FIXED POINT METHODS AND OPTIMIZATION
Volume 2 (2025), No. 2, 127–148
https://doi.org/10.69829/fpmo-025-0202-ta03
Tulipa Opera Scholarum

A CLASS OF NOVEL THREE-BLOCK BREGMAN-TYPE PEACEMAN–RACHFORD SPLITTING METHODS WITH LINEAR APPROXIMATION FOR SOLVING SPARSE SIGNAL RECONSTRUCTION PROBLEMS

YING ZHAO[1,2] AND HENG-YOU LAN[1,3,*]

[1] *College of Mathematics and Statistics, Sichuan University of Science & Engineering, Zigong, 643000, Sichuan, PR China*
[2] *Chengdu Qingbaijiang Experimental Primary School, Chengdu, 610300, Sichuan, PR China*
[3] *Sichuan Province University Key Laboratory of Bridge Non-destruction Detecting and Engineering Computing, Zigong, 643000, Sichuan, PR China*

Dedicated to Professor Hari Mohan Srivastava on the Occasion of His 85th Birthday

Abstract. Exploring three-block nonconvex optimization with a nonseparable structure has substantial theoretical significance and potential applications in nonconvex background or foreground extraction problems such as image and signal processing, phase retrieval, and so on. A class of novel generalized three-block Bregman-type Peaceman–Rachford splitting methods are proposed, which integrates the inexact concepts of linear approximation. Under some generalization assumptions, the optimality condition is used to establish global convergence. Furthermore, via constituting Cauchy sequence, strong convergence is proved when the augmented Lagrangian function for the three-block nonconvex and nonseparable optimizations satisfies Kurdyka–Łojasiewicz property. Lastly, a preliminary numerical application experiment associated with sparse signal reconstruction confirms the effectiveness.

Keywords. Convergence analysis, Novel generalized Peaceman–Rachford splitting method, Three-block nonconvex and nonseparable optimization, Linear approximation under Bregman distance, Kurdyka-Łojasiewicz property.

1. Introduction

Let $f\colon \mathbb{R}^{n_1} \to \mathbb{R}\bigcup\{+\infty\}$, $g\colon \mathbb{R}^{n_2} \to \mathbb{R}\bigcup\{+\infty\}$ and $h\colon \mathbb{R}^{n_3} \to \mathbb{R}\bigcup\{+\infty\}$ be proper and lower semicontinuous, $l\colon \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\mathbb{R}^{n_3} \to \mathbb{R}$ be continuously differentiable, and $A \in \mathbb{R}^{m\times n_1}$, $B \in \mathbb{R}^{m\times n_2}$, $C \in \mathbb{R}^{m\times n_3}$ and $b \in \mathbb{R}^m$ be given matrix or vector. In this paper, we shall tackle the following nonconvex and nonseparable optimization problem with a general linear constraint:

$$\begin{aligned} \min \quad & f(x) + g(y) + h(z) + l(x, y, z), \\ \text{s.t.} \quad & Ax + By + Cz = b. \end{aligned} \tag{1.1}$$

Significantly, the problem (1.1) finds numerous applications across different domains, e.g. image and signal processing [26, 31]. It includes techniques like fused Lasso [11, 29], group Lasso [38], and total variation regularization [32]. Additionally, it is relevant in areas such as phase retrieval [35] and nonconvex background/foreground extraction problems [37], specifically when addressing large-scale issues [27].

*Corresponding author.
E-mail address: hengyoulan@163.com (H.-Y. Lan), z_zhaoying@163.com (Y. Zhao)
2020 Mathematics Subject Classification: 90C26, 65K05, 49K35, 41A25.
Accepted: June 24, 2025.

In particular, Zhao et al. [39] addressed a specific instance of (1.1) where $n_2 = n_3 = m$, and B and C are identity matrices. The modified problem is as follows

$$\begin{aligned} \min \quad & f(x) + g(y) + h(z) + l(x, y, z), \\ \text{s.t.} \quad & Ax + y + z = b. \end{aligned} \tag{1.2}$$

Further, if $C = O$, null matrix, and $h(z) + l(x, y, z) \equiv \Im(x, y)$ for any $(x, y, z) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$, then (1.1) reduces to the following two nonseparable problem:

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

which includes two separable optimization problems. Thus, unlike the two separable/nonseparable problem (1.3) discussed in [18, 19] or the problem (1.2), the problem (1.1) that we are examining is more.

As everyone knows, several methods are available for solving the nonseparable optimization problems with linear constraints, which including alternating direction method of multipliers (ADMM) [16], Douglas–Rachford splitting method [3, 14, 22], and Peaceman–Rachford splitting method (PRSM) [12, 30]. We note that the main method currently employed to solve the forms of (1.1) and (1.3) is ADMM, and recent investigations have shown that incorporating appropriate Bregman distances into ADMM can significantly simplify subproblem calculations or enable closed-form solutions, thereby enhancing numerical efficiency. For instance, Li and Pong [21] introduced a variant of ADMM by integrating a Bregman distance parameter into the second subproblem. Wang et al. [34] extended this idea by incorporating Bregman distances into two subproblems to relax constraints on the objective function. Substantial progress has been made over the past few years in enhancing ADMM and its comprehensive reviews, refer to [5, 20, 33] and references therein.

While ADMM has been extensively used to address the problem (1.1), recent research has increasingly focused on PRSM as well. PRSM differs from ADMM primarily due to its inclusion of an additional intermediate update step, resembling a symmetric variant of ADMM. As noted by Gabay [15], PRSM typically achieves faster convergence than ADMM, although it may require stricter conditions for convergence, striking a balance between speed and robustness. Moreover, PRSM has been successfully extended to handle multi-block problems [8, 25], highlighting its versatility across various optimization contexts. Thus, there is a natural interest in adapting PRSM to address (1.1). For instance, Liu et al. [23] proposed Bregman-PRSM for (1.3). In the context of the three-block nonseparable the problem (1.1), Chao et al. [9] introduced Linear Bregman ADMM (LBADMM). However, it's worth noting that direct extensions of multi-block convex optimization problems, as shown by Chen et al. [10], do not always guarantee convergence, which has sparked significant interest and discussion in the research community. Explorations into three-block nonseparable problems are still relatively new and evolving.

Using splitting algorithms like ADMM or PRSM, particularly in the subproblem iteration of x for separable optimization, that is, the nonseparable term $l(x, y, z) = 0$ in (1.1) for each $(x, y, z) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$, the iterative step is as follows

$$\begin{aligned} x^{k+1} &\in \arg\min_x \left\{ \mathcal{L}_\beta(x, y^k, z^k, \Lambda^k) \right\} \\ &= \arg\min_x \left\{ f(x) + \frac{\beta}{2} \| Ax + By^k + Cz^k - b - \frac{1}{\beta}\Lambda^k \|^2 \right\}, \end{aligned} \tag{1.4}$$

where

$$\begin{aligned} \mathcal{L}_\beta(x, y, z, \Lambda) = {} & f(x) + g(y) + h(z) + l(x, y, z) - \langle \Lambda, Ax + By + Cz - b \rangle \\ & + \frac{\beta}{2} \| Ax + By + Cz - b \|^2 \end{aligned} \tag{1.5}$$

is the augmented Lagrange function (ALF) with the Lagrange multiplier $\Lambda \in \mathbb{R}^m$ for (1.1). In practical scenarios, solving (1.4) as $n_1 = m$ becomes straightforward when A is a unit matrix. However, this is not always the case when A is not a unit matrix. To address this, one approach is to linearize the Lagrangian penalty term $\frac{\beta}{2}\left\|Ax + By^k + Cz^k - b\right\|^2$ by introducing a proximity term $\frac{1}{2}\|x - x^k\|_T^2$ during the subproblem update. Specifically, when $T = \alpha I - \beta A^\top A$, the subproblem for x is formulated as

$$x^{k+1} \in \arg\min_x \left\{ \mathcal{L}_\beta(x, y^k, z^k, \Lambda^k) + \frac{1}{2}\left\|x - x^k\right\|_T^2 \right\} = \arg\min_x \left\{ f(x) + \frac{\alpha}{2}\left\|x - b^k\right\|^2 \right\}, \quad (1.6)$$

here b^k is a known quantity. Clearly, solving (1.6) is often simplified compared to the original the problem (1.4). The choice of an appropriate matrix T plays a crucial role in streamlining the solution process, expediting iterative algorithms, and improving overall numerical effectiveness.

However, in many optimization scenarios, even with the inclusion of proximal terms in subproblems, solving (1.6) remains challenging, especially if $f(x)$ is difficult to minimize. To address this, an alternative approach involves replacing the original function $f(x)$ with a linear approximation at each iteration, leveraging its differentiability. This linear approximation, $f(x) \approx f(x^k) + (x - x^k)^\top \nabla f(x^k)$, differs significantly from the proximity linearization used for the Lagrangian penalty term $\frac{\beta}{2}\left\|Ax + By^k + Cz^k - b\right\|^2$. While this inexact computation of x updates may increase the total number of iterations and potentially slow down convergence, it significantly reduces the cost per iteration, leading to overall computational savings. This technical approach of using linear approximations to simplify the nonseparable structure in objective functions is explored in various previous work. See, for example, [4, 17, 24] and their references. In summary, while choosing an appropriate matrix T can simplify the solution process in iterative algorithms, the use of linear approximations for nonseparable objective functions offers a practical compromise between computational efficiency and convergence speed in challenging optimization problems.

Based on the above analysis, in this paper, we consider linear approximation of $l(x, y, z)$ in the problem (1.1) as follows

$$\begin{aligned}
&l(x, y^k, z^k) \approx l(x^k, y^k, z^k) + (x - x^k)^\top \nabla_x l(x^k, y^k, z^k),\\
&l(x^{k+1}, y, z^k) \approx l(x^{k+1}, y^k, z^k) + (y - y^k)^\top \nabla_y l(x^{k+1}, y^k, z^k),\\
&l(x^{k+1}, y^{k+1}, z) \approx l(x^{k+1}, y^{k+1}, z^k) + (z - z^k)^\top \nabla_z l(x^{k+1}, y^{k+1}, z^k),
\end{aligned} \quad (1.7)$$

here $\nabla_x l$ denotes the gradient of the vector x for the function l. Aiming at the problem (1.1), combining with Bregman distance, (1.7) and PRSM, the following novel generalized three-block Bregman Peaceman–Rachford splitting method (3-GBPRSM) is proposed:

$$\begin{cases} x^{k+1} \in \arg\min_x \{f(x) + (x - x^k)^\top \nabla_x l(x^k, y^k, z^k) - \langle \Lambda^k, Ax\rangle \\ \qquad + \frac{\beta}{2}\left\|Ax + By^k + Cz^k - b\right\|^2 + \triangle_{\phi_1}(x, x^k)\}, \\ \Lambda^{k+\frac{1}{2}} = \Lambda^k - r\beta(Ax^{k+1} + By^k + Cz^k - b), \\ y^{k+1} \in \arg\min_y \{g(y) + (y - y^k)^\top \nabla_y l(x^{k+1}, y^k, z^k) - \langle \Lambda^{k+\frac{1}{2}}, By\rangle \\ \qquad + \frac{\beta}{2}\left\|Ax^{k+1} + By + Cz^k - b\right\|^2 + \triangle_{\phi_2}(y, y^k)\}, \\ z^{k+1} \in \arg\min_z \{h(z) + (z - z^k)^\top \nabla_z l(x^{k+1}, y^{k+1}, z^k) - \langle \Lambda^{k+\frac{1}{2}}, Cz\rangle \\ \qquad + \frac{\beta}{2}\left\|Ax^{k+1} + By^{k+1} + Cz - b\right\|^2 + \triangle_{\phi_3}(z, z^k)\}, \\ \Lambda^{k+1} = \Lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} + By^{k+1} + Cz^{k+1} - b), \end{cases} \tag{1.8}$$

where $\beta > 0$ is the penalty parameter, r and s are constants, $\triangle_\phi$ is Bregman distance with respect to the differentiable convex function $\phi : \mathbb{R}^n \to \mathbb{R}$, namely (see [7])

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

Specifically, when $\phi(x) = \|x\|^2$, Bregman distance simplifies to $\|x - y\|^2$, which corresponds to the classical Euclidean distance [2].

In this paper, the primary focus and key contribution revolve around the introduction and proof of convergence for the novel 3-GBPRSM. This method is specifically designed to tackle nonconvex and nonseparable optimization problems as defined in (1.1).

