a constant on Ω .*

Proof. (i) The conclusion (i) holds by the definition of Ω and the boundedness of $\{\omega^k\}$.

(ii) Let $\omega^* = (x^*, y^*, z^*, \lambda^*) \in \Omega$. Then there exists a subsequence $\{\omega_j^k\}$ of $\{\omega^k\}$ converges to ω^* . $\lim_{k \rightarrow +\infty} \|\omega^{k+1} - \omega^k\| = 0$ is known from Lemma 3.7; thus, $\lim_{k_j \rightarrow +\infty} \omega_j^{k_j+1} = \omega^*$. Then letting $k = k_j \rightarrow +\infty$ in (3.2), we can know that $\{\lambda_j^{k_j+\frac{1}{2}}\}$ is bounded, so we have $\lim_{k_j \rightarrow +\infty} \lambda_j^{k_j+\frac{1}{2}} = \lambda^{**}$. Let $k = k_j \rightarrow +\infty$ and take the limit in (3.2) and (3.5). Then

$$\lambda^{**} = \lambda^* - r\beta(Ax^* + y^* + z^* - b), \quad \lambda^* = \lambda^{**} - s\beta(Ax^* + y^* + z^* - b).$$

And because $r + s > 0$, one can obtain from the aforementioned equation that

$$Ax^* + y^* + z^* - b = 0, \quad \lambda^{**} = \lambda^*.$$

Thus, $v^* = (x^*, y^*, z^*)$ is the feasible point for the problem (1.3). In addition, noting that x^{k_j+1} is the optimal solution of (3.1), then the following inequality holds:

$$\begin{aligned} f(x^{k_j+1}) + l(x^{k_j+1}, y^{k_j}, z^{k_j}) - \langle \lambda^{k_j}, Ax^{k_j+1} \rangle + \frac{\beta}{2} \|Ax^{k_j+1} + y^{k_j} + z^{k_j} - b\|^2 + \Delta_\phi(x^{k_j+1}, x^{k_j}) \\ \leq f(x^*) + h(x^*, y^{k_j}, z^{k_j}) - \langle \lambda^{k_j}, Ax^* \rangle + \frac{\beta}{2} \|Ax^* + y^{k_j} + z^{k_j} - b\|^2 + \Delta_\phi(x^*, x^{k_j}). \end{aligned}$$

Combining $\lim_{k_j \rightarrow +\infty} \omega^{k_j} = \lim_{k_j \rightarrow +\infty} \omega^{k_j+1} = \omega^*$ and the continuous differentiability of ϕ , by the aforementioned inequality, now we know that $\limsup_{k_j \rightarrow +\infty} f(x^{k_j+1}) \leq f(x^*)$ holds. Also noticing that the lower semicontinuity of f , it follows that $f(x^*) \leq \lim_{k_j \rightarrow +\infty} f(x^{k_j+1})$, so we can obtain

$$\lim_{k_j \rightarrow +\infty} f(x^{k_j+1}) = f(x^*). \quad (3.25)$$

In addition, $\lim_{k_j \rightarrow +\infty} \|\nabla \phi(x^{k_j+1}) - \nabla \phi(x^{k_j})\| = 0$ according to Assumption 3.2 and $\lim_{k_j \rightarrow +\infty} \|x^{k_j+1} - x^{k_j}\| = 0$. Hence, taking into account the closeness of ∂f , the continuity of ∇_g and ∇_h , as well as (3.25), letting $k = k_j \rightarrow +\infty$ and taking the limit in (3.9), then we obtain

$$\begin{aligned} A^\top \lambda^* - \nabla_x l(v^*) &\in \partial f(x^*), \quad \lambda^* - \nabla_y l(v^*) = \nabla g(y^*), \\ \lambda^* - \nabla_z l(v^*) &= \nabla h(z^*), \quad Ax^* + y^* + z^* - b = 0, \end{aligned}$$

and $\omega^* \in \text{crit } \mathcal{L}_\beta$ by Lemma 3.1.

(iii) Let $\omega^* \in \Omega$. There exists at least one subsequence $\{\omega^{k_j}\}$ of $\{\omega^k\}$ converges to ω^* . According to (3.6) and (3.11), $\lim_{k_j \rightarrow +\infty} \mathcal{L}_\beta(\omega^{k_j+1}) = \mathcal{L}_\beta(\omega^*)$ holds. Combining with the monotonicity of $\{\mathcal{L}_\beta(\omega^k)\}$ and Lemma 2.14, we can see that the whole sequence $\{\mathcal{L}_\beta(\omega^k)\}$ is convergent. Thus, $+\infty > \mathcal{L}_\beta(\omega^0) \geq \lim_{k \rightarrow +\infty} \mathcal{L}_\beta(\omega^k) = \inf_k \mathcal{L}_\beta(\omega^k) = \mathcal{L}_\beta(\omega^*)$ saves by $\mathcal{L}_\beta(\omega^k) \leq \mathcal{L}_\beta(\omega^0) < +\infty$. Hence, $\mathcal{L}_\beta(\omega^*) \equiv \lim_{k \rightarrow +\infty} \mathcal{L}_\beta(\omega^k) < +\infty$ holds for all $\omega^* \in \Omega$. The proof is completed. \square

Lemma 3.9. *Let*

$$\begin{cases} \mathcal{E}_1^{k+1} = \beta A^\top (y^{k+1} - y^k) + \beta A^\top (z^{k+1} - z^k) - A^\top (\lambda^{k+1} - \lambda^k) \\ \quad - [\nabla \phi(x^{k+1}) - \nabla \phi(x^k)] + \nabla_x l(v^{k+1}) - \nabla_x l(x^{k+1}, y^k, z^k), \\ \mathcal{E}_2^{k+1} = -\frac{s}{r+s}(\lambda^{k+1} - \lambda^k) + \frac{rs\beta}{r+s}(y^{k+1} - y^k + z^{k+1} - z^k) \\ \quad - \nabla_y l(x^{k+1}, y^{k+1}, z^k) + \nabla_y l(v^{k+1}) + \beta(z^{k+1} - z^k), \\ \mathcal{E}_3^{k+1} = \frac{s}{r+s}(\lambda^{k+1} - \lambda^k) + \frac{rs\beta}{r+s}(y^{k+1} - y^k + z^{k+1} - z^k), \\ \mathcal{E}_4^{k+1} = \frac{1}{(r+s)\beta}(\lambda^{k+1} - \lambda^k) - \frac{r}{r+s}(y^{k+1} - y^k + z^{k+1} - z^k). \end{cases} \quad (3.26)$$

