Journal of Applied Analysis and Computation
Volume 16, Number 5, October 2026, 2366–2391

Website:http://www.jaac-online.com
DOI:10.11948/20250166

A SEQUENTIAL COMPLETE INERTIAL BREGMAN ADMM FOR MULTI-BLOCK NONCONVEX PROBLEMS*

Lina Xia[1], Yujie Zhang[1] and Hongwei Li[1,†]

Abstract In this paper, a sequential complete inertial Bregman alternating direction method of multipliers (SCIB-ADMM) is proposed for multi-block nonconvex problems. The iterative method is established by utilizing the inertial strategy and Bregman distance to enhance its processing speed and efficiency. At each iteration, the SCIB-ADMM method utilizes two different relaxation factors and updates the Lagrange multiplier twice. The convergence of the SCIB-ADMM can be established under appropriate assumptions. Moreover, numerical experiments are presented on smoothly clipped absolute deviation (SCAD) and robust principal component analysis (PCA) problems to show the effectiveness of the SCIB-ADMM method.

Keywords Multi-block nonconvex problems, inertial strategy, Bregman distance, alternating direction method of multipliers, global convergence.

MSC(2010) 90C25.

1. Introduction

In this paper, a special class of multi-block optimization problems with nonlinear constraints is discussed. The problem comprises l decision variables $\boldsymbol{u}_i$ $(i = 1, ..., l)$ and a decision variable $\boldsymbol{v}$, along with an equality constraint. The multi-block nonconvex problem is formulated as

$$\min \sum_{i=1}^{l} f_i(\boldsymbol{u}_i) + g(\boldsymbol{v}) \quad \text{s.t.} \sum_{i=1}^{l} \boldsymbol{P}_i\boldsymbol{u}_i + \boldsymbol{Q}\boldsymbol{v} = \boldsymbol{r}, \tag{1.1}$$

where $l \geq 1$, $f_i : \mathbb{R}^{p_i} \to \mathbb{R} \cup \{+\infty\}$ $(i = 1, 2, \cdots, l)$ and $g : \mathbb{R}^q \to \mathbb{R}$ are variables, $\boldsymbol{P}_i \in \mathbb{R}^{m\times p_i}$, $\boldsymbol{Q} \in \mathbb{R}^{m\times q}$, $\boldsymbol{u}_i \in \mathbb{R}^{p_i}$, $\boldsymbol{v} \in \mathbb{R}^q$, $\boldsymbol{r} \in \mathbb{R}^m$.

For the i and j $(j \geq i,\ i, j \in \mathrm{N}^+)$, denote

$$\boldsymbol{P}_{[i,j]}\boldsymbol{u}_{[i,j]} = \sum_{n=i}^{j} \boldsymbol{P}_n\boldsymbol{u}_n,$$

where $\boldsymbol{P}_{[i,j]} = (\boldsymbol{P}_i, \boldsymbol{P}_{i+1}, \cdots, \boldsymbol{P}_j)$ and $\boldsymbol{u}_{[i,j]} = (\boldsymbol{u}_i^{\mathrm{T}}, \boldsymbol{u}_{i+1}^{\mathrm{T}}, \cdots, \boldsymbol{u}_j^{\mathrm{T}})^{\mathrm{T}}$.

The augmented Lagrangian function for problem (1.1) is

$$\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}, \boldsymbol{v}, \boldsymbol{w}) = \sum_{i=1}^{l} f_i(\boldsymbol{u}_i) + g(\boldsymbol{v}) - \langle \boldsymbol{w}, \boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]} + \boldsymbol{Q}\boldsymbol{v} - \boldsymbol{r}\rangle$$

[†]The corresponding author.
[1]School of Mathematics and Physics, China University of Geosciences, Wuhan 430074, China
[*]The authors were supported by National Natural Science Foundation of China (42374174).
Email: xln@cug.edu.cn(L. Xia), zhangyujie@cug.edu.cn(Y. Zhang), hwli@cug.edu.cn(H. Li)

$$+\frac{\beta}{2}\|\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}+\boldsymbol{Q}\boldsymbol{v}-\boldsymbol{r}\|^2, \tag{1.2}$$

where $\boldsymbol{w}\in\mathbb{R}^m$ and $\beta>0$. The ADMM is an effective method for problem (1.1) [10, 12], and its iteration form is

$$\begin{cases}\boldsymbol{u}_1^{k+1}=\arg\min_{\boldsymbol{u}_1}\{\mathcal{L}_\beta(\boldsymbol{u}_1,\boldsymbol{u}_{[2,l]}^k,\boldsymbol{v}^k,\boldsymbol{w}^k)\}, & (1.3a)\\ \boldsymbol{u}_2^{k+1}=\arg\min_{\boldsymbol{u}_2}\{\mathcal{L}_\beta(\boldsymbol{u}_1^{k+1},\boldsymbol{u}_2,\boldsymbol{u}_{[3,l]}^k,\boldsymbol{v}^k,\boldsymbol{w}^k)\}, & (1.3b)\\ \quad\vdots & \\ \boldsymbol{u}_l^{k+1}=\arg\min_{\boldsymbol{u}_l}\{\mathcal{L}_\beta(\boldsymbol{u}_{[1,l-1]}^{k+1},\boldsymbol{u}_l,\boldsymbol{v}^k,\boldsymbol{w}^k)\}, & (1.3c)\\ \boldsymbol{v}^{k+1}=\arg\min_{\boldsymbol{v}}\{\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1},\boldsymbol{v},\boldsymbol{w}^k)\}, & (1.3d)\\ \boldsymbol{w}^{k+1}=\boldsymbol{w}^k-\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{Q}\boldsymbol{v}^{k+1}-\boldsymbol{r}). & (1.3e)\end{cases}$$

When $l=1$, the convergence analysis of ADMM and its variants has been gradually mature and perfect in cases where both f_1 and g are convex [18, 27, 28, 45]. When at least one of f_1 and g is a nonconvex function, the convergence cannot be guaranteed. Extensive research has been conducted to identify the conditions in order to nonconvex ADMM converges [6, 7, 11, 17, 21, 30, 43]. Guo et al. [17] conducted a convergence analysis of the (1.3) applied to the two-block problem. Xu et al. [43] proposed an inertial Bregman ADMM that introduces a proximal term via Bregman distance, enabling the derivation of new proximal splitting methods for two-block nonconvex problems. Without relying on smoothness or differentiability, Barber et al. [6] obtained the convergence of the nonconvex ADMM and applied it to CT image processing. In addition, Liu et al. [30] explored the convergence of the ADMM for the nonconvex cosparse problem.

When $l\geq 2$, the ADMM or other splitting methods are well studied [15, 19, 20, 31] in cases where both f_i and g are convex. When at least one of f_i and g is a nonconvex function, for multi-block optimization problems, the convergence of the corresponding ADMM still needs further research, and in certain cases, it may even fail to converge. However, its outstanding performance in certain practical engineering applications has motivated scholars to investigate the improved version ADMM. Hong et al. [22] demonstrated the convergence of the multi-block ADMM when resolving the sharing and consensus issue. Additionally, when $\boldsymbol{Q}=\boldsymbol{I}$, Guo et al. [17] further developed the two-block variant [16] into a three-block version and verified its convergence. Unfortunately, in many cases, the $\boldsymbol{u}_i$-subproblems (1.3a)-(1.3c) and $\boldsymbol{v}$-subproblem (1.3d) of ADMM cannot be easily solved [8, 26]. In order to resolve the aforementioned drawback, Wang et al. [36] added the Bregman distance to the $\boldsymbol{u}_i$-subproblems (1.3a)-(1.3c) and $\boldsymbol{v}$-subproblem (1.3d). Then they developed these results to the multi-block Bregman alternating direction method of multipliers (BADMM). Furthermore, for multi-block problems, Jian et al. [24] established a partially symmetric regularized alternating direction method of multipliers (PSRADMM) and proved the subsequential convergence of the PSRADMM. In addition, the PSRADMM globally converges under the Kurdyka-Łojasiewicz property. For multi-block problem, combining inertial strategy and proximal term, Wang et al. [39] proposed an inertial proximal partially symmetric alternating direction method of multipliers (IPPS-ADMM) and obtained its global convergence. As a result, many improved ADMM variants have been proposed, see [4, 5, 23, 25, 29, 38] and their associated literatures.

Recently, some scholars have devoted their research to exploring the inertial strategy of the ADMM. The inertial technique has its origins in the time discretization of some second-order differential inclusions [33, 34], similarly sharing the feature that the new iterate is defined by using the previous two iterates. Subsequently, this inertial technique was extended to solve the maximal monotone operator inclusion problem in [1, 2]. An increasing number of scholars are focusing on studying inertial strategies to solve optimization problems, including the alternated multi-step inertial iterative algorithm [37], the inertial proximal gradient methods [40, 41] and the inertial proximal alternating algorithm [13].

The aforementioned works provide clear evidence that the inertial strategy effectively enhances the computational performance of the method, while the Bregman distance indeed simplifies the original subproblems. However, there has been limited attention in the literature to applying the inertial strategy and Bregman distance for solving multi-block nonconvex optimization problems. This motivates us to develop an ADMM framework that leverages the benefits of both the inertial strategy and the Bregman distance to improve computational efficiency. In this paper, a sequential complete inertial Bregman ADMM is proposed, which has the following contributions.

Firstly, a sequential complete inertial Bregman ADMM is proposed for solving the problem (1.1). This iterative method is an extension of numerous variants of well-known ADMM, such as the ADMM [16, 17], the algorithm in [26] and the BADMM [36]. Secondly, the inertial strategy and the Bregman distance are incorporated into all the $\boldsymbol{u}_i$-subproblems and $\boldsymbol{v}$-subproblem. The appropriate Bregman distance can be chosen to simplify the $\boldsymbol{u}_i$-subproblems and $\boldsymbol{v}$-subproblem at each iteration. Thirdly, for the SCIB-ADMM, the augmented Lagrangian function does not exhibit inherent monotonic nonincrease. For this purpose, an auxiliary function, with the augmented Lagrangian function as its foundation, is established to promote the analysis of the global convergence.

The framework of the paper is outlined below. Section 2 summarizes preliminaries. Section 3 introduces the SCIB-ADMM and prove its convergence. Section 4 validates the performance of the SCIB-ADMM through numerical results. In Section 5, some conclusions are presented.

2. Preliminaries

For two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, let $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$, $\|\boldsymbol{u}\| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u}\rangle}$. For a subset $\boldsymbol{S} \subseteq \mathbb{R}^n$ and a point $\boldsymbol{u} \in \mathbb{R}^n$, if $\boldsymbol{S}$ is nonempty, the distance from $\boldsymbol{u}$ to $\boldsymbol{S}$ is defined as $d(\boldsymbol{u}, \boldsymbol{S}) = \inf_{\boldsymbol{v}\in\boldsymbol{S}} \|\boldsymbol{v} - \boldsymbol{u}\|$. When $\boldsymbol{S} = \emptyset$, $d(\boldsymbol{u}, \boldsymbol{S}) = +\infty$. For function $f : \mathbb{R}^n \to (-\infty, +\infty]$, the domain of f is defined by $\operatorname{dom} f = \{\boldsymbol{u} \in \mathbb{R}^n : f(\boldsymbol{u}) < +\infty\}$.

Definition 2.1. [35] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper and lower semicontinuous.

(i) The Fréchet subdifferential denoted by $\hat{\partial} f(\boldsymbol{u})$, of f at $\boldsymbol{u} \in \operatorname{dom} f$, is

$$\hat{\partial} f(\boldsymbol{u}) = \{\boldsymbol{u} \in \mathbb{R}^n : \liminf_{\boldsymbol{v}\neq\boldsymbol{u}\ \boldsymbol{v}\to\boldsymbol{u}} \frac{f(\boldsymbol{v}) - f(\boldsymbol{u}) - \langle \boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle}{\|\boldsymbol{v} - \boldsymbol{u}\|} \geq 0\},$$

where $\hat{\partial} f(\boldsymbol{u}) = \emptyset$ if $\boldsymbol{u} \notin \operatorname{dom} f$.

(ii) If $\boldsymbol{u} \in \mathbb{R}^n$ is the minimum value of f, then $0 \in \partial f(\boldsymbol{u})$. $\boldsymbol{u}$ is a critical point of f if it satisfies $0 \in \partial f(\boldsymbol{u})$. The critical set of f can be denoted by $\operatorname{crit} f$.

Lemma 2.1. [14] *Let $\boldsymbol{Q} \in \mathbb{R}^{m\times q}$ denotes the nonzero matrix, $\mu_{\boldsymbol{Q}}$ represents the smallest eigenvalue that is positive among all the eigenvalues of $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}$, $\boldsymbol{P_Q}(\cdot)$ represents the orthogonal* *projection to denotes the Im$\boldsymbol{Q}$. Then*

$$\|\boldsymbol{P}_{\boldsymbol{Q}^{\mathrm{T}}}\boldsymbol{u}\| = \frac{1}{\sqrt{\mu_{\boldsymbol{Q}}}}\|\boldsymbol{Q}\boldsymbol{u}\|, \quad \forall \boldsymbol{u} \in \mathbb{R}^q.$$

Lemma 2.2. [32] *If g is continuously differentiable and ∇g is L_g-Lipschitz continuous. Then*

$$|g(\boldsymbol{v}) - g(\boldsymbol{u}) - \langle \nabla g(\boldsymbol{u}), \boldsymbol{v} - \boldsymbol{u}\rangle| \leq \frac{L_g}{2}\|\boldsymbol{v} - \boldsymbol{u}\|^2, \ \forall \ \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n.$$

The following lemma characterizes key properties of the Bregman distance.

Lemma 2.3. [36] *Let $\triangle_\phi(\boldsymbol{u}, \boldsymbol{v})$, $\forall \ \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ be the Bregman distance, it is formulated as*

$$\triangle_\phi(\boldsymbol{u}, \boldsymbol{v}) = \phi(\boldsymbol{u}) - \phi(\boldsymbol{v}) - \langle \nabla\phi(\boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v}\rangle,$$

where ϕ is convex differentiable. Then
(i) $\triangle_\phi(\boldsymbol{u}, \boldsymbol{v}) \geq 0$,
(ii) $\triangle_\phi(\boldsymbol{u}, \boldsymbol{v})$ may not be convex at $\boldsymbol{v}$, but it remains convex for all values of $\boldsymbol{u}$,
(iii) $\triangle_\phi(\boldsymbol{u}, \boldsymbol{v}) \geq \frac{\ell}{2}\|\boldsymbol{u} - \boldsymbol{v}\|^2$ when ϕ is ℓ-strongly convex.

Definition 2.2. [3] Let $\hbar : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper and lower semicontinuous. There is a positive e and a neighborhood U of $\hat{\boldsymbol{\varpi}}$ satisfying
(i) $\xi(0) = 0$,
(ii) ξ is continuously differential and continuous at 0,
(iii) $\xi'(s) > 0$, $\forall s \in (0, e)$,
(iv) $\xi'(\hbar(\boldsymbol{\varpi}) - \hbar(\hat{\boldsymbol{\varpi}}))d(0, \partial\hbar(\boldsymbol{\varpi})) \geq 1$, for all $\boldsymbol{\varpi} \in U \cap [\hbar(\hat{\boldsymbol{\varpi}}) + e > \hbar(\boldsymbol{\varpi}) > \hbar(\hat{\boldsymbol{\varpi}})]$,
where $\xi : [0, e) \to \mathbb{R}_+$ is continuous and concave. Then $\hbar$ has the Kurdyka-Łojasiewicz (KL) property at $\hat{\boldsymbol{\varpi}}$.