The novelties of this paper lie in the following aspects:

- The problems (1.1) are generalised and no longer special cases. Three-block nonconvex and nonseparable optimization problem (1.1) is solved under the constraint that B and C are not identity matrices.
- Expanding the range of values for the relaxation factors (r, s) broadens the representativeness of 3-GBPRSM.
- Proposing an innovative linear approximation technique in the algorithm design to tackle the presence of nonseparable structures in complex optimization problems.
- It presents a fresh perspective to introduce a novel optimization strategy via incorporating Bregman distance into the computation of subproblems.

As a result, 3-GBPRSM offers a versatile splitting algorithm framework for addressing the problem (1.1), allowing for a comprehensive analysis of the theoretical properties applicable to a wide range of Bregman-type PRSM in a unified manner.

The paper is organised as follows: In Section 2, we describe the basics and establish the foundational framework for subsequent discussions. Then, we shall delve into convergence properties of 3-GBPRSM for the three-block nonconvex and nonseparable problem (1.1) in Section 3. In Section 4, we present a sparse signal reconstruction numerical example to demonstrate and validate our key findings. Finally, we offer conclusions and potential directions for future research in Section 5.

2. Preliminaries

Denote $\|\cdot\|$ as Euclidean norm of a vector. Let $\mathrm{dom}f = \{x \in \mathbb{R}^n : f(x) < +\infty\}$ be the domain of function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. For a symmetric semi-positive (positive) definite matrix $H \succeq (\succ)\ 0$, we know that $\|x\|_H^2 = x^\top Hx$ and $\Lambda_{\min}(H)\|x\|^2 \le x^\top Hx \le \Lambda_{\max}(H)\|x\|^2$ 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 H, respectively.

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

Definition 2.2. The distance from a point $x \in \mathbb{R}^n$ to a set $\mathbb{S} \subseteq \mathbb{R}^n$ is defined as $d(x, \mathbb{S}) = \inf_{y\in\mathbb{S}} \|y-x\|$. Specifically, if $\mathbb{S} = \emptyset$, then $d(x, \mathbb{S}) = +\infty$.

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

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha)f(x_2)$$

holds, then f is termed convex on $\mathbb{S}$. Sepcifically, if $x_1 \neq x_2$, and the strict inequality

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

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

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

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

then f is said to be strongly α-convex.

Definition 2.4. A differentiable function f is defined on a nonempty open convex set $\mathbb{S} \subset \mathbb{R}^n$. Then

(i) f is convex if and only if either

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

or

$$\langle \nabla f(x) - \nabla f(y), x - y\rangle \ge 0,\ \forall x, y \in \mathbb{S};$$

(ii) f is strongly σ-convex if and only if there exists a constant $\sigma > 0$ such that

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

Definition 2.5. The subdifferential of a proper lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R}$ exhibits the following several fundamental properties:

(i) For each $x \in \mathbb{R}^n$, $\hat{\partial} f(x) \subseteq \partial f(x)$, and $\partial f(x)$ is closed if $\hat{\partial} f(x)$ is a closed convex set.
(ii) If $x_k^* \in \partial f(x_k)$ and $\lim_{k\to\infty}(x_k, x_k^*) = (x, x^*)$, then $x^* \in \partial f(x)$, ensuring $\partial f(x)$ is closed.
(iii) $0 \in \partial f(\hat{x})$ when $\hat{x} \in \mathbb{R}^n$ is a local minimum of f. $\hat{x}$ is a critical point of f whenever $0 \in \partial f(\hat{x})$, the set of critical points of f is denoted by $\mathrm{crit} f$.
(iv) For a continuously differentiable function $g : \mathbb{R}^n \to \mathbb{R}$, the subdifferential of the sum $\partial(f + g)(x)$ satisfies $\partial(f + g)(x) = \partial f(x) + \nabla g(x)$ for all $x \in \mathrm{dom}f$.

The set Φ_η consists of all concave functions $\varphi : [0, \eta) \to [0, +\infty)$ that adhere to the following criteria: (i) $\varphi(0) = 0$; (ii) φ is continuous at 0 and differentiable on $(0, \eta)$; (iii) $\varphi'(t) > 0$ for every $t \in (0, \eta)$.

Lemma 2.6. ([1]) (Kurdyka–Łojasiewicz property, KLP) *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function, and let $\bar{x} \in \mathrm{dom}(\partial f)$, where $dom(\partial f)$:=$\{x \in \mathbb{R}^n | \partial f(x) \neq \emptyset\}$. Define $[\eta_1 < f < \eta_2] := \{x \in \mathbb{R}^n | \eta_1 < f(x) < \eta_2\}$. 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 holds 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 referred to as satisfying KLP *at $\bar{x}$. Meanwhile, φ is the associate function of f with* KLP.

Lemma 2.7. ([7]) *Let $\triangle_\phi(x, y)$ denote Bregman distance corresponding to a differentiable convex function $\phi : \mathbb{R}^n \to \mathbb{R}$. The properties of $\triangle_\phi(x, y)$ are listed:*

(i) *Nonnegative: For all $x, y \in \mathbb{R}^n$, $\triangle_\phi(x, y) \geq 0$ and $\triangle_\phi(x, x) = 0$;*
(ii) *Convexity: $\triangle_\phi(x, y)$ is convex with respect to x, although it may not be convex at y;*
(iii) *Strong convexity: If ϕ is strongly σ-convex, then for all $x, y \in \mathbb{R}^n$, $\triangle_\phi(x, y) \geq \frac{\sigma}{2} \|x - y\|^2$ holds.*

Lemma 2.8. ([1]) *Assume that $H(x, y, z) = p(x) + q(y) + h(z)$, here $p : \mathbb{R}^{n_1} \to \mathbb{R} \cup \{+\infty\}$, $q : \mathbb{R}^{n_2} \to \mathbb{R} \cup \{+\infty\}$ and $h : \mathbb{R}^{n_3} \to \mathbb{R} \cup \{+\infty\}$ are proper lower semicontinuous functions. Then for all $(x, y, z) \in dom\ H = dom\ p \times dom\ q \times dom\ h$, the following equality holds:*

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

Lemma 2.9. ([6]) (Uniform KLP) *Suppose Ω is a compact set and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper and lower semicontinuous function. If f is constant on Ω and at any point of Ω f satisfies* KLP *what is stated in Lemma 2.6. Then, there exist $\zeta > 0$, $\eta > 0$ and $\varphi \in \Phi_\eta$ 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 2.10. ([28]) *Suppose that $p : \mathbb{R}^u \to \mathbb{R}$ is continuous and differentiable, and gradient ∇p is L_p-Lipschitz continuous. Then*

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

Lemma 2.11. (i) *The sequence $\{e^k\}$ converges if it is monotonically decreasing and bounded below.*
(ii) *If $\{e^k\}$ is monotonic and there exists an infinite subsequence $e^{k_j} \to e$, then* $\lim_{k\to\infty} e^{k_j} = e$.

3. Convergence Analysis

Letting $\omega := (x, y, z)$ and $\imath := (x, y, z, \Lambda)$. Using ALF (1.5), we define

$$\begin{cases} \partial_x \mathcal{L}_\beta(\imath) = \partial f(x) + \nabla_x l(\omega) - A^\top \Lambda + \beta A^\top (Ax + By + Cz - b), \\ \partial_y \mathcal{L}_\beta(\imath) = \partial g(y) + \nabla_y l(\omega) - B^\top \Lambda + \beta B^\top (Ax + By + Cz - b), \\ \partial_z \mathcal{L}_\beta(\imath) = \partial h(z) + \nabla_z l(\omega) - C^\top \Lambda + \beta C^\top (Ax + By + Cz - b), \\ \partial_\Lambda \mathcal{L}_\beta(\imath) = -(Ax + By + Cz - b). \end{cases} \tag{3.1}$$

From (3.1), we obtain the following result.

Lemma 3.1. *Let $(x^\star, y^\star, z^\star, \Lambda^\star)$ be a stable point of $\mathcal{L}_\beta(x, y, z, \Lambda)$. Then $0 \in \partial \mathcal{L}_\beta(x^\star, y^\star, z^\star, \Lambda^\star)$ if and only if the following equalities own by* Lemma 2.8*:*

$$A^\top \Lambda^\star - \nabla_x l\,(x^\star, y^\star, z^\star) \in \partial f(x^\star), \ B^\top \Lambda^\star - \nabla_y l(x^\star, y^\star, z^\star) \in \partial g(y^\star),$$
$$C^\top \Lambda^\star - \nabla_z l\,(x^\star, y^\star, z^\star) \in \partial h(z^\star), \ Ax^\star + By^\star + Cz^\star - b = 0.$$

The ALF $\mathcal{L}_\beta(x, y, z, \Lambda)$ satisfies the following property for all $\theta \in \mathbb{R}$ and $(x, y, z, \Lambda) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \times \mathbb{R}^m$:

$$\mathcal{L}_\beta(x, y, z, \Lambda - \theta(Ax + By + Cz - b)) = \mathcal{L}_\beta(x, y, z, \Lambda) + \theta\|Ax + By + Cz - b\|^2. \tag{3.2}$$

To simplify the analysis, we introduce the following symbols:

$$\begin{aligned}
&\omega^\star := (x^\star, y^\star, z^\star), \quad \tilde{\omega} := (\tilde{x}, \tilde{y}, \tilde{z}), \quad \omega^k := (x^k, y^k, z^k),\\
&\imath^\star := (x^\star, y^\star, z^\star, \Lambda^\star), \quad \imath^k := (x^k, y^k, z^k, \Lambda^k),\\
&\hat{\imath} := (x, y, z, \Lambda, \hat{z}), \quad \hat{\imath}^\star := (x^\star, y^\star, z^\star, \Lambda^\star, \hat{z}^\star), \quad \hat{\imath}^k := (x^k, y^k, z^k, \Lambda^k, z^{k-1}).
\end{aligned}$$

Some basic assumptions of problem (1.1) are given in order to analyze convergence.

Assumption 3.2. (i) *$h(z)$ and $l(x, y, z)$ are differentiable. The gradient ∇h and ∇l are Lipschitz continuous with constants L_h and L_l, respectively. That is, for all $(x, y, z),\ (\tilde{x}, \tilde{y}, \tilde{z}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$,*

$$\begin{aligned}
&\|\nabla h(z) - \nabla h(\tilde{z})\| \le L_h\|z - \tilde{z}\|,\\
&\|(\nabla_x l(\omega) - \nabla_x l(\tilde{\omega}), \nabla_y l(\omega) - \nabla_y l(\tilde{\omega}), \nabla_z l(\omega) - \nabla_z l(\tilde{\omega}))\| \le L_l\|(x - \tilde{x}, y - \tilde{y}, z - \tilde{z})\|.
\end{aligned}$$

(ii) *$\nabla\phi_i$ is Lipshitz continuous with constant L_{ϕ_i} for $i = 1, 2, 3$.*

(iii) *The parameters r and s, as well as the strong convex coefficient σ_{ϕ_i} of function ϕ_i for $i = 1, 2, 3$, satisfy the following conditions:*

$$\begin{aligned}
&r + s > 0,\\
&\sigma_{\phi_1} > \frac{18\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + \beta^2(1-s)^2\Lambda_{\max}(A^\top A)\big)}{(r+s)\beta} + L_l,\\
&\sigma_{\phi_2} > \frac{18\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + \beta^2(1-s)^2\Lambda_{\max}(B^\top B)\big)}{(r+s)\beta} + \frac{2rs\beta}{r+s} + L_l,\\
&\sigma_{\phi_3} > \frac{18\big(L_h^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + \beta^2(1-s)^2\Lambda_{\max}(C^\top C) + L_{\phi_3}^2\big)}{(r+s)\beta} + \frac{2rs\beta}{r+s} + L_l.
\end{aligned}$$

(iv) 3-GBPRSM *generates a bounded sequence of points $\{\imath^k\}$.*

Remark 3.3. According to the optimality conditions of 3-GBPRSM, one can get

$$\begin{cases}
0 \in \partial f(x^k) + \nabla_x l(x^k, y^k, z^k) - A^\top\Lambda^k + \beta A^\top(Ax^{k+1} + By^k\\
\qquad + Cz^k - b) + \nabla\phi_1(x^{k+1}) - \nabla\phi_1(x^k),\\
0 \in \partial g(y^{k+1}) + \nabla_y l(x^{k+1}, y^k, z^k) - B^\top\Lambda^{k+\frac{1}{2}} + \beta B^\top(Ax^{k+1}\\
\qquad + By^{k+1} + Cz^k - b) + \nabla\phi_2(y^{k+1}) - \nabla\phi_2(y^k),\\
0 \in \partial h(z^{k+1}) + \nabla_z l(x^{k+1}, y^{k+1}, z^k) - C^\top\Lambda^{k+\frac{1}{2}} + \beta C^\top(Ax^{k+1}\\
\qquad + By^{k+1} + Cz^{k+1} - b) + \nabla\phi_3(z^{k+1}) - \nabla\phi_3(z^k).
\end{cases} \tag{3.3}$$

Since $\{\mathcal{L}_\beta(\imath^k)\}$ does not necessarily have good monotonicity, we need to construct a appropriate benefit function with decreasing properties. Thus, we consider the following benefit function:

$$\hat{\mathcal{L}}_\beta(\hat{\imath}) = \mathcal{L}_\beta(\imath^k) + \frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + L_{\phi_3}^2\big)}{(r+s)\beta}\|z - \hat{z}\|. \tag{3.4}$$

Also for 3-GBPRSM, assume that the initial point $\imath^0$ makes $\mathcal{L}_\beta(\imath^0) < +\infty$.