If both $\nabla \phi$ and $\nabla \phi$ are Lipschitz continuous, then $\mathcal{E}^{k+1} := (\mathcal{E}_1^{k+1}, \mathcal{E}_2^{k+1}, \mathcal{E}_3^{k+1}, \mathcal{E}_4^{k+1}) \in \partial \mathcal{L}_\beta(\omega^{k+1})$ and there exists a constant $\zeta > 0$ such that

$$(0, \partial \mathcal{L}_\beta(\omega^{k+1})) \leq \zeta \|v^{k+1} - v^k\|, \quad \forall k \geq 0. \quad (3.27)$$

Proof. First, according to (3.9), there exists $\xi^{k+1} \in \partial f(x^{k+1})$ such that

$$\xi^{k+1} + \nabla_x h(x^{k+1}, y^k) - A^\top \lambda^k + \beta A^\top (Ax^{k+1} + y^k - b) + \nabla \phi(x^{k+1}) - \nabla \phi(x^k) = 0.$$

Combining (3.26) and (3.7), one has

$$\begin{aligned} \varepsilon_1^{k+1} &= \beta A^\top (y^{k+1} - y^k) + \beta A^\top (z^{k+1} - z^k) - A^\top (\lambda^{k+1} - \lambda^k) - [\nabla \phi(x^{k+1}) - \nabla \phi(x^k)] + \nabla_x l(v^{k+1}) - \nabla_x l(x^{k+1}, y^k, z^k) \\ &\quad + \{\xi^{k+1} + \nabla_x h(x^{k+1}, y^k) - A^\top \lambda^k + \beta A^\top (Ax^{k+1} + y^k - b) + \nabla \phi(x^{k+1}) - \nabla \phi(x^k)\} \\ &= \xi^{k+1} + \nabla_x l(v^{k+1}) - A^\top \lambda^{k+1} + \beta A^\top (Ax^{k+1} + y^{k+1} + z^{k+1} - b) \in \partial_x \mathcal{L}_\beta(\omega^{k+1}). \end{aligned}$$

Second, we substitute $\lambda^{k+\frac{1}{2}}$ from (3.5) into (3.9), which yields

$$s\beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b) = \nabla g(y^{k+1}) + \nabla_y l(v^{k+1}) - \lambda^{k+1} + \beta(Ax^{k+1} + y^{k+1} + z^k - b).$$

Combining (3.7), (3.19), and (3.26), one has

$$\begin{aligned}\varepsilon_2^{k+1} &= -\frac{s}{r+s}(\lambda^{k+1} - \lambda^k) + \frac{rs\beta}{r+s}(y^{k+1} - y^k + z^{k+1} - z^k) - \nabla_y l(x^{k+1}, y^{k+1}, z^k) + \nabla_y l(v^{k+1}) + \beta(z^{k+1} - z^k) \\ &= s\beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b) - \nabla_y l(x^{k+1}, y^{k+1}, z^k) + \nabla_y l(v^{k+1}) + \beta(z^{k+1} - z^k) \\ &= \nabla g(y^{k+1}) + \nabla_y l(v^{k+1}) - \lambda^{k+1} + \beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b) \in \partial_y \mathcal{L}_\beta(\omega^{k+1}).\end{aligned}$$

Similarly, by substituting (3.4) for $\lambda^{k+\frac{1}{2}}$ in (3.5), we obtain

$$s\beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b) = \nabla h(z^{k+1}) + \nabla_z l(v^{k+1}) - \lambda^{k+1} + \beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b).$$

By combining with (3.7), (3.19), and (3.26), we have

$$\begin{aligned}\varepsilon_3^{k+1} &= s\beta \frac{1}{(r+s)\beta}(\lambda^{k+1} - \lambda^k) + \frac{r}{r+s}(y^{k+1} - y^k + z^{k+1} - z^k) \\ &= s\beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b) \\ &= \nabla h(z^{k+1}) + \nabla_z l(v^{k+1}) - \lambda^{k+1} + \beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b) \in \partial_z \mathcal{L}_\beta(\omega^{k+1}).\end{aligned}$$

Again, from (3.7), (3.19), and Lemma 3.9, we know that

$$\begin{aligned}\varepsilon_4^{k+1} &= \frac{1}{(r+s)\beta}(\lambda^{k+1} - \lambda^k) - \frac{r}{r+s}(y^{k+1} - y^k + z^{k+1} - z^k) \\ &= -(Ax^{k+1} + y^{k+1} - b) \stackrel{(3.7)}{=} \partial_\lambda \mathcal{L}_\beta(\omega^{k+1}).\end{aligned}$$

So, $\varepsilon^{k+1} \in \partial \mathcal{L}_\beta(\omega^{k+1})$.

Finally, combining (3.26) and the Lipschitz continuity of $\nabla \phi$ and ∇l , there exists $\zeta_1 > 0$ such that

$$\|\varepsilon^{k+1}\| \leq \zeta_1(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\| + \|z^{k+1} - z^k\| + \|\lambda^{k+1} - \lambda^k\|).$$

In addition, by (3.21), there exists $\zeta_2 > 0$ such that

$$\|\lambda^{k+1} - \lambda^k\| \leq \zeta_2(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\| + \|z^{k+1} - z^k\|).$$

Considering the aforementioned two inequalities and $\varepsilon^{k+1} \in \partial \mathcal{L}_\beta(\omega^{k+1})$, similar to the proof of Lemma 9 in [20], now one knows that

$$d(0, \partial \mathcal{L}_\beta(\omega^{k+1})) \leq \|\varepsilon^{k+1}\| \leq \zeta_1(1 + \zeta_2)(\|x^{k+1} - x^k\| + \|y^{k+1} - y^k\| + \|z^{k+1} - z^k\|) \leq \zeta \|v^{k+1} - v^k\|,$$

where $\zeta = \sqrt{2}\zeta_1(1 + \zeta_2)$. The proof is completed. \square

Theorem 3.10. (Strong convergence) *If both Assumptions 3.2 and 3.5 hold, $\{\omega^k\}$ is bounded, and $\mathcal{L}_\beta(\omega)$ is a function with KLP, then*

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

Furthermore, $\{\omega^k\}$ converges to the critical point of $\mathcal{L}_\beta(\cdot)$.

Proof. According to Theorem 3.8 (iii), we have $\mathcal{L}_\beta(\omega^*) = \lim_{k \rightarrow +\infty} \mathcal{L}_\beta(\omega^k) = \inf_k \mathcal{L}_\beta(\omega^k)$ for all $\omega^* \in \Omega$. Next, we consider two cases.

