Journal of Applied Mathematics and Physics, 2026, 14(4), 1714-1732
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A Symmetric Alternating Direction Method of Multipliers with Two Different Relaxation Factors for Solving Non-Separable Nonconvex Minimization Problems

Mei Lu, Zidan Wang*

Key Laboratory of Optimization Theory and Applications, School of Mathematics and Information, China West Normal University, Nanchong, China
Email: 1339975549@qq.com, *1265130810@qq.com

How to cite this paper: Lu, M., and Wang, Z.D. (2026) A Symmetric Alternating Direction Method of Multipliers with Two Different Relaxation Factors for Solving Non-Separable Nonconvex Minimization Problems. *Journal of Applied Mathematics and Physics*, **14**, 1714-1732.
https://doi.org/10.4236/jamp.2026.144081

Received: April 1, 2026
Accepted: April 25, 2026
Published: April 28, 2026

Abstract

This paper proposes a symmetric alternating direction method of multipliers with two different relaxation factors for solving nonconvex optimization problems with linear constraints and a non-separable structure. Although many studies have proposed variants of symmetric ADMM with a single relaxation factor, incorporating techniques such as Bregman distances, inertial terms, regularization terms, or linearization, the most basic form of symmetric ADMM with two different relaxation factors for solving non-separable problems has not yet been resolved. The introduction of two different relaxation factors in this method yields a broader range of parameters, making it applicable to more practical problems, and also provides a fundamental theoretical basis for accelerating the algorithm with other techniques. Based on the Kurdyka-Łojasiewicz property, we establish the convergence of the sequences generated by the proposed algorithm and analyze its convergence rate.

Keywords

Symmetric Alternating Direction Method of Multipliers, Non-Separable Structure, Nonconvex Optimization, Kurdyka-Łojasiewicz Inequality

1. Introduction

In many practical problems such as image processing, machine learning, and statistical modeling [1]-[3], the objective function often contains a coupling term $H(x, y)$. That is, for the nonconvex and nonseparable problem considered in this paper, the iterative scheme is as follows:

$$\min_{x,y} f(x)+g(y)+H(x,y),\quad \text{s.t. } Ax+y=b. \tag{1}$$

where $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is a proper lower semicontinuous function, possibly nonsmooth and nonconvex, both $g:\mathbb{R}^m\to\mathbb{R}$ and $H:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ are continuously differentiable and possibly nonconvex functions, the gradient ∇g is L_g-Lipschitz continuous and the gradient ∇H is L_h-Lipschitz continuous. $A\in\mathbb{R}^{m\times n}$ is a matrix and $b\in\mathbb{R}^m$ is a vector. In particular, when $H(x,y)=0$, problem (1) reduces to a separable optimization problem of the following form:

$$\min_{x,y} f(x)+g(y),\quad \text{s.t. } Ax+y=b. \tag{2}$$

The alternating direction method of multipliers (ADMM) is one of the most effective approaches for solving problem (2). The convergence of ADMM has been extensively studied [4]-[9], and the convergence rate analysis is relatively well-established for problem (2). However, when $H(x,y)\neq 0$, for problem (1), the convergence results for ADMM are relatively limited. Gao *et al.* [10] proved the convergence of the proximal ADMM under the setting where H is smooth, f and g are convex, combined with the assumption that ∇H is Lipschitz continuous. Chen *et al.* [11] proved the convergence of the extended ADMM when the coupling term H is a quadratic function. In recent years, researchers have begun to focus on solving (1) by using ADMM in non-convex settings, and have employed tools such as the Kurdyka-Łojasiewicz (KL) inequality to prove the convergence of the algorithm. Guo *et al.* [12] [13] proved the convergence of the classic ADMM and extended it to the generalized ADMM (GADMM). Liu *et al.* [14] proposed a linearized ADMM and proved the convergence of the algorithm.

Regarding problem (2), it has been established in the literature that under convergence occurs, the symmetric ADMM often converges faster than the classical ADMM. Wu *et al.* [15] proposed a symmetric ADMM with one relaxation factor and verified its convergence. On this basis, some scholars have studied the symmetric ADMM with a relaxation factor and its variants for solving the nonseparable problem (1). For example, Dang *et al.* [16] proposed a linear proximal symmetric ADMM and verified that the sequence generated by the algorithm converges to a stationary point of the problem. The iterative scheme is as follows:

$$\begin{cases} x^{k+1}\in\arg\min\limits_{x}\Big\{ f(x)+\big\langle x-x^k,\nabla_x H\big(x^k,y^k\big)+\beta A^{\mathrm{T}}\big(Ax^k+By^k-b\big)\big\rangle \\ \qquad\qquad -\big\langle \lambda^k,Ax\big\rangle+\dfrac{\alpha_x}{2}\big\|x-x^k\big\|^2\Big\}, \\ \lambda^{k+\frac{1}{2}}=\lambda^k-\tau\beta\big(Ax^{k+1}+By^k-b\big), \\ y^{k+1}\in\arg\min\limits_{y}\left\{ g(y)+\big\langle y-y^k,\nabla_y H\big(x^{k+1},y^k\big)\big\rangle+\dfrac{\beta}{2}\left\|Ax^{k+1}+By-b-\dfrac{\lambda^{k+\frac{1}{2}}}{\beta}\right\|^2\right\}, \\ \lambda^{k+1}=\lambda^{k+\frac{1}{2}}-\beta\big(Ax^{k+1}+By^{k+1}-b\big). \end{cases}$$

The parameters are selected as follows:

$$\tau \in (-1,0), \alpha_y \geq l_h,$$

$$\alpha_x \geq \frac{8\mu^2 l_h^2}{(\tau+1)\beta} + \beta\lambda_{\max}\left(A^{\mathrm{T}}A\right) + l_h, \beta \geq \frac{\mu^2\left(b-\sqrt{b^2-64\tau c}\right)}{4\tau}.$$

Dang *et al.* [17] proposed an inertial Bregman symmetric ADMM algorithm and established its global convergence. The iterative scheme is as follows:

$$\begin{cases} x^{k+1} \in \arg\min_x \left\{ L_\rho\left(x, y^k, \lambda^k\right) + \Delta_{\phi_1}\left(x, x^k\right) + \theta_{1k}\left\langle x, x^{k-1}-x^k \right\rangle + \theta_{2k}\left\langle x, x^{k-2}-x^{k-1}\right\rangle \right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\rho\left(Ax^{k+1}+By^k-b\right) + \theta_{1k}B\left(y^{k-1}-y^k\right) + \theta_{2k}B\left(y^{k-2}-y^{k-1}\right), \\ y^{k+1} \in \arg\min_y \left\{ L_\rho\left(x^{k+1}, y, \lambda^{k+\frac{1}{2}}\right) + \Delta_{\phi_2}\left(y, y^k\right) + \theta_{1k}\left\langle By, B\left(y^{k-1}-y^k\right)\right\rangle \right. \\ \qquad\qquad \left. + \theta_{2k}\left\langle By, B\left(y^{k-2}-y^{k-1}\right)\right\rangle \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \rho\left(Ax^{k+1}+By^{k+1}-b\right) + \theta_{1k}B\left(y^k-y^{k-1}\right) + \theta_{2k}B\left(y^{k-1}-y^{k-2}\right). \end{cases}$$

where $\Delta_{\phi_i}(s,t)$, $i=1,2$ represents the Bregman distances, and $r \in (-1,1)$ is a relaxation factor. However, the convergence of the sequence generated by the algorithm depends on the strong convexity of the kernel function associated with the Bregman distance.