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

$$\begin{cases}\delta_1 := \delta_1(r,s,\beta,\sigma_{\phi_1}) \\ \quad = \dfrac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + \beta^2(1-s)^2\Lambda_{\max}(A^\top A)\big)}{(r+s)\beta} - \dfrac{\sigma_{\phi_1}-L_l}{2}, \\ \delta_2 := \delta_2(r,s,\beta,\sigma_{\phi_2}) \\ \quad = \dfrac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + \beta^2(1-s)^2\Lambda_{\max}(B^\top B)\big)}{(r+s)\beta} + \dfrac{rs\beta}{r+s} - \dfrac{\sigma_{\phi_2}-L_l}{2}, \\ \delta_3 := \delta_3(r,s,\beta,\sigma_{\phi_3}) \\ \quad = \dfrac{9\big(L_h^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + \beta^2(1-s)^2\Lambda_{\max}(C^\top C) + L_{\phi_3}^2\big)}{(r+s)\beta} + \dfrac{rs\beta}{r+s} - \dfrac{\sigma_{\phi_1}-L_l}{2}, \\ \delta := \delta(r,s,\beta) = \min\{\delta_1,\delta_2,\delta_3\} \\ \quad = \min\{\delta_1(r,s,\beta,\sigma_{\phi_1}), \delta_2(r,s,\beta,\sigma_{\phi_2}), \delta_3(r,s,\beta,\sigma_{\phi_3})\}. \end{cases} \tag{3.5}$$

Lemma 3.4. *Suppose Assumption 3.2 holds, then for each $k \geq 0$,*

$$\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}) - \hat{\mathcal{L}}_\beta(\hat{\imath}^k) \leq -\delta(\|x^{k+1}-x^k\|^2 + \|y^{k+1}-y^k\|^2 + \|z^{k+1}-z^k\|^2), \tag{3.6}$$

where the definition of δ is shown in (3.5).

Proof. First of all, x^{k+1} is the optimal solution to the subproblem of x. Then,

$$\begin{aligned} &f(x^{k+1}) + (x^{k+1}-x^k)^\top \nabla_x l(x^k,y^k,z^k) - \langle \Lambda^k, Ax^{k+1}\rangle + \frac{\beta}{2}\|Ax^{k+1}+By^k+Cz^k-b\|^2 \\ &\leq f(x^k) - \langle \Lambda^k, Ax^k\rangle + \frac{\beta}{2}\|Ax^k+By^k+Cz^k-b\|^2 - \triangle_\phi(x^{k+1},x^k). \end{aligned}$$

According to the definition of $\mathcal{L}_\beta(\imath)$ in (1.5), one has

$$\begin{aligned} \mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^k) - \mathcal{L}_\beta(\imath^k) \leq\ & l(x^{k+1},y^k,z^k) - l(x^k,y^k,z^k) \\ & - (x^{k+1}-x^k)^\top\nabla_x l(x^k,y^k,z^k) - \triangle_{\phi_1}\big(x^{k+1},x^k\big). \end{aligned}$$

By the Lipschitz continuity of ∇l, the strong convexity of $\triangle_{\phi_1}$ and Lemma 2.7, one can obtain that

$$\begin{aligned} \mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^k) - \mathcal{L}_\beta(\imath^k) &\leq \frac{L_l}{2}\|x^{k+1}-x^k\|^2 - \triangle_{\phi_1}(x^{k+1},x^k) \\ &\leq -\frac{\sigma_{\phi_1}-L_l}{2}\|x^{k+1}-x^k\|^2. \end{aligned} \tag{3.7}$$

Similarly, y^{k+1} is the optimal solution to the subproblem of y, thus

$$\begin{aligned} &g(y^{k+1}) + (y^{k+1}-y^k)^\top\nabla_y l(x^{k+1},y^k,z^k) - \langle \Lambda^{k+\frac{1}{2}}, By^{k+1}\rangle + \frac{\beta}{2}\|Ax^{k+1}+By^{k+1}+Cz^k-b\|^2 \\ &\leq g(y^k) - \langle \Lambda^{k+\frac{1}{2}}, By^k\rangle + \frac{\beta}{2}\|Ax^{k+1}+By^k+Cz^k-b\|^2 - \triangle_{\phi_2}(y^{k+1},y^k). \end{aligned}$$

Further, combining with the definition of $\mathcal{L}_\beta(\imath)$, one gets

$$\begin{aligned} &\mathcal{L}_\beta(x^{k+1},y^{k+1},z^k,\Lambda^{k+\frac{1}{2}}) - \mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^{k+\frac{1}{2}}) \\ &\leq l(x^{k+1},y^{k+1},z^k) - l(x^{k+1},y^k,z^k) - (y^{k+1}-y^k)^\top\nabla_y l(x^{k+1},y^k,z^k) - \triangle_{\phi_2}\big(y^{k+1},y^k\big). \end{aligned}$$

From the Lipschitz continuity of ∇l, the strong convexity of $\triangle_{\phi_2}$ and Lemma 2.10, it follows that

$$\begin{aligned} \mathcal{L}_\beta(x^{k+1},y^{k+1},z^k,\Lambda^{k+\frac{1}{2}}) - \mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^{k+\frac{1}{2}}) &\leq \frac{L_l}{2}\|y^{k+1}-y^k\|^2 - \triangle_{\phi_2}(y^{k+1},y^k) \\ &\leq -\frac{\sigma_{\phi_2}-L_l}{2}\|y^{k+1}-y^k\|^2. \end{aligned} \tag{3.8}$$

In the same way, the following inequality holds:

$$\begin{aligned}\mathcal{L}_\beta(\omega^{k+1},\Lambda^{k+\frac{1}{2}})-\mathcal{L}_\beta(x^{k+1},y^{k+1},z^k,\Lambda^{k+\frac{1}{2}}) &\le \frac{L_l}{2}\|z^{k+1}-z^k\|^2-\triangle_{\phi_3}(z^{k+1},z^k)\\ &\le -\frac{\sigma_{\phi_3}-L_l}{2}\|z^{k+1}-z^k\|^2.\end{aligned} \tag{3.9}$$

Secondly, it can deduce from (1.8) and (3.3) that

$$\begin{aligned}\Lambda^{k+1}=&\left[\partial h(z^{k+1})+\nabla_z l(x^{k+1},y^{k+1},z^k)+\nabla\phi_3(z^{k+1})-\nabla\phi_3(z^k)\right](C^\top)^{-1}\\ &+\beta(1-s)(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b),\end{aligned}$$

and combining with the Lipschitz continuity of ∇h and ∇l, we can derive

$$\begin{aligned}\|\Lambda^{k+1}-\Lambda^k\|\le&\|(C^\top)^{-1}\|(L_h\|z^{k+1}-z^k\|+L_l\|x^{k+1}-x^k\|+L_l\|y^{k+1}-y^k\|+L_l\|z^k-z^{k-1}\|)\\ &+L_{\phi_3}\|z^{k+1}-z^k\|+L_{\phi_3}\|z^k-z^{k-1}\|\\ &+|1-s|\beta\left[\|A(x^{k+1}-x^k)\|+\|B(y^{k+1}-y^k)\|+\|C(z^{k+1}-z^k)\|\right]\\ \le&(L_l\|(C^\top)^{-1}\|+\beta|1-s|\|A\|)\|x^{k+1}-x^k\|\\ &+(L_l\|(C^\top)^{-1}\|+\beta|1-s|\|B\|)\|y^{k+1}-y^k\|\\ &+\left[L_h\|(C^\top)^{-1}\|+\beta|1-s|\|C\|+L_{\phi_3}\right]\|z^{k+1}-z^k\|\\ &+\big(L_l\|(C^\top)^{-1}+L_{\phi_3}\big)\|z^k-z^{k-1}\|.\end{aligned} \tag{3.10}$$

By Cauchy inequality, (3.10) means that

$$\begin{aligned}\frac{1}{9}\|\Lambda^{k+1}-\Lambda^k\|^2\le&\Big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(A^\top A)\Big)\|x^{k+1}-x^k\|\\ &+\Big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(B^\top B)\Big)\|y^{k+1}-y^k\|\\ &+\Big(L_h^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(C^\top C)+L_{\phi_3}^2\Big)\|z^{k+1}-z^k\|\\ &+\Big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\Big)\|z^k-z^{k-1}\|.\end{aligned} \tag{3.11}$$

Besides, by using multipliers Λ^{k+1} and $\Lambda^{k+\frac{1}{2}}$ to update the formulas, we have

$$\begin{aligned}Ax^{k+1}+By^k+Cz^k-b&=-\frac{1}{(r+s)\beta}(\Lambda^{k+1}-\Lambda^k)-\frac{s}{r+s}(y^{k+1}-y^k)-\frac{s}{r+s}(z^{k+1}-z^k),\\ Ax^{k+1}+By^{k+1}+Cz^{k+1}-b&=-\frac{1}{(r+s)\beta}(\Lambda^{k+1}-\Lambda^k)+\frac{r}{r+s}(y^{k+1}-y^k)+\frac{r}{r+s}(z^{k+1}-z^k).\end{aligned} \tag{3.12}$$

According to (3.11), (3.12) and $r+s>0$, the following inequality owns:

$$\begin{aligned}&r\beta\|Ax^{k+1}+By^k+Cz^k-b\|^2+s\beta\|Ax^{k+1}+By^{k+1}+Cz^{k+1}-b\|^2\\ &=\frac{1}{(r+s)\beta}\|\Lambda^{k+1}-\Lambda^k\|^2+\frac{rs\beta}{r+s}\|y^{k+1}-y^k\|^2+\frac{rs\beta}{r+s}\|z^{k+1}-z^k\|^2\\ &\le\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(A^\top A)\big)}{(r+s)\beta}\|x^{k+1}-x^k\|\\ &\quad+\left(\frac{9\left(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(B^\top B)\right)}{(r+s)\beta}+\frac{rs\beta}{r+s}\right)\|y^{k+1}-y^k\|\\ &\quad+\left(\frac{9\big(L_h^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(C^\top C)+L_{\phi_3}^2\big)}{(r+s)\beta}+\frac{rs\beta}{r+s}\right)\|z^{k+1}-z^k\|\end{aligned}$$

$$+\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}\|z^k-z^{k-1}\|. \tag{3.13}$$

Finally, from (1.8) and (3.2), it is easy to know that

$$\mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^{k+\frac{1}{2}})-\mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^k)=r\beta\|Ax^{k+1}+By^k+Cz^k-b\|^2,$$
$$\mathcal{L}_\beta(\imath^{k+1})-\mathcal{L}_\beta(\omega^{k+1},\Lambda^{k+\frac{1}{2}})=s\beta\|Ax^{k+1}+By^{k+1}+Cz^{k+1}-b\|^2.$$