The lemma presented below holds significant importance for demonstrating the global convergence of the SCIB-ADMM.

Lemma 2.4. [7] *If Θ denotes the compact set, Θ_e is the set of concave functions Θ that fulfill conditions (i), (ii), and (iii) as outlined in Definition 2.2. If $\hbar(\boldsymbol{\varpi}) \equiv t$ for all $\boldsymbol{\varpi} \in \Theta$ and $\hbar$ has the KL property for any Θ. Then*

$$\xi'(\hbar(\boldsymbol{\varpi}) - t)d(0, \partial\hbar(\boldsymbol{\varpi})) \geq 1,$$

where $\boldsymbol{\varpi} \in \{\boldsymbol{\varpi} \in \mathbb{R}^n | d(\boldsymbol{\varpi}, \Theta) < \varepsilon, \ t < \hbar(\boldsymbol{\varpi}) < t + e\}$, $\varepsilon > 0$, $e > 0$, $\xi \in \Theta_e$.

Definition 2.3. If the following (2.1) holds,

$$\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{w}^* \in \partial f_i(\boldsymbol{u}_i^*), \ i = 1, 2, \cdots, l, \quad \nabla g(\boldsymbol{v}^*) = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{w}^*, \quad \boldsymbol{P}_{[1,l]}\boldsymbol{u}^*_{[1,l]} + \boldsymbol{Q}\boldsymbol{v}^* - \boldsymbol{r} = \boldsymbol{0}, \tag{2.1}$$

then $(\boldsymbol{u}^*_{[1,l]}, \boldsymbol{v}^*, \boldsymbol{w}^*) \in \mathrm{crit}\,\mathcal{L}_\beta$.

3. Convergence analysis

In this section, the SCIB-ADMM is proposed and its global convergence is proved.

3.1. SCIB-ADMM description

Motivated by the inertial strategy and the Bregman distance, the following method for problem (1.1) is proposed. Note that the following iterative schemes are sequentially updated using the inertial strategy and the Bregman distance, therefore the proposed method is named sequential complete inertial Bregman ADMM.

SCIB-ADMM. The iterative scheme of the sequential complete inertial Bregman ADMM is as follows.

$$
\begin{cases}
\bar{\boldsymbol{u}}_i^k = \boldsymbol{u}_i^k + \sigma_i(\boldsymbol{u}_i^k - \boldsymbol{u}_i^{k-1}),\ i = 1, 2, \cdots, l, & (3.1a)\\
\boldsymbol{u}_1^{k+1} = \arg\min\limits_{\boldsymbol{u}_1} \{\mathcal{L}_\beta(\boldsymbol{u}_1, \boldsymbol{u}_{[2,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k) + \triangle_{\phi_1}(\boldsymbol{u}_1, \bar{\boldsymbol{u}}_1^k)\}, & (3.1b)\\
\boldsymbol{u}_2^{k+1} = \arg\min\limits_{\boldsymbol{u}_2} \{\mathcal{L}_\beta(\boldsymbol{u}_1^{k+1}, \boldsymbol{u}_2, \boldsymbol{u}_{[3,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k) + \triangle_{\phi_2}(\boldsymbol{u}_2, \bar{\boldsymbol{u}}_2^k)\}, & (3.1c)\\
\quad\vdots & \\
\boldsymbol{u}_l^{k+1} = \arg\min\limits_{\boldsymbol{u}_l} \{\mathcal{L}_\beta(\boldsymbol{u}_{[1,l-1]}^{k+1}, \boldsymbol{u}_l, \boldsymbol{v}^k, \boldsymbol{w}^k) + \triangle_{\phi_l}(\boldsymbol{u}_l, \bar{\boldsymbol{u}}_l^k)\}, & (3.1d)\\
\boldsymbol{w}^{k+\frac{1}{2}} = \boldsymbol{w}^k - \tau\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^k - \boldsymbol{r}), & (3.1e)\\
\boldsymbol{v}^{k+1} = \arg\min\limits_{\boldsymbol{v}} \{\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}, \boldsymbol{w}^{k+\frac{1}{2}}) + \triangle_\varphi(\boldsymbol{v}, \boldsymbol{v}^k) + \rho\langle \boldsymbol{v}, \boldsymbol{v}^{k-1} - \boldsymbol{v}^k\rangle\}, & (3.1f)\\
\boldsymbol{w}^{k+1} = \boldsymbol{w}^{k+\frac{1}{2}} - \eta\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^{k+1} - \boldsymbol{r}). & (3.1g)
\end{cases}
$$

The SCIB-ADMM method has the following characteristics.

Remark 3.1. (i) The inertial step of the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d) is $\bar{\boldsymbol{u}}_i^k = \boldsymbol{u}_i^k + \sigma_i(\boldsymbol{u}_i^k - \boldsymbol{u}_i^{k-1})$. The inertial step is derived from the previous two iterations in order to avoid the large difference between two consecutive iteration points. The functions defined in (3.1b)-(3.1d) possess the term $\triangle_{\phi_i}(\boldsymbol{u}_i, \bar{\boldsymbol{u}}_i^k)$. By incorporating the inertial term into the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d), the new iteration point tends to the direction $\boldsymbol{u}_i^k - \boldsymbol{u}_i^{k-1}$, which is a "descending direction" of the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d).

(ii) The term $\rho\langle \boldsymbol{v}, \boldsymbol{v}^{k-1} - \boldsymbol{v}^k\rangle$ is called the inertial term. By incorporating $\rho\langle \boldsymbol{v}, \boldsymbol{v}^{k-1} - \boldsymbol{v}^k\rangle$ into the $\boldsymbol{v}$-subproblem (3.1f), the new iteration point tends to the direction $\boldsymbol{v}^k - \boldsymbol{v}^{k-1}$.

(iii) Choosing an appropriate Bregman distance will be able to simplify the subproblem. For the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d), if $\phi_i(\boldsymbol{u}_i) = 0$, then

$$
f_i(\boldsymbol{u}_i) + \frac{\beta}{2}\left\| \boldsymbol{P}_{[1,i-1]}\boldsymbol{u}_{[1,i-1]}^{k+1} + \boldsymbol{P}_i\boldsymbol{u}_i + \boldsymbol{P}_{[i,l]}\boldsymbol{u}_{[i,l]}^k + \boldsymbol{Q}\boldsymbol{v}^k - \boldsymbol{r} - \frac{\boldsymbol{w}^k}{\beta}\right\|^2.
$$

Solving the above problem is difficult. If $\phi_i(\boldsymbol{u}_i) = \frac{\mu}{2}\|\boldsymbol{u}_i\|^2 - \frac{\beta}{2}\left\| \boldsymbol{P}_{[1,i-1]}\boldsymbol{u}_{[1,i-1]}^{k+1} + \boldsymbol{P}_i\boldsymbol{u}_i + \boldsymbol{P}_{[i,l]}\boldsymbol{u}_{[i,l]}^k + \boldsymbol{Q}\boldsymbol{v}^k - \boldsymbol{r} - \frac{\boldsymbol{w}^k}{\beta}\right\|^2$ with $\mu > \|\boldsymbol{Q}\|^2$, then the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d) is transformed into minimizing function

$$
f_i(\boldsymbol{u}_i) + \frac{\mu\beta}{2}\|\boldsymbol{u}_i - c_{\boldsymbol{u}}\|^2,
$$

where $c_{\boldsymbol{u}} = \bar{\boldsymbol{u}}_i^k - \mu^{-1}\boldsymbol{Q}^{\mathrm{T}}\left(\boldsymbol{P}_{[1,i-1]}\boldsymbol{u}_{[1,i-1]}^{k+1} + \boldsymbol{P}_i\bar{\boldsymbol{u}}_i + \boldsymbol{P}_{[i,l]}\boldsymbol{u}_{[i,l]}^k + \boldsymbol{Q}\boldsymbol{v}^k - \boldsymbol{r} - \frac{\boldsymbol{w}^k}{\beta}\right)$ and $c_{\boldsymbol{u}}$ is a certain known vector. The functions f_i are proper, lower semicontinuous, and bounded below, from the proximal behavior in [9, 44], the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d) has a closed form solution.

(iv) The SCIB-ADMM incorporate two relaxation factors τ and η, whose appropriate selection across diverse applications effectively enhance experimental performance.

(v) The SCIB-ADMM is an extension of many variants of well-known ADMM. For example, when $\tau = 0,\ \eta = 1,\ \phi_i = 0,\ \varphi = 0,\ \sigma_i = 0,\ \rho = 0$, the SCIB-ADMM becomes the ADMM [16,17]; when $l = 1,\ \tau = 0,\ \eta = 1,\ \phi_i = 0,\ \sigma_i = 0,\ \rho = 0$, the SCIB-ADMM becomes the algorithm in [26]; when $\tau = 0,\ \eta = 1,\ \rho = 0,\ \sigma_i = 0$, the SCIB-ADMM becomes the BADMM [36].

Based on the problem (1.1) and the SCIB-ADMM (3.1), the following assumptions are provided.

Assumption 3.1. *(i) f_i $(i = 1, 2, \cdots, l)$ are proper and lower semicontinuous, g is continuously differentiable, and the gradient ∇g is L_g-Lipschitz continuous;*

(ii) $\triangle_{\phi_i}$ $(i = 1, 2, \cdots, l)$ and $\triangle_{\varphi}$ denote the Bregman distance corresponding to the functions ϕ_i and φ, respectively;

(iii) ϕ_i $(i = 1, 2, \cdots, l)$ are strongly convex continuously differentiable and $\nabla\phi_i$ are L_{ϕ_i}-Lipschitz continuous;

(iv) $\nabla\varphi$ is L_{φ}-Lipschitz continuous.

The global convergence of the SCIB-ADMM will be proved below: (i) $\boldsymbol{Q} = \boldsymbol{I}$; (ii) $\boldsymbol{Q} \neq \boldsymbol{I}$ and $\eta = 1$. For convenience, the following notations that will be utilized throughout this paper are introduced.

$$\begin{aligned} &\boldsymbol{u}_{[1,l]} = (\boldsymbol{u}_1^{\mathrm{T}}, \boldsymbol{u}_2^{\mathrm{T}}, \cdots, \boldsymbol{u}_l^{\mathrm{T}})^{\mathrm{T}} = (\boldsymbol{u}_{[1,i]}, \boldsymbol{u}_{[i+1,l]}), \quad i = 1, 2, \cdots, l, \\ &\triangle\boldsymbol{u}_i^k = \boldsymbol{u}_i^k - \boldsymbol{u}_i^{k-1}, \quad \triangle\boldsymbol{v}^k = \boldsymbol{v}^k - \boldsymbol{v}^{k-1}, \quad \triangle\boldsymbol{w}^k = \boldsymbol{w}^k - \boldsymbol{w}^{k-1}. \end{aligned}$$

3.2. Convergence analysis when $\boldsymbol{Q} = \boldsymbol{I}$

Based on the optimality condition of the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d) and $\boldsymbol{v}$-subproblem (3.1f), one has

$$\begin{cases} 0 \in \partial f_i(\boldsymbol{u}_i^{k+1}) - \boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{w}^k + \beta\boldsymbol{P}_i^{\mathrm{T}}\Big(\boldsymbol{P}_{[1,i]}\boldsymbol{u}_{[1,i]}^{k+1} + \boldsymbol{P}_{[i+1,l]}\boldsymbol{u}_{[i+1,l]}^{k} + \boldsymbol{v}^k - \boldsymbol{r}\Big) \\ \qquad +\nabla\phi_i(\boldsymbol{u}_i^{k+1}) - \nabla\phi_i(\bar{\boldsymbol{u}}_i^k), \quad i = 1, 2, \cdots, l, & (3.2\text{a}) \\ 0 = \nabla g(\boldsymbol{v}^{k+1}) - \boldsymbol{w}^{k+\frac{1}{2}} + \beta\big(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}\big) + \rho(\boldsymbol{v}^{k-1} - \boldsymbol{v}^k) \\ \qquad +\nabla\varphi(\boldsymbol{v}^{k+1}) - \nabla\varphi(\boldsymbol{v}^k). & (3.2\text{b}) \end{cases}$$

In this subsection, $\{\boldsymbol{\varpi}^k = (\boldsymbol{u}_{[1,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k)\}$ is generated by the SCIB-ADMM. Denote

$$\hat{\boldsymbol{\varpi}}^k = (\boldsymbol{u}_{[1,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k, \boldsymbol{u}_{[1,l]}^{k-1}, \boldsymbol{v}^{k-1}, \boldsymbol{v}^{k-2}), \qquad \hat{\boldsymbol{\varpi}} = (\boldsymbol{u}_{[1,l]}, \boldsymbol{v}, \boldsymbol{w}, \hat{\boldsymbol{u}}_{[1,l]}, \hat{\boldsymbol{v}}, \widetilde{\boldsymbol{v}}).$$

An auxiliary function is constructed with the augmented Lagrangian function as its foundation, namely

$$\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) = \mathcal{L}_\beta(\boldsymbol{\varpi}^k) + \sum_{i=1}^{l}\frac{\sigma_i^2 L_{\phi_i}^2}{2}\|\triangle\boldsymbol{u}_i^k\|^2 + \kappa_1\|\triangle\boldsymbol{v}^k\|^2 + \kappa_2\|\triangle\boldsymbol{v}^{k-1}\|^2, \tag{3.3}$$

where $\kappa_1 = \frac{(l+6)(2\rho^2+L_\varphi^2)}{(\tau+\eta)\beta} + \frac{\rho^2}{2}$ and $\kappa_2 = \frac{(l+6)\rho^2}{(\tau+\eta)\beta}$.

In order to analyze the monotonicity of the auxiliary function $\{\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)\}$, denote

$$\begin{cases} \delta_i = \dfrac{\ell_i - 1 - \sigma_i^2 L_{\phi_i}^2}{2} - \dfrac{(l+6)\beta(1-\eta)^2}{\tau+\eta}\lambda_{\max}(\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{P}_i),\ i = 1,2,\cdots,l, \\ \delta_{l+1} = \dfrac{\left[\tau(1-2\eta)+\eta-2(l+6)(1-\eta)^2\right]\beta^2 - \left(L_g + L_\varphi^2+\rho^2+2\right)(\tau+\eta)\beta - 2(l+6)\left(L_g^2+2L_\varphi^2+2\rho^2\right)}{2(\tau+\eta)\beta}, \\ \delta = \min\{\delta_1,\delta_2,\cdots,\delta_{l+1}\}. \end{cases} \tag{3.4}$$

Based on (3.4), the following assumptions are provided.