(Case I) If there exists an integer k_0 such that $\mathcal{L}_\beta(\omega^{k_0}) = \mathcal{L}_\beta(\omega^*)$, then from (3.11) and Lemma 3.6, one can obtain

$$\delta \|v^{k+1} - v^k\|^2 \leq \mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^{k+1}) \leq \mathcal{L}_\beta(\omega^{k_0}) - \mathcal{L}_\beta(\omega^*) = 0, \quad \forall k \geq k_0.$$

Thereby, $v^{k+1} = v^k$ owns for all $k \geq k_0$. This together with (3.21), it follows that $\lambda^{k+1} = \lambda^k$ for all $k \geq k_0$. Further, $\omega^{k+1} = \omega^{k_0} \in \Omega$ holds for all $k \geq k_0$. The conclusion saves.

(Case II) Suppose that $\mathcal{L}_\beta(\omega^k) > \mathcal{L}_\beta(\omega^*)$ holds for all $k \geq 0$. As is known from Theorem 3.8 (iii), $\mathcal{L}_\beta(\cdot)$ takes a constant on Ω . Thus, $\mathcal{L}_\beta(\cdot)$ satisfies the uniform **KLP** by Lemma 2.12. For the parameters ζ and η in Lemma 2.12, since $d(\omega^k, \Omega) \rightarrow 0$ and $\mathcal{L}_\beta(\omega^*) = \inf_k \mathcal{L}_\beta(\omega^k)$, there exists a positive integer \tilde{k} such that

$$d(\omega^k, \Omega) < \zeta, \mathcal{L}_\beta(\omega^*) < \mathcal{L}_\beta(\omega^k) < \mathcal{L}_\beta(\omega^*) + \eta, \quad \forall k > \tilde{k}.$$

Thus, by combining with Lemma 2.12, we have

$$\varphi'(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*))d(0, \partial\mathcal{L}_\beta(\omega^k)) \geq 1, \quad \forall k > \tilde{k}.$$

In addition, the following inequality owns by the concavity of φ :

$$\varphi(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*)) \geq \varphi'(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*))(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^{k+1})).$$

Simultaneously, because $\frac{1}{\varphi'(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*))} \leq d(0, \partial\mathcal{L}_\beta(\omega^k)) \leq \zeta\|v^k - v^{k-1}\|$ and $\varphi'(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*)) > 0$, one has

$$\begin{aligned} \mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^{k+1}) &\leq \frac{\varphi(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))}{\varphi'(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*))} \\ &\leq \zeta\|v^k - v^{k-1}\|[\varphi(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))]. \end{aligned}$$

Combining Lemma 3.4 with the aforementioned inequalities, we have

$$\delta\|v^{k+1} - v^k\|^2 \leq \zeta\|v^k - v^{k-1}\|[\varphi(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))], \quad \forall k > \tilde{k},$$

further implies,

$$\|v^{k+1} - v^k\| \leq \sqrt{\|v^k - v^{k-1}\| \frac{\zeta}{\delta} [\varphi(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))]}, \quad \forall k > \tilde{k}.$$

According to the aforementioned inequality and $2\sqrt{ab} \leq a + b$ for all $a, b \geq 0$, we can obtain

$$2\|v^{i+1} - v^i\| \leq \|v^i - v^{i-1}\| + \frac{\zeta}{\delta} [\varphi(\mathcal{L}_\beta(\omega^i) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{i+1}) - \mathcal{L}_\beta(\omega^*))], \quad \forall i > \tilde{k}. \quad (3.28)$$

Then, summing up (3.28) from $i = k + 1$ ($\geq \tilde{k} + 1$) to $i = q$ and rearranging terms, we have

$$2 \sum_{i=k+1}^q \|v^{i+1} - v^i\| \leq \sum_{i=k+1}^q \|v^i - v^{i-1}\| + \frac{\zeta}{\delta} [\varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*)) - \varphi(\mathcal{L}_\beta(\omega^{q+1}) - \mathcal{L}_\beta(\omega^*))].$$

According to $\varphi(\mathcal{L}_\beta(\omega^{q+1}) - \mathcal{L}_\beta(\omega^*)) \geq 0$, further one has

$$\sum_{i=k+1}^q \|v^{i+1} - v^i\| \leq \sum_{i=k+1}^q \|v^{i+1} - v^i\| + \|v^{i+1} - v^q\| \leq \|v^{k+1} - v^k\| + \frac{\zeta}{\delta} [\varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))].$$

And taking the limit $q \rightarrow +\infty$ in the aforementioned inequality, then one has

$$\sum_{i=k+1}^{+\infty} \|v^{i+1} - v^i\| \leq \|v^{k+1} - v^k\| + \frac{\zeta}{\delta} [\varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))] < +\infty, \quad \forall k \geq \tilde{k}. \quad (3.29)$$

When $k = \tilde{k}$, we have

$$\sum_{i=\tilde{k}+1}^{+\infty} \|v^{i+1} - v^i\| \leq \|v^{\tilde{k}+1} - v^{\tilde{k}}\| + \frac{\zeta}{\delta} [\varphi(\mathcal{L}_\beta(\omega^{\tilde{k}+1}) - \mathcal{L}_\beta(\omega^*))] < +\infty.$$

So $\sum_{k=0}^{+\infty} \|v^{k+1} - v^k\| < +\infty$, and it follows from (3.21) that $\sum_{k=0}^{+\infty} \|\lambda^{k+1} - \lambda^k\| < +\infty$. Finally, from $\omega^k = (v^k, \lambda^k)$, we obtain

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

Hence, $\{w^k\}$ is Cauchy sequence, which is convergent. From Theorem 3.8, it can be seen that $\{w^k\}$ converges to the critical point of $\mathcal{L}_\beta(\cdot)$. The proof is completed. \square

Theorem 3.11. (Convergence rate analysis) Suppose that both Assumptions 3.2 and 3.5 hold, $\{\omega^k\}$ generated by 3-BPRSM is bounded, $\mathcal{L}_\beta(\cdot)$ satisfies **KLP**, and the association function is $\varphi(t) = ct^{1-\theta}$, $\theta \in [0, 1)$, $c > 0$. Then the following four conclusions own:

- (i) $\lim_{k \rightarrow +\infty} \omega^k = \omega^* \in \text{crit} \mathcal{L}_\beta$.
- (ii) If $\theta = 0$, then ω^k converges after a finite number of iterations, i.e., k exists such that $\omega^k = \omega^*$.
- (iii) If $\theta \in (0, \frac{1}{2}]$, then there is a constant $\tau \in [0, 1)$ such that $\|\omega^k - \omega^*\| = O(\tau^k)$, and thus, 3-BPRSM converges linearly.
- (iv) If $\theta \in (\frac{1}{2}, 1]$, then $\|\omega^k - \omega^*\| = O(k^{\frac{1-\theta}{1-2\theta}})$, and so 3-BPRSM is sublinearly convergent.