Then, by (3.7)-(3.9) and (3.13), one can get

$$\begin{aligned}
&\mathcal{L}_\beta(\imath^{k+1})-\mathcal{L}_\beta(\imath^k)\\
\le&\mathcal{L}_\beta(x^{k+1},y^{k+1},z^k,\Lambda^{k+\frac{1}{2}})-\frac{\sigma_{\phi_3}-L_l}{2}\|z^{k+1}-z^k\|^2\\
&+s\beta\|Ax^{k+1}+By^{k+1}+Cz^{k+1}-b\|^2-\mathcal{L}_\beta(\imath^k)\\
\le&\mathcal{L}_\beta(x^{k+1},y^k,z^k,\Lambda^{k+\frac{1}{2}})-\frac{\sigma_{\phi_2}-L_l}{2}\|y^{k+1}-y^k\|^2-\frac{\sigma_{\phi_3}-L_l}{2}\|z^{k+1}-z^k\|^2\\
&+s\beta\|Ax^{k+1}+By^{k+1}+Cz^{k+1}-b\|^2-\mathcal{L}_\beta(\imath^k)\\
\le&-\frac{\sigma_{\phi_1}-L_l}{2}\|x^{k+1}-x^k\|^2-\frac{\sigma_{\phi_2}-L_l}{2}\|y^{k+1}-y^k\|^2-\frac{\sigma_{\phi_3}-L_l}{2}\|z^{k+1}-z^k\|^2\\
&+r\beta\|Ax^{k+1}+By^k+Cz^k-b\|^2+s\beta\|Ax^{k+1}+By^{k+1}+Cz^{k+1}-b\|^2\\
\le&\left(\frac{9\left(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(A^\top A)\right)}{(r+s)\beta}-\frac{\sigma_{\phi_1}-L_l}{2}\right)\|x^{k+1}-x^k\|\\
&+\left(\frac{9\left(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(B^\top B)\right)}{(r+s)\beta}+\frac{rs\beta}{r+s}-\frac{\sigma_{\phi_2}-L_l}{2}\right)\|y^{k+1}-y^k\|\\
&+\left(\frac{9\left(L_h^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(C^\top C)+L_{\phi_3}^2\right)}{(r+s)\beta}+\frac{rs\beta}{r+s}-\frac{\sigma_{\phi_1}-L_l}{2}\right)\|z^{k+1}-z^k\|\\
&+\frac{9\left(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\right)}{(r+s)\beta}\|z^k-z^{k-1}\|,
\end{aligned}$$

which implies according to the definition of δ in (3.5) that

$$\begin{aligned}
&\left[\mathcal{L}_\beta(\imath^{k+1})+\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}\|z^{k+1}-z^k\|\right]\\
&-\left[\mathcal{L}_\beta(\imath^k)+\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}\|z^k-z^{k-1}\|\right]\\
\le&\left(\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(A^\top A)\big)}{(r+s)\beta}-\frac{\sigma_{\phi_1}-L_l}{2}\right)\|x^{k+1}-x^k\|\\
&+\left(\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(B^\top B)\big)}{(r+s)\beta}+\frac{rs\beta}{r+s}-\frac{\sigma_{\phi_2}-L_l}{2}\right)\|y^{k+1}-y^k\|\\
&+\left(\frac{9\big(L_h^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+\beta^2(1-s)^2\Lambda_{\max}(C^\top C)+L_{\phi_3}^2\big)}{(r+s)\beta}+\frac{rs\beta}{r+s}-\frac{\sigma_{\phi_1}-L_l}{2}\right)\|z^{k+1}-z^k\|\\
\le&-\delta_1(r,s,\beta,\sigma_{\phi_1})\|x^{k+1}-x^k\|^2-\delta_2(r,s,\beta,\sigma_{\phi_2})\|y^{k+1}-y^k\|^2-\delta_3(r,s,\beta,\sigma_{\phi_3})\|z^{k+1}-z^k\|^2\\
\le&-\delta\big(\|x^{k+1}-x^k\|^2+\|y^{k+1}-y^k\|^2+\|z^{k+1}-z^k\|^2\big).
\end{aligned}$$

Thus, $\{\hat{\mathcal{L}}_\beta(\hat{\imath}^k)\}$ has sufficient descent property when $\delta>0$. The proof is completed. □

Lemma 3.5. *If Assumption 3.2 holds, then*

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

Proof. Since $\left\{\imath^k\right\}$ is bounded, the sequence $\left\{\hat{\imath}^k\right\}$ is also bounded. Then, there exists a subsequence $\left\{\hat{\imath}^{k_j}\right\}$ such that $\lim_{j\to+\infty}(\hat{\imath}^{k_j})=\hat{\imath}^\star$. Given the lower semicontinuity of f, g and h. We know that the Lipschitz differentiability of $\hat{\mathcal{L}}_\beta$ is lower semicontinuous. Thus

$$\hat{\mathcal{L}}_\beta(\hat{\imath}^\star)\le\lim_{k_j\to+\infty}\hat{\mathcal{L}}_\beta(\hat{\imath}^{k_j})\le\hat{\mathcal{L}}_\beta(\hat{\imath}^0)=\mathcal{L}_\beta(\imath^0)<+\infty.$$

From Lemma 3.5, it follows that $\left\{\hat{\mathcal{L}}_\beta(\hat{\imath}^{k_j})\right\}$ is bounded. Due to the monotonicity of $\left\{\hat{\mathcal{L}}_\beta(\hat{\imath}^{k_j})\right\}$, $\left\{\hat{\mathcal{L}}_\beta(\hat{\imath}^{k})\right\}$ is also covergent. And for each k, $\hat{\mathcal{L}}_\beta\left(\hat{\imath}^k\right)\ge\hat{\mathcal{L}}_\beta\left(\hat{\imath}^\star\right)$. By Lemma 3.5, we have

$$\delta(\left\|x^{k+1}-x^k\right\|^2+\left\|y^{k+1}-y^k\right\|^2+\left\|z^{k+1}-z^k\right\|^2)\le\hat{\mathcal{L}}_\beta(\hat{\imath}^k)-\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}).$$

Summing this inequality from $k=1$ to t, we obtain

$$\begin{aligned}&\delta\sum_{k=1}^{t}\left(\left\|x^{k+1}-x^k\right\|^2+\left\|y^{k+1}-y^k\right\|^2+\left\|z^{k+1}-z^k\right\|^2\right)\\&\le\hat{\mathcal{L}}_\beta(\hat{\imath}^0)-\hat{\mathcal{L}}_\beta(\hat{\imath}^{q+1})\le\hat{\mathcal{L}}_\beta(\hat{\imath}^0)-\hat{\mathcal{L}}_\beta(\hat{\imath}^\star)<+\infty.\end{aligned}$$

Since $\delta>0$, it follows that

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

In view of (3.11), $\sum_{k=0}^{+\infty}\|\Lambda^{k+1}-\Lambda^k\|^2<+\infty$ holds and so one has $\sum_{k=0}^{+\infty}\|\imath^{k+1}-\imath^k\|^2<+\infty$. □

Theorem 3.6. (Global convergent) Denote Ω and $\hat{\Omega}$ as the set of all clusters of the sequence $\{\imath^k\}$ and $\{\hat{\imath}^k\}$, respectively. Let Assumption 3.2 hold. Then:

(i) Ω and $\hat{\Omega}$ are nonempty convex, and $d(\imath^k,\Omega)\to 0$, $d(\hat{\imath}^k,\hat{\Omega})\to 0$ as $k\to+\infty$.
(ii) $\Omega\subseteq\mathrm{crit}\mathcal{L}_\beta$.
(iii) $\hat{\Omega}=\{(x^\star,y^\star,z^\star,\Lambda^\star,z^\star):(x^\star,y^\star,z^\star,\Lambda^\star)\in\Omega\}$.
(iv) The entire sequence of $\{\hat{\mathcal{L}}_\beta(\hat{\imath}^k)\}$ is convergent. Moreover $\hat{\mathcal{L}}_\beta(\hat{\imath}^\star)=\lim_{k\to+\infty}\hat{\mathcal{L}}_\beta(\hat{\imath}^k)=\inf_k\hat{\mathcal{L}}_\beta(\hat{\imath}^k)$ for any $\hat{\imath}^\star\in\hat{\Omega}$. Thus, $\hat{\mathcal{L}}_\beta$ is finite and constant on $\hat{\Omega}$.

Proof. we will prove each of the results one by one as follows.

(i) The conclusion holds due to the definition of Ω, $\hat{\Omega}$ and $\hat{\imath}^k$, and the boundedness of $\{\imath^k\}$.

(ii) Letting $\imath^*=(x^\star,y^\star,z^\star,\Lambda^\star)\in\Omega$, then there exists a subsequence $\{\imath^{k_j}\}$ of $\{\imath^k\}$ such that $\{\imath^{k_j}\}$ converges to $\imath^\star$. It is known that $\lim_{k\to+\infty}\|\imath^{k+1}-\imath^k\|=0$ and $\lim_{k_j\to+\infty}\imath^{k_j+1}=\imath^\star$ by Lemma 3.5. Taking $k=k_j\to+\infty$ in (1.8), then one can know that $\{\Lambda^{k_j+\frac{1}{2}}\}$ is bounded, so we have $\lim_{k_j\to+\infty}\Lambda^{k_j+\frac{1}{2}}=\Lambda^{\star\star}$. Let $k=k_j\to+\infty$ take the limit in (1.8). Then one gets

$$\Lambda^{\star\star}=\Lambda^\star-r\beta(Ax^\star+By^\star+Cz^\star-b),\ \Lambda^\star=\Lambda^{\star\star}-s\beta(Ax^\star+By^\star+Cz^\star-b).$$

And because $r+s>0$, we can obtain from the above equation that $Ax^\star+By^\star+Cz^\star-b=0$, $\Lambda^{\star\star}=\Lambda^\star$. Thus, $\omega^\star$ is the optimal solution of (1.8), and the following inequality holds:

$$\begin{aligned} & f(x^{k_j+1})+(x^{k_j+1}-x^{k_j})^\top\nabla_x l(x^{k_j},y^{k_j},z^{k_j})-\langle\Lambda^{k_j},Ax^{k_j+1}\rangle \\ &\quad +\frac{\beta}{2}\|Ax^{k_j+1}+By^{k_j}+Cz^{k_j}-b\|^2+\triangle_{\phi_1}(x^{k_j+1},x^{k_j}) \\ &\leq f(x^\star)+(x^\star-x^{k_j})^\top\nabla_x l(x^\star,y^{k_j},z^{k_j})-\langle\Lambda^{k_j},Ax^\star\rangle \\ &\quad +\frac{\beta}{2}\|Ax^*+By^{k_j}+Cz^{k_j}-b\|^2+\triangle_{\phi_1}(x^\star,x^{k_j}). \end{aligned}$$

Combining $\lim_{k_j\to+\infty}\imath^{k_j}=\lim_{k_j\to+\infty}\imath^{k_j+1}=\imath^\star$ and the continuous differentiability of ϕ, this inequality means that $\lim sup_{k_j\to+\infty}f(x^{k_j+1})\leq f(x^\star)$. Noting the lower semicontinuity of f, $f(x^\star)\leq\lim_{k_j\to+\infty}f(x^{k_j+1})$ holds. So we can get

$$\lim_{k_j\to+\infty}f(x^{k_j+1})=f(x^\star). \tag{3.14}$$

What is more, it is known that $\lim_{k_j\to+\infty}\|x^{k_j+1}-x^{k_j}\|=0$, $\lim_{k_j\to+\infty}\|y^{k_j+1}-y^{k_j}\|=0$ and $\lim_{k_j\to+\infty}\|z^{k_j+1}-z^{k_j}\|=0$ from Lemma 3.5. According to Assumption 3.2, one has

$$\begin{aligned} &\lim_{k_j\to+\infty}\left\|\nabla\phi_1\big(x^{k_j+1}\big)-\nabla\phi_1\big(x^{k_j}\big)\right\|=0,\quad \lim_{k_j\to+\infty}\left\|\nabla\phi_2\big(y^{k_j+1}\big)-\nabla\phi_2\big(y^{k_j}\big)\right\|=0,\\ &\lim_{k_j\to+\infty}\left\|\nabla\phi_3\big(z^{k_j+1}\big)-\nabla\phi_3\big(z^{k_j}\big)\right\|=0. \end{aligned}$$

Hence, taking into account the closeness of ∂f, ∂g and ∂h, the continuity of ∇_l, as well as (3.14), letting $k=k_j\to+\infty$ take the limit in (3.3), then one gets $A^\top\Lambda^\star-\nabla_x l(\omega^\star)\in\partial f(x^\star)$, $B^\top\Lambda^\star-\nabla_y l(\omega^\star)\in\partial g(y^\star)$, $C^\top\Lambda^\star-\nabla_z l(\omega^\star)\in\partial h(z^\star)$ and $Ax^\star+By^\star+Cz^\star-b=0$, and so $\imath^\star\in\operatorname{crit}\mathcal{L}_\beta$ by Lemma 3.1.