In recent studies on solving problem (2), we observe that Lu *et al.* [18] proposed a symmetric ADMM with two different relaxation factors, without introducing the Bregman distance, they proved its convergence with a broader range of parameters, numerical experiments verified that their algorithm converges faster than the algorithm proposed by [15]. Its iterative scheme is as follows:

$$\begin{cases} x^{k+1} \in \arg\min_x \left\{ \mathcal{L}_\beta^s\left(x, y^k, \lambda^k\right) \right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta\left(Ax^{k+1}+y^k-b\right), \\ y^{k+1} \in \arg\min_y \left\{ \mathcal{L}_\beta^s\left(x^{k+1}, y, \lambda^{k+\frac{1}{2}}\right) \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta\left(Ax^{k+1}+y^{k+1}-b\right). \end{cases} \tag{3}$$

Among them, $\mathcal{L}_\beta^s$ is the augmented Lagrangian function associated with problem (2), defined as follows:

$$\mathcal{L}_\beta^s(x,y,\lambda) = f(x) + g(y) - \langle \lambda, Ax+y-b \rangle + \frac{s\beta}{2}\|Ax+y-b\|^2. \tag{4}$$

where $\beta > 0$ is the penalty parameter, α and s are two different relaxation factors.

The above research on symmetric ADMM algorithms for solving problem (1) is limited to those containing only one relaxation factor. However, inspired by the algorithm proposed by the work of Lu *et al.* [18], we consider introducing two relaxation factors to achieve wider applicability, faster convergence, and a more concise and unified algorithmic framework, this paper proposes using a symmetric ADMM with two different relaxation factors to solve the non-separable (1), the iterative scheme is as follows:

$$\begin{cases} x^{k+1} \in \arg\min_x \left\{ f(x) + H(x, y^k) - \langle \lambda^k, Ax \rangle + \frac{s\beta}{2} \|Ax + y^k - b\|^2 \right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + y^k - b), \\ y^{k+1} \in \arg\min_y \left\{ g(y) + H(x^{k+1}, y) - \left\langle \lambda^{k\frac{1}{2}}, y \right\rangle + \frac{s\beta}{2} \|Ax^{k+1} + y - b\|^2 \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} + y^{k+1} - b). \end{cases} \tag{5}$$

Compared with the algorithm proposed by Dang *et al.* [16], our algorithm has a wider range of parameter values, allowing it to be applied to more practical scenarios through appropriate parameter tuning. Moreover, unlike the algorithm proposed by Dang *et al.* [17], our method does not require the introduction of the Bregman distance, thereby providing a unified iterative framework for solving non-separable problems via ADMM. The optimality conditions are as follows:

$$\begin{cases} 0 \in \partial f(x^{k+1}) + \nabla_x H(x^{k+1}, y^k) - A^{\mathrm{T}}\lambda^k + s\beta A^{\mathrm{T}}(Ax^{k+1} + y^k - b), \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + y^k - b), \\ 0 = \nabla g(y^{k+1}) + \nabla_y H(x^{k+1}, y^{k+1}) - \lambda^{k+\frac{1}{2}} + s\beta(Ax^{k+1} + y^{k+1} - b), \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(Ax^{k+1} + y^{k+1} - b). \end{cases} \tag{6}$$

In the case where $H = 0$, when $a = 0$ and $s = 1$, the proposed algorithm reduces to the classical ADMM. When $a \neq 0$ and $s = 1$, the algorithm reduces to the symmetric ADMM with a scaling factor. This demonstrates that our method provides an improved basic iterative framework for symmetric ADMM with relaxation factors.

2. Preliminaries

In this section, we provide the necessary definitions and properties required for the following study.

Notations:

- $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$: real numbers, n-dimensional real vectors, $m \times n$-real matrices.
- $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$: inner product and associated norm.

Set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$:

- Domain: $\mathrm{dom}F := \{x \in \mathbb{R}^n \mid F(x) \neq \varnothing\}$.
- Graph: $\mathrm{Gra}\,F := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in F(x)\}$.

Distance: For $S \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$, $d(x,S) := \inf\{\|y-x\| \mid y \in S\}$, with $d(x,S) := +\infty$.

Definition 1. [19] The domain of function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is denoted by

$$\operatorname{dom} f = \{x \in R^n : f(x) < +\infty\},$$

then $f(x)$ is called a proper function.

Definition 2. [19] A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be lower semicontinuous at $x \in \mathbb{R}^n$ if it satisfies

$$f(x) \le \liminf_{k\to\infty} f(x_k).$$

If this holds for every point in $\operatorname{dom} f$, then f is said to be a lower semicontinuous function.

Definition 3. [20] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function.

(i) The Fréchet subdifferential of f at $x \in \operatorname{dom} f$, written $\hat{\partial} f(x)$, is the set of vectors $x^* \in \mathbb{R}^n$ that satisfy

$$\liminf_{y\ne x\ y\to x} \frac{f(y) - f(x) - \langle x^*, y-x\rangle}{\|y-x\|} \ge 0.$$

When $x \notin \operatorname{dom} f$, we set $\hat{\partial} f(x) = \varnothing$.