Proof. First, the conclusion (i) holds from Theorem 3.10, and $\mathcal{L}_\beta(\cdot)$ satisfies **KLP** at ω^* . Now let us analyze conclusions (ii)–(iv).

For case $\theta = 0$, one has $\varphi(t) = ct$ and $\varphi'(t) = c$. If there exists a positive integer k_0 such that $\mathcal{L}_\beta(\omega^{k_0}) = \mathcal{L}_\beta(\omega^*)$, combining with the proof of (Case I) in Theorem 3.10, $\{\omega^k\}$ terminates the iteration at k_0 . If $\mathcal{L}_\beta(\omega^k) > \mathcal{L}_\beta(\omega^*)$ holds for all k , then we have $\varphi'(\mathcal{L}_\beta(\omega^k) - \mathcal{L}_\beta(\omega^*))d(0, \partial \mathcal{L}_\beta(\omega^k)) = cd(0, \partial \mathcal{L}_\beta(\omega^k)) \geq 1$ for all $k > \tilde{k}$ by the proof process of (Case II) in Theorem 3.10 and (3.26), so this contradicts (3.27). Thus, there exists a positive integer $\hat{k} (> \tilde{k})$ for all $k > \hat{k}$ such that $\mathcal{L}_\beta(\omega^k) = \mathcal{L}_\beta(\omega^*)$, and conclusion (iii) holds from (Case I) of Theorem 3.10.

Next, let us discuss the case of $\theta > 0$. Taking $\Xi_k = \sum_{i=k+1}^{+\infty} \|v^{i+1} - v^i\|$, then from (3.29), it follows that

$$\Xi_{k+1} \leq (\Xi_k - \Xi_{k+1}) + \frac{\zeta}{\delta} \varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*)), \quad \forall k \geq \tilde{k}. \quad (3.30)$$

At the same time, it is noted that $\mathcal{L}_\beta(\cdot)$ satisfies **KLP** at ω^* , and one can immediately known that

$$\varphi'(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))d(0, \partial \mathcal{L}_\beta(\omega^{k+1})) \geq 1.$$

In combination with $\varphi(t) = ct^{1-\theta}$, the aforementioned inequality is equivalent to

$$(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))^\theta \leq c(1 - \theta)d(0, \partial \mathcal{L}_\beta(\omega^{k+1})). \quad (3.31)$$

From (3.27), we can also know

$$d(0, \partial \mathcal{L}_\beta(\omega^{k+1})) \leq \zeta \|v^{k+1} - v^k\| = \zeta(\Xi_k - \Xi_{k+1}). \quad (3.32)$$

Combining with (3.31) and (3.32), there exists $\gamma > 0$ such that

$$\varphi(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*)) = c(\mathcal{L}_\beta(\omega^{k+1}) - \mathcal{L}_\beta(\omega^*))^{1-\theta} \leq \gamma(\Xi_k - \Xi_{k+1})^{\frac{1-\theta}{\theta}}.$$

According to (3.30), one has

$$\Xi_{k+1} \leq (\Xi_k - \Xi_{k+1}) + \frac{\zeta}{\delta} \gamma (\Xi_k - \Xi_{k+1})^{\frac{1-\theta}{\theta}}, \quad \forall k \geq \tilde{k}. \quad (3.33)$$

Based on (3.33), the remaining proof are completed with the help of the relevant conclusions of Attouch and Bolte [45].

By [45], if $\theta \in (0, \frac{1}{2}]$, then one knows that there exist $c_1 > 0$ and $\tau \in [0, 1)$ such that

$$\|v^k - v^*\| \leq c_1 \tau^k \Rightarrow \|x^k - x^*\| = O(\tau^k), \|y^k - y^*\| = O(\tau^k), \|z^k - z^*\| = O(\tau^k). \quad (3.34)$$

Further, there exist $c_2 > 0$ such that

$$\|v^k - v^*\| \leq c_2 k^{\frac{1-\theta}{1-2\theta}} \Rightarrow \|x^k - x^*\| = O(k^{\frac{1-\theta}{1-2\theta}}), \|y^k - y^*\| = O(k^{\frac{1-\theta}{1-2\theta}}), \|z^k - z^*\| = O(k^{\frac{1-\theta}{1-2\theta}}). \quad (3.35)$$

when $\theta \in (\frac{1}{2}, 1]$.

In addition, from Lemma 3.1, (3.20), and the Lipschitz continuity of ∇g and ∇h , we obtain

$$\begin{aligned} \|\lambda^k - \lambda^*\| &= \|\nabla g(y^k) - \nabla g(y^*) + \nabla h(z^k) - \nabla h(z^*) + \nabla_y l(v^k) - \nabla_y l(v^*) + \nabla_z l(v^k) - \nabla_z l(v^*) \\ &\quad + \beta(1-s)[A(x^k - x^*) + (y^k - y^*) + (z^k - z^*)]\| \\ &= O(\|v^k - v^*\|). \end{aligned} \quad (3.36)$$

Finally, combining with (3.34)–(3.36), now we know that conclusions (iii) and (iv) are verified. \square

4 Numerical simulations

Sparse signal reconstruction from incomplete observation data sets is a prominent research area in compressed sensing. The objective is to find the sparsest solution to a system of linear equations represented by [15, p. 14]:

$$\begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & Cx = b, \end{aligned} \quad (4.1)$$

where $C \in \mathbb{R}^{m \times n}$ is a measurement matrix, $b \in \mathbb{R}^m$ is observation data, $e > 0$ is a regularization parameter and $\|x\|_0$ is the number of nonzero elements of x . How to recover an n -dimensional sparse signal x ? Wang et al. described in [15] that the system (4.1) and its linear transformation or its regularization version are NP-hard problems. To surmount this difficulty, l_0 regularization is often relaxed to $l_{1/2}$ regularization. Instead of directly solving (4.1), researchers commonly solve the following constrained optimization problem [10,15] (also [34, the problem (79), p. 11]):

$$\begin{aligned} \min \quad & e\|x\|_{\frac{1}{2}}^2 + \frac{1}{2}\|y\|^2 \\ \text{s.t.} \quad & Cx - y = b, \end{aligned} \quad (4.2)$$

where $\|x\|_{(1/2)} = \left(\sum_{i=1}^n |x_i|^{(1/2)} \right)^2$.

Based on (4.2) and the objective existence of nonseparable structure, now we generate the following optimization problem with line constraint:

$$\begin{aligned} \min \quad & e\|x\|_{\frac{1}{2}}^2 + \frac{1}{2}\|y\|^2 + \frac{1}{2}\|D_1x + D_2y + z\|^2 \\ \text{s.t.} \quad & Ax + y + z = b, \end{aligned} \quad (4.3)$$

which was studied in [34, (81), p. 12] by Chao et al. when $A_2 = E$, the identity matrix.