(iii) The conclusion holds because of Lemma 3.5 and the definition of $\hat\imath^k$.

(iv) Let $\hat\imath^\star\in\hat\Omega$, and consider a subsequence $\{\hat\imath^{k_j}\}$ that $\{\hat\imath^{k_j}\}$ converges to $\hat\imath^\star$. According to (1.5), (3.4) and (3.14), we have $\lim_{k_j\to+\infty}\hat{\mathcal{L}}_\beta(\hat\imath^{k_j+1})=\hat{\mathcal{L}}_\beta(\hat\imath^\star)$. Given the monotonicity of $\{\hat{\mathcal{L}}_\beta(\hat\imath^k)\}$ and Lemma 2.11, the entire sequence $\{\hat{\mathcal{L}}_\beta(\hat\imath^k)\}$ converges. Therefore, $+\infty>\hat{\mathcal{L}}_\beta(\hat\imath^0)\geq\lim_{k\to+\infty}\hat{\mathcal{L}}_\beta(\hat\imath^k)=\inf_k\hat{\mathcal{L}}_\beta(\hat\imath^k)=\hat{\mathcal{L}}_\beta(\hat\imath^\star)$. This follows because $\hat{\mathcal{L}}_\beta(\hat\imath^k)\leq\hat{\mathcal{L}}_\beta(\hat\imath^0)<+\infty$. Hence, for all $\hat\imath^\star\in\Omega$, $\hat{\mathcal{L}}_\beta(\hat\imath^\star)\equiv\lim_{k\to+\infty}\hat{\mathcal{L}}_\beta(\hat\imath^k)<+\infty$ holds. The proof is completed. □

Remark 3.7. From $\hat\imath^{k+1}=(x^{k+1},y^{k+1},z^{k+1},\Lambda^{k+1},\hat z^{k+1})$ and the definition of $\hat{\mathcal{L}}_\beta(\hat\imath)$ in (3.4), the following limiting subdifferential result is obtained

$$\begin{cases} \partial_x\hat{\mathcal{L}}_\beta(\hat\imath^{k+1})=\partial f(x^{k+1})+\nabla_x l(x^{k+1},y^{k+1},z^{k+1})-A^\top\Lambda^{k+1}\\ \qquad\qquad +\beta A^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b),\\ \partial_y\hat{\mathcal{L}}_\beta(\hat\imath^{k+1})=\partial g(y^{k+1})+\nabla_y l(x^{k+1},y^{k+1},z^{k+1})-B^\top\Lambda^{k+1}\\ \qquad\qquad +\beta B^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b),\\ \partial_z\hat{\mathcal{L}}_\beta(\hat\imath^{k+1})=\partial h(z^{k+1})+\nabla_z l(x^{k+1},y^{k+1},z^{k+1})-C^\top\Lambda^{k+1}\\ \qquad\qquad +\beta C^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b)\\ \qquad\qquad +\dfrac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}(z^{k+1}-z^k),\\ \partial_\Lambda\hat{\mathcal{L}}_\beta(\hat\imath^{k+1})=-(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b),\\ \partial_{\hat z}\hat{\mathcal{L}}_\beta(\hat\imath^{k+1})=-\dfrac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}(z^{k+1}-z^k). \end{cases} \tag{3.15}$$

To prove strong convergence of sequences generated by 3-GBPRSM, the following result is needed.

Lemma 3.8. *Let*

$$\begin{cases} \mathcal{E}_1^{k+1} = \beta A^\top(By^{k+1}-By^k)+\beta A^\top(Cz^{k+1}-Cz^k)-A^\top(\Lambda^{k+1}-\Lambda^k) \\ \qquad -\Big[\nabla\phi_1(x^{k+1})-\nabla\phi_2(x^k)\Big]+\nabla_x l(\omega^{k+1})-\nabla_x l(x^k,y^k,z^k), \\ \mathcal{E}_2^{k+1} = B^\top\big(-\frac{s}{r+s}(\Lambda^{k+1}-\Lambda^k)+\frac{rs\beta}{r+s}(y^{k+1}-y^k+z^{k+1}-z^k)\big) \\ \qquad +\nabla_y l(\omega^{k+1})-\nabla_y l(x^{k+1},y^{k+1},z^k)+\beta B^\top(Cz^{k+1}-Cz^k) \\ \qquad -\Big[\nabla\phi_2(y^{k+1})-\nabla\phi_2(y^k)\Big], \\ \mathcal{E}_3^{k+1} = \nabla_z l(\omega^{k+1})-\nabla_z l(x^{k+1},y^{k+1},z^k)-\frac{sC^\top}{r+s}(\Lambda^{k+1}-\Lambda^k) \\ \qquad +\frac{rs\beta C^\top}{r+s}(y^{k+1}-y^k)-\Big[\nabla\phi_3(z^{k+1})-\nabla\phi_3(z^k)\Big] \\ \qquad +\frac{rs\beta^2C^\top+9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}(z^{k+1}-z^k), \\ \mathcal{E}_4^{k+1} = \frac{1}{(r+s)\beta}(\Lambda^{k+1}-\Lambda^k)-\frac{r}{r+s}(y^{k+1}-y^k)-\frac{r}{r+s}(z^{k+1}-z^k), \\ \mathcal{E}_5^{k+1} = -\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}(z^{k+1}-z^k). \end{cases} \tag{3.16}$$

If Assumption 3.2 *holds, then* $\mathcal{E}^{k+1} := (\mathcal{E}_1^{k+1},\mathcal{E}_2^{k+1},\mathcal{E}_3^{k+1},\mathcal{E}_4^{k+1},\mathcal{E}_5^{k+1}) \in \partial\mathcal{L}_\beta(\imath^{k+1})$ *holds. For all* $k\geq 0$*, there exists* $\zeta>0$ *such that*

$$(0,\partial\mathcal{L}_\beta(\imath^{k+1}))\leq\zeta(\|x^{k+1}-x^k\|+\|y^{k+1}-y^k\|+\|z^{k+1}-z^k\|+\|z^k-z^{k-1}\|). \tag{3.17}$$

Proof. Firstly, let ${\xi_1}^{k+1}\in\partial f(x^{k+1})$. According to (3.3), we have

$${\xi_1}^{k+1}+\nabla_x l(x^k,y^k,z^k)+\beta A^\top(Ax^{k+1}+By^k+Cz^k-b)-A^\top\Lambda^k+\nabla\phi_1(x^{k+1})-\nabla\phi_1(x^k)=0.$$

Combining (3.16) and (3.1), we have

$$\begin{aligned} \mathcal{E}_1^{k+1} &= \beta A^\top(By^{k+1}-By^k)+\beta A^\top(Cz^{k+1}-Cz^k)-A^\top(\Lambda^{k+1}-\Lambda^k) \\ &\quad -\Big[\nabla\phi_1(x^{k+1})-\nabla\phi_2(x^k)\Big]+\nabla_x l(\omega^{k+1})-\nabla_x l(x^k,y^k,z^k)+\Big\{{\xi_1}^{k+1}+\nabla_x l(x^k,y^k,z^k) \\ &\quad -A^\top\Lambda^k+\beta A^\top(Ax^{k+1}+By^k+Cz^k-b)+\nabla\phi_1(x^{k+1})-\nabla\phi_1(x^k)\Big\} \\ &= {\xi_1}^{k+1}+\nabla_x l(x^{k+1},y^{k+1},z^{k+1})-A^\top\Lambda^{k+1}+\beta A^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b) \\ &\in \partial_x\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}). \end{aligned}$$

Secondly, according to (3.3), taking ${\xi_2}^{k+1}\in\partial g(y^{k+1})$, then

$$\begin{aligned} &{\xi_2}^{k+1}+\nabla_y l(x^{k+1},y^k,z^k)+\beta B^\top(Ax^{k+1}+By^{k+1}+Cz^k-b) \\ &-B^\top\Lambda^{k+\frac{1}{2}}+\nabla\phi_2(y^{k+1})-\nabla\phi_2(y^k)=0, \end{aligned}$$

and it follows from (3.16) and (3.1) that

$$
\begin{aligned}
\mathcal{E}_2^{k+1} &= B^\top\left(-\frac{s}{r+s}(\Lambda^{k+1}-\Lambda^k)+\frac{rs\beta}{r+s}(y^{k+1}-y^k+z^{k+1}-z^k)\right)+\nabla_y l(\omega^{k+1})\\
&\quad -\nabla_y l(x^{k+1},y^{k+1},z^k)+\beta B^\top(Cz^{k+1}-Cz^k)-\Big[\nabla\phi_2(y^{k+1})-\nabla\phi_2(y^k)\Big]\\
&= B^\top\left(s\beta(Ax^{k+1}+y^{k+1}+z^{k+1}-b)\right)+\nabla_y l(\omega^{k+1})-\nabla_y l(x^{k+1},y^{k+1},z^k)\\
&\quad +\beta B^\top(Cz^{k+1}-Cz^k)-\Big[\nabla\phi_2(y^{k+1})-\nabla\phi_2(y^k)\Big]\\
&= B^\top(\Lambda^{k+\frac{1}{2}}-\Lambda^{k+1})+\nabla_y l(\omega^{k+1})-\nabla_y l(x^{k+1},y^k,z^k)+\beta B^\top(Cz^{k+1}\\
&\quad -Cz^k)-\Big[\nabla\phi_2(y^{k+1})-\nabla\phi_2(y^k)\Big]+\Big\{{\xi_2}^{k+1}+\nabla_y l(x^{k+1},y^k,z^k)-B^\top\Lambda^{k+\frac{1}{2}}\\
&\quad +\beta B^\top(Ax^{k+1}+By^{k+1}+Cz^k-b)+\nabla\phi_2(y^{k+1})-\nabla\phi_2(y^k)\Big\}\\
&= {\xi_2}^{k+1}+\nabla_y l(x^{k+1},y^{k+1},z^{k+1})-B^\top\Lambda^{k+1}+\beta B^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b)\\
&\in \partial_y\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}).
\end{aligned}
$$

According to (3.3), there exists ${\xi_3}^{k+1}\in\partial h(z^{k+1})$ such that

$$
\begin{aligned}
&{\xi_3}^{k+1}+\nabla_z l(x^{k+1},y^{k+1},z^k)+\beta C^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b)\\
&-C^\top\Lambda^{k+\frac{1}{2}}+\nabla\phi_3(z^{k+1})-\nabla\phi_3(z^k)=0.
\end{aligned}
$$

Similarly, substituting (3.3) for $\Lambda^{k+\frac{1}{2}}$ in (1.8), it holds

$$
\begin{aligned}
s\beta C^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b) =&{\xi_3}^{k+1}+\nabla_z l(x^{k+1},y^{k+1},z^k)-C^\top\Lambda^{k+1}\\
&+\beta C^\top\left(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b\right)\\
&+\nabla\phi_3(z^{k+1})-\nabla\phi_3(z^k).
\end{aligned}
$$

Combining with (3.12), (3.15) and (3.16), the following equation holds:

$$
\begin{aligned}
\mathcal{E}_3^{k+1} &= \nabla_z l(\omega^{k+1})-\nabla_z l(x^{k+1},y^{k+1},z^k)-\frac{sC^\top}{r+s}(\Lambda^{k+1}-\Lambda^k)\\
&\quad +\frac{rs\beta C^\top}{r+s}(y^{k+1}-y^k)-\Big[\nabla\phi_3(z^{k+1})-\nabla\phi_3(z^k)\Big]\\
&\quad +\frac{rs\beta^2C^\top+9\left(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\right)}{(r+s)\beta}(z^{k+1}-z^k)\\
&= {\xi_3}^{k+1}+\nabla_z l(\omega^{k+1})-C^\top\Lambda^{k+1}+\beta C^\top(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b)\\
&\quad +\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top+L_{\phi_3}^2\big)}{(r+s)\beta}(z^{k+1}-z^k)\\
&\in \partial_z\mathcal{L}_\beta(\imath^{k+1}).
\end{aligned}
$$

Further, it can be deduced that

$$
\begin{aligned}
\mathcal{E}_4^{k+1} &= \frac{1}{(r+s)\beta}(\Lambda^{k+1}-\Lambda^k)-\frac{r}{r+s}(y^{k+1}-y^k)-\frac{r}{r+s}(z^{k+1}-z^k)\\
&= -(Ax^{k+1}+By^{k+1}+Cz^{k+1}-b)\\
&= \partial_\Lambda\mathcal{L}_\beta\big(\imath^{k+1}\big).
\end{aligned}
$$

Again, it follows from (3.15) and (3.16) that

$$\mathcal{E}_5^{k+1} = -\frac{9\big(L_l^2\Lambda_{\max}(C^{-1})(C^{-1})^\top + L_{\phi_3}^2\big)}{(r+s)\beta}(z^{k+1}-z^k) = \partial_{\hat{z}}\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}).$$

Hence, we have $\mathcal{E}^{k+1} \in \partial\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1})$.