Assumption 3.2. *The parameters $\tau,\ \eta,\ \beta$ and ℓ_i satisfy:*

(i) The solution set of τ and η can be denoted by $(\tau,\eta)\in S = S_1\cup S_2$, where

$$S_1 = \left(\frac{2(1-\eta)^2(l+6)-\eta}{1-2\eta}, +\infty\right)\times\left(-\infty,\frac{1}{2}\right),$$
$$S_2 = \left(-\eta, \frac{2(1-\eta)^2(l+6)-\eta}{1-2\eta}\right)\times\left(\frac{l+6-\sqrt{l+6}}{l+6}, \frac{l+6+\sqrt{l+6}}{l+6}\right);$$

(ii) $\beta > \frac{(\tau+\eta)(L_g+L_\varphi^2+\rho^2+2)}{2[\tau(1-2\eta)+\eta-2(l+6)(1-\eta)^2]}$
$+\frac{\sqrt{(\tau+\eta)^2(L_g+L_\varphi^2+\rho^2+2)^2+8(l+6)(L_g^2+2L_\varphi^2+2\rho^2)[\tau(1-2\eta)+\eta-2(l+6)(1-\eta)^2]}}{2[\tau(1-2\eta)+\eta-2(l+6)(1-\eta)^2]}$;

(iii) $\ell_i \geq \frac{2(l+6)\beta(1-\eta)^2}{\tau+\eta}\lambda_{\max}(\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{P}_i) + \sigma_i^2 L_{\phi_i}^2 + 1,\ i = 1,2,\ldots,l.$

Remark 3.2. From (3.4) and Assumption 3.2, it holds that $\delta > 0$.

The following lemma establishes that (3.3) is monotonically nonincreasing.

Lemma 3.1. *If Assumptions* 3.1 *and* 3.2 *hold. Then*

$$\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) + \delta\left(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\|^2 + \|\triangle \boldsymbol{v}^{k+1}\|^2\right) \leq \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k), \quad \forall k \geq 2.$$

Furthermore, the function referenced by (3.3) *is monotonically nonincreasing.*

Proof. From (1.2) and (3.2b), one has

$$\begin{aligned} &\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^{k+1}, \boldsymbol{w}^{k+\frac{1}{2}}) - \mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^{k}, \boldsymbol{w}^{k+\frac{1}{2}}) \\ \leq& -\frac{\beta}{2}\|\triangle \boldsymbol{v}^{k+1}\|^2 + \frac{L_g}{2}\|\triangle \boldsymbol{v}^{k+1}\|^2 \\ &+ \left\langle \boldsymbol{v}^{k+1}-\boldsymbol{v}^k, \nabla\varphi(\boldsymbol{v}^k) - \nabla\varphi(\boldsymbol{v}^{k+1}) + \rho(\boldsymbol{v}^k - \boldsymbol{v}^{k-1})\right\rangle \\ \leq& \frac{\rho^2}{2}\|\triangle \boldsymbol{v}^k\|^2 - \frac{\beta - L_g - L_\varphi^2 - 2}{2}\|\triangle \boldsymbol{v}^{k+1}\|^2, \end{aligned} \tag{3.5}$$

where the two inequalities above utilize Lemma 2.2, Assumptions 3.1 (i) and (iv).

In view of (3.1a)-(3.1d), it holds that

$$\begin{aligned} &\mathcal{L}_\beta(\boldsymbol{u}_{[1,i]}^{k+1}, \boldsymbol{u}_{[i+1,l]}^{k}, \boldsymbol{v}^k, \boldsymbol{w}^k) - \mathcal{L}_\beta(\boldsymbol{u}_{[1,i-1]}^{k+1}, \boldsymbol{u}_{[i,l]}^{k}, \boldsymbol{v}^k, \boldsymbol{w}^k) \\ \leq& \triangle_{\phi_i}(\boldsymbol{u}_i^k, \bar{\boldsymbol{u}}_i^k) - \triangle_{\phi_i}(\boldsymbol{u}_i^{k+1}, \bar{\boldsymbol{u}}_i^k) \\ =& \langle \nabla\phi_i(\bar{\boldsymbol{u}}_i^k) - \nabla\phi_i(\boldsymbol{u}_i^k), \boldsymbol{u}_i^{k+1}-\boldsymbol{u}_i^k\rangle - \left(\phi_i(\boldsymbol{u}_i^{k+1}) - \phi_i(\boldsymbol{u}_i^k) - \langle\nabla\phi_i(\boldsymbol{u}_i^k), \boldsymbol{u}_i^{k+1}-\boldsymbol{u}_i^k\rangle\right) \end{aligned}$$

$$
\begin{aligned}
&\le \frac{1}{2}\|\nabla\phi_i(\bar{\boldsymbol{u}}_i^k)-\nabla\phi_i(\boldsymbol{u}_i^k)\|^2+\frac{1}{2}\|\boldsymbol{u}_i^{k+1}-\boldsymbol{u}_i^k\|^2-\frac{\ell_i}{2}\|\boldsymbol{u}_i^{k+1}-\boldsymbol{u}_i^k\|^2\\
&\le \frac{\sigma_i^2L_{\phi_i}^2}{2}\|\triangle\boldsymbol{u}_i^k\|^2-\frac{\ell_i-1}{2}\|\triangle\boldsymbol{u}_i^{k+1}\|^2,
\end{aligned}
$$

where the second inequality above utilizes Lemma 2.3 (iii). Then, one has

$$
\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1},\boldsymbol{v}^k,\boldsymbol{w}^k)-\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k},\boldsymbol{v}^k,\boldsymbol{w}^k)\le-\sum_{i=1}^{l}\Big(\frac{\ell_i-1}{2}\|\triangle\boldsymbol{u}_i^{k+1}\|^2-\frac{\sigma_i^2L_{\phi_i}^2}{2}\|\triangle\boldsymbol{v}_i^k\|^2\Big). \tag{3.6}
$$

From (3.1g), one has $\boldsymbol{w}^{k+1}=\boldsymbol{w}^{k+\frac{1}{2}}-\eta\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^{k+1}-\boldsymbol{r})$. This, together with (3.2b), one has

$$
\begin{aligned}
\|\boldsymbol{w}^{k+1}-\boldsymbol{w}^k\|^2\le{}&(l+6)\beta^2(1-\eta)^2\sum_{i=1}^{l}\lambda_{\max}(\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{P}_i)\|\triangle\boldsymbol{u}_i^{k+1}\|^2\\
&+(l+6)\Big(L_g^2+L_\varphi^2+\beta^2(1-\eta)^2\Big)\|\triangle\boldsymbol{v}^{k+1}\|^2\\
&+(l+6)(\rho^2+L_\varphi^2)\|\triangle\boldsymbol{v}^k\|^2+(l+6)\rho^2\|\triangle\boldsymbol{v}^{k-1}\|^2.
\end{aligned}\tag{3.7}
$$

In view of (3.1e) and (3.1g), one has

$$
\mathcal{L}_\beta(\boldsymbol{u}^{k+1},\boldsymbol{v}^k,\boldsymbol{w}^{k+\frac{1}{2}})-\mathcal{L}_\beta(\boldsymbol{u}^{k+1},\boldsymbol{v}^k,\boldsymbol{w}^k)=\tau\beta\|\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^k-\boldsymbol{r}\|^2, \tag{3.8}
$$

$$
\mathcal{L}_\beta(\boldsymbol{\varpi}^{k+1})-\mathcal{L}_\beta(\boldsymbol{u}^{k+1},\boldsymbol{v}^{k+1},\boldsymbol{w}^{k+\frac{1}{2}})=\eta\beta\|\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^{k+1}-\boldsymbol{r}\|^2, \tag{3.9}
$$

$$
\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^{k+1}-\boldsymbol{r}=\frac{\tau}{\tau+\eta}\triangle\boldsymbol{v}^{k+1}-\frac{1}{(\tau+\eta)\beta}\triangle\boldsymbol{w}^{k+1}, \tag{3.10}
$$

and

$$
\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^{k}-\boldsymbol{r}=-\frac{\eta}{\tau+\eta}\triangle\boldsymbol{v}^{k+1}-\frac{1}{(\tau+\eta)\beta}\triangle\boldsymbol{w}^{k+1}. \tag{3.11}
$$

According to (3.7), (3.10), (3.11) and $\tau+\eta>0$, one has

$$
\begin{aligned}
&\tau\beta\|\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^k-\boldsymbol{r}\|^2+\eta\beta\|\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1}+\boldsymbol{v}^{k+1}-\boldsymbol{r}\|^2\\
={}&\frac{\tau\eta\beta}{\tau+\eta}\|\triangle\boldsymbol{v}^{k+1}\|^2+\frac{1}{(\tau+\eta)\beta}\|\triangle\boldsymbol{w}^{k+1}\|^2\\
\le{}&\frac{(l+6)\beta(1-\eta)^2}{\tau+\eta}\sum_{i=1}^{l}\lambda_{\max}(\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{P}_i)\|\triangle\boldsymbol{u}_i^{k+1}\|^2+\frac{(l+6)\rho^2}{(\tau+\eta)\beta}\|\triangle\boldsymbol{v}^{k-1}\|^2\\
&+\frac{(l+6)(\rho^2+L_\varphi^2)}{(\tau+\eta)\beta}\|\triangle\boldsymbol{v}^k\|^2\\
&+\frac{\tau\eta\beta^2+(l+6)\big(L_g^2+L_\varphi^2+\beta^2(1-\eta)^2\big)}{(\tau+\eta)\beta}\|\triangle\boldsymbol{v}^{k+1}\|^2.
\end{aligned}\tag{3.12}
$$

Combining (3.5), (3.6), (3.8), (3.9) and (3.12), one has

$$
\begin{aligned}
&\mathcal{L}_\beta(\boldsymbol{\varpi}^{k+1})\\
\le{}&\mathcal{L}_\beta(\boldsymbol{\varpi}^k)+\sum_{i=1}^{l}\Big(\frac{\sigma_i^2L_{\phi_i}^2}{2}\|\triangle\boldsymbol{u}_i^k\|^2-\frac{\ell_i-1}{2}\|\triangle\boldsymbol{u}_i^{k+1}\|^2\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{(l+6)\beta(1-\eta)^2}{\tau+\eta}\sum_{i=1}^{l}\lambda_{\max}(\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{P}_i)\|\triangle\boldsymbol{u}_i^{k+1}\|^2\\
&+\frac{(l+6)\rho^2}{(\tau+\eta)\beta}\|\triangle\boldsymbol{v}^{k-1}\|^2+\left(\frac{(l+6)(\rho^2+L_\varphi^2)}{(\tau+\eta)\beta}+\frac{\rho^2}{2}\right)\|\triangle\boldsymbol{v}^k\|^2\\
&-\left(\frac{\beta-L_g-L_\varphi^2-2}{2}-\frac{\tau\eta\beta^2+(l+6)\big(L_g^2+L_\varphi^2+\beta^2(1-\eta)^2\big)}{(\tau+\eta)\beta}\right)\|\triangle\boldsymbol{v}^{k+1}\|^2.
\end{aligned}
$$

This, together with (3.3), one has

$$
\begin{aligned}
&\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})\\
\le&\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k})-\sum_{i=1}^{l}\left(\frac{\ell_i-1-\sigma_i^2L_{\phi_i}^2}{2}-\frac{(l+6)\beta(1-\eta)^2}{\tau+\eta}\lambda_{\max}(\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{P}_i)\right)\|\triangle\boldsymbol{u}_i^{k+1}\|^2\\
&+\frac{\tau\eta\beta^2+(l+6)\Big(L_g^2+2L_\varphi^2+\beta^2(1-\eta)^2+2\rho^2\Big)}{(\tau+\eta)\beta}\|\triangle\boldsymbol{v}^{k+1}\|^2\\
&-\frac{\beta-L_g-L_\varphi^2-2-\rho^2}{2}\|\triangle\boldsymbol{v}^{k+1}\|^2.
\end{aligned}
$$

It follows from (3.4) that

$$
\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})+\delta\big(\|\triangle\boldsymbol{u}_{[1,l]}^{k+1}\|^2+\|\triangle\boldsymbol{v}^{k+1}\|^2\big)\le\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k}). \tag{3.13}
$$

From $\delta>0$, the auxiliary function $\{\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k})\}$ is monotonically nonincreasing. So, the lemma holds. □

Lemma 3.2. *If Assumptions* 3.1 *and* 3.2 *hold and* $\{\boldsymbol{\varpi}^k\}$ *is bounded. Then*

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

Proof. Since $\{\boldsymbol{\varpi}^k\}$ is bounded, $\{\hat{\boldsymbol{\varpi}}^k\}$ is also bounded. Then $\lim\limits_{k_j\to+\infty}\hat{\boldsymbol{\varpi}}^{k_j}=\hat{\boldsymbol{\varpi}}^*$, where $\{\hat{\boldsymbol{\varpi}}^{k_j}\}$ is a subsequence of $\hat{\boldsymbol{\varpi}}^*$. From Assumption 3.1 (i), $\liminf\limits_{k_j\to+\infty}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k_j})\ge\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)$. Thus, $\{\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k_j})\}$ is bounded below. Combining Lemma 3.1,

$$
\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\le\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k).
$$

By summing (3.13) for $k\ge 2$, one has

$$
\delta\sum_{k=2}^{+\infty}\big(\|\triangle\boldsymbol{u}_{[1,l]}^{k+1}\|^2+\|\triangle\boldsymbol{v}^{k+1}\|^2\big)\le\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^2)-\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)<+\infty.
$$

Since $\delta>0$, one has $\sum\limits_{k=0}^{+\infty}\|\triangle\boldsymbol{u}_{[1,l]}^{k+1}\|^2<+\infty$, $\sum\limits_{k=0}^{+\infty}\|\triangle\boldsymbol{v}^{k+1}\|^2<+\infty$. This, together with (3.7), shows that $\sum\limits_{k=0}^{+\infty}\|\triangle\boldsymbol{w}^{k+1}\|^2<+\infty$.

Thus, $\sum\limits_{k=0}^{+\infty}\|\boldsymbol{\varpi}^{k+1}-\boldsymbol{\varpi}^k\|^2<+\infty$. □

The following subsequential convergence can be proved.

Theorem 3.1. *If Assumptions* 3.1 *and* 3.2 *hold and* $\{\boldsymbol{\varpi}^k\}$ *is bounded. The sequence* $\{\boldsymbol{\varpi}^k\}$ *possesses a cluster point set of* Θ*, while* $\{\hat{\boldsymbol{\varpi}}^k\}$ *possesses a cluster point set of* $\hat{\Theta}$*. Then*

(i) $\lim\limits_{k\to+\infty} d(\hat{\boldsymbol{\varpi}}^k, \hat{\Theta}) = \lim\limits_{k\to+\infty} d(\boldsymbol{\varpi}^k, \Theta) = 0$.

(ii) $\Theta \subseteq \operatorname{crit} \mathcal{L}_\beta$.

(iii) The $\{\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)\}$ *is convergent, and* $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) = \lim\limits_{k\to+\infty} \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)$.

Proof. (i) From the characteristics of Θ and $\hat{\Theta}$, the assertion (i) holds.