To evaluate the effectiveness, by applying 3-BPRSM, we construct the following algorithm to solve the nonconvex optimization problem (4.3) with $\phi(x) = \frac{1}{2}x_{\mu_1 I_1}^2 - \beta A^T A$:

$$\begin{aligned} x^{k+1} &= S \left(\frac{1}{\mu_1} \left[G_1 x^k - D_1^T (D_1 x^k + D_2 y^k + z^k) - \beta A^T \left(y^k - z^k + b + \frac{\lambda^k}{\beta} \right) \right]; \frac{2e}{\mu_1} \right), \\ \lambda^{k+\frac{1}{2}} &= \lambda^k - r\beta(Ax^{k+1} + y^k + z^k - b), \\ y^{k+1} &= \frac{1}{1+\beta} \left[-D_2^T (D_1 x^k + D_2 y^k + z^k) - \beta \left(Ax^{k+1} - z^k + b + \frac{\lambda^k}{\beta} \right) \right], \\ z^{k+1} &= \frac{1}{1+\beta} \left[-D_1 x^{k+1} + D_2 y^{k+1} + \beta \left(Ax^{k+1} + y^{k+1} - b - \frac{\lambda^k}{\beta} \right) \right], \\ \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} + y^{k+1} + z^{k+1} - b), \end{aligned} \quad (4.4)$$

Table 1: Comparison of iteration effect between 3-BPRSM and LBADMM

(n, m, m)	Alg		Iter (k)					$\min \ r\ _2 / f\text{-val}$
			30	60	90	120	150	
(1,500, 1,500, 1,500)	LBADMM	$\ r\ _2$	0.0149	0.0095	0.005	—	—	0.0038
		$f\text{-val}$	426.08	346.53	324.89	—	—	318.91
	3-BPRSM	$\ r\ _2$	0.6265	0.0569	0.0245	0.0177	0.0146	0.0039
		$f\text{-val}$	438.59	319.35	290.67	280.04	274.69	260.78
(3,000, 3,000, 3,000)	LBADMM	$\ r\ _2$	0.0198	0.0096	0.0062	—	—	0.0054
		$f\text{-val}$	776.68	642.71	606.32	—	—	603.88
	3-BPRSM	$\ r\ _2$	0.8995	0.1094	0.0392	0.0266	0.0235	0.0054
		$f\text{-val}$	799.43	588.99	536.66	514.98	503.62	484.49
(6,000, 6,000, 6,000)	LBADMM	$\ r\ _2$	0.0295	0.0161	0.0103	—	—	0.0075
		$f\text{-val}$	1482.94	1212.92	1135.46	—	—	1121.64
	3-BPRSM	$\ r\ _2$	0.6311	0.0728	0.0455	0.0326	0.0295	0.0077
		$f\text{-val}$	1518.09	1126.23	1043.22	1010.47	991.68	959.42

where $S(\cdot, \cdot)$ represents the half shrinkage operator [10], which is defined as $S(x, \tau) = \{s_\tau(x_1), \dots, s_\tau(x_n)\}^T$ with

$$s_\tau(x_i) = \begin{cases} \frac{2x_i}{3} \left(1 + \cos \frac{2}{3}(\pi - \wp(x)) \right), & |x_i| > \frac{\sqrt[3]{54}}{4} \tau^{\frac{2}{3}}, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\wp(x) = \arccos \left(\frac{\tau}{8} \left(\frac{|x|}{3} \right)^{-\frac{3}{2}} \right).$$

In this example, the measurement matrix A is generated from a normal distribution $N(0, 1)$ and normalized to have unit $(l_{\frac{1}{2}})$ norm for its columns. The variables x and y are generated with 100 nonzero entries a Gaussian distribution. The initial values x^0 , y^0 , z^0 , and λ^0 are all initialized to be 0. The vector $b = Ax^0 + y^0 + z^0 + v$, where $v \sim N(0, 10^{-3}I)$. We set $\mu_1 = 30$, $\beta = 20$ and the regularization parameters $c = 0.1$. At the iteration point of k , the residual at iteration k is defined as $r_k = Ax^k + y^k + z^k - b$. The termination criterion is defined as follows:

$$\|r_k\|_2 \leq \sqrt{m} 10^{-4}.$$

To validate the effectiveness, we compare 3-BPRSM with LBADMM [34]. The numerical results are presented in Table 1. The code is implemented using MATLAB R2022a, and the computations are performed on a computer running Windows 10, Intel (R) Core (TM) i7-8550U 1.80GHz CPU, with 8GB of memory. The reported results include the number of iterations (Iter (k)) and the objective function value ($f\text{-val}$). When $n = 1,500$ and $m = 1,500$, we conduct a total of 435 iterations, resulting in a convergence value of 260.78. As the values of n and m increase, our iteration time extends, but we can still ensure the convergence of the algorithm and guarantee that 3-BPRSM outperforms LBADMM.

A part of computational results are presented in Figures 2–4, which show the trends of the objective value (objective-value) and the residual defined as $\|r_k\| = \|Ax^k + y^k + z^k - b\|$ ($\|r\|_2$).

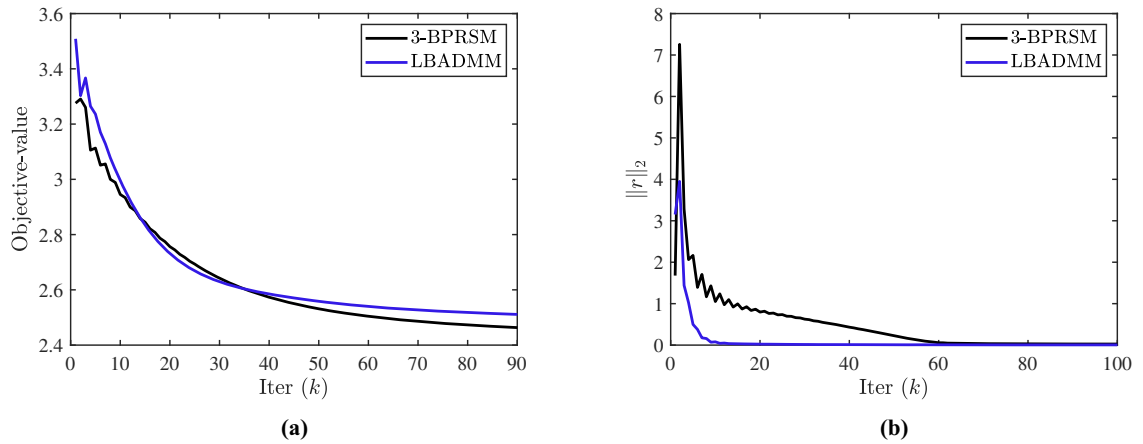


Figure 2: Comparison of convergence when $n = 1,500$ and $m = 1,500$. (a) The objective value when $n = 1,500$ and $m = 1,500$ and (b) $\|r\|_2$ under $n = 1,500$ and $m = 1,500$.