Finally, combining (3.16) and the Lipschitz continuity of $\nabla\phi_1$, $\nabla\phi_2$, $\nabla\phi_3$ and ∇l, there exists $\varsigma_1 > 0$ such that

$$\|\mathcal{E}^{k+1}\| \le \varsigma_1\left(\|x^{k+1}-x^k\| + \|y^{k+1}-y^k\| + \|z^{k+1}-z^k\| + \|\Lambda^{k+1}-\Lambda^k\|\right).$$

Additionally, by (3.10), there exists $\varsigma_2 > 0$ such that

$$\|\Lambda^{k+1}-\Lambda^k\| \le \varsigma_2\left(\|x^{k+1}-x^k\| + \|y^{k+1}-y^k\| + \|z^{k+1}-z^k\| + \|z^k-z^{k-1}\|\right).$$

Considering the above two inequalities and $\mathcal{E}^{k+1} \in \partial\mathcal{L}_\beta(\imath^{k+1})$, one has

$$\begin{aligned} d(0,\partial\mathcal{L}_\beta(\imath^{k+1})) \le& \|\mathcal{E}^{k+1}\| \\ \le& \varsigma_1(1+\varsigma_2)\left(\|x^{k+1}-x^k\| + \|y^{k+1}-y^k\| + \|z^{k+1}-z^k\|\right) + \varsigma_1\varsigma_2\|z^k-z^{k-1}\| \\ \le& \zeta\|x^{k+1}-x^k\| + \|y^{k+1}-y^k\| + \|z^{k+1}-z^k\| + \|z^k-z^{k-1}\|, \end{aligned}$$

where $\zeta := \max\big\{\varsigma_1\big(1+\varsigma_2\big), \varsigma_1\varsigma_2\big\}$. □

Theorem 3.9. (Strong convergence) If Assumption 3.2 is satisfied and $\hat{\mathcal{L}}_\beta$ satisfies KLP, then

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

Additionally, $\{\imath^k\}$ converges to a critical point of $\mathcal{L}_\beta$.

Proof. According to Theorem 3.6 (iv), $\hat{\mathcal{L}}_\beta(\hat{\imath}^*) = \lim_{k\to+\infty}\hat{\mathcal{L}}_\beta(\hat{\imath}^k) = \inf_k \hat{\mathcal{L}}_\beta(\hat{\imath}^k)$ for all $\hat{\imath}^* \in \hat{\Omega}$. Now we examine two scenarios.

Firstly, if an integer k_0 satisfy $\hat{\mathcal{L}}_\beta\left(\hat{\imath}^{k_0}\right) = \hat{\mathcal{L}}_\beta\left(\hat{\imath}^*\right)$. For all $k \ge k_0$, we can obtain from (3.6) that

$$\delta(\|x^{k+1}-x^k\|^2 + \|y^{k+1}-y^k\|^2 + \|z^{k+1}-z^k\|^2) \le \hat{\mathcal{L}}_\beta(\hat{\imath}^k) - \hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}) \le \hat{\mathcal{L}}_\beta(\hat{\imath}^{k_0}) - \hat{\mathcal{L}}_\beta(\hat{\imath}^*) = 0.$$

Thus, $x^{k+1} = x^k$, $y^{k+1} = y^k$ and $z^{k+1} = z^k$ hold for all $k \ge k_0$. This together with (3.6), it follows that $\Lambda^{k+1} = \Lambda^k$ for all $k \ge k_0$. Further, $\imath^{k+1} = \imath^{k_0} \in \Omega$ holds for all $k \ge k_0$. The conclusion holds.

Secondly, if $\hat{\mathcal{L}}_\beta(\hat{\imath}^k) > \hat{\mathcal{L}}_\beta(\hat{\imath}^\star)$ for all $k \ge 0$. As is known from Theorem 3.6 (iv), $\hat{\mathcal{L}}_\beta$ takes a constant on $\hat{\Omega}$. Thus, $\hat{\mathcal{L}}_\beta$ satisfies the uniform KLP by Lemma 2.9. For ζ and η in Lemma 2.9, since $d(\hat{\imath}^k,\hat{\Omega}) \to 0$ and $\hat{\mathcal{L}}_\beta(\hat{\imath}^\star) = \inf_k \hat{\mathcal{L}}_\beta(\hat{\imath}^k)$, there exists a positive integer $\tilde{k}$ such that

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

Combining with Lemma 2.9, one can get

$$\varphi'(\hat{\mathcal{L}}_\beta(\hat{\imath}^k) - \hat{\mathcal{L}}_\beta(\hat{\imath}^\star))d(0,\partial\hat{\mathcal{L}}_\beta(\hat{\imath}^k)) \ge 1,\ \forall k > \tilde{k}. \tag{3.18}$$

In addition, the following inequality holds:

$$\varphi(\hat{\mathcal{L}}_\beta(\hat{\imath}^k) - \hat{\mathcal{L}}_\beta(\hat{\imath}^\star)) - \varphi(\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}) - \hat{\mathcal{L}}_\beta(\hat{\imath}^\star)) \ge \varphi'(\hat{\mathcal{L}}_\beta(\hat{\imath}^k) - \hat{\mathcal{L}}_\beta(\hat{\imath}^\star))(\hat{\mathcal{L}}_\beta(\hat{\imath}^k) - \hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1})). \tag{3.19}$$

Simultaneously, for any $k > \tilde{k}$, it can be derived from (3.18) and (3.17) in Lemma 3.8 that

$$\begin{aligned} \frac{1}{\varphi'(\hat{\mathcal{L}}_\beta(\hat{\imath}^k) - \hat{\mathcal{L}}_\beta(\hat{\imath}^\star))} &\le d\big(0,\partial\hat{\mathcal{L}}_\beta\big(\hat{\imath}^k\big)\big) \\ &\le \zeta(\|x^k-x^{k-1}\| + \|y^k-y^{k-1}\| + \|z^k-z^{k-1}\| + \|z^{k-1}-z^{k-2}\|). \end{aligned}$$

We can get from (3.19) and $\varphi'\left(\hat{\mathcal{L}}_\beta\left(\hat{\imath}^k\right)-\hat{\mathcal{L}}_\beta\left(\hat{\imath}^*\right)\right)>0$ that

$$\begin{aligned}\hat{\mathcal{L}}_\beta(\hat{\imath}^k)-\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1}) &\leq \frac{\varphi\left(\hat{\mathcal{L}}_\beta\left(\hat{\imath}^k\right)-\hat{\mathcal{L}}_\beta\left(\hat{\imath}^k\right)\right)-\varphi\left(\hat{\mathcal{L}}_\beta\left(\hat{\imath}^{k-1}\right)-\hat{\mathcal{L}}_\beta\left(\hat{\imath}^\star\right)\right)}{\varphi\left(\hat{\mathcal{L}}_\beta\left(\hat{\imath}^k\right)-\hat{\mathcal{L}}_\beta\left(\hat{\imath}^\star\right)\right)} \\ &\leq \zeta\left(\|x^k-x^{k-1}\|+\|y^k-y^{k-1}\|+\|z^k-z^{k-1}\|+\|z^{k-1}-z^{k-2}\|\right) \\ &\quad\cdot\left[\varphi(\hat{\mathcal{L}}_\beta(\hat{\imath}^k)-\hat{\mathcal{L}}_\beta(\hat{\imath}^\star))-\varphi(\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1})-\hat{\mathcal{L}}_\beta(\hat{\imath}^\star))\right].\end{aligned}$$

Let $\triangle_{p,q}=\varphi(\hat{\mathcal{L}}_\beta(\hat{\imath}^p)-\hat{\mathcal{L}}_\beta(\hat{\imath}^\star))-\varphi(\hat{\mathcal{L}}_\beta(\hat{\imath}^q)-\hat{\mathcal{L}}_\beta(\hat{\imath}^\star))$. Combining the above inequalities with Lemma 3.5, for each $k>\tilde{k}$, one has

$$\begin{aligned}&\delta(\|x^{k+1}-x^k\|^2+\|y^{k+1}-y^k\|^2+\|z^{k+1}-z^k\|^2) \\ \leq&\zeta(\|x^k-x^{k-1}\|+\|y^k-y^{k-1}\|+\|z^k-z^{k-1}\|+\|z^{k-1}-z^{k-2}\|)\triangle_{k,k+1},\end{aligned}$$

further implies that

$$\begin{aligned}&\left(2\|x^{k+1}-x^k\|^2+2\|y^{k+1}-y^k\|^2+2\|z^{k+1}-z^k\|^2\right)^{\frac{1}{2}} \\ \leq&\left(\|x^k-x^{k-1}\|+\|y^k-y^{k-1}\|+\|z^k-z^{k-1}\|+\|z^{k-1}-z^{k-2}\|\right)^{\frac{1}{2}}\sqrt{\frac{2\zeta}{\delta}\triangle_{k,k+1}}\,.\end{aligned}$$

According to $(a+b)\leq\sqrt{2(a^2+b^2)}$ for all $a,b\geq 0$ and the above inequality, we can obtain that for any $i>\tilde{k}$

$$\begin{aligned}&\|x^{i+1}-x^i\|+\|y^{i+1}-y^i\|+\|z^{i+1}-z^i\| \\ \leq&\left(\|x^i-x^{i-1}\|+\|y^i-y^{i-1}\|+\|z^i-z^{i-1}\|+\|z^{i-1}-z^{i-2}\|\right)^{\frac{1}{2}}\sqrt{\frac{2\zeta}{\delta}\triangle_{k,k+1}}\,.\end{aligned}$$

Thus,

$$\begin{aligned}&4\left(\|x^{i+1}-x^i\|+\|y^{i+1}-y^i\|+\|z^{i+1}-z^i\|\right) \\ \leq&2\left(\|x^i-x^{i-1}\|+\|y^i-y^{i-1}\|+\|z^i-z^{i-1}\|+\|z^{i-1}-z^{i-2}\|\right)^{\frac{1}{2}}\sqrt{\frac{8\zeta}{\delta}\triangle_{k,k+1}} \\ \leq&\|x^i-x^{i-1}\|+\|y^i-y^{i-1}\|+\|z^i-z^{i-1}\|+\|z^{i-1}-z^{i-2}\|+\frac{8\zeta}{\delta}\triangle_{k,k+1}\,.\end{aligned}\tag{3.20}$$

The final inequality holds due to $2\sqrt{ab}\leq a+b$ for all $a,b\geq 0$. By summing up (3.20) from $i=k+1(\geq\tilde{k}+1)$ to $i=q$ and rearranging terms, one can get

$$\begin{aligned}&3\sum_{i=k+1}^{q}\|x^{i+1}-x^i\|+3\sum_{i=k+1}^{q}\|y^{i+1}-y^i\|+2\sum_{i=k+1}^{q}\|z^{i+1}-z^i\| \\ \leq&2\|z^{k+1}-z^k\|-2\|z^{q+1}-z^q\|+\|x^{k+1}-x^k\|-\|x^{q+1}-x^q\|+\|y^{k+1}-y^k\| \\ &-\|y^{q+1}-y^q\|+\|z^k-z^{k-1}\|-\|z^q-z^{q-1}\|+\frac{8\zeta}{\delta}\triangle_{k+1,q+1} \\ \leq&2\|z^{k+1}-z^k\|+\|x^{k+1}-x^k\|+\|y^{k+1}-y^k\|+\|z^k-z^{k-1}\|+\frac{8\zeta}{\delta}\triangle_{k+1,q+1} \\ \leq&2\|z^{k+1}-z^k\|+\|x^{k+1}-x^k\|+\|y^{k+1}-y^k\|+\|z^k-z^{k-1}\|+\frac{8\zeta}{\delta}\varphi\left(\hat{\mathcal{L}}_\beta(\hat{\imath}^{k+1})-\hat{\mathcal{L}}_\beta(\hat{\imath}^\star)\right).\end{aligned}\tag{3.21}$$