(ii) The limiting-subdifferential of f at $x \in \operatorname{dom} f$, written $\partial f(x)$, is defined as follows

$$\partial f(x) = \left\{x^* \in \mathbb{R}^n, \exists x_n \to x, f(x_n) \to f(x), x_n^* \in \hat{\partial} f(x_n), \text{with } x_n^* \to x^*\right\}.$$

Lemma 1. [21] Let $g : R^m \to R$ be a continuously differentiable function and suppose that ∇g is L-Lipschitz continuous. Then

$$\|g(u) - g(v) - \langle u-v, \nabla g(v)\rangle\| \le \frac{L}{2}\|u-v\|^2,\ \forall u, v \in R^m.$$

Definition 4. ([22], Kurdyka-Lojasiewicz inequality) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function. For $-\infty < \eta_1 < \eta_1 \le +\infty$, set

$$[\eta_1 < f < \eta_1] = \{x \in \mathbb{R}^n : \eta_1 < f(x) < \eta_2\}.$$

We say that function f has the KL property at $x^* \in \operatorname{dom}(\partial f)$ if there exist $\eta \in (0, +\infty]$, a neighbourhood U of x^*, and a continuous concave function $\varphi : [0, \eta) \to \mathbb{R}_+$, such that

(i) $\varphi(0) = 0$;

(ii) φ is C^1 on $(0, \eta)$ and continuous at 0;

(iii) $\varphi'(x) > 0, \forall x \in (0, \eta)$;

(iv) for all x in $U \cap \left[f(x^*) < f < f(x^*) + \eta\right]$, the Kurdyka-Lojasiewicz inequality holds

$$\varphi'\left(f(x) - f(x^*)\right) d\left(0, \partial f(x)\right) \ge 1,$$

where $d(x,\partial f(x)) = \inf_{y\in\partial f(x)} \|y - x\|$ is the distance from x to $\partial f(x)$.

Lemma 2. ([23], Uniformized KL property) Let Ω be a compact set and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function. Assume that f is constant on Ω and satisfies the KL property at each point of Ω. Then, there exist $\epsilon > 0, \eta > 0$, and $\varphi \in \Phi_\eta$ such that for all $\bar{x} \in \Omega$ and for all x in the following intersection:

$$\{x \in \mathbb{R}^n : d(x,\Omega) < \epsilon\} \cap [f(\bar{x}) < f < f(\bar{x}) + \eta],$$

one has

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

Lemma 3. [24] Suppose that $F(x,y) = f(x) + g(x)$, where $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $g : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ are proper lower semicontinuous functions. Then for all $(x,y) \in \operatorname{dom} F = \operatorname{dom} f \times \operatorname{dom} g$, we have

$$\partial F(x,y) = \partial_x F(x,y) \times \partial_y F(x,y).$$

Definition 5. ([24], Kurdyka-Lojasiewicz function) If f satisfies the KL property at each point of $\operatorname{dom}(\partial f)$, then f is called a KL function.

Definition 6. (x^*, y^*, λ^*) is a stationary point of the augmented Lagrangian function $\mathcal{L}_\beta^s(\cdot)$ for problem (1), if and only if

$$\begin{cases} -\nabla_x H(x^*, y^*) + A^{\mathrm{T}}\lambda^* \in \partial f(x^*), \\ -\nabla_y H(x^*, y^*) + \lambda^* = \nabla g(y^*), \\ Ax^* + y^* = b. \end{cases} \tag{7}$$

Lemma 4. Let the iterative sequence generated by the algorithm (5) be denoted as $\{w^k : (x^k, y^k, \lambda^k)\}$. Then, the following holds

$$\begin{cases} 0 \in \partial f(x^{k+1}) + \nabla_x H(x^{k+1}, y^k) - A^{\mathrm{T}}\lambda^{k+1} + \dfrac{\alpha}{\alpha+s} A^{\mathrm{T}}(\lambda^{k+1} - \lambda^k) \\ \qquad - \dfrac{s^2\beta}{\alpha+s} A^{\mathrm{T}}(y^{k+1} - y^k), \\ \nabla g(y^{k+1}) = \lambda^{k+1} - \nabla_y H(x^{k+1}, y^{k+1}), \\ \lambda^{k+1} = \lambda^k - \beta\left[(\alpha+s)(Ax^{k+1} + y^k - b) + s(y^{k+1} - y^k)\right]. \end{cases}$$

Proof. Combining the second and fourth equations in (6) yields

$$\lambda^{k+1} = \lambda^k - \alpha\beta(Ax^{k+1} + y^k - b) - s\beta(x^{k+1} + y^{k+1} - b),$$

$$\lambda^{k+1} = \lambda^k - (\alpha+s)\beta(Ax^{k+1} + y^k - b) - s\beta(y^{k+1} - y^k),$$

subtracting the above two equations, we obtain

$$\lambda^{k+1} - \lambda^k = -(\alpha+s)\beta(Ax^{k+1} + y^k - b) - s\beta(y^{k+1} - y^k), \tag{8}$$

and thus

$$Ax^{k+1} + y^k - b = -\frac{1}{(\alpha+s)\beta}(\lambda^{k+1} - \lambda^k) - \frac{s}{\alpha+s}(y^{k+1} - y^k), \tag{9}$$

$$Ax^{k+1}+y^{k+1}-b=-\frac{1}{(\alpha+s)\beta}\left(\lambda^{k+1}-\lambda^{k}\right)+\frac{\alpha}{\alpha+s}\left(y^{k+1}-y^{k}\right). \tag{10}$$

According to the optimality condition (6) of the x-subproblem, we have

$$\begin{aligned}0&\in\partial f\left(x^{k+1}\right)+\nabla_x H\left(x^{k+1},y^{k}\right)-A^{\mathrm{T}}\lambda^{k}+s\beta A^{\mathrm{T}}\left(Ax^{k+1}+y^{k}-b\right)\\&=\partial f\left(x^{k+1}\right)+\nabla_x H\left(x^{k+1},y^{k}\right)-A^{\mathrm{T}}\lambda^{k+1}+A^{\mathrm{T}}\left(\lambda^{k+1}-\lambda^{k}\right)\\&\quad+s\beta A^{\mathrm{T}}\left(Ax^{k+1}+y^{k}-b\right),\end{aligned} \tag{11}$$

putting (9) into (11), we get

$$\begin{aligned}0&\in\partial f\left(x^{k+1}\right)+\nabla_x H\left(x^{k+1},y^{k}\right)-A^{\mathrm{T}}\lambda^{k+1}+\frac{\alpha}{\alpha+s}A^{\mathrm{T}}\left(\lambda^{k+1}-\lambda^{k}\right)\\&\quad-\frac{s^{2}\beta}{\alpha+s}A^{\mathrm{T}}\left(y^{k+1}-y^{k}\right).\end{aligned} \tag{12}$$

According to the third and fourth equations in (6), we obtain

$$\nabla g\left(y^{k+1}\right)=\lambda^{k+1}-\nabla_y H\left(x^{k+1},y^{k+1}\right). \tag{13}$$