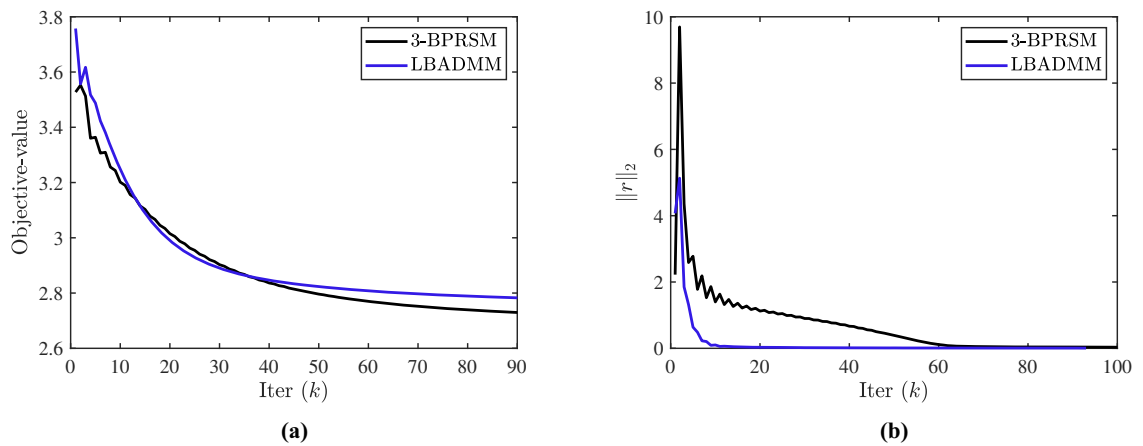


Figure 3: Comparison of convergence when $n = 3,000$ and $m = 3,000$. (a) The objective value when $n = 3,000$ and $m = 3,000$. and (b) $\|r\|_2$ under $n = 3,000$ and $m = 3,000$.

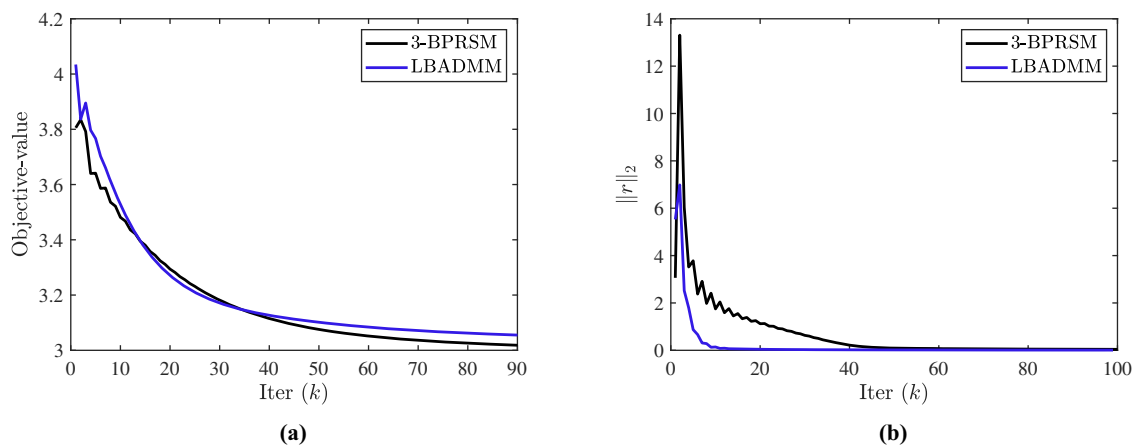


Figure 4: Comparison of convergence when $n = 6,000$ and $m = 6,000$. (a) The objective value when $n = 6,000$ and $m = 6,000$. and (b) $\|r\|_2$ under $n = 6,000$ and $m = 6,000$.

5 Concluding remarks

In this article, we propose a novel 3-BPRSM, which tailored for three-block optimization problems characterized by a nonconvex and nonseparable structure. The efficacy of 3-BPRSM for solving the problem of sparse signal reconstruction in compressed sensing models based on incomplete observation datasets is demonstrated, which is able to deal with nonconvexity and indivisibility and is therefore particularly suited to the case of incomplete datasets. By integrating 3-BPRSM, we aim to improve the reconstruction of sparse signals and contribute to the wider field of compressed sensing. Finally, we perform a comparative analysis to demonstrate the superior performance of 3-BPRSM over LBADMM.

In summary, we have not only theoretically verified the convergence and convergence rates of 3-BPRSM but also validated its effectiveness through numerical examples in practical applications. However, despite the encouraging achievements of 3-BPRSM, it is evident that further refinement is needed, particularly in the context of sparse signal reconstruction from an incomplete observed dataset in compressive sensing. We aim to identify examples that more closely align with real-world scenarios to articulate the practical significance of our algorithm more clearly.

Moreover, we think that the following open questions for reference by compare with [46–48] will play a positive guiding role in the future research:

- (a) In the problem (1.3), the vectors y and z have the same dimension m , and this choice is a special case. Can the problem (1.3) be generalized such that y and z have different dimensions? If the constraint is $Ax + By + Cz = 0$, then it can be extended to a more general constrained optimization. Can convergence be guaranteed in this case?
- (b) In the method proposed in this article, only Bregman distance is added to the subproblem of x . If Bregman distance is added to the subproblems of y and z , can the numerical effect be further improved?
- (c) Is it possible to deal the nonseparable structure in the objective function with linear approximation to solve some cases, which are difficult to solve subproblems because of the nonseparable structure?
- (d) In validating the effectiveness of Algorithm 1, we considered one synthetic numerical example. How to solve statement of real practical problems via using 3-BPRSM proposed in this article? This will help to demonstrate the applicability of the algorithm in real scenarios and provide the reader with a more practical validation in the future work.

Acknowledgements: The authors are grateful for the reviewer's valuable comments and constructive remarks that improved the manuscript.

Funding information: This work was partially supported by the Innovation Fund of Postgraduate, Sichuan University of Science & Engineering (Y2023333) and the Scientific Research and Innovation Team Program of Sichuan University of Science and Technology (SUSE652B002).

Author contributions: All authors have read and agreed to the final manuscript. YZ, LH, and XH designed the conceptualization and models; YZ and XH formally analyzed, wrote the main manuscript text, and developed the model code and performed the simulations; LH supervised, reviewed, and edited the manuscript.

Conflict of interest: No potential conflict of interest was reported by the authors.