(ii) Setting $\boldsymbol{\varpi}^* \in \Theta$, a subsequence $\{\boldsymbol{\varpi}^{k_j}\}$ of $\{\boldsymbol{\varpi}^k\}$ satisfies the condition $\lim\limits_{k_j\to+\infty} \boldsymbol{\varpi}^{k_j} = \boldsymbol{\varpi}^*$. From Lemma 3.2, $\lim\limits_{k_j\to+\infty} \boldsymbol{\varpi}^{k_j+1} = \boldsymbol{\varpi}^*$. Therefore, it follows from (3.1e) that $\{\boldsymbol{w}^{k_j+\frac{1}{2}}\}$ is convergent. Let $\lim\limits_{k_j\to+\infty} \boldsymbol{w}^{k_j+\frac{1}{2}} = \boldsymbol{w}^{**}$. Then by taking limit $k_j \to +\infty$ in (3.1e) and (3.1g), it follows that

$$\boldsymbol{w}^{**} = \boldsymbol{w}^* - \tau\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}^*_{[1,l]} + \boldsymbol{v}^* - \boldsymbol{r}), \qquad \boldsymbol{w}^* = \boldsymbol{w}^{**} - \eta\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}^*_{[1,l]} + \boldsymbol{v}^* - \boldsymbol{r}).$$

Considering the fact that $\tau + \eta > 0$, one has

$$\boldsymbol{w}^{**} = \boldsymbol{w}^*, \qquad \boldsymbol{P}_{[1,l]}\boldsymbol{u}^*_{[1,l]} + \boldsymbol{v}^* - \boldsymbol{r} = \boldsymbol{0}. \tag{3.14}$$

Thus, problem (1.1) has a valid solution, which is $(\boldsymbol{u}^*, \boldsymbol{v}^*)$. Since $\boldsymbol{u}_i^{k_j+1}$ $(i = 1, 2, \cdots, l)$ minimize the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d), it follows that

$$\begin{aligned}
& f_i(\boldsymbol{u}_i^{k_j+1}) - \langle \boldsymbol{w}^{k_j}, \boldsymbol{P}_i\boldsymbol{u}_i^{k_j+1}\rangle + \frac{\beta}{2}\|\boldsymbol{P}_{[1,i]}\boldsymbol{u}^{k_j+1}_{[1,i]} + \boldsymbol{P}_{[i+1,l]}\boldsymbol{u}^{k_j+1}_{[i+1,l]} + \boldsymbol{v}^{k_j} - \boldsymbol{r}\|^2 \\
& + \triangle_{\phi_i}(\boldsymbol{u}_i^{k_j+1}, \bar{\boldsymbol{u}}_i^{k_j}) \\
\le\ & f_i(\boldsymbol{u}_i^*) - \langle \boldsymbol{w}^{k_j}, \boldsymbol{P}_i\boldsymbol{u}_i^*\rangle + \frac{\beta}{2}\|\boldsymbol{P}_{[1,i-1]}\boldsymbol{u}^{k_j+1}_{[1,i-1]} + \boldsymbol{P}_i\boldsymbol{u}_i^* + \boldsymbol{P}_{[i+1,l]}\boldsymbol{u}^{k_j+1}_{[i+1,l]} + \boldsymbol{v}^{k_j} - \boldsymbol{r}\|^2 \\
& + \triangle_{\phi_i}(\boldsymbol{u}_i^*, \bar{\boldsymbol{u}}_i^{k_j}).
\end{aligned}$$

Associated with $\lim\limits_{k\to+\infty} \hat{\boldsymbol{\varpi}}^{k_j} = \lim\limits_{k\to+\infty} \hat{\boldsymbol{\varpi}}^{k_j+1} = \hat{\boldsymbol{\varpi}}^*$, one has

$$f_i(\boldsymbol{u}_i^*) \ge \limsup_{k_j\to+\infty} f_i(\boldsymbol{u}_i^{k_j+1}).$$

Moreover, From Assumption 3.1 (i), one has $\lim\limits_{k_j\to+\infty} f_i(\boldsymbol{u}_i^{k_j+1}) \ge f_i(\boldsymbol{u}_i^*)$. Hence

$$\lim_{k_j\to+\infty} f_i(\boldsymbol{u}_i^{k_j+1}) = f_i(\boldsymbol{u}_i^*). \tag{3.15}$$

In view of (3.1a), Assumption 3.1 (iii) and $\lim\limits_{k_j\to+\infty} \|\triangle\boldsymbol{u}_i^{k_j+1}\| = 0$ $(i = 1, 2, \cdots, l)$, one has

$$\begin{aligned}
\lim_{k_j\to+\infty} \|\nabla\phi_i(\boldsymbol{u}_i^{k_j+1}) - \nabla\phi_i(\bar{\boldsymbol{u}}_i^{k_j})\| &= \lim_{k_j\to+\infty} L_{\phi_i}\|\boldsymbol{u}_i^{k_j+1} - \bar{\boldsymbol{u}}_i^{k_j}\| \\
&\le \lim_{k_j\to+\infty} L_{\phi_i}\|\triangle\boldsymbol{u}_i^{k_j+1}\| + \sigma_i L_{\phi_i}\|\triangle\boldsymbol{u}_i^{k_j}\| \\
&= 0.
\end{aligned} \tag{3.16}$$

Combining the closeness of ∂f_i, the continuity of ∇g, and (3.14)-(3.16), by taking the limit $k = k_j \to +\infty$ in (3.2), one has

$$\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{w}^* \in \partial f_i(\boldsymbol{u}_i^*), \quad \nabla g(\boldsymbol{v}^*) = \boldsymbol{w}^*, \quad \boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^* + \boldsymbol{v}^* - \boldsymbol{r} = \boldsymbol{0}.$$

From Definition 2.3, $\boldsymbol{\varpi}^* \in \operatorname{crit}\mathcal{L}_\beta$. Thus, $\Theta \subseteq \operatorname{crit}\mathcal{L}_\beta$

(iii) Setting $\hat{\boldsymbol{\varpi}}^* \in \hat{\Theta}$, a subsequence $\{\hat{\boldsymbol{\varpi}}^{k_j}\}$ of $\{\hat{\boldsymbol{\varpi}}^k\}$ satisfies the condition $\lim\limits_{k_j\to+\infty} \hat{\boldsymbol{\varpi}}^{k_j+1} = \hat{\boldsymbol{\varpi}}^*$. According to (3.3), (3.15) and continuity of g, one has $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) = \lim\limits_{k_j\to+\infty} \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k_j+1})$. Associated with the monotonicity of $\{\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)\}$, one has $\{\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)\}$ is convergent. Thus, $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) = \lim\limits_{k\to+\infty} \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)$, $\forall\ \hat{\boldsymbol{\varpi}}^* \in \hat{\Theta}$. □

Lemma 3.3. *If Assumptions* 3.1 *and* 3.2 *hold. For all* $k \geq 2$, *define*

$$\begin{cases} \epsilon_i^{k+1} = -\boldsymbol{P}_i^{\mathrm{T}}\triangle\boldsymbol{w}^{k+1} + \beta\boldsymbol{P}_i^{\mathrm{T}}\Big(\boldsymbol{P}_{[i+1,l]}\triangle\boldsymbol{u}_{[i+1,l]}^{k+1}\Big) + \beta\boldsymbol{P}_i^{\mathrm{T}}\triangle\boldsymbol{v}^{k+1} \\ \qquad\quad +\sigma_i^2 L_{\phi_i}^2\triangle\boldsymbol{u}_i^{k+1} - \Big(\nabla\phi_i(\boldsymbol{u}_i^{k+1}) - \nabla\phi_i(\bar{\boldsymbol{u}}_i^k)\Big),\ i=1,2,\cdots,l, \\ \epsilon_{l+1}^{k+1} = \eta\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}) + \rho\triangle\boldsymbol{v}^k + 2\kappa_1\triangle\boldsymbol{v}^{k+1} - \Big(\nabla\varphi(\boldsymbol{v}^{k+1}) - \nabla\varphi(\boldsymbol{v}^k)\Big), \\ \epsilon_{l+2}^{k+1} = -(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}), \\ \epsilon_{l+2+i}^{k+1} = -\sigma_i^2 L_{\phi_i}^2\triangle\boldsymbol{u}_i^{k+1},\ i=1,2,\cdots,l, \\ \epsilon_{2l+3}^{k+1} = 2\kappa_2\triangle\boldsymbol{v}^k - 2\kappa_1\triangle\boldsymbol{v}^{k+1}, \quad \epsilon_{2l+4}^{k+1} = -2\kappa_2\triangle\boldsymbol{v}^k. \end{cases}$$

Thus, $\boldsymbol{\epsilon}^{k+1} = (\boldsymbol{\epsilon}_{[1,l]}^{k+1}, \epsilon_{l+1}^{k+1}, \epsilon_{l+2}^{k+1}, \boldsymbol{\epsilon}_{[l+3,2l+2]}^{k+1}, \epsilon_{2l+3}^{k+1}, \epsilon_{2l+4}^{k+1}) \in \partial\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})$, *where* $\boldsymbol{\epsilon}_{[i,j]}^{k+1} = (\epsilon_i^{k+1}, \epsilon_{i+1}^{k+1}, \cdots, \epsilon_j^{k+1})$ *for* $j \geq i$. *Then*

$$d(0, \partial\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})) \leq \chi\Big(\|\triangle\boldsymbol{u}_{[1,l]}^k\| + \|\triangle\boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle\boldsymbol{v}^{k-1}\| + \|\triangle\boldsymbol{v}^k\| + \|\triangle\boldsymbol{v}^{k+1}\|\Big), \quad \chi > 0.$$

Proof. By the characteristics of $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})$ and $\hat{\boldsymbol{\varpi}}^{k+1}$, one has

$$\begin{cases} \partial_{\boldsymbol{u}_i}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) = \partial f_i(\boldsymbol{u}_i^{k+1}) - \boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{w}^{k+1} + \beta\boldsymbol{P}_i^{\mathrm{T}}(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}) \\ \qquad\qquad\qquad\qquad +\sigma_i^2 L_{\phi_i}^2\triangle\boldsymbol{u}_i^{k+1},\ i=1,2,\cdots,l, \\ \partial_{\boldsymbol{v}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) = \nabla g(\boldsymbol{v}^{k+1}) - \boldsymbol{w}^{k+1} + 2\kappa_1\triangle\boldsymbol{v}^{k+1} + \beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}), \\ \partial_{\boldsymbol{w}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) = -(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}), \\ \partial_{\hat{\boldsymbol{u}}_i}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) = -\sigma_i^2 L_{\phi_i}^2\triangle\boldsymbol{u}_i^{k+1},\ i=1,2,\cdots,l, \\ \partial_{\hat{\boldsymbol{v}}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) = 2\kappa_2\triangle\boldsymbol{v}^k - 2\kappa_1\triangle\boldsymbol{v}^{k+1}, \quad \partial_{\tilde{\boldsymbol{v}}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) = -2\kappa_2\triangle\boldsymbol{v}^k. \end{cases}$$

From this together with (3.2), one has

$$\begin{cases} -\boldsymbol{P}_i^{\mathrm{T}}\triangle\boldsymbol{w}^{k+1} + \beta\boldsymbol{P}_i^{\mathrm{T}}\triangle\boldsymbol{v}^{k+1} + \beta\boldsymbol{P}_i^{\mathrm{T}}\Big(\boldsymbol{P}_{[i+1,l]}\triangle\boldsymbol{u}_{[i+1,l]}^{k+1}\Big) + \sigma_i^2 L_{\phi_i}^2\triangle\boldsymbol{u}_i^{k+1} \\ -\Big(\nabla\phi_i(\boldsymbol{u}_i^{k+1}) - \nabla\phi_i(\bar{\boldsymbol{u}}_i^k)\Big) \in \partial_{\boldsymbol{u}_i}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}),\ i=1,2,\cdots,l, \\ \eta\beta(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}) - \Big(\nabla\varphi(\boldsymbol{v}^{k+1}) - \nabla\varphi(\boldsymbol{v}^k)\Big) + \rho\triangle\boldsymbol{v}^k \\ +2\kappa_1\triangle\boldsymbol{v}^{k+1} \in \partial_{\boldsymbol{v}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}), \\ -(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{v}^{k+1} - \boldsymbol{r}) \in \partial_{\boldsymbol{w}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}), \\ -\sigma_i^2 L_{\phi_i}^2\triangle\boldsymbol{u}_i^{k+1} \in \partial_{\hat{\boldsymbol{u}}_i}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}),\ i=1,2,\cdots,l, \\ 2\kappa_2\triangle\boldsymbol{v}^k - 2\kappa_1\triangle\boldsymbol{v}^{k+1} \in \partial_{\hat{\boldsymbol{v}}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}), \quad -2\kappa_2\triangle\boldsymbol{v}^k \in \partial_{\tilde{\boldsymbol{v}}}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}), \end{cases} \tag{3.17}$$

which implies $\boldsymbol{\epsilon}^{k+1} \in \partial\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})$. From (3.17) and the Lipschitz continuities of $\nabla\phi_i$ $(i = 1, 2, \cdots, l)$, there is a positive χ_0 satisfying

$$\|\boldsymbol{\epsilon}^{k+1}\| \leq \chi_0\Big(\|\triangle \boldsymbol{u}_{[1,l]}^k\| + \|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{w}^{k+1}\| + \|\triangle \boldsymbol{v}^k\| + \|\triangle \boldsymbol{v}^{k+1}\|\Big). \tag{3.18}$$

From (3.7), there is a positive χ_1 satisfying

$$\|\triangle \boldsymbol{w}^{k+1}\| \leq \chi_1\Big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k-1}\| + \|\triangle \boldsymbol{v}^k\| + \|\triangle \boldsymbol{v}^{k+1}\|\Big). \tag{3.19}$$

For $k \geq 2$, combining (3.18), (3.19) and $\boldsymbol{\epsilon}^{k+1} \in \partial\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})$, one has

$$\begin{aligned}&d(0, \partial\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}))\\ \leq&\chi_0(1+\chi_1)\big(\|\triangle \boldsymbol{u}_{[1,l]}^k\| + \|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k-1}\| + \|\triangle \boldsymbol{v}^k\| + \|\triangle \boldsymbol{v}^{k+1}\|\big).\end{aligned}$$

So, setting $\chi = \chi_0(1+\chi_1)$, the lemma holds. □

Next, the following theorem proves the global convergence of the SCIB-ADMM.

Theorem 3.2. *If* $\{\boldsymbol{\varpi}^k\}$ *is bounded and Assumptions* 3.1 *and* 3.2 *hold. Suppose that* $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}})$ *has the KL property, then* $\sum\limits_{k=0}^{+\infty}\|\boldsymbol{\varpi}^{k+1} - \boldsymbol{\varpi}^k\| < +\infty$.

Proof. According to Theorem 3.1 (iii), one has

$$\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) = \lim_{k\to+\infty}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k),\ \forall\ \hat{\boldsymbol{\varpi}}^* \in \hat{\Theta}.$$

The proof will be completed by considering two cases.

(i) If $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k_0}) = \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)$ for an integer k_0, then by Lemma 3.1, one has

$$\delta\big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\|^2 + \|\triangle \boldsymbol{v}^{k+1}\|^2\big) \leq \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k_0}) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) = 0,\ \forall\ k > k_0.$$

Thus, $\boldsymbol{u}_i^{k+1} = \boldsymbol{u}_i^k$ and $\boldsymbol{v}^{k+1} = \boldsymbol{v}^k$, $\forall\ k > k_0$. Associated with (3.7), it follows that $\boldsymbol{w}^{k+1} = \boldsymbol{w}^k$, $\forall\ k > k_0$. Then $\boldsymbol{\varpi}^{k+1} = \boldsymbol{\varpi}^k$, and $\sum\limits_{k=0}^{+\infty}\|\boldsymbol{\varpi}^{k+1} - \boldsymbol{\varpi}^k\| < +\infty$.