This completes the proof. □

3. Convergence Analysis

Assumption A. let $f:\mathbb{R}^{n}\to\mathbb{R}\cup\{+\infty\}$ be a proper lower semicontinuous function, and let $g:\mathbb{R}^{m}\to\mathbb{R}$ be a continuously differentiable function with an L_g-Lipschitz continuous gradient ∇g, and $H:\mathbb{R}^{m}\times\mathbb{R}^{n}\to\mathbb{R}$ be a continuously differentiable function with an L_h-Lipschitz continuous gradient ∇H. Also assume that

- $\left\{(\alpha,s)\in\mathbb{R}^{2}\mid\alpha<s<0\right\}$,
- $\left\{\beta\in\mathbb{R}\mid\beta>-\frac{1}{s}\left(L_g+L_h\right)\right\}$

Lemma 5. Let the iterative sequence generated by the algorithm (5) be denoted as $\left\{w^{k}:\left(x^{k},y^{k},\lambda^{k}\right)\right\}$. Suppose that this sequence is bounded and that Assumption A holds. Then the following conclusion holds

$$\mathcal{L}_{\beta}^{s}\left(w^{k+1}\right)-\mathcal{L}_{\beta}^{s}\left(w^{k}\right)\le-\eta\left(\left\|x^{k+1}-x^{k}\right\|^{2}+\left\|y^{k+1}-y^{k}\right\|^{2}\right), \tag{14}$$

where $\eta>0$ (see Remark (0.1)).

Proof. The polarization identity implies

$$\begin{aligned}&\frac{s\beta}{2}\left\|Ax^{k+1}+y^{k+1}-b\right\|^{2}-\frac{s\beta}{2}\left\|Ax^{k+1}+y^{k}-b\right\|^{2}\\&=-\frac{s\beta}{2}\left\|y^{k}-y^{k+1}\right\|^{2}+s\beta\left\langle Ax^{k+1}+y^{k+1}-b,y^{k+1}-y^{k}\right\rangle.\end{aligned} \tag{15}$$

Due to ∇g is L_g-Lipschitz continuous and ∇H is L_h-Lipschitz continuous, we combining the lemma 1, then we have

$$g\left(y^{k+1}\right)-g\left(y^{k}\right)\le\left\langle\nabla g\left(y^{k+1}\right),y^{k+1}-y^{k}\right\rangle+\frac{L}{2}\left\|y^{k}-y^{k+1}\right\|^{2}. \tag{16}$$

$$\begin{aligned}&H\left(x^{k+1},y^{k+1}\right)-H\left(x^{k+1},y^{k}\right)\\&\le\left\langle\nabla_y H\left(x^{k+1},y^{k+1}\right),y^{k+1}-y^{k}\right\rangle+\frac{L}{2}\left\|y^{k}-y^{k+1}\right\|^2.\end{aligned}\tag{17}$$

From the fourth equation in the optimality condition (6) of the problem, we obtain

$$\nabla g\left(y^{k+1}\right)=\lambda^{k+\frac{1}{2}}-s\beta\left(Ax^{k+1}+y^{k+1}-b\right).\tag{18}$$

We obtain $\lambda^{k+1}=\nabla g\left(y^{k+1}\right)+\nabla_y H\left(x^{k+1},y^{k+1}\right)$ in (13) and combining the Lipschitz continuities of ∇g and ∇H, we have

$$\begin{aligned}\left\|\lambda^{k+1}-\lambda^{k}\right\|^2&=\left\|\nabla g\left(y^{k+1}\right)-\nabla g\left(y^{k}\right)+\nabla_y H\left(x^{k+1},y^{k+1}\right)-\nabla_y H\left(x^{k},y^{k}\right)\right\|^2\\&\le 2\left\|\nabla g\left(y^{k+1}\right)-\nabla g\left(y^{k}\right)\right\|^2+2\left\|\nabla_y H\left(x^{k+1},y^{k+1}\right)-\nabla_y H\left(x^{k},y^{k}\right)\right\|^2\\&\le 2L_g^2\left\|y^{k+1}-y^{k}\right\|^2+2L_h^2\left\|\left(x^{k+1},y^{k+1}\right)-\left(x^{k},y^{k}\right)\right\|^2\\&=2\left(L_g^2+L_h^2\right)\left\|y^{k+1}-y^{k}\right\|^2+2L_h^2\left\|x^{k+1}-x^{k}\right\|^2.\end{aligned}\tag{19}$$

Recall (9) and (10), we get

$$\begin{aligned}&\alpha\beta\left\|Ax^{k+1}+y^{k}-b\right\|^2+s\beta\left\|Ax^{k+1}+y^{k+1}-b\right\|^2\\&=\frac{1}{(\alpha+s)\beta}\left\|\lambda^{k}-\lambda^{k+1}\right\|^2+\frac{\alpha s\beta}{\alpha+s}\left\|y^{k}-y^{k+1}\right\|^2.\end{aligned}\tag{20}$$

From the definition of the augmented Lagrangian function $\mathcal{L}_\beta^s(\cdot)$ in (4), and in combination with (15), we have

$$\begin{aligned}&\mathcal{L}_\beta^s\left(x^{k+1},y^{k+1},\lambda^{k+\frac{1}{2}}\right)-\mathcal{L}_\beta^s\left(x^{k+1},y^{k},\lambda^{k+\frac{1}{2}}\right)\\&=g\left(y^{k+1}\right)-g\left(y^{k}\right)-\left\langle\lambda^{k+\frac{1}{2}},y^{k+1}-y^{k}\right\rangle+H\left(x^{k+1},y^{k+1}\right)-H\left(x^{k+1},y^{k}\right)\\&\quad+\frac{s\beta}{2}\left\|Ax^{k+1}+y^{k+1}-b\right\|^2-\frac{s\beta}{2}\left\|Ax^{k+1}+y^{k}-b\right\|^2\\&=g\left(y^{k+1}\right)-g\left(y^{k}\right)-\left\langle\lambda^{k+\frac{1}{2}},y^{k+1}-y^{k}\right\rangle+H\left(x^{k+1},y^{k+1}\right)-H\left(x^{k+1},y^{k}\right)\\&\quad-\frac{s\beta}{2}\left\|y^{k}-y^{k+1}\right\|^2+s\beta\left\langle Ax^{k+1}+y^{k+1}-b,y^{k+1}-y^{k}\right\rangle.\end{aligned}\tag{21}$$