Ethical approval: The conducted research is not related to either human or animal use.

Data availability statement: Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

References

- [1] M. Umar, Z. Sabir, M. A. Z. Raja, H. M. Baskonus, M. R. Ali, and N. A. Shah, *Heuristic computing with sequential quadratic programming for solving a nonlinear hepatitis B virus model*, *Math. Comput. Simulat.* **212** (2023), 234–248, DOI: <https://doi.org/10.1016/j.matcom.2023.04.034>.
- [2] Z. Sabir, M. A. Z. Raja, H. M. Baskonus, and A. Ciancio, *Numerical performance using the neural networks to solve the nonlinear biological quarantined based COVID-19 model*, *Atti Accad. Peloritana Pericolanti Cl. Sci. Fis. Mat. Natur.* **101** (2023), no. 1, A10, 18.
- [3] A. Ayub, Z. Sabir, S. B. Said, H. M. Baskonus, R. Sadat, and M. R. Ali, *Nature analysis of Cross fluid flow with inclined magnetic dipole*, *Microsyst. Technol.* **29** (2023), 697–714, DOI: <https://doi.org/10.1007/s00542-023-05438-5>.
- [4] Q. L. Chen, Z. Sabir, M. A. Z. Raja, W. Gao, and H. M. Baskonus, *A fractional study based on the economic and environmental mathematical model*, *Alexandria Eng. J.* **65** (2023), 761–770, DOI: <https://doi.org/10.1016/j.aej.2022.09.033>.
- [5] Z. Sabir, M. Umar, M. A. Z. Raja, H. M. Baskonus, and W. Gao, *Designing of Morlet wavelet as a neural network for a novel prevention category in the HIV system*, *Int. J. Biomath.* **15** (2022), no. 4, Paper no. 2250012, 22 pp, DOI: <https://doi.org/10.1142/S1793524522500127>.
- [6] Z. Sabir, H. A. Wahab, S. Javeed, and H. M. Baskonus, *An efficient stochastic numerical computing framework for the nonlinear higher order singular models*, *Fractal Fract.* **5** (2021), no. 4, Paper No. 176, 14 pp, DOI: <https://doi.org/10.3390/fractalfract5040176>.
- [7] M. Umar, Z. Sabir, M. A. Z. Raja, H. M. Baskonus, S. W. Yao, and E. Ilhan, *A novel study of Morlet neural networks to solve the nonlinear HIV infection system of latently infected cells*, *Results Phys.* **25** (2021), Paper No. 104235, 13 pp, DOI: <https://doi.org/10.1016/j.rinp.2021.104235>.
- [8] Y. G. Sánchez, Z. Sabir, H. Günerhan, and H. M. Baskonus, *Analytical and approximate solutions of a novel nervous stomach mathematical model*, *Discrete Dyn. Nat. Soc.* **2020** (2020), no. 1, 5063271, DOI: <https://doi.org/10.1155/2020/5063271>.
- [9] W. Bian and X. J. Chen, *Linearly constrained non-Lipschitz optimization for image restoration*, *SIAM J. Imaging Sci.* **8** (2015), 2294–2322, DOI: <https://doi.org/10.1137/140985639>.
- [10] Z. B. Xu, X. Y. Chang, F. M. Xu, and H. Zhang, *$L_{1/2}$ regularization: A thresholding representation theory and a fast solver*, *IEEE Trans. Neural Network Learn. Sys.* **23** (2012), no. 7, 1013–1027, DOI: <https://doi.org/10.1109/TNNLS.2012.2197412>.
- [11] Z. C. Lin, R. S. Liu, and H. Li, *Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning*, *Mach. Learn.* **99** (2015), 287–325, DOI: <https://doi.org/10.1007/s10994-014-5469-5>.
- [12] B. P. W. Ames and M. Y. Hong, *Alternating direction method of multipliers for penalized zero-variance discriminant analysis*, *Comput. Optim. Appl.* **64** (2016), 725–754, DOI: <https://doi.org/10.1007/s10589-016-9828-y>.
- [13] F. Lin, M. Fardad, and M. R. Jovanovic, *Design of optimal sparse feedback gains via the alternating direction method of multipliers*, *IEEE Trans. Autom. Control* **58** (2013), no. 9, 2426–2431, DOI: <https://doi.org/10.1109/TAC.2013.2257618>.
- [14] K. Guo, D. R. Han, and T. T. Wu, *Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints*, *Int. J. Comput. Math.* **94** (2017), 1653–1669, DOI: <https://doi.org/10.1080/00207160.2016.1227432>.
- [15] F. H. Wang, Z. B. Xu, and H. K. Xu, *Convergence of alternating direction method with multipliers for non-convex composite problems*, *arXiv preprint* (2014), <https://arxiv.org/abs/1410.8625>.
- [16] M. Ohishia, K. Fukuib, K. Okamurac, Y. Itohc, and H. Yanagihara, *Coordinate optimization for generalized fused Lasso*, *Commun. Stat. Theory Methods* **50** (2021), 5955–5973, DOI: <https://doi.org/10.1080/03610926.2021.1931888>.
- [17] X. L. Lu and X. B. Lü, *ADMM for image restoration based on nonlocal simultaneous sparse Bayesian coding*, *Signal Process Image Commun.* **70** (2019), 157–173, DOI: <https://doi.org/10.1016/j.image.2018.09.012>.
- [18] B. Wahlberg, S. Boyd, M. Annergren, and Y. Wang, *An ADMM algorithm for a class of total variation regularized estimation problems*, *IFAC Proc.* **45** (2012), 83–88, DOI: <https://doi.org/10.3182/20120711-3-BE-2027.00310>.
- [19] M. Meselhi, R. Sarker, D. Essam, and S. Elsayed, *A decomposition approach for large-scale non-separable optimization problems*, *Appl. Soft. Comput.* **115** (2022), 108168, DOI: <https://doi.org/10.1016/j.asoc.2021.108168>.
- [20] P. J. Liu, J. B. Jian, B. Heeee, and X. Z. Jiang, *Convergence of Bregman Peaceman-Rachford splitting method for nonconvex nonseparable optimization*, *J. Oper. Res. Soc. China* **11** (2023), 707–733, DOI: <https://doi.org/10.1007/s40305-022-00411-x>.
- [21] B. S. He and X. M. Yuan, *A class of ADMM-based algorithms for three-block separable convex programming*, *Comput. Optim. Appl.* **70** (2018), no. 3, 791–826, DOI: <https://doi.org/10.1007/s10589-018-9994-1>.
- [22] L. M. Zeng and J. Xie, *Group variable selection via SCAD- L_2* , *Statistics* **48** (2014), no. 1, 49–66, DOI: <https://doi.org/10.1080/02331888.2012.719513>.
- [23] C. Zhang, Y. Z. Song, X. J. Cai, and D. R. Han, *An extended proximal ADMM algorithm for three-block nonconvex optimization problems*, *J. Comput. Appl. Math.* **398** (2021), Paper no. 113681, pp. 1–14, DOI: <https://doi.org/10.1016/j.cam.2021.113681>.
- [24] L. Yang, T. K. Pong, and X. J. Chen, *Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction*, *SIAM J. Imaging Sci.* **10** (2017), 74–110, DOI: <https://doi.org/10.1137/15M1027528>.
- [25] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, *Comput. Math. Appl.* **2** (1976), 17–40, DOI: [https://doi.org/10.1016/0898-1221\(76\)90003-1](https://doi.org/10.1016/0898-1221(76)90003-1).
- [26] J. Eckstein and D. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, *Math. Program.* **55** (1992), 293–318, DOI: <https://doi.org/10.1007/BF01581204>.
- [27] D. W. Peaceman and J. H. H. Rachford, *The numerical solution of parabolic and elliptic differential equations*, *J. Soc. Indust. Appl. Math.* **3** (1955), 28–41, DOI: <https://doi.org/10.1137/0103003>.