(ii) If $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) > \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)$. From Theorem 3.1 (i), given $\varepsilon > 0$, there is k_1 satisfying $d(\hat{\boldsymbol{\varpi}}^k, \hat{\Theta}) < \varepsilon$, $\forall\ k > k_1 > 0$ holds. In addition, considering that $\lim\limits_{k\to\infty}\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) = \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)$, given $e > 0$, there is k_2 satisfying $\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) + e > \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)$, $\forall\ k > k_2 > 0$ holds. Thus, given $\varepsilon, e > 0$, one has

$$d(\hat{\boldsymbol{\varpi}}^k, \hat{\Theta}) < \varepsilon, \quad \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) < \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) < \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*) + e, \quad \forall k > \tilde{k} = \max\{k_1, k_2\}.$$

From Lemma 2.4, one has $\xi'\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\big)d(0, \partial\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k)) \geq 1,\ \forall\ k > \tilde{k}$.

Considering that ξ is a concave function, it holds that

$$\begin{aligned}&\xi'\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\big)\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1})\big)\\ \leq&\xi\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\big) - \xi\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\big).\end{aligned}$$

This, together with $\xi'(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)) > 0$ and Lemma 3.3, one has

$$\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^k) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{k+1}) \leq \chi\Lambda_k\triangle_{k,k+1}, \tag{3.20}$$

where $\Lambda_k = \|\triangle \boldsymbol{u}_{[1,l]}^k\| + \|\triangle \boldsymbol{u}_{[1,l]}^{k-1}\| + \|\triangle \boldsymbol{v}^k\| + \|\triangle \boldsymbol{v}^{k-1}\| + \|\triangle \boldsymbol{v}^{k-2}\|$ and $\triangle_{p,q} = \xi\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^p) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\big) - \xi\big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^q) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\big)$. In view of (3.20) and Lemma 3.1, one has

$$\delta\big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\|^2 + \|\triangle \boldsymbol{v}^{k+1}\|^2\big) \le \chi \Lambda_k \triangle_{k,k+1}.$$

Then,

$$4\big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\big) \le 2\Lambda_k^{\frac{1}{2}} \times \left(\sqrt{\frac{8\chi}{\delta}} \triangle_{k,k+1}^{\frac{1}{2}}\right). \tag{3.21}$$

Summing up (3.21) and using $2ab \le a^2 + b^2$ for all $a, b \ge 0$, one has

$$4 \sum_{k=\tilde{k}+1}^{m} \big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\big) \le \sum_{k=\tilde{k}+1}^{m} \left(\Lambda_k + \|\triangle \boldsymbol{u}_{[1,l]}^{k-2}\| + \frac{8\chi}{\delta} \triangle_{k,k+1}\right).$$

Noting $\xi(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{\tilde{k}+1}) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)) \ge 0$, letting $m \to +\infty$ yield

$$\begin{aligned}
&\sum_{k=\tilde{k}+1}^{m} \big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\big) \\
\le& 3\|\triangle \boldsymbol{u}_{[1,l]}^{\tilde{k}+1}\| + 2\|\triangle \boldsymbol{u}_{[1,l]}^{\tilde{k}}\| + \|\triangle \boldsymbol{u}_{[1,l]}^{\tilde{k}-1}\| + 3\|\triangle \boldsymbol{v}^{\tilde{k}+1}\| \\
&+ 2\|\triangle \boldsymbol{v}^{\tilde{k}}\| + \|\triangle \boldsymbol{v}^{\tilde{k}-1}\| + \frac{8\chi}{\delta} \xi\Big(\hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^{\tilde{k}+1}) - \hat{\mathcal{L}}_\beta(\hat{\boldsymbol{\varpi}}^*)\Big).
\end{aligned}$$

Thus, $\sum\limits_{k=0}^{+\infty} \big(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\big) < +\infty$.

According to (3.7), one has $\sum\limits_{k=0}^{+\infty} \|\triangle \boldsymbol{w}^{k+1}\| < +\infty$.

Moreover, from $\boldsymbol{\varpi}^k = (\boldsymbol{u}_{[1,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k)$, one has

$$\|\boldsymbol{\varpi}^{k+1} - \boldsymbol{\varpi}^k\| \le \|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\| + \|\triangle \boldsymbol{w}^{k+1}\|.$$

Thus, $\sum\limits_{k=0}^{+\infty} \|\boldsymbol{\varpi}^{k+1} - \boldsymbol{\varpi}^k\| < +\infty$. Finally, it follows from Theorem 3.1 that $\{\boldsymbol{\varpi}^k\}$ converges to a critical point of $\mathcal{L}_\beta(\cdot)$. □

3.3. Convergence analysis when $\boldsymbol{Q} \ne \boldsymbol{I}$ and $\eta = 1$

In this section, $\{\boldsymbol{o}^k = (\boldsymbol{u}_{[1,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k)\}$ is generated by the SCIB-ADMM. Denote

$$\hat{\boldsymbol{o}}^k = (\boldsymbol{u}_{[1,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k, \boldsymbol{u}_{[1,l]}^{k-1}, \boldsymbol{v}^{k-1}, \boldsymbol{v}^{k-2}), \qquad \hat{\boldsymbol{o}} = (\boldsymbol{u}_{[1,l]}, \boldsymbol{v}, \boldsymbol{w}, \hat{\boldsymbol{u}}_{[1,l]}, \hat{\boldsymbol{v}}, \tilde{\boldsymbol{v}}).$$

According to the $\boldsymbol{u}_i$-subproblems (3.1b)-(3.1d) and the $\boldsymbol{v}$-subproblem (3.1f), one has

$$\begin{cases}
0 \in \partial f_i(\boldsymbol{u}_i^{k+1}) - \boldsymbol{P}_i^{\mathrm{T}} \boldsymbol{w}^k + \beta \boldsymbol{P}_i^{\mathrm{T}} \Big(\boldsymbol{P}_{[1,i]} \boldsymbol{u}_{[1,i]}^{k+1} + \boldsymbol{P}_{[i+1,l]} \boldsymbol{u}_{[i+1,l]}^k + \boldsymbol{Q}\boldsymbol{v}^k - \boldsymbol{r}\Big) \\
\qquad + \nabla \phi_i(\boldsymbol{u}_i^{k+1}) - \nabla \phi_i(\bar{\boldsymbol{u}}_i^k), \ i = 1, 2, \cdots, l, & (3.22\text{a}) \\
0 = \nabla g(\boldsymbol{v}^{k+1}) - \boldsymbol{Q}^{\mathrm{T}} \boldsymbol{w}^{k+1} + \rho(\boldsymbol{v}^{k-1} - \boldsymbol{v}^k) + \nabla \varphi(\boldsymbol{v}^{k+1}) - \nabla \varphi(\boldsymbol{v}^k). & (3.22\text{b})
\end{cases}$$

The following assumptions are provided to analysis the convergence of the SCIB-ADMM.

Assumption 3.3. *(i)* $\boldsymbol{H} = \left(\frac{\beta}{2} - \frac{\tau\beta}{\tau+1}\right)\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} - \left(\frac{5L_g^2+10L_\varphi^2+10\rho^2}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}} + \frac{L_g+L_\varphi^2+2+\rho^2}{2}\right)\boldsymbol{I}_p > \boldsymbol{0}$, $\ell_i > \sigma_i L_{\phi_i}^2 + 1$ *and* $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} > \boldsymbol{0}$*.* $\boldsymbol{I}_p$ *is an identity matrix and* $\mu_{\boldsymbol{Q}^{\mathrm{T}}}$ *denotes the smallest positive eigenvalue of* $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}$*;*

(ii) $\tau \in (-1, 1)$ *is a relaxation factor;*

(iii) $\mathrm{Im}\boldsymbol{Q} \supset \mathrm{Im}\boldsymbol{P} \cup \{\boldsymbol{r}\}$.

Remark 3.3. If $\beta > \frac{(L_g+L_\varphi^2+2+\rho^2)(\tau+1)+\sqrt{(L_g+L_\varphi^2+2+\rho^2)^2(\tau+1)^2+(1-\tau)(40L_g^2+80L_\varphi^2+80\rho^2)}}{2(1-\tau)\mu_{\boldsymbol{Q}^{\mathrm{T}}}}$, it follows from Assumption 3.3 (i) that $\boldsymbol{H} > \boldsymbol{0}$.

Lemma 3.4. *If Assumptions* 3.1 *and* 3.3 *hold, then*

$$\|\triangle \boldsymbol{w}^{k+1}\|^2 \leq \frac{5(L_g^2 + L_\varphi^2)}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k+1}\|^2 + \frac{5(\rho^2 + L_\varphi^2)}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k}\|^2 + \frac{5\rho^2}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k-1}\|^2.$$

Proof. From (3.22b), one has

$$\begin{aligned}&\|\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{w}^{k+1} - \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{w}^{k}\|^2\\ \leq & 5\Big((L_g^2 + L_\varphi^2)\|\triangle \boldsymbol{v}^{k+1}\|^2 + (L_\varphi^2 + \rho^2)\|\triangle \boldsymbol{v}^{k}\|^2 + \rho^2\|\triangle \boldsymbol{v}^{k-1}\|^2\Big).\end{aligned} \tag{3.23}$$

It follows from Lemma 2.1, one has $\|\triangle \boldsymbol{w}^{k+1}\| \leq \frac{1}{\sqrt{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}}\|\boldsymbol{Q}^{\mathrm{T}}\triangle \boldsymbol{w}^{k+1}\|$.

This, together (3.23), yields

$$\begin{aligned}\|\triangle \boldsymbol{w}^{k+1}\|^2 &\leq \frac{1}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\boldsymbol{Q}^{\mathrm{T}}\triangle \boldsymbol{w}^{k+1}\|^2\\ &\leq \frac{5(L_g^2 + L_\varphi^2)}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k+1}\|^2 + \frac{5(\rho^2 + L_\varphi^2)}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k}\|^2 + \frac{5\rho^2}{\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k-1}\|^2.\end{aligned}$$

Thus, the lemma is proved. □

An auxiliary function is given, namely

$$\Gamma_\beta(\hat{\boldsymbol{o}}^k) = \mathcal{L}_\beta(\boldsymbol{o}^k) + \sum_{i=1}^{l}\frac{\sigma_i^2 L_{\phi_i}^2}{2}\|\triangle \boldsymbol{u}_i^k\|^2 + \varsigma_1\|\triangle \boldsymbol{v}^k\|^2 + \varsigma_2\|\triangle \boldsymbol{v}^{k-1}\|^2, \tag{3.24}$$

where $\varsigma_1 = \frac{10\rho^2+5L_\varphi^2}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}} + \frac{\rho^2}{2}$ and $\varsigma_2 = \frac{5\rho^2}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}}$.

The following lemma establishes that the auxiliary function $\{\Gamma_\beta(\hat{\boldsymbol{o}}^k)\}$ is monotonically nonincreasing.

Lemma 3.5. *If Assumptions* 3.1 *and* 3.3 *hold. For any* $k \geq 1$*, it holds that*

$$\Gamma_\beta(\hat{\boldsymbol{o}}^{k+1}) - \Gamma_\beta(\hat{\boldsymbol{o}}^k) \leq -\|\triangle \boldsymbol{v}^{k+1}\|_{\boldsymbol{H}}^2 - \sum_{i=1}^{l}\frac{\ell_i - 1 - \sigma_i^2 L_{\phi_i}^2}{2}\|\triangle \boldsymbol{u}_i^{k+1}\|^2.$$

Furthermore, the function referenced by (3.24) *is monotonically nonincreasing.*

Proof. From (1.2) and (3.22b), one has

$$\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^{k+1}, \boldsymbol{w}^{k+\frac{1}{2}}) - \mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^{k}, \boldsymbol{w}^{k+\frac{1}{2}})$$

$$\leq \frac{L_g + L_\varphi^2 + 2}{2}\|\triangle \boldsymbol{v}^{k+1}\|^2 + \frac{\rho^2}{2}\|\triangle \boldsymbol{v}^k\|^2 - \frac{\beta}{2}\|\boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}\|^2, \tag{3.25}$$

where the inequalities above utilize Lemma 2.2, Assumptions 3.1 (i) and (iv). From (3.1e) and (3.1g), it holds that

$$\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^{k+1} - \boldsymbol{r} = -\frac{1}{(\tau+1)\beta}\triangle \boldsymbol{w}^{k+1} + \frac{\tau}{\tau+1}\boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}, \tag{3.26}$$

and

$$\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^{k} - \boldsymbol{r} = -\frac{1}{(\tau+1)\beta}\triangle \boldsymbol{w}^{k+1} - \frac{1}{\tau+1}\boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}. \tag{3.27}$$

Furthermore, according to (3.1e), (3.1g), (3.26), (3.27) and Lemma 3.4, one has

$$\begin{aligned}
&\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^k, \boldsymbol{w}^{k+\frac{1}{2}}) - \mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^k, \boldsymbol{w}^k) + \mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^{k+1}, \boldsymbol{w}^{k+1}) \\
&- \mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^{k+1}, \boldsymbol{w}^{k+\frac{1}{2}}) \\
=&(\boldsymbol{w}^k - \boldsymbol{w}^{k+\frac{1}{2}})^{\mathrm{T}}(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^k - \boldsymbol{r}) - (\boldsymbol{w}^{k+1} - \boldsymbol{w}^{k+\frac{1}{2}})^{\mathrm{T}}(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^{k+1} - \boldsymbol{r}) \\
=&\frac{\tau\beta}{\tau+1}\|\boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}\|^2 + \frac{1}{(\tau+1)\beta}\|\triangle \boldsymbol{w}^{k+1}\|^2 \\
\leq&\frac{5(L_g^2 + L_\varphi^2)}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k+1}\|^2 + \frac{5(\rho^2 + L_\varphi^2)}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k}\|^2 + \frac{5\rho^2}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k-1}\|^2 \\
&+ \frac{\tau\beta}{\tau+1}\|\boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}\|^2.
\end{aligned} \tag{3.28}$$

From the $\boldsymbol{u}_i$-subproblems (3.1a)-(3.1d) and Lemma 2.3, one has

$$\begin{aligned}
&\mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k+1}, \boldsymbol{v}^k, \boldsymbol{w}^k) - \mathcal{L}_\beta(\boldsymbol{u}_{[1,l]}^{k}, \boldsymbol{v}^k, \boldsymbol{w}^k) \\
&\leq -\sum_{i=1}^{l}\Big(\frac{\ell_i - 1}{2}\|\triangle \boldsymbol{u}_i^{k+1}\|^2 - \frac{\sigma_i^2 L_{\phi_i}^2}{2}\|\triangle \boldsymbol{u}_i^{k}\|^2\Big).
\end{aligned} \tag{3.29}$$

Summing up (3.25), (3.28) and (3.29), one has

$$\begin{aligned}
&\mathcal{L}_\beta(\boldsymbol{o}^{k+1}) - \mathcal{L}_\beta(\boldsymbol{o}^k) \\
\leq& -\Big(\frac{\beta}{2} - \frac{\tau\beta}{\tau+1}\Big)\|\boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}\|^2 + \Big(\frac{L_g + L_\varphi^2 + 2}{2} + \frac{5(L_g^2 + L_\varphi^2)}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\Big)\|\triangle \boldsymbol{v}^{k+1}\|^2 \\
&+ \Big(\frac{5(\rho^2 + L_\varphi^2)}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}} + \frac{\rho^2}{2}\Big)\|\triangle \boldsymbol{v}^{k}\|^2 + \frac{5\rho^2}{(\tau+1)\beta\mu_{\boldsymbol{Q}^{\mathrm{T}}}}\|\triangle \boldsymbol{v}^{k-1}\|^2 \\
&- \sum_{i=1}^{l}\Big(\frac{\ell_i - 1}{2}\|\triangle \boldsymbol{u}_i^{k+1}\|^2 - \frac{\sigma_i^2 L_{\phi_i}^2}{2}\|\triangle \boldsymbol{u}_i^{k}\|^2\Big).
\end{aligned}$$