Then, combining it with (16) and (17) yields

$$\begin{aligned}&\mathcal{L}_\beta^s\left(x^{k+1},y^{k+1},\lambda^{k+\frac{1}{2}}\right)-\mathcal{L}_\beta^s\left(x^{k+1},y^{k},\lambda^{k+\frac{1}{2}}\right)\\&\le\left\langle\nabla g\left(y^{k+1}\right),y^{k+1}-y^{k}\right\rangle+\frac{L_g}{2}\left\|y^{k}-y^{k+1}\right\|^2-\left\langle\lambda^{k+\frac{1}{2}},y^{k+1}-y^{k}\right\rangle\\&\quad+\left\langle\nabla_y H\left(x^{k+1},y^{k+1}\right),y^{k+1}-y^{k}\right\rangle+\frac{L_h}{2}\left\|y^{k}-y^{k+1}\right\|^2-\frac{s\beta}{2}\left\|y^{k}-y^{k+1}\right\|^2\\&\quad+s\beta\left\langle Ax^{k+1}+y^{k+1}-b,y^{k+1}-y^{k}\right\rangle.\end{aligned}\tag{22}$$

Substituting (13) into the above expression and simplifying, we obtain

$$\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k+1},\lambda^{k+\frac{1}{2}}\right)-\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k+\frac{1}{2}}\right)\leq\frac{L_g+L_h-s\beta}{2}\left\|y^{k+1}-y^{k}\right\|^2. \tag{23}$$

Note that, by using (4), (6), (20) and (19), we have

$$\begin{aligned}&\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k+1},\lambda^{k+1}\right)-\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k+1},\lambda^{k+\frac{1}{2}}\right)\\&+\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k+\frac{1}{2}}\right)-\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k}\right)\\&=\left\langle\lambda^{k+1}-\lambda^{k+\frac{1}{2}},Ax^{k+1}+y^{k+1}-b\right\rangle+\left\langle\lambda^{k+\frac{1}{2}}-\lambda^{k},Ax^{k+1}+y^{k}-b\right\rangle\\&=s\beta\left\|Ax^{k+1}+y^{k+1}-b\right\|^2+\alpha\beta\left\|Ax^{k+1}+y^{k}-b\right\|^2\\&=\frac{1}{(\alpha+s)\beta}\left\|\lambda^{k+1}-\lambda^{k}\right\|^2+\frac{\alpha s\beta}{\alpha+s}\left\|y^{k+1}-y^{k}\right\|^2\\&\leq\frac{2\left(L_g^2+L_h^2\right)+\alpha s\beta^2}{(\alpha+s)\beta}\left\|y^{k+1}-y^{k}\right\|^2+\frac{2L_h^2}{(\alpha+s)\beta}\left\|x^{k+1}-x^{k}\right\|^2.\end{aligned} \tag{24}$$

Summing the inequalities (23), (24) and the first equation in Section (5) we obtain

$$\begin{aligned}&\mathcal{L}_{\beta}^{s}\left(w^{k+1}\right)-\mathcal{L}_{\beta}^{s}\left(w^{k}\right)\\&=\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k+1},\lambda^{k+1}\right)-\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k+1},\lambda^{k+\frac{1}{2}}\right)+\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k+1},\lambda^{k+\frac{1}{2}}\right)\\&\quad-\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k+\frac{1}{2}}\right)+\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k+\frac{1}{2}}\right)-\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k}\right)\\&\quad+\mathcal{L}_{\beta}^{s}\left(x^{k+1},y^{k},\lambda^{k}\right)-\mathcal{L}_{\beta}^{s}\left(x^{k},y^{k},\lambda^{k}\right)\\&\leq\left[\frac{L_g+L_h-s\beta}{2}+\frac{2\left(L_g^2+L_h^2\right)+\alpha s\beta^2}{(\alpha+s)\beta}\right]\left\|y^{k+1}-y^{k}\right\|^2\\&\quad+\frac{2L_h^2}{(\alpha+s)\beta}\left\|x^{k+1}-x^{k}\right\|^2.\end{aligned} \tag{25}$$

Thus

$$\mathcal{L}_{\beta}^{s}\left(w^{k+1}\right)-\mathcal{L}_{\beta}^{s}\left(w^{k}\right)\leq-\eta\left(\left\|x^{k+1}-x^{k}\right\|^2+\left\|y^{k+1}-y^{k}\right\|^2\right), \tag{26}$$

there $\eta:=\min\left\{\frac{s\beta-L_g-L_h}{2}-\frac{2\left(L_g^2+L_h^2\right)+\alpha s\beta^2}{(\alpha+s)\beta},-\frac{2L_h^2}{(\alpha+s)\beta}\right\}$. This completes the proof. □

Remark 3.1. Thus, as long as $\eta>0$, the sequence $\mathcal{L}_{\beta}^{s}\left(w_{k+1}\right)$ has sufficient descent properties, which means that $\mathcal{L}_{\beta}^{s}\left(w_{k+1}\right)$ is monotonically nonincreasing. Therefore, we can achieve $\eta>0$ by appropriately choosing the parameters (α,s,β). The constraints on these parameters are as follows:

- $\{(\alpha, s) \in \mathbb{R}^2 \mid \alpha < s < 0\}$,
- $\left\{\beta \in \mathbb{R} \mid \beta > -\frac{1}{s}(L_g + L_h)\right\}$

Lemma 6. Let the iterative sequence generated by the algorithm (5) be denoted as $\{w^k : (x^k, y^k, \lambda^k)\}$. Suppose that this sequence is bounded, that Assumption A holds and $\eta > 0$. Then we have

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

Proof. The boundedness of $\{w^k\}$ implies the existence of a subsequence $\{w^{k_j}\}$ converging to w^*. Moreover, the lower semicontinuity of $f(x)$ together with the continuity of $g(y)$ ensures that $\mathcal{L}_\beta^s(\cdot)$ is lower semicontinuous. Hence,

$$\mathcal{L}_\beta^s(w^*) \le \liminf_{j \to +\infty} \mathcal{L}_\beta^s(w^{k_j}).$$

Hence, $\{\mathcal{L}_\beta^s(w^{k_j})\}$ is bounded below. The fact that it is also nonincreasing implies its convergence. Since $\{\mathcal{L}_\beta^s(w^k)\}$ is monotonic and contains a convergent subsequence, it follows that $\{\mathcal{L}_\beta^s(w^k)\}$ itself converges, satisfying $\mathcal{L}_\beta^s(w^k) \ge \mathcal{L}_\beta^s(w^*)$. Finally, invoking Equation (14) yields

$$\eta\left(\left\| x^{k+1} - x^k \right\|^2 + \left\| y^{k+1} - y^k \right\|^2\right) \le \mathcal{L}_\beta^s(w^k) - \mathcal{L}_\beta^s(w^{k+1}), \ \forall k.$$