The final inequality holds according to $\varphi(\mathcal{L}_\beta(\imath^{q+1}) - \mathcal{L}_\beta(\imath^\star)) \geq 0$. When $k = \tilde{k}$, by (3.21) one has

$$\begin{aligned}
&3\sum_{i=\tilde{k}+1}^{q} \|x^{i+1} - x^i\| + 3\sum_{i=\tilde{k}+1}^{q} \|y^{i+1} - y^i\| + 2\sum_{i=\tilde{k}+1}^{q} \|z^{i+1} - z^i\| \\
\leq& 2\|z^{\tilde{k}+1} - z^{\tilde{k}}\| + \|x^{\tilde{k}+1} - x^{\tilde{k}}\| + \|y^{\tilde{k}+1} - y^{\tilde{k}}\| + \|z^{\tilde{k}} - z^{\tilde{k}-1}\| + \frac{8\zeta}{\delta}\varphi\left(\hat{\mathcal{L}}_\beta(\hat{\imath}^{\tilde{k}+1}) - \hat{\mathcal{L}}_\beta(\hat{\imath}^\star)\right) \\
\leq& +\infty,
\end{aligned}$$

and so $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\| < +\infty$, $\sum_{k=0}^{+\infty} \|y^{k+1} - y^k\| < +\infty$ and $\sum_{k=0}^{+\infty} \|z^{k+1} - z^k\| < +\infty$. Thus, it follows from (3.10) that $\sum_{k=0}^{+\infty} \|\Lambda^{k+1} - \Lambda^k\| < +\infty$. Hence, one gets

$$\begin{aligned}
\sum_{k=0}^{+\infty} \|\imath^{k+1} - \imath^k\| \leq& \sum_{k=0}^{+\infty} \|x^{k+1} - x^k\| + \sum_{k=0}^{+\infty} \|y^{k+1} - y^k\| + \sum_{k=0}^{+\infty} \|z^{k+1} - z^k\| + \sum_{k=0}^{+\infty} \|\Lambda^{k+1} - \Lambda^k\| \\
<& +\infty.
\end{aligned}$$

Therefore, $\{\imath^k\}$ forms a Cauchy sequence, demonstrating convergence. According to Theorem 3.6, $\{\imath^k\}$ converges to the critical point of $\mathcal{L}_\beta$. This completes the proof. □

4. An Application and Numerical Simulation

Sparse signal reconstruction from incomplete observation data sets is a significant focus within the field of compressed sensing. The primary goal is to identify the most compact representation of a solution to a set of linear equations expressed the problem as [13]:

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

where $E \in \mathbb{R}^{m\times n}$ is a measurement matrix, $b \in \mathbb{R}^m$ represents observation data, and $\|x\|_0$ denotes the number of nonzero elements of x. In brief, the aforementioned models are characterized as NP-hard models. To address this challenge, researchers frequently replace l_0 regularization with $l_{\frac{1}{2}}$ regularization. Instead of directly tackling (4.1), it is common for researchers to work on the following problem, as indicated by Xu et al. [36]:

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

where $\|x\|_{\frac{1}{2}} = \left(\sum_{i=1}^{n} |x_i|^{\frac{1}{2}}\right)^2$ and $e > 0$ is a regularization parameter.

Based on (4.2) and the inherent nonseparable structure, we formulate the following optimization problem with a linear constraint:

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

which was studied by Chao et al. [9].

To assess the effectiveness of our approach, we employ 3-GBPRSM (1.8), where we define $\phi_1(x) = \frac{1}{2}\left(x_{\mu_1 I_1 - \beta A^TA}\right)^2$, $\phi_2(y) = \frac{1}{2}\left(y_{\mu_2 I_2 - \beta B^TB}\right)^2$ and $\phi_3(z) = \frac{1}{2}\left(z_{\mu_3 I_3 - \beta}\right)^2$. The algorithm can be outlined as follows

$$\begin{cases} x^{k+1} = S(\frac{1}{\mu_1}\left[G_1x^k - D_1^T(D_1x^k + D_2y^k + z^k) - \beta A^T(By^k + z^k - b - \frac{\Lambda^k}{\beta})\right]; \frac{2e}{\mu_1}), \\ \Lambda^{k+\frac{1}{2}} = \Lambda^k - r\beta(Ax^{k+1} + By^k + z^k - b), \\ y^{k+1} = \frac{1}{1+\mu_2}\left[G_2y^k - D_2^T(D_1x^{k+1} + D_2y^k + z^k) - \beta B^T(Ax^{k+1} + z^k - b - \frac{\Lambda^{k+\frac{1}{2}}}{\beta})\right], \\ z^{k+1} = \frac{1}{1+\mu_3+\beta}\left[\mu_3z^k - (D_1x^{k+1} + D_2y^{k+1}) - \beta(Ax^{k+1} + By^{k+1} - b - \frac{\Lambda^{k+\frac{1}{2}}}{\beta})\right], \\ \Lambda^{k+1} = \Lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} + By^{k+1} + z^{k+1} - b), \end{cases}$$

where S represents the half shrinkage operator by [36], which is defined as
$S(x;\tau) = \{s_\tau(x_1), \dots, s_\tau(x_n)\}^\top$ with

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

and

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

for any $x = (x_1, x_1, \cdots, x_n) \in \mathbb{R}^n$ and each $\tau \in \mathbb{R}$.

In this example, the measurement matrix A and B are generated from a standard normal distribution with mean 0 and variance 1. After generating these matrices, they are further adjusted to ensure that their columns possess a unit $l_{\frac{1}{2}}$ norm. The variables x and y are created with 100 nonzero entries, following a Gaussian distribution. Initially, all variables x^0, y^0, z^0, and Λ^0 are initialized to 0. The vector b is defined as $b = Ax^0 + By^0 + z^0 + \nu$, where ν is sampled from a normal distribution with mean 0 and a covariance matrix scaled by $10^{-3}I$. The regularization parameters are set as follows: $\mu_1 = 30$, $\mu_2 = 30$, $\mu_3 = 1$, $\beta = 20$, and $e = 0.1$. At the k-th iteration, the residual is defined as $r_k = Ax^k + By^k + z^k - b$. The termination criterion for stopping the algorithm is defined as $\|r_k\|_2 \leq \sqrt{m}10^{-4}$.

To evaluate the effectiveness of our approach 3-GBPRSM, we conducted a comparative analysis between 3-GBPRSM and LBADMM as proposed by Chao et al. [9]. The numerical results are summarized in Table 1. The implementation was performed using MATLAB R2022a on a computer running Windows 10, Intel Core i7-8550U 1.80GHz CPU with 8GB of memory. The reported results include two key metrics: the number of iterations ("Iter") and objective function value ("f-val"). The findings unequivocally demonstrate the superior performance of 3-GBPRSM when compared to LBADMM. Additionally, a subset of computational results has been visualized in Figs. 1-3, depicting trends in the objective function value ("Objective-value") and the residual $\|r_k\| = \|Ax^k + By^k + z^k - b\|$ ($\|r\|_2$).

5. Concluding Remarks

In this paper, we introduced a class of novel generalized algorithms, known as 3-GBPRSMs, which employ a Bregman-type linear approximation and incorporate the concept of inexactness. Under general assumptions, global convergence of 3-GBPRSMs was established by leveraging optimality conditions. Furthermore, we provided rigorous proof of strong convergence for 3-GBPRSMs when augmented Lagrangian function adheres to KLP. Notably, when correlation function within KLP exhibits a specific structural pattern, we can guarantee both linear and sublinear convergence rates. Finally, the