Rearranging the above relations, one has

$$\Gamma_\beta(\hat{\boldsymbol{o}}^{k+1}) - \Gamma_\beta(\hat{\boldsymbol{o}}^k) \leq -\|\triangle \boldsymbol{v}^{k+1}\|_{\boldsymbol{H}}^2 - \sum_{i=1}^{l}\frac{\ell_i - 1 - \sigma_i^2 L_{\phi_i}^2}{2}\|\triangle \boldsymbol{u}_i^{k+1}\|^2.$$

From $\boldsymbol{H} > 0$ and $\ell_i > \sigma_i L_{\phi_i}^2 + 1$, it holds that $\{\Gamma_\beta(\hat{\boldsymbol{o}}^k)\}$ is monotonically nonincreasing. □

Lemma 3.6. *If Assumptions* 3.1 *and* 3.3 *hold. For all* $k \geq 2$, $i = 1, 2, \cdots, l$, *define*

$$\begin{cases} \boldsymbol{\epsilon}_i^{k+1} = \beta \boldsymbol{P}_i^{\mathrm{T}}\Big(\boldsymbol{Q}_{[i+1,l]}\triangle \boldsymbol{u}_{[i+1,l]}^{k+1} + \boldsymbol{Q}\triangle \boldsymbol{v}^{k+1}\Big) - \boldsymbol{P}_i^{\mathrm{T}}\triangle \boldsymbol{w}^{k+1} \\ \qquad\quad + \sigma_i^2 L_{\phi_i}^2 \triangle \boldsymbol{u}_i^{k+1} - \Big(\nabla\phi_i(\boldsymbol{u}_i^{k+1}) - \nabla\phi_i(\bar{\boldsymbol{u}}_i^k)\Big), \\ \boldsymbol{\epsilon}_{l+1}^{k+1} = \beta \boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^{k+1} - \boldsymbol{r}) + \rho\triangle \boldsymbol{v}^k + 2\kappa_1\triangle \boldsymbol{v}^{k+1} \\ \qquad\quad - \Big(\nabla\varphi(\boldsymbol{v}^{k+1}) - \nabla\varphi(\boldsymbol{v}^k)\Big), \\ \boldsymbol{\epsilon}_{l+2}^{k+1} = -(\boldsymbol{P}_{[1,l]}\boldsymbol{u}_{[1,l]}^{k+1} + \boldsymbol{Q}\boldsymbol{v}^{k+1} - \boldsymbol{r}), \\ \boldsymbol{\epsilon}_{l+2+i}^{k+1} = -\sigma_i^2 L_{\phi_i}^2 \triangle \boldsymbol{u}_i^{k+1}, \\ \boldsymbol{\epsilon}_{2l+3}^{k+1} = -2\varsigma_1\triangle \boldsymbol{v}^{k+1} + 2\varsigma_2\triangle \boldsymbol{v}^k, \\ \boldsymbol{\epsilon}_{2l+4}^{k+1} = -2\varsigma_2\triangle \boldsymbol{v}^k. \end{cases}$$

Thus, $\big(\boldsymbol{\epsilon}_{[1,l]}^{k+1}, \boldsymbol{\epsilon}_{l+1}^{k+1}, \boldsymbol{\epsilon}_{l+2}^{k+1}, \boldsymbol{\epsilon}_{[l+3,2l+2]}^{k+1}, \boldsymbol{\epsilon}_{2l+3}^{k+1}, \boldsymbol{\epsilon}_{2l+4}^{k+1}\big) \in \partial\Gamma_\beta(\hat{\boldsymbol{o}}^{k+1})$. *Then*

$$\begin{aligned} & d(0, \partial\Gamma_\beta(\hat{\boldsymbol{o}}^{k+1})) \\ \leq & \hat{\chi}\Big(\|\triangle \boldsymbol{u}_{[1,l]}^k\| + \|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k-1}\| + \|\triangle \boldsymbol{v}^k\| + \|\triangle \boldsymbol{v}^{k+1}\|\Big), \quad \hat{\chi} > 0. \end{aligned} \tag{3.30}$$

Proof. Utilizing (3.22), (3.24) and a similar method outlined in Lemma 3.3, the proof of (3.30) can be established. □

Under Assumptions 3.1 and 3.3, from Lemma 3.5, the convergence of the SCIB-ADMM can be established by the similar method outlined in Theorem 3.2.

Theorem 3.3. *If* $\{\boldsymbol{o}^k\}$ *is bounded and Assumptions* 3.1 *and* 3.3 *hold. The cluster point set of the sequence* $\{\hat{\boldsymbol{o}}^k\}$ *is* $\hat{\Upsilon}$. *Suppose that* $\Gamma_\beta(\hat{\boldsymbol{o}})$ *satisfies the KL property, then* $\sum\limits_{k=0}^{+\infty}\|\boldsymbol{o}^{k+1} - \boldsymbol{o}^k\| < +\infty$.

Proof. From the similar approach in Theorem 3.2, there exists $\varepsilon, e > 0$, when $k > \overline{k}$, one has $d(\hat{\boldsymbol{o}}^k, \hat{\Upsilon}) < \varepsilon$, $\Gamma_\beta(\hat{\boldsymbol{o}}^*) < \Gamma_\beta(\hat{\boldsymbol{o}}^k) < \Gamma_\beta(\hat{\boldsymbol{o}}^*) + e$. Consequently, it follows from Lemma 2.4 that $\xi'\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big)d\big(0, \partial\Gamma_\beta(\hat{\boldsymbol{o}}^k)\big) \geq 1$, $\forall\ k > \overline{k}$. From the concavity of ξ, one has

$$\begin{aligned} & \xi'\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big)\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^{k+1})\Big) \\ \leq & \xi\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big) - \xi\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^{k+1}) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big). \end{aligned}$$

Combining this with $\xi'(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)) > 0$ and Lemma 3.6, one has

$$\begin{aligned} & \Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^{k+1}) \\ \leq & \frac{\xi\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big) - \xi\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^{k+1}) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big)}{\xi'\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^k) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big)} \\ \leq & \hat{\chi}\Lambda_k \triangle_{k,k+1}, \end{aligned} \tag{3.31}$$

where $\Lambda_k = \|\triangle \boldsymbol{u}_{[1,l]}^k\| + \|\triangle \boldsymbol{u}_{[1,l]}^{k-1}\| + \|\triangle \boldsymbol{v}^k\| + \|\triangle \boldsymbol{v}^{k-1}\| + \|\triangle \boldsymbol{v}^{k-2}\|$ and $\triangle_{p,q} = \xi\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^p) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big) - \xi\Big(\Gamma_\beta(\hat{\boldsymbol{o}}^q) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\Big)$.

Let $\gamma = \max\left\{\lambda_{\min}(H), \frac{\ell_i-1}{2} - \frac{\sigma_i^2 L_{\phi_i}^2}{2}\right\}$, $i = 1, 2, \cdots, l$. Combining Lemma 3.5 and (3.31), one has $\gamma\left(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\|^2 + \|\triangle \boldsymbol{v}^{k+1}\|^2\right) \le \hat{\chi}\Lambda_k \triangle_{k,k+1}$. Then,

$$4\left(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\right) \le 2\Lambda_k^{\frac{1}{2}} \times \left(\sqrt{\frac{8\hat{\chi}}{\gamma}} \triangle_{k,k+1}^{\frac{1}{2}}\right). \tag{3.32}$$

Summing up (3.32) and using $2ab \le a^2 + b^2$ for all $a, b \ge 0$, letting $m \to +\infty$ yield

$$\sum_{k=\bar{k}+1}^{m} \left(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\right) \le 3\|\triangle \boldsymbol{u}_{[1,l]}^{\bar{k}+1}\| + 2\|\triangle \boldsymbol{u}_{[1,l]}^{\bar{k}}\| + \|\triangle \boldsymbol{u}_{[1,l]}^{\bar{k}-1}\| + 3\|\triangle \boldsymbol{v}^{\bar{k}+1}\|$$
$$+ 2\|\triangle \boldsymbol{v}^{\bar{k}}\| + \|\triangle \boldsymbol{v}^{\bar{k}-1}\| + \frac{8\hat{\chi}}{\gamma}\xi\big(\Gamma_\beta(\hat{\boldsymbol{o}}^{\bar{k}+1}) - \Gamma_\beta(\hat{\boldsymbol{o}}^*)\big).$$

Thus, $\sum_{k=0}^{+\infty} \left(\|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\|\right) < +\infty$. According to Lemma 3.4, one has

$$\sum_{k=0}^{+\infty} \|\triangle \boldsymbol{w}^{k+1}\| < +\infty.$$

Moreover, from $\boldsymbol{o}^k = (\boldsymbol{u}_{[1,l]}^k, \boldsymbol{v}^k, \boldsymbol{w}^k)$, one has

$$\|\boldsymbol{o}^{k+1} - \boldsymbol{o}^k\| \le \|\triangle \boldsymbol{u}_{[1,l]}^{k+1}\| + \|\triangle \boldsymbol{v}^{k+1}\| + \|\triangle \boldsymbol{w}^{k+1}\|.$$

Therefore, $\{\boldsymbol{o}^k\}$ is convergent and $\sum_{k=0}^{+\infty} \|\boldsymbol{o}^{k+1} - \boldsymbol{o}^k\| < +\infty$. □

4. Numerical experiment

To evaluate the numerical efficacy of the SCIB-ADMM, one applies it to the SCAD and robust PCA nonconvex optimization, and compares the SCIB-ADMM with the BADMM [36], the PSRADMM [24] and the IPPS-ADMM [39]. The experiments are executed with the software MATLAB R2021a, Intel(R) Core(TM) i5-8250 CPU@1.60GHz and 8.00GB of RAM.

4.1. Effects of the inertial parameters

The SCIB-ADMM is applied to the SCAD model, which arises in statistical learning [43]. The SCAD model is formulated as

$$\min_{\boldsymbol{u},\boldsymbol{v}} \sum_{i=1}^{p} g_\kappa(|u_i|) + \frac{1}{2}\|\boldsymbol{v}\|^2 \quad \text{s.t. } \boldsymbol{P}\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{r}, \tag{4.1}$$

where $\boldsymbol{P} \in \mathbb{R}^{m\times p}$, $\boldsymbol{u} \in \mathbb{R}^p$, $\boldsymbol{v} \in \mathbb{R}^m$, $\boldsymbol{r} \in \mathbb{R}^m$, and $g_\kappa(\cdot)$ can be generated as

$$g_\kappa(\nu) = \begin{cases} \dfrac{(c+1)\kappa^2}{2}, & \nu > c\kappa, \\[2mm] \dfrac{-\nu^2 + 2c\kappa\nu - \kappa^2}{2(c+1)}, & \kappa < \nu \le c\kappa, \\[2mm] \kappa\nu, & \nu \le \kappa, \end{cases}$$

where $\kappa > 0$ and $c > 2$. Note that the SCAD model (4.1) is a particular instance of (1.1) where $f(\boldsymbol{u}) = \sum_{i=1}^{p} g_\kappa(|u_i|),\ g(\boldsymbol{v}) = \frac{1}{2}\|\boldsymbol{v}\|^2$. The solution to the problem mentioned below

$$\min_{\boldsymbol{u}} \sum_{i=1}^{p} g_\kappa(|u_i|) + \frac{1}{2v}\|\boldsymbol{u} - \boldsymbol{s}\|^2, \tag{4.2}$$

where $\boldsymbol{s} \in \mathbb{R}^p$ and $1 + v \leq c$, can be generated from [41]

$$u_i = \begin{cases} s_i, & |s_i| > c\kappa, \\ \dfrac{(c-1)s_i - \operatorname{sign}(s_i)c\kappa v}{c-1-v}, & (1+v)\kappa < |s_i| \leq c\kappa, \\ \operatorname{sign}(s_i)(|s_i| - \kappa v)_+, & |s_i| \leq (1+v)\kappa. \end{cases}$$

Let $\triangle_\phi(\boldsymbol{u}, \bar{\boldsymbol{u}}^k) = \frac{1}{2}\|\boldsymbol{u} - \bar{\boldsymbol{u}}^k\|^2_{\mu_1 \boldsymbol{I} - \beta \boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}}$ and $\triangle_\varphi(\boldsymbol{v}, \bar{\boldsymbol{v}}^k) = \frac{\mu_2}{2}\|\boldsymbol{v} - \bar{\boldsymbol{v}}^k\|^2$. Applying the SCIB-ADMM to the SCAD model (4.1) yields

$$\begin{cases} \bar{\boldsymbol{u}}^k = \boldsymbol{u}^k + \sigma(\boldsymbol{u}^k - \boldsymbol{u}^{k-1}),\ \sigma \in [0,1), \\ \boldsymbol{u}^{k+1} = \arg\min_{\boldsymbol{u}} \left\{ \sum_{i=1}^{p} g_\kappa(|u_i|) + \dfrac{\mu_1}{2}\left\| \boldsymbol{u} - \dfrac{\beta \boldsymbol{P}^{\mathrm{T}}(\boldsymbol{v}^k + \boldsymbol{b} + \frac{\boldsymbol{w}^k}{\beta} - \boldsymbol{P}\bar{\boldsymbol{u}}^k) + \mu_1 \bar{\boldsymbol{u}}^k}{\mu_1} \right\|^2 \right\}, \\ \boldsymbol{w}^{k+\frac{1}{2}} = \boldsymbol{w}^k - \tau\beta(\boldsymbol{P}\boldsymbol{u}^{k+1} - \boldsymbol{v}^k - \boldsymbol{r}), \\ \boldsymbol{v}^{k+1} = \dfrac{\beta(\boldsymbol{P}\boldsymbol{u}^{k+1} - \boldsymbol{r}) - \boldsymbol{w}^{k+\frac{1}{2}} + \rho(\boldsymbol{v}^k - \boldsymbol{v}^{k-1}) + \mu_2 \boldsymbol{v}^k}{1 + \beta + \mu_2}, \\ \boldsymbol{w}^{k+1} = \boldsymbol{w}^{k+\frac{1}{2}} - \eta\beta(\boldsymbol{P}\boldsymbol{u}^{k+1} - \boldsymbol{v}^{k+1} - \boldsymbol{r}). \end{cases}$$

For the SCIB-ADMM, the effect of the parameter σ on its performance is tested. In particular, our experiments set σ to the values $\sigma = \{0,\ 0.3,\ 0.6,\ 0.9,\ 0.9999\}$. In our experiments, $\boldsymbol{b} = \boldsymbol{P}\boldsymbol{u}^* + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \frac{100}{p}\boldsymbol{I})$, $\boldsymbol{P}_{ij} \sim \mathcal{N}(0,1)$, and $\boldsymbol{u}^*$ with the density of $\frac{100}{p}$. For the remaining parameters in the SCIB-ADMM, let $\rho = 0.1,\ \mu_1 = 25,\ \mu_2 = 1,\ (\tau, \eta) = (0.2, 1.1),\ \beta = 6.52,\ \kappa = 0.1,\ \nu = 0.99,\ c = 3.7$.