By summing over $k = 0, \cdots, n$, and observing that $\mathcal{L}_\beta^s(w^0) < +\infty$, we arrive at

$$\begin{aligned} \eta \sum_{k=0}^{n}\left(\left\| x^{k+1} - x^k \right\|^2 + \left\| y^{k+1} - y^k \right\|^2\right) &\le \mathcal{L}_\beta^s(w^0) - \mathcal{L}_\beta^s(w^{n+1}) \\ &\le \mathcal{L}_\beta^s(w^0) - \mathcal{L}_\beta^s(w^*) < +\infty. \end{aligned}$$

The condition $\eta > 0$ yields $\sum_{k=0}^{\infty} \left\| y^{k+1} - y^k \right\|^2 < +\infty$ and $\sum_{k=0}^{\infty} \left\| x^{k+1} - x^k \right\|^2 < +\infty$. Using (19), we further obtain $\sum_{k=0}^{\infty} \left\| \lambda^{k+1} - \lambda^k \right\|^2 < +\infty$. Hence, $\sum_{k=0}^{+\infty} \left\| w^{k+1} - w^k \right\|^2 < +\infty$. This completes the proof. □

Lemma 7. Let the iterative sequence generated by the algorithm (5) be denoted as $\{w^k : (x^k, y^k, \lambda^k)\}$. Suppose that this sequence is bounded and that Assumption A holds. We define

$$\begin{cases} \varepsilon_1^{k+1} = A^{\mathrm{T}}(\lambda^k - \lambda^{k+1}) + \nabla_x H(x^{k+1}, y^{k+1}) - \nabla_x H(x^{k+1}, y^k) + s\beta(y^{k+1} - y^k), \\ \varepsilon_2^{k+1} = -\frac{s}{\alpha + s}(\lambda^{k+1} - \lambda^k) + \frac{\alpha s \beta}{\alpha + s}(y^{k+1} - y^k), \\ \varepsilon_3^{k+1} = \frac{1}{(\alpha + s)\beta}(\lambda^{k+1} - \lambda^k) - \frac{\alpha}{\alpha + s}(y^{k+1} - y^k). \end{cases} \tag{27}$$

Hence, it holds that $\varepsilon^{k+1} := (\varepsilon_1^{k+1}, \varepsilon_2^{k+1}, \varepsilon_3^{k+1}) \in \partial \mathcal{L}_\beta^s(w^{k+1})$, and there exists a constant $\delta > 0$ such that

$$d\left(0,\partial\mathcal{L}_\beta^s\left(w^{k+1}\right)\right)\le\delta\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right). \tag{28}$$

Proof. From the definition of the function $\mathcal{L}_\beta^s(\cdot)$ in (4), the following system of equations holds

$$\begin{cases}\partial_x\mathcal{L}_\beta^s\left(w^{k+1}\right)=\partial f\left(x^{k+1}\right)+\nabla_x H\left(x^{k+1},y^{k+1}\right)-A^{\mathrm{T}}\lambda^{k+1}+s\beta A^{\mathrm{T}}\left(Ax^{k+1}+y^{k+1}-b\right),\\ \partial_y\mathcal{L}_\beta^s\left(w^{k+1}\right)=\nabla g\left(y^{k+1}\right)+\nabla_y H\left(x^{k+1},y^{k+1}\right)-\lambda^{k+1}+s\beta\left(Ax^{k+1}+y^{k+1}-b\right),\\ \partial_\lambda\mathcal{L}_\beta^s\left(w^{k+1}\right)=-\left(Ax^{k+1}+y^{k+1}-b\right).\end{cases} \tag{29}$$

From the optimality condition (6), after rearrangement, we have

$$\begin{cases}A^{\mathrm{T}}\lambda^k-\nabla_x H\left(x^{k+1},y^k\right)-s\beta A^{\mathrm{T}}\left(Ax^{k+1}+y^k-b\right)\in\partial f\left(x^{k+1}\right),\\ \lambda^{k+\frac{1}{2}}-\nabla_y H\left(x^{k+1},y^{k+1}\right)-s\beta A^{\mathrm{T}}\left(Ax^{k+1}+y^{k+1}-b\right)=\nabla g\left(y^{k+1}\right),\\ \dfrac{1}{(\alpha+s)\beta}\left(\lambda^{k+1}-\lambda^k\right)-\dfrac{\alpha}{\alpha+s}\left(y^{k+1}-y^k\right)=-\left(Ax^{k+1}+y^{k+1}-b\right).\end{cases}$$

Then, by substituting the above into (29), we obtain

$$\begin{cases}A^{\mathrm{T}}\left(\lambda^k-\lambda^{k+1}\right)+\nabla_x H\left(x^{k+1},y^{k+1}\right)-\nabla_x H\left(x^{k+1},y^k\right)+s\beta\left(y^{k+1}-y^k\right)\in\partial_x\mathcal{L}_\beta^s\left(w^{k+1}\right),\\ -\dfrac{s}{\alpha+s}\left(\lambda^{k+1}-\lambda^k\right)+\dfrac{\alpha s\beta}{\alpha+s}\left(y^{k+1}-y^k\right)\in\partial_y\mathcal{L}_\beta^s\left(w^{k+1}\right),\\ \dfrac{1}{(\alpha+s)\beta}\left(\lambda^{k+1}-\lambda^k\right)-\dfrac{\alpha}{\alpha+s}\left(y^{k+1}-y^k\right)\in\partial_\lambda\mathcal{L}_\beta^s\left(w^{k+1}\right).\end{cases}$$

Consequently, applying Lemma 3 yields $\left(\varepsilon_1^{k+1},\varepsilon_2^{k+1},\varepsilon_3^{k+1}\right)\in\partial\mathcal{L}_\beta^s\left(w^{k+1}\right)$. Based on the preceding relation, we can find δ_1,δ_2 such that

$$\begin{aligned}\left\|\varepsilon^{k+1}\right\|^2&=\left\|\left(\varepsilon_1^{k+1},\varepsilon_2^{k+1},\varepsilon_3^{k+1}\right)\right\|^2\\&\le\left\|\varepsilon_1^{k+1}\right\|^2+\left\|\varepsilon_2^{k+1}\right\|^2+\left\|\varepsilon_3^{k+1}\right\|^2\\&\le\delta_1^2\left\|y^{k+1}-y^k\right\|^2+\delta_2^2\left\|\lambda^{k+1}-\lambda^k\right\|^2.\end{aligned} \tag{30}$$

Applying (19), there exists δ for which the following holds

$$d^2\left(0,\partial\mathcal{L}_\beta^s\left(w^{k+1}\right)\right)\le\left\|\varepsilon^{k+1}\right\|^2\le\delta^2\left(\left\|x^{k+1}-x^k\right\|^2+\left\|y^{k+1}-y^k\right\|^2\right),$$

then

$$d\left(0,\partial\mathcal{L}_\beta^s\left(w^{k+1}\right)\right)\le\delta\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right). \tag{31}$$