- [28] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *SIAM J. Multiscale Model. Simul.* **4** (2005), 1168–1200, DOI: <https://doi.org/10.1137/050626090>.
- [29] D. Gabay, *Chapter IX applications of the method of multipliers to variational inequalities*, *Stud. Math. Appl.* **15** (1983), 299–331, DOI: [https://doi.org/10.1016/S0168-2024\(08\)70034-1](https://doi.org/10.1016/S0168-2024(08)70034-1)
- [30] G. Y. Li and T. K. Pong, *Global convergence of splitting methods for nonconvex composite optimization*, *SIAM J. Optim.* **25** (2015), 2434–2460, DOI: <https://doi.org/10.1137/140998135>.
- [31] Z. H. Jia, X. Gao, X. Cai, and D. R. Han, *Local linear convergence of the alternating direction method of multipliers for nonconvex separable optimization problems*, *J. Optim. Theory Appl.* **188** (2021), 1–25, DOI: <https://doi.org/10.1007/s10957-020-01782-y>.
- [32] R. I. Bot and D. K. Nguyen, *The proximal alternating direction method of multipliers in the nonconvex setting: convergence analysis and rates*, *Math. Oper. Res.* **45** (2020), 682–712, DOI: <https://doi.org/10.1287/moor.2019.1008>.
- [33] F. H. Wang, W. F. Cao, and Z. B. Xu, *Convergence of multiblock Bregman ADMM for nonconvex composite problems*, *Sci. China Inf. Sci.* **61** (2018), no. 12, Paper No. 122101, pp. 1–12, DOI: <https://doi.org/10.1007/s11432-017-9367-6>.
- [34] M. T. Chao, Z. Deng, and J. B. Jian, *Convergence of linear Bregman ADMM for nonconvex and nonsmooth problems with nondivisible structure*, *Complexity* **2020** (2020), Paper No. 6237942, pp. 1–14, DOI: <https://doi.org/10.1155/2020/6237942>.
- [35] C. H. Chen, B. S. He, Y. Y. Ye, and X. M. Yuan, *The direct extension of ADMM for multiblock convex minimization problems is not necessarily convergent*, *Math. Program.* **155** (2016), 57–79, DOI: <https://doi.org/10.1007/s10107-014-0826-5>.
- [36] A. Bnouhachem and M. T. Rassias, *A Bregman proximal Peaceman-Rachford splitting method for convex programming*, *Appl. Set-Valued Anal. Optim.* **4** (2022), 129–143, DOI: <https://doi.org/10.23952/asvao.4-2022.2.01>.
- [37] P. Li, Y. Shen, S. Jiang, Z. Liu, and C. H. Chen, *Convergence study on strictly contractive Peaceman-Rachford splitting method for nondivisible convex minimization models with quadratic coupling terms*, *Comput. Optim. Appl.* **78** (2021), 87–124, DOI: <https://doi.org/10.1007/s10589-020-00229-4>.
- [38] F. X. Liu, L. L. Xu, Y. H. Sun, and D. R. Han, *A proximal alternating direction method for multiblock coupled convex optimization*, *J. Ind. Manag. Optim.* **15** (2018), no. 2, 723–737, DOI: <https://doi.org/10.3934/jimo.201-8067>.
- [39] M. T. Chao, C. Z. Cheng, and D. Y. Liang, *A proximal block minimization method of multipliers with a substitution procedure*, *Optim. Methods Softw.* **30** (2015), 825–842, DOI: <https://doi.org/10.1080/10556788.2014.992432>.
- [40] L. M. Bregman, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, *USSR Comput. Math. Math. Phys.* **7** (1967), no. 3, 200–217, DOI: [https://doi.org/10.1016/0041-5553\(67\)90040-7](https://doi.org/10.1016/0041-5553(67)90040-7).
- [41] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, *Clustering with bregman divergences*, *J. Mach. Learn. Res.* **6** (2005), 1705–1749, DOI: <https://dl.acm.org/doi/pdf/10.5555/1046920.1194902>.
- [42] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality*, *Math. Oper. Res.* **35** (2010), 438–457, DOI: <https://doi.org/10.48550/arXiv.0801.1780>.
- [43] J. Bolte, S. Sabach, and M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, *Math. Program.* **146** (2014), 459–494, DOI: <https://doi.org/10.1007/s10107-013-0701-9>.
- [44] Y. Nesterov, *Introduction Lectures on Convex Optimization: A Basic Course*, Springer Science & Business Media, Berlin, 2013.
- [45] H. Attouch and J. Bolte, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*, *Math. Program.* **116** (2009), 5–16, DOI: <https://doi.org/10.1007/s10107-007-0133-5>.
- [46] M. N. Raihen and S. Akter, *Prediction modeling using deep learning for the classification of grape-type dried fruits*, *Int. J. Math. Comput. Eng.* **2** (2024), no. 1, 1–12, DOI: <https://doi.org/10.2478/ijmce-2024-0001>.
- [47] M. Omar and D. Burrell, *From text to threats: A language model approach to software vulnerability detection*, *Int. J. Math. Comput. Eng.* **2** (2024), no. 1, 23–34, DOI: <https://doi.org/10.2478/ijmce-2024-0003>.
- [48] D. Dalal, P. Kumar, and C. Cattani, *Optimizing industrial growth through alternative forest biomass resources: A mathematical model using DDE*, *Int. J. Math. Comput. Eng.* **1** (2023), no. 2, 187–200, DOI: <https://doi.org/10.2478/ijmce-2023-0015>.