For different σ, the variables $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ are initialized with zero-vectors. The termination condition is $\max\{\|\triangle \boldsymbol{u}^{k+1}\|,\ \|\triangle \boldsymbol{v}^{k+1}\|\} \leq 10^{-4}$.

The constraint error is defined as error $= \|\boldsymbol{P}\boldsymbol{u}^{k+1} - \boldsymbol{v}^{k+1} - \boldsymbol{r}\|$.

In Figure 1, the trends of $\log_{10}$(error) and objective values are displayed for various σ values. From Figure 1, it is evident that the SCIB-ADMM with $\sigma = 0.9$ outperforms the other cases.

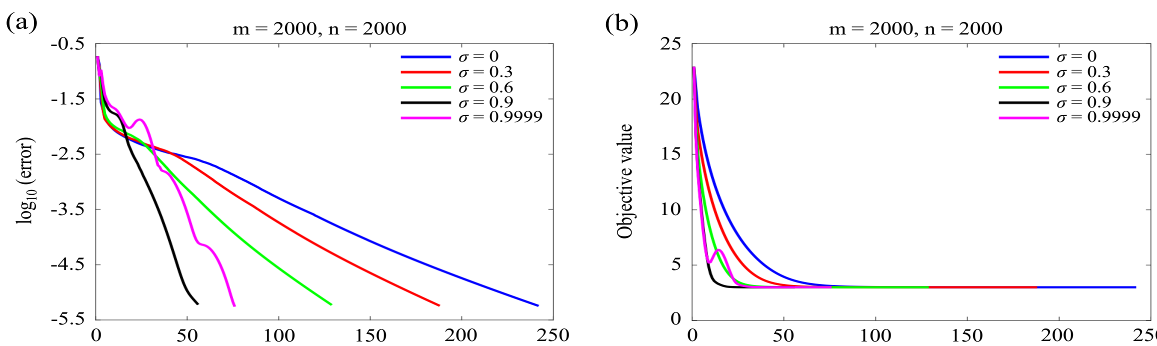

Figure 1. Effect of the parameter σ on SCAD. (a) Shows the constraint errors of different σ. (b) Shows the objective function values of different σ.

4.2. Numerical comparisons on SCAD

In this subsection, the numerical results on SCAD are presented for the BADMM [36], PSRADMM [24], IPPS-ADMM [39] and SCIB-ADMM. The parameter settings are consistent among the four methods, namely $\beta = 6.52,\ \kappa = 0.1,\ \nu = 0.99,\ c = 3.7$. In addition, for the BADMM, let $\triangle_{\varphi_1}(\boldsymbol{u}, \boldsymbol{u}^k) = \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}^k\|^2_{\varrho \boldsymbol{I} - \beta \boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}}$, $\varrho = 25$ and $\triangle_{\varphi_2}(\boldsymbol{v}, \boldsymbol{v}^k) = \frac{1}{2}\|\boldsymbol{v} - \boldsymbol{v}^k\|^2$. For the PSRADMM, let $\boldsymbol{F}_1 = \bar{\mu}\boldsymbol{I} - \beta \boldsymbol{P}^{\mathrm{T}}\boldsymbol{P},\ (\tau, \eta) = (0.85, 0.85)$ and $\bar{\mu} = 25$. For the IPPS-ADMM, let $\boldsymbol{F}_1 = \hat{\mu}\boldsymbol{I} - \beta \boldsymbol{P}^{\mathrm{T}}\boldsymbol{P},\ \hat{\mu} = 25,\ (\tau, \eta) = (0.2, 0.5)$ and $\alpha_1 = 0.4$. For the SCIB-ADMM, the values selected for the parameters are the same as in Section 4.1, where $\sigma = 0.9$.

The termination condition of the all methods is

$$\max\{\|\triangle \boldsymbol{u}^{k+1}\|,\ \|\triangle \boldsymbol{v}^{k+1}\|\} \leq 10^{-4}.$$

Table 1 displays the computational results of the four methods relative to Iter (number of iterations), Time (computation time in seconds), $\log_{10}$(error)(constraint error) and Obj. value (objective function values). The bold data indicates the best results among the four methods. It is evident that SCIB-ADMM requires fewer iterations and the shortest computing time compared to the other three methods.

Figure 2 depicts the trends of $\log_{10}$(error) and objective values of four methods on this experiment. It is clear that the SCIB-ADMM is the most effective method, due to its incorporation of the inertial strategy and the Bregman distance, both of which are proven to be effective on SCAD.

4.3. Numerical comparisons on robust PCA

The SCIB-ADMM is applied to the three-block robust PCA model, which arises in matrix factorization applications [42]

$$\min_{\boldsymbol{U}_1, \boldsymbol{U}_2, \boldsymbol{V}} \|\boldsymbol{U}_1\|_* + \gamma \|\boldsymbol{U}_2\|_1 + \frac{w}{2}\|\boldsymbol{V} - \boldsymbol{M}\|_F^2 \quad \text{s.t.}\, \boldsymbol{U}_1 + \boldsymbol{U}_2 - \boldsymbol{V} = \boldsymbol{0}, \tag{4.3}$$

where $\boldsymbol{U}_1,\ \boldsymbol{U}_2,\ \boldsymbol{V} \in \mathbb{R}^{m\times q}$ are the decision variables, $\|\boldsymbol{U}_1\|_* = \sum_{i=1}^{\min(m,q)} |\sigma_i(\boldsymbol{U}_1)|^{1/2}$ is the low-rank term, $\|\boldsymbol{U}_2\|_1 = \sum_{i=1}^{m}\sum_{j=1}^{q} |U_{ij}|$ is the sparse term, γ is the trade-off parameter, w is the penalty factor related to the noise level. Note that the robust PCA model (4.3) is a particular instance of (1.1) with $f_1(\boldsymbol{U}_1) = \|\boldsymbol{U}_1\|_*,\ f_2(\boldsymbol{U}_2) = \gamma\|\boldsymbol{U}_2\|_1,\ g(\boldsymbol{V}) = \frac{w}{2}\|\boldsymbol{V} - \boldsymbol{M}\|_F^2$, and $\boldsymbol{P}_1 = \boldsymbol{P}_2 = \boldsymbol{I},\ \boldsymbol{r} = \boldsymbol{0}$. The augmented Lagrangian function of problem (4.3) is defined as follows

$$\begin{aligned}&\mathcal{L}_\beta(\boldsymbol{U}_1, \boldsymbol{U}_2, \boldsymbol{V}, \boldsymbol{W})\\ =&\|\boldsymbol{U}_1\|_* + \gamma\|\boldsymbol{U}_2\|_1 + \frac{w}{2}\|\boldsymbol{V} - \boldsymbol{M}\|_F^2 - \langle \boldsymbol{W}, \boldsymbol{U}_1 + \boldsymbol{U}_2 - \boldsymbol{V}\rangle + \frac{\beta}{2}\|\boldsymbol{U}_1 + \boldsymbol{U}_2 - \boldsymbol{V}\|_F^2,\end{aligned}$$

where $\boldsymbol{W} \in \mathbb{R}^{m\times q}$ and $\beta > 0$ are the Lagrangian multiplier and penalty factor, respectively.

Table 1. Numerical results on SCAD.

Method	m	n	Iter	Time	$\log_{10}$(error)	Obj. value
BADMM	500	1000	563	7.3757	−4.9740	2.1281
	1000	2000	368	12.3948	−4.9341	2.4507
	2000	2000	253	15.1416	−4.8875	2.9956
	2000	3000	296	22.9452	−4.8906	2.9575
	3000	3000	245	35.7363	−4.8731	3.4553
PSRADMM	500	1000	550	6.9508	−5.3294	2.1281
	1000	2000	356	12.6379	−5.2877	2.4507
	2000	2000	242	15.3777	−5.2351	2.9956
	2000	3000	283	23.3602	−5.2405	2.9575
	3000	3000	234	32.8352	−5.2215	3.4553
IPPS-ADMM	500	1000	425	5.1680	−5.3291	2.1281
	1000	2000	275	9.4314	−5.2778	2.4507
	2000	2000	191	11.5668	−5.2300	2.9956
	2000	3000	223	18.2581	−5.2359	2.9575
	3000	3000	185	30.1296	−5.2203	3.4553
SCIB-ADMM	500	1000	**119**	**1.3969**	−5.3095	2.1281
	1000	2000	**74**	**2.8303**	−5.2188	2.4507
	2000	2000	**56**	**3.4377**	−5.2228	2.9956
	2000	3000	**62**	**5.0446**	−5.2360	2.9575
	3000	3000	**51**	**6.9824**	−5.2205	3.4553

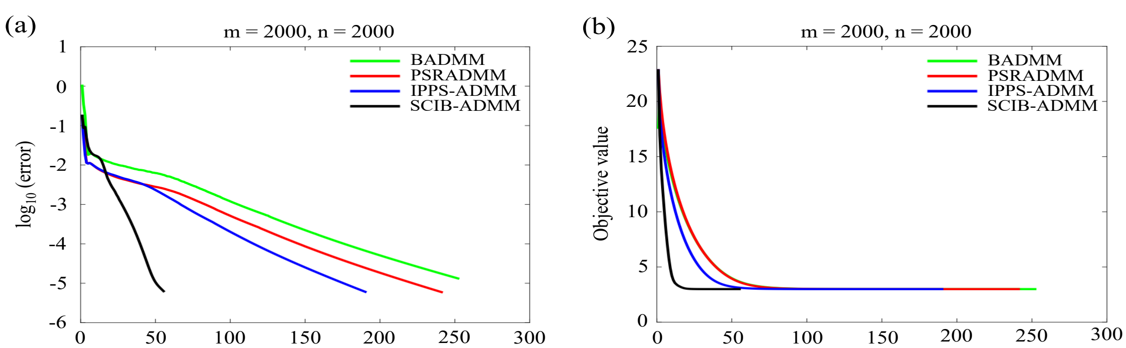

Figure 2. Comparison of the four methods on SCAD. (a) Shows the constraint errors of the four methods. (b) Shows the objective function values of the four methods.

Applying the SCIB-ADMM to the problem (4.3) yields

$$\begin{cases} \bar{\boldsymbol{U}}_1^k = \boldsymbol{U}_1^k + \sigma_1(\boldsymbol{U}_1^k - \boldsymbol{U}_1^{k-1}), \qquad \bar{\boldsymbol{U}}_2^k = \boldsymbol{U}_2^k + \sigma_2(\boldsymbol{U}_2^k - \boldsymbol{U}_2^{k-1}), \\ \boldsymbol{U}_1^{k+1} = \mathcal{H}\Big(\dfrac{\beta(\boldsymbol{V}^k - \boldsymbol{U}_2^k) + \bar{\boldsymbol{U}}_1^k + \boldsymbol{W}^k}{\beta+1}, \dfrac{1}{\beta+1}\Big), \\ \boldsymbol{U}_2^{k+1} = \mathcal{S}\Big(\dfrac{\beta(\boldsymbol{V}^k - \boldsymbol{U}_1^{k+1}) + \bar{\boldsymbol{U}}_2^k + \boldsymbol{W}^k}{\beta+1}, \dfrac{\gamma}{\beta+1}\Big), \\ \boldsymbol{W}^{k+\frac{1}{2}} = \boldsymbol{W}^k - \tau\beta(\boldsymbol{U}_1^{k+1} + \boldsymbol{U}_2^{k+1} - \boldsymbol{V}^k), \\ \boldsymbol{V}^{k+1} = \dfrac{\beta(\boldsymbol{U}_1^{k+1} + \boldsymbol{U}_2^{k+1}) + w\boldsymbol{M} - \boldsymbol{W}^{k+\frac{1}{2}} + \rho(\boldsymbol{V}^k - \boldsymbol{V}^{k-1})}{\beta + w}, \\ \boldsymbol{W}^{k+1} = \boldsymbol{W}^{k+\frac{1}{2}} - \eta\beta(\boldsymbol{U}_1^{k+1} + \boldsymbol{U}_2^{k+1} - \boldsymbol{V}^{k+1}), \end{cases}$$

where $\triangle_{\phi_1}(\boldsymbol{U}_1, \bar{\boldsymbol{U}}_1^k) = \frac{1}{2}\|\boldsymbol{U}_1 - \bar{\boldsymbol{U}}_1^k\|_{\boldsymbol{I}}^2$, $\triangle_{\phi_2}(\boldsymbol{U}_2, \bar{\boldsymbol{U}}_2^k) = \frac{1}{2}\|\boldsymbol{U}_2 - \bar{\boldsymbol{U}}_2^k\|_{\boldsymbol{I}}^2$ and $\triangle_{\varphi}(\boldsymbol{V}, \boldsymbol{V}^k) = 0$. $\mathcal{H}(\cdot; \frac{1}{\beta+1})$ represents the half shrinkage operator [44], while $\mathcal{S}(\cdot; \frac{\gamma}{\beta+1})$ represents the soft shrinkage operator [9].

For the SCIB-ADMM, let $(\tau, \eta) = (0.2, 1.1)$, $\rho = 0.1$ and $\sigma_1 = \sigma_2 = 0.9$. For the IPPS-ADMM, let $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{I}$, $(\tau, \eta) = (0.2, 0.5)$ and $\alpha_1 = \alpha_2 = 0.9$. For the PSRADMM, let $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{I}$ and $(\tau, \eta) = (0.85, 0.85)$. For the BADMM, let $\triangle_{\phi_1}(\boldsymbol{U}_1, \boldsymbol{U}_1^k) = \frac{1}{2}\|\boldsymbol{U}_1 - \boldsymbol{U}_1^k\|_{\boldsymbol{I}}^2$, $\triangle_{\phi_2}(\boldsymbol{U}_2, \boldsymbol{U}_2^k) = \frac{1}{2}\|\boldsymbol{U}_2 - \boldsymbol{U}_2^k\|_{\boldsymbol{I}}^2$ and $\triangle_{\varphi}(\boldsymbol{V}, \boldsymbol{V}^k) = 0$. The matrices $\boldsymbol{U}_1$, $\boldsymbol{U}_2$ and $\boldsymbol{V}$ are initialized with zero-matrix. Specifically, let $p = n = 100$, $\gamma = \frac{0.1}{\sqrt{p}}$, $w = 10^3$, $\beta = 7.09$, and test 6 different "rank" and "spr", which denote the matrix rank of the $\boldsymbol{U}_1$ and the sparsity ratio of the matrix $\boldsymbol{U}_2$, respectively. Let $\boldsymbol{M} = \boldsymbol{V} + \boldsymbol{N}$, the MATLAB script to generate $\boldsymbol{M}$ is:

$K = \text{round}(\text{spr} * m * q)$; $p = \text{randperm}(m * q)$; $\boldsymbol{U}_1 = \text{randn}(m, \text{rank}) * \text{randn}(\text{rank}, q)$;
$\boldsymbol{U}_2 = \text{zeros}(m, q)$; $\boldsymbol{U}_2(p(1:K)) = \text{randn}(K, 1)$; $\vartheta = 0$; $\vartheta = 0.01$;
$\boldsymbol{N} = \text{randn}(m, q) * \vartheta$; $\boldsymbol{V} = \boldsymbol{U}_1 + \boldsymbol{U}_2$; $\boldsymbol{M} = \boldsymbol{V} + \boldsymbol{N}$.