This completes the proof. □

Lemma 8. Let the iterative sequence generated by the algorithm (5) be denoted as $\left\{w^k:\left(x^k,y^k,\lambda^k\right)\right\}$. Suppose that this sequence is bounded and that Assumption A holds. Let Ω represent the set of all cluster points of the sequence $\left\{w^k\right\}$. Then the following statement is true

(i) Ω is a nonempty compact set, and

$$d\left(w^k,\Omega\right)\to 0,\text{ as }k\to+\infty;$$

(ii) $\Omega \subset \operatorname{crit}\mathcal{L}_\beta$, where $\operatorname{crit}\mathcal{L}_\beta$ denotes the set of all stationary points of $\mathcal{L}_\beta^s$;

(iii) $\mathcal{L}_\beta^s(\cdot)$ is finite and constant on Ω, which equals to

$$\inf_{k\in\mathbb{N}} \mathcal{L}_\beta^s\left(w^k\right) = \lim_{k\to+\infty} \mathcal{L}_\beta^s\left(w^k\right).$$

Proof. We now verify the above results one by one.

(i) This is immediate from the definition of limit points.

(ii) Suppose $w^* \in \Omega$. Then there exists a subsequence $\left\{w^{k_j}\right\}$ of $\left\{w^k\right\}$ such that $w^{k_j} \to w^*$. Applying Lemma 6 yields

$$\lim_{k\to+\infty} \left\|w^{k+1} - w^k\right\| = 0, \tag{32}$$

thus, $w^{k_j+1} \to w^*$. Noting that x^{k+1} minimize $\mathcal{L}_\beta^s\left(x, y^k, \lambda^k\right)$ with respect to x, we get

$$\mathcal{L}_\beta^s\left(x^{k+1}, y^k, \lambda^k\right) \le \mathcal{L}_\beta^s\left(x^*, y^k, \lambda^k\right). \tag{33}$$

With respect to the variables y, λ, and $\left(y^{k_j}, \lambda^{k_j}\right) \to \left(y^*, \lambda^*\right)$. Hence, it follows that

$$\limsup_{j\to+\infty} \mathcal{L}_\beta^s\left(x^{k_j+1}, y^{k_j}, \lambda^{k_j}\right) = \limsup_{j\to+\infty} \mathcal{L}_\beta^s\left(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}\right). \tag{34}$$

Furthermore, applying (33) yields

$$\limsup_{j\to+\infty} \mathcal{L}_\beta^s\left(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}\right) \le \mathcal{L}_\beta^s\left(x^*, y^*, \lambda^*\right). \tag{35}$$

Since $\mathcal{L}_\beta^s(\cdot)$ is lower semicontinuous, we know that

$$\limsup_{j\to+\infty} \mathcal{L}_\beta^s\left(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}\right) \ge \mathcal{L}_\beta^s\left(x^*, y^*, \lambda^*\right). \tag{36}$$

From (34), (35) and (36), we obtain

$$\lim_{j\to+\infty} f\left(x^{k_j+1}\right) = f\left(x^*\right).$$

Taking the limit in (6) along the subsequence $\left\{\left(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}\right)\right\}$ using (32) again, we obtain

$$\begin{cases} -\nabla_x H\left(x^*, y^*\right) + A^{\mathrm{T}}\lambda^* \in \partial f\left(x^*\right), \\ -\nabla_y H\left(x^*, y^*\right) + \lambda^* = \nabla g\left(y^*\right), \\ Ax^* + y^* - b = 0. \end{cases}$$

Therefore, $\left(x^*, y^*, \lambda^*\right)$ satisfies the critical point condition of (4), which implies that $w^* \in \operatorname{crit}\mathcal{L}_\beta^s$. Thus, $\Omega \subset \operatorname{crit}\mathcal{L}_\beta^s$.

(iii) Take any $\left(x^*, y^*, \lambda^*\right) \in \Omega$. Then there exists a subsequence $\left\{\left(x^{k_j}, y^{k_j}, \lambda^{k_j}\right)\right\}$, such that $\left\{\left(x^{k_j}, y^{k_j}, \lambda^{k_j}\right)\right\} \to \left(x^*, y^*, \lambda^*\right)$. Since $\mathcal{L}_\beta^s\left(w^k\right)$ is nonincreasing and has a convergent subsequence, the entire sequence $\mathcal{L}_\beta^s\left(w^k\right)$

converges, we have

$$\lim_{k\to+\infty}\mathcal{L}_\beta^s\left(x^k,y^k,\lambda^k\right)=\mathcal{L}_\beta^s\left(x^*,y^*,\lambda^*\right).$$

That is, $\mathcal{L}_\beta^s(\cdot)$ takes a constant value on Ω. Clearly,

$$\inf_{k\in\mathbb{N}}\mathcal{L}_\beta^s\left(w^k\right)=\lim_{k\to+\infty}\mathcal{L}_\beta^s\left(w^k\right).$$

This completes the proof. □

Theorem 9. Let the iterative sequence generated by the algorithm (5) be denoted as $\left\{w^k:\left(x^k,y^k,\lambda^k\right)\right\}$. Suppose that this sequence is bounded, that Assumption A holds and $\eta>0$. When $\mathcal{L}_\beta^s(\cdot)$ is a KL function, then $\left\{w^k\right\}$ has finite length, that is

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

Moreover, it converges to a critical point of $\mathcal{L}_\beta^s(\cdot)$.

Proof. Lemma 8 implies that $\mathcal{L}_\beta^s\left(w^k\right)\to\mathcal{L}_\beta^s\left(w^*\right)$ for any $w^*\in\Omega$. Next, we examine two cases.

(i) Suppose there exists k_0 with $\mathcal{L}_\beta^s\left(w^{k_0}\right)=\mathcal{L}_\beta^s\left(w^*\right)$, then using (14) and Remark 0.1, for every $k>k_0$, we have

$$\begin{aligned}\eta\left(\left\|x^{k+1}-x^k\right\|^2+\left\|y^{k+1}-y^k\right\|^2\right)&\le\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^{k+1}\right)\\&\le\mathcal{L}_\beta^s\left(w^{k_0}\right)-\mathcal{L}_\beta^s\left(w^*\right)=0.\end{aligned}$$

Hence, $y^{k+1}=y^k$ and $x^{k+1}=x^k$ for any $k>k_0$. Then, by (19), we further obtain $\lambda^{k+1}=\lambda^k$ for any $k>k_0+1$, which means $w^{k+1}=w^k$.

(ii) If $\mathcal{L}_\beta^s\left(w^k\right)>\mathcal{L}_\beta^s\left(w^*\right)$ holds for all k, then the following convergence properties hold:

- Since $d\left(w^k,\Omega\right)\to0$, for any $\varepsilon_1>0$ there exists $k_1>0$ such that for all $k>k_1$, it holds $d\left(w^k,\Omega\right)<\varepsilon_1$ is true.
- Since $\mathcal{L}_\beta^s\left(w^k\right)\to\mathcal{L}_\beta^s\left(w^*\right)$, for any $\varepsilon_2>0$ there exists $k_2>0$ such that for all $k>k_2$, it holds that $\mathcal{L}_\beta^s\left(w^k\right)<\mathcal{L}_\beta^s\left(w^*\right)+\varepsilon_2$ is true.

Now set $k>\tilde{k}=\max\left\{k_1,k_2\right\}$ and any $\varepsilon_1,\varepsilon_2>0$, we have

$$d\left(w^k,\Omega\right)<\varepsilon_1,\ \mathcal{L}_\beta^s\left(w^*\right)<\mathcal{L}_\beta^s\left(w^k\right)<\mathcal{L}_\beta^s\left(w^*\right)+\varepsilon_2.$$

By Lemma 8, we have established that Ω is a nonempty compact set and that $\mathcal{L}_\beta^s(\cdot)$ is constant Ω. Consequently, Lemma 2 implies that

$$\varphi'\left(\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)d\left(0,\partial\mathcal{L}_\beta^s\left(w^k\right)\right)\ge1,\ \forall k>\tilde{k}.\tag{37}$$

Drawing on what has been clearly established, namely that

$$\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^{k+1}\right)=\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)-\left(\mathcal{L}_\beta^s\left(w^{k+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right),$$

and the concavity of $\varphi(\cdot)$, it follows that

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

Now, taking the above inequality together with

$$d\left(0,\partial\mathcal{L}_\beta^s\left(w^k\right)\right)\leq \xi\left\|y^k-y^{k-1}\right\|,\varphi'\left(\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)>0,$$

and relation (37), it follows that

$$\begin{aligned}
&\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^{k+1}\right)\\
&\leq \frac{\varphi\left(\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)-\varphi\left(\mathcal{L}_\beta^s\left(w^{k+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)}{\varphi'\left(\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)}\\
&\leq d\left(0,\partial\mathcal{L}_\beta^s\left(w^k\right)\right)\left[\varphi\left(\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)-\varphi\left(\mathcal{L}_\beta^s\left(w^{k+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)\right]\\
&\leq \delta\left(\left\|x^k-x^{k-1}\right\|+\left\|y^k-y^{k-1}\right\|\right)\left[\varphi\left(\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)-\varphi\left(\mathcal{L}_\beta^s\left(w^{k+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)\right].
\end{aligned}$$

For convenience, we define

$$\Delta_{p,q}:=\varphi\left(\mathcal{L}_\beta^s\left(w^p\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)-\varphi\left(\mathcal{L}_\beta^s\left(w^q\right)-\mathcal{L}_\beta^s\left(w^*\right)\right).$$

Then the above equation is equivalent to

$$\mathcal{L}_\beta^s\left(w^k\right)-\mathcal{L}_\beta^s\left(w^{k+1}\right)\leq \delta\left(\left\|x^k-x^{k-1}\right\|+\left\|y^k-y^{k-1}\right\|\right)\Delta_{k,k+1}. \tag{38}$$

Together, Lemma 5 and (38), imply that

$$\eta\left(\left\|x^{k+1}-x^k\right\|^2+\left\|y^{k+1}-y^k\right\|^2\right)\leq \delta\left(\left\|x^k-x^{k-1}\right\|+\left\|y^k-y^{k-1}\right\|\right)\Delta_{k,k+1},\ k>\tilde{k},$$

together with

$$\frac{1}{2}\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right)^2\leq\left(\left\|x^{k+1}-x^k\right\|^2+\left\|y^{k+1}-y^k\right\|^2\right),$$

and thereby

$$2\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right)\leq 2\sqrt{\left\|x^k-x^{k-1}\right\|+\left\|y^k-y^{k-1}\right\|}\sqrt{\frac{2\delta}{\eta}\Delta_{k,k+1}},\ k>\tilde{k}.$$

Using the fact that $2\sqrt{\alpha\beta}\leq\alpha+\beta$, we have

$$2\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right)\leq\left\|x^k-x^{k-1}\right\|+\left\|y^k-y^{k-1}\right\|+\frac{2\delta}{\eta}\Delta_{k,k+1}. \tag{39}$$

Taking the sum of (39) over $k=\tilde{k}+1,\cdots,N$ gives

$$2\sum_{k=\tilde{k}+1}^{N}\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right)\leq\sum_{k=\tilde{k}+1}^{N}\left(\left\|x^k-x^{k-1}\right\|+\left\|y^k-y^{k-1}\right\|\right)+\frac{2\delta}{\eta}\Delta_{\tilde{k}+1,N+1}.$$

From Definition, we have $\varphi\left(\mathcal{L}_\beta^s\left(w^{N+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)>0$ from Definition 4. Rearranging terms and taking $N\to+\infty$ gives

$$\begin{aligned}
&\sum_{k=\tilde{k}+1}^{+\infty}\left(\left\|x^{k+1}-x^k\right\|+\left\|y^{k+1}-y^k\right\|\right)\\
&\leq\left\|x^{\tilde{k}+1}-x^{\tilde{k}}\right\|+\left\|y^{\tilde{k}+1}-y^{\tilde{k}}\right\|+\frac{2\delta}{\eta}\varphi\left(\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right).
\end{aligned} \tag{40}$$

Thus,

$$\sum_{k=0}^{+\infty}\left\|y^{k+1}-y^{k}\right\|<+\infty, \tag{41}$$

and

$$\sum_{k=0}^{+\infty}\left\|x^{k+1}-x^{k}\right\|<+\infty. \tag{42}$$

Substituting into (19) yields

$$\sum_{k=0}^{+\infty}\left\|\lambda^{k+1}-\lambda^{k}\right\|<+\infty. \tag{43}$$

Additionally, we note that

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

From (41), (42) and (43), we arrive at

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

Finally, $\left\{w^{k}\right\}$ converges to a critical point of $\mathcal{L}_{\beta}^{s}(\cdot)$ by Lemma 8. This completes the proof. □

Lemma 10. Let $\left\{w^{k}:=\left(x^{k}, y^{