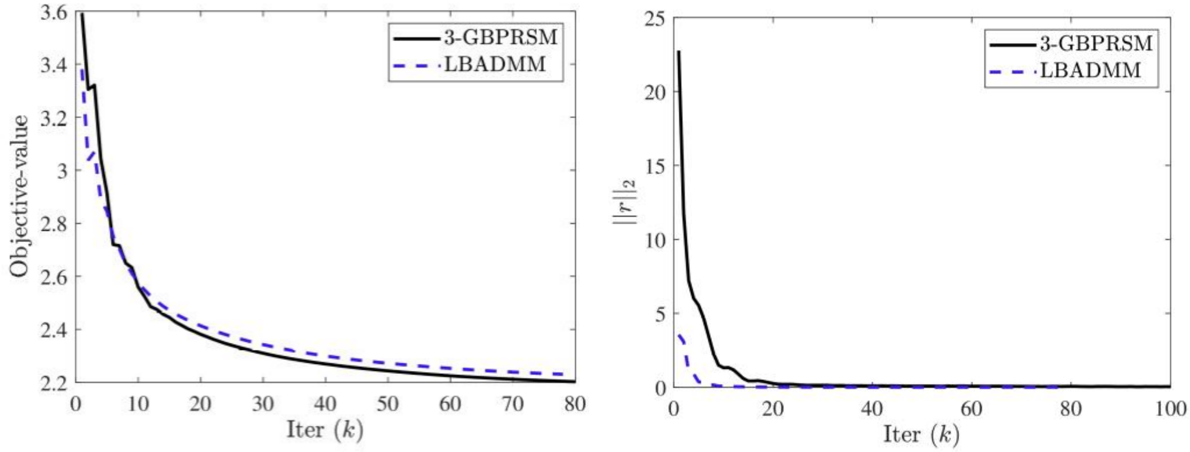

FIGURE 1. Comparison of convergence when $n_1 = n_2 = m = 1500$

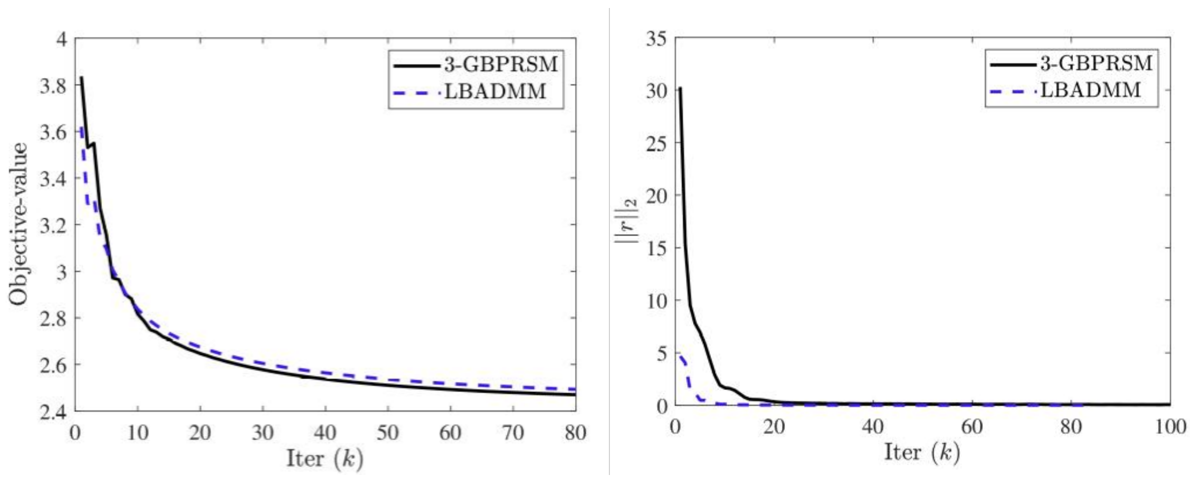

FIGURE 2. Comparison of convergence as $n_1 = n_2 = m = 3000$

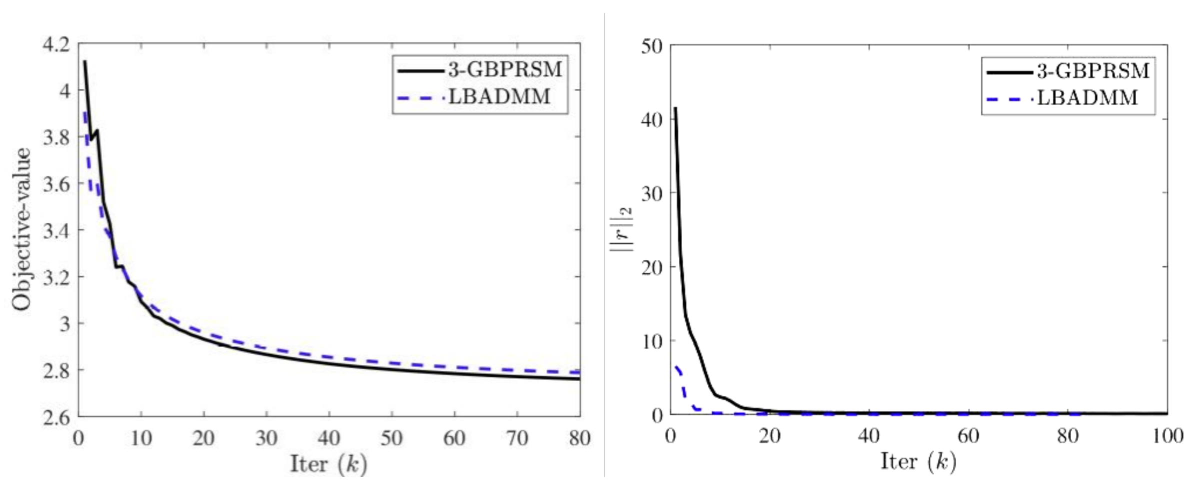

FIGURE 3. Comparison of convergence with $n_1 = n_2 = m = 6000$

Table 1. Comparison of iteration effect between 3-GBPRSM and LBADMM

(n_1, n_2, m)	Alg		Iter					min $\|r\|_2/f$-val
			30	60	90	120	150	
(1500,1500,1500)	LBADMM	$\|r\|_2$	0.0149	0.0095	0.005	-	-	0.0038
		f-val	426.08	346.53	324.89	-	-	318.91
	3-GBPRSM	$\|r\|_2$	0.6265	0.0569	0.0245	0.0177	0.0146	0.0039
		f-val	438.59	319.35	290.67	280.04	274.69	260.78
(1500,1500,2000)	LBADMM	$\|r\|_2$	0.0115	0.0038	-	-	-	0.0038
		f-val	358.19	313.47	-	-	-	313.47
	3-GBPRSM	$\|r\|_2$	0.1075	0.0591	0.0359	0.0282	0.0239	0.0044
		f-val	344.30	301.48	287.84	281.86	278.48	273.04
(2000,1500,2000)	LBADMM	$\|r\|_2$	0.0128	0.0071	-	-	-	0.0041
		f-val	323.08	268.37	-	-	-	250.43
	3-GBPRSM	$\|r\|_2$	0.1288	0.0752	0.0533	0.0303	0.0223	0.0045
		f-val	300.34	248.74	232.10	224.65	220.37	212.37
(1500,2000,2000)	LBADMM	$\|r\|_2$	0.0120	0.0052	-	-	-	0.0034
		f-val	307.54	258.88	-	-	-	256.21
	3-GBPRSM	$\|r\|_2$	0.1065	0.0615	0.0362	0.0279	0.0244	0.0044
		f-val	295.90	251.53	238.73	233.02	229.92	224.68
(2000,2000,1500)	LBADMM	$\|r\|_2$	0.0103	0.0059	0.0043	-	-	0.0036
		f-val	191.75	155.40	142.17	-	-	138.76
	3-GBPRSM	$\|r\|_2$	0.1377	0.0580	0.0586	0.0391	0.0237	0.0038
		f-val	187.60	151.49	138.46	131.93	127.82	114.30
(2000,2000,2000)	LBADMM	$\|r\|_2$	0.0138	0.0078	-	-	-	0.0045
		f-val	290.18	238.95	-	-	-	223.16
	3-GBPRSM	$\|r\|_2$	0.1539	0.0783	0.0594	0.0427	0.0274	0.0044
		f-val	271.56	223.43	207.81	200.38	196.29	187.47
(3000,3000,3000)	LBADMM	$\|r\|_2$	0.0153	0.0083	-	-	-	0.0046
		f-val	402.24	329.20	-	-	-	309.03
	3-GBPRSM	$\|r\|_2$	0.1842	0.1131	0.0898	0.0665	0.0414	0.0054
		f-val	377.90	310.96	290.38	281.13	275.90	264.72
(4000,4000,4000)	LBADMM	$\|r\|_2$	0.0181	0.0111	-	-	-	0.0062
		f-val	518.49	429.72	-	-	-	404.94
	3-GBPRSM	$\|r\|_2$	0.2006	0.1143	0.0898	0.0880	0.0684	0.0061
		f-val	490.57	406.26	378.64	365.34	357.64	343.14
(5000,5000,5000)	LBADMM	$\|r\|_2$	0.0202	0.0107	0.0092	-	-	0.0068
		f-val	650.80	536.43	500.10	-	-	495.73
	3-GBPRSM	$\|r\|_2$	0.2320	0.1243	0.0948	0.0705	0.0672	0.0069
		f-val	610.59	505.52	472.83	457.62	448.87	432.42
(6000,6000,6000)	LBADMM	$\|r\|_2$	0.0215	0.0129	-	-	-	0.0076
		f-val	782.04	647.29	-	-	-	609.35
	3-GBPRSM	$\|r\|_2$	0.2463	0.1382	0.0898	0.0705	0.0734	0.0075
		f-val	733.79	608.00	566.64	546.59	535.86	513.78

efficacy of 3-GBPRSMs was validated through numerical application experiments in connection with sparse signal reconstruction.

Furthermore, the following open questions, offered as points of reference, are poised to provide valuable guidance for future research endeavors:

(a) Whether the problem (1.1) can be equivalently transformed by introducing indicator functions with auxiliary variables?
(b) How to improve the Lagrange multiplier updating technique in traditional splitting algorithms to make them converge faster?
(c) Whether the novel algorithm 3-GBPRSM (1.8) can be generalised to multi-block problems?

Statements and Declarations

The authors declare that they have no conflict of interest, and the manuscript has no associated data.

Acknowledgments

We thank Prof. Zai-Yun Peng for recommendation and valuable comments and suggestions to improve our paper. This work was partially supported by the Scientific Research and Innovation Team Program of Sichuan University of Science and Engineering (SUSE652B002) and the Opening Project of Sichuan Province University Key Laboratory of Bridge Non-destruction Detecting and Engineering Computing (2024QZJ01).

References

[1] 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. *Mathematics of Operations Research*, 35:438–457, 2010.
[2] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh. Clustering with bregman divergences. *Journal of Machine Learning Research*, 6(4):1705–1749, 2005.
[3] F. M. Bian and X. Q. Zhang. A parameterized Douglas-Rachford splitting algorithm for nonconvex optimization. *Applied Mathematics and Computation*, 410:Article ID 126425, 2021.
[4] R. I. Bot, E. R. Csetnek, and D. K. Nguyen. A proximal minimization algorithm for structured nonconvex and nonsmooth problems. *SIAM Journal on Optimization*, 29(2):1300–1328, 2019.
[5] R. I. Bot and D. K. Nguyen. The proximal alternating direction method of multipliers in the nonconvex setting: Convergence analysis and rates. *Mathematics of Operations Research*, 45(2):682–712, 2020.
[6] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, 146:459–494, 2014.
[7] 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 Computational Mathematics and Mathematical Physics*, 7(3):200–217, 1967.
[8] M. T. Chao, C. Z. Cheng, and D. Y. Liang. A proximal block minimization method of multipliers with a substitution procedure. *Optimization Methods and Software*, 30:825–842, 2015.
[9] M. T. Chao, Z. Deng, and J. B. Jian. Convergence of linear Bregman ADMM for nonconvex and nonsmooth problems with nonseparable structure. *Complexity*, 2020:Article ID 6237942, 2020.
[10] C. H. Chen, B. S. He, Y. Y. Ye, and X. M. Yuan. The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. *Mathematical Programming*, 155:57–79, 2016.
[11] L. X. Cui, L. Bai, Y. Wang, P. S. Yu, and E. R. Hancock. Fused lasso for feature selection using structural information. *Pattern Recognition*, 119:Article ID 108058, 2021.
[12] Z. Deng and S. Y. Liu. Inertial generalized proximal Peaceman-Rachford splitting method for separable convex programming. *Calcolo*, 58:Article ID 10, 2021.

[13] D. L. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization. *Proceedings of the National Academy of Sciences of the United States of America*, 100(5):2197–2202, 2003.

[14] J. Eckstein and D. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(3):293-318, 1992.

[15] D. Gabay. Chapter IX applications of the method of multipliers to variational inequalities. *Studies in Mathematics and Its Applications*, 15:299–331, 1983.

[16] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.

[17] X. Gao, X. J. Cai, and D. R. Han. A Gauss-Seidel type inertial proximal alternating linearized minimization for a class of nonconvex optimization problems. *Journal of Global Optimization*, 76:863–887, 2020.

[18] K. Guo, D. R. Han, and T. T. Wu. Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints. *International Journal of Computer Mathematics*, 94(8):1653–1669, 2017.

[19] B. S. He and X. M. Yuan. A class of ADMM-based algorithms for three-block separable convex programming. *Computational Optimization and Applications*, 70:791–826, 2018.

[20] Z. H. Jia, X. Gao, X. J. Cai, and D. R. Han. Local linear convergence of the alternating direction method of multipliers for nonconvex separable optimization problems. *Journal of Optimization Theory and Applications*, 188:1–25, 2021.

[21] G. Y. Li and T. K. Pong. Global convergence of splitting methods for nonconvex composite optimization. *SIAM Journal on Optimization*, 25(4):2434–2460, 2015.

[22] G. Y. Li and T. K. Pong. Douglas-Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Mathematical Programming*, 159: 371–401, 2016.

[23] P. J. Liu, J. B. Jian, B. He, and X. Z. Jiang. Convergence of Bregman Peaceman-Rachford splitting method for nonconvex nonseparable optimization. *Journal of the Operations Research Society of China*, 11:707–733, 2023.

[24] Q. H. Liu, X. Y. Shen, and Y. T. Gu. Linearized ADMM for nonconvex nonsmooth optimization with convergence analysis. *IEEE Access*, 7:76131–76144, 2019.

[25] F. X. Liu, L. L. Xu, Y. H. Sun, and D. R. Han. A proximal alternating direction method for multi-block coupled convex optimization. *Journal of Industrial and Management Optimization*, 15:723–737, 2019.

[26] X. L. Lu and X. B. Lü. ADMM for image restoration based on nonlocal simultaneous sparse Bayesian coding. *Signal Processing: Image Communication*, 70:157–173, 2019.

[27] M. Meselhi, R. Sarker, D. Essam, and S. Elsayed. A decomposition approach for large-scale nonseparable optimization problems. *Applied Soft Computing*, 115:Article ID 108168, 2022.

[28] Y. Nesterov. *Introduction Lectures on Convex Optimization: A Basic Course*. 1st edition, Applied Optimization, Volume 87, Springer, New York, 2004.

[29] M. Ohishia, K. Fukuib, K. Okamurac, Y. Itohc, and H. Yanagihara. Coordinate optimization for generalized fused Lasso. *Communications in Statistics-Theory and Methods*, 50(24):5955–5973, 2021.

[30] D. W. Peaceman and J. H. H. Rachford. The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3(1):28–41, 1955.

[31] S. Setzer. Operator splittings, Bregman methods and frame shrinkage in image processing. *International Journal of Computer Vision*, 92(3):265–280, 2011.

[32] B. Wahlberg, S. Boyd, M. Annergren, and Y. Wang. An ADMM algorithm for a class of total variation regularized estimation problems. *IFAC Proceedings Volumes*, 45(16):83–88, 2012.

[33] H. H. Wang and A. Banerjee. Bregman alternating direction method of multipliers. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*, Volume 2, Pages

2816–2824, Cambridge, United States, 2014.

[34] F. H. Wang, W. F. Cao, and Z. B. Xu. Convergence of multi-block Bregman ADMM for nonconvex composite problems. *Science China Information Sciences*, 61:Article ID 122101, 2018.

[35] Z. W. Wen, C. Yang, X. Liu, and S. Marchesini. Alternating direction methods for classical and ptychographic phase retrieval. *Inverse Problems*, 28(11):Article ID 115010, 2012.

[36] 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 Transactions on Neural Networks and Learning Systems*, 23(7):1013–1027, 2012.

[37] 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 Journal on Imaging Sciences*, 10(1):74–110, 2017.

[38] L. M. Zeng and J. Xie. Group variable selection via SCAD-L_2. *Statistics*, 48:49–66, 2014.

[39] Y. Zhao, H. Y. Lan, and H. Y. Xu. Convergence of Peaceman-Rachford splitting method with Bregman distance for three-block nonconvex nonseparable optimization. *Demonstratio Mathematica*, 57(1):Article ID 20240036, 2024.