The termination condition is $\text{RelChg} = \frac{\|(\boldsymbol{U}_1^{k+1}, \boldsymbol{U}_2^{k+1}, \boldsymbol{V}^{k+1}) - (\boldsymbol{U}_1^k, \boldsymbol{U}_2^k, \boldsymbol{V}^k)\|_F}{\|(\boldsymbol{U}_1^k, \boldsymbol{U}_2^k, \boldsymbol{V}^k)\|_F + 1} \leq 10^{-7}$.

The relative error (RelErr) is $\text{RelErr} = \frac{\|(\boldsymbol{U}_1^*, \boldsymbol{U}_2^*, \boldsymbol{V}^*) - (\hat{\boldsymbol{U}}_1, \hat{\boldsymbol{U}}_2, \hat{\boldsymbol{V}})\|_F}{\|(\boldsymbol{U}_1^*, \boldsymbol{U}_2^*, \boldsymbol{V}^*)\|_F + 1}$, where $\boldsymbol{U}_1^*$, $\boldsymbol{U}_2^*$, $\boldsymbol{V}^*$ and $\hat{\boldsymbol{U}}_1$, $\hat{\boldsymbol{U}}_2$, $\hat{\boldsymbol{V}}$ represent the original matrices and the solution of (4.3).

In Tables 2 and 3, the convergence results of the four methods are reported for the ϑ=0 (noiseless) and ϑ=0.01 (Gaussian noise). In addition, Tables 2 and 3 display the numerical results of the above four method relative to Iter, RelErr (relative error) and RelChg, respectively. Bold numerical value highlight the superior performance achieved by our method.

In Figure 3, the RelErr for the ($\text{rank} = 10$, $\text{spr} = 0.05$) is given for the ϑ=0 and ϑ=0.01.

As can be seen from Tables 2 and 3 and Figure 3, the SCIB-ADMM has the following main advantages over the other three methods.

(i) The SCIB-ADMM method uses fewer iterations than the other three methods.

(ii) The relative errors (RelErr) of the proposed SCIB-ADMM tend to zero faster than the other three methods with the noiseless case and Gaussian noise case.

Our numerical experiments indicate that the SCIB-ADMM outperforms other three methods. Therefore, the inertial strategy and the Bregman distance are effective on robust PCA.

Table 2. Numerical results on robust PCA with $\vartheta = 0$.

	(rank, spr)	(1, 0.05)	(1, 0.1)	(10, 0.05)	(10, 0.1)	(20, 0.05)	(20, 0.1)
BADMM	Iter	468	556	1220	1349	2035	2780
	$\log_{10}$(RelChg)	−7.2456	−7.4575	−7.1239	−7.1635	−7.1245	−7.1507
	$\log_{10}$(RelErr)	−5.6142	−5.4977	−5.9826	−5.8562	−5.9712	−5.8270
PSRADMM	Iter	463	549	1201	1330	1996	2727
	$\log_{10}$(RelChg)	−7.0615	−7.2305	−7.3293	−7.1457	−7.1379	−7.0690
	$\log_{10}$(RelErr)	−5.5819	−5.4885	−5.9816	−5.8588	−5.9737	−5.8251
IPPS-ADMM	Iter	385	458	1014	1116	1705	2324
	$\log_{10}$(RelChg)	−7.6624	−7.2406	−7.1364	−7.2185	−7.0371	−7.0841
	$\log_{10}$(RelErr)	−5.5609	−5.4745	−5.9790	−5.8561	−5.9711	−5.8274
SCIB-ADMM	Iter	**136**	**209**	**402**	**485**	**622**	**955**
	$\log_{10}$(RelChg)	−7.1008	−7.2102	−7.2075	−7.0737	−7.0834	−7.1240
	$\log_{10}$(RelErr)	−5.3835	−5.3557	−5.9502	−5.8297	−5.9599	−5.8260

Table 3. Numerical results on robust PCA with $\vartheta = 0.01$.

	(rank, spr)	(1, 0.05)	(1, 0.1)	(10, 0.05)	(10, 0.1)	(20, 0.05)	(20, 0.1)
BADMM	Iter	1083	1184	2022	2293	3160	4104
	$\log_{10}$(RelChg)	−7.0024	−7.0016	−7.0005	−7.0002	−7.0005	−7.0004
	$\log_{10}$(RelErr)	−2.0072	−2.0182	−2.4729	−2.4618	−2.5567	−2.5056
PSRADMM	Iter	1234	1287	2017	2279	3129	4054
	$\log_{10}$(RelChg)	−7.0000	−7.0030	−7.0003	−7.0010	−7.0003	−7.0001
	$\log_{10}$(RelErr)	−2.0101	−2.0211	−2.4738	−2.4628	−2.5574	−2.5064
IPPS-ADMM	Iter	1146	1166	1823	2061	2810	3655
	$\log_{10}$(RelChg)	−7.0022	−7.0010	−7.0001	−7.0008	−7.0001	−7.0002
	$\log_{10}$(RelErr)	−2.0072	−2.0182	−2.4729	−2.4618	−2.5567	−2.5057
SCIB-ADMM	Iter	**1062**	**1009**	**1151**	**1398**	**1690**	**2156**
	$\log_{10}$(RelChg)	−7.0006	−7.0015	−7.0018	−7.0008	−7.0002	−7.0003
	$\log_{10}$(RelErr)	−2.0071	−2.0182	−2.4728	−2.4619	−2.5568	−2.5058

5. Conclusion

In conclusion, a sequential complete inertial Bregman ADMM is proposed for multi-block non-convex problems. The SCIB-ADMM combines the inertial term and the Bregman distance to improve the computational efficiency. The convergence of the SCIB-ADMM has been proven under appropriate conditions. Finally, numerical experiments on SCAD and robust PCA show that the SCIB-ADMM outperforms some other methods.

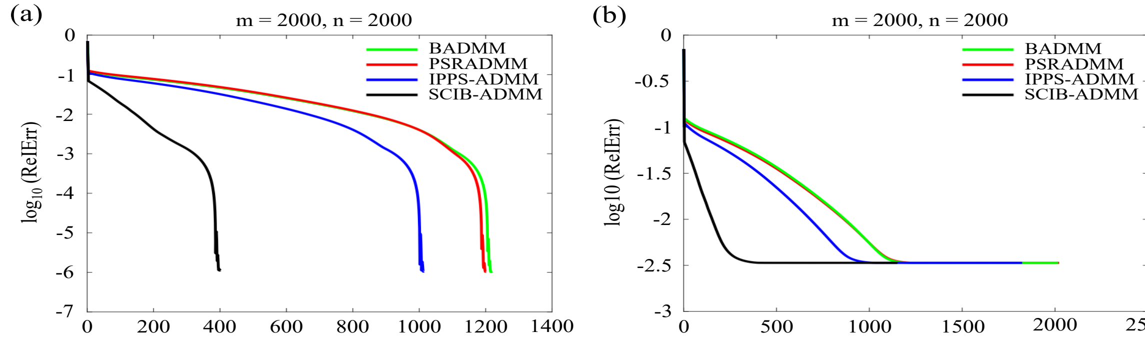

Figure 3. Comparison of the four methods on robust PCA. (a) Shows the noiseless case ($\vartheta = 0$). (b) Shows Gaussian noise case ($\vartheta = 0.01$).

Conflict of interest

The authors declared no conflict of interest.

References

[1] M. Abdellatif and E. Elissabeth, *Approximate inertial proximal methods using the enlargement of maximal monotone operators*, International Journal of Pure and Applied Mathematics, 2013, 5(3), 283–299.

[2] F. Alvarez, *Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space*, SIAM Journal on Optimization, 2004, 14(3), 773–782.

[3] H. Attouch, J. Bolte and P. Redont, *A proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality*, Mathematics of Operations Research, 2010, 35(2), 438–457.

[4] R. F. Barber and E. Y. Sidky, *Convergence for nonconvex ADMM, with applications to CT imaging*, Journal of Machine Learning Research, 2024, 25(38), 1–46.

[5] J. Bolte, S. Sabach and M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Mathematical Programming, 2014, 146(1), 459–494.

[6] M. Chao, Y. Geng and Y. Zhao, *A method of inertial regularized ADMM for separable nonconvex optimization problems*, Soft Computing, 2023, 27(22), 16741–16757.

[7] C. H. Chen, R. H. Chan, S. Q. Ma, et al., *Inertial proximal ADMM for linearly constrained separable convex optimization*, SIAM Journal on Imaging Sciences, 2015, 8(4), 2239–2267.

[8] L. Cheng, S. R. Zhang and Y. C. Wang, *Distributed optimal capacity allocation of integrated energy system via modified ADMM*, Applied Mathematics and Computation, 2024, 465, 128369.

[9] I. Daubechies, M. Defrise and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Communications on Pure and Applied Mathematics, 2003, 57(11), 1413–1457.

[10] J. Douglas and H. H. Rachford, *On the numerical solution of the heat conduction problem*

in two and three space variables, Transactions of the American Mathematical Society, 1956, 82(2), 421–439.

[11] Z. Ebrahimi and M. Deghat, *Distributed pose graph optimization using the splitting method based on the alternating direction method of multipliers*, Journal of Intelligent & Robotic Systems, 2025, 111, 56.

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

[13] 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, 2020, 76, 863–887.

[14] M. L. Goncalves, J. G. Melo and R. D. Monteiro, *Convergence rate bounds for a proximal ADMM with over-relaxation stepsize parameter for solving nonconvex linearly constrained problems*, Pacific Journal of Optimization, 2017, 15(3), 379–398.

[15] Y. Gu and N. Yamashita, *An alternating direction method of multipliers with the BFGS update for structured convex quadratic optimization*, Computational and Applied Mathematics, 2021, 40, 81.

[16] K. Guo, D. R. Han, D. Z. W. Wang, et al., *Convergence of ADMM for multi-block nonconvex separable optimization models*, Frontiers of Mathematics in China, 2017, 12(5), 1139–1162.

[17] 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, 2017, 94(8), 1653–1669.

[18] B. He, F. Ma and X. Yuan, *Convergence study on the symmetric version of ADMM with larger step sizes*, SIAM Journal on Imaging Sciences, 2016, 9(3), 1467–1501.

[19] B. He, M. Tao and X. Yuan, *Alternating direction method with Gaussian back substitution for separable convex programming*, SIAM Journal on Optimization, 2012, 22(2), 313–340.

[20] B. He, M. Tao and X. Yuan, *A splitting method for separable convex programming*, IMA Journal of Numerical Analysis, 2015, 35(1), 394–426.

[21] L. T. K. Hien, D. N. Phan and N. Gillis, *Inertial alternating direction method of multipliers for nonconvex nonsmooth optimization*, Computational Optimization and Applications, 2022, 83(1), 247–285.

[22] M. Hong, Z. Luo and M. Razaviyayn, *Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems*, SIAM Journal on Optimization, 2016, 26(1), 337–364.

[23] J. Jian, D. Cui and G. Ma, *A partially symmetric linearized bregman ADMM for three-block nonconvex nonsmooth optimization*, Journal of Computational and Applied Mathematics, 2026, 483, 117433.

[24] J. B. Jian, P. J. Liu and X. Z. Jiang, *A partially symmetric regularized alternating direction method of multipliers for nonconvex multi-block optimization*, Acta Mathematica Sinica, 2021, 64(6), 1005–1026.

[25] J. B. Jian, C. Zhang and P. J. Liu, *A superlinearly convergent splitting feasible sequential quadratic optimization method for two-block large-scale smooth optimization*, Acta Mathematica Scientia, 2023, 43B(1), 1–24.

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

[27] L. Li, L. Wang, G. Q. Wang, et al., *Linearized alternating direction method of multipliers for separable convex optimization of real functions in complex domain*, Journal of Applied Analysis and Computation, 2019, 9(5), 1686–1705.

[28] M. Li, D. Sun and K. Toh, *A majorized ADMM with indefinite proximal terms for linearly constrained convex composite optimization*, SIAM Journal on Optimization, 2016, 26(2), 922–950.

[29] K. Liu, A. Zhou and J. Fan, *A generalized alternating direction method of multipliers for tensor complementarity problems*, Computational Optimization and Applications, 2024, 88, 903–921.

[30] Z. S. Liu and T. Zhang, *Convergence analysis on the alternating direction method of multipliers for the cosparse optimization problem*, Journal of Applied Analysis and Computation, 2024, 14(6), 3061–3077.

[31] R. L. Luo and Z. S. Yu, *The proximal alternating direction method of multipliers for a class of nonlinear constrained optimization problems*, Mathematics, 2025, 13(3), 407.

[32] Y. Nesterov, *Introduction Lectures on Convex Optimization: A Basic Course*, Springer, Berlin, 2004.

[33] P. Ochs, T. Brox and T. Pock, *iPiasco: Inertial proximal algorithm for strongly convex optimization*, Journal of Mathematical Imaging and Vision, 2015, 53, 171–181.

[34] P. Ochs, X. J. Chen, T. Brox, et al., *iPiano: Inertial proximal algorithm for nonconvex optimization*, SIAM Journal on Imaging Sciences, 2014, 7(2), 1388–1419.

[35] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, Springer, Berlin, 2009.

[36] F. H. Wang, W. F. Cao and Z. B. Xu, *Convergence of multi-block Bregman ADMM for nonconvex composite problems*, Science China Information Sciences, 2018, 61(12), 1–12.

[37] M. Wang, H. Liu and J. Yang, *Alternated multi-step inertial iterative algorithm for solving the split feasibility problem in Hilbert spaces*, Computational and Applied Mathematics, 2025, 44(1), 1–19.

[38] P. Wang, C. D. Lin, X. B. Yang, et al., *Low-rank and sparse matrix recovery from noisy observations via 3-block ADMM algorithm*, Journal of Applied Analysis and Computation, 2020, 10(3), 1024–1037.

[39] X. Q. Wang, H. Shao, P. J. Liu, et al., *An inertial proximal partially symmetric ADMM-based algorithm for linearly constrained multi-block nonconvex optimization problems with applications*, Journal of Computational and Applied Mathematics, 2023, 420, 114821.

[40] Z. M. Wu, C. Li, M. Li, et al., *Inertial proximal gradient methods with Bregman regularization for a class of nonconvex optimization problems*, Journal of Global Optimization, 2021, 79, 617–644.

[41] Z. M. Wu and M. Li, *General inertial proximal gradient method for a class of nonconvex nonsmooth optimization problems*, Computational Optimization and Applications, 2019, 73(1), 129–158.

[42] H. Xu, C. Caramanis and S. Mannor, *Outlier-robust PCA: The high-dimensional case*, IEEE Transactions on Information Theory, 2013, 59(1), 546–572.

[43] J. W. Xu and M. T. Chao, *An inertial Bregman generalized alternating direction method of multipliers for nonconvex optimization*, Journal of Applied Mathematics and Computing, 2022, 68(3), 1757–1783.

[44] Z. B. Xu, X. Y. Chang, F. M. Xu, et al., $l_{\frac{1}{2}}$ *regularization a thresholding representation theory and a fast solve*, IEEE Transactions on Neural Networks and Learning Systems, 2012, 23(7), 1013–1027.

[45] L. Zhang, J. Wu and L. Zhang, *A linearly convergent majorized ADMM with indefinite proximal terms for convex composite programming and its applications*, Mathematics of Computation, 2020, 89(324), 1867–1894.

Received May 2025; Accepted March 2026; Available online April 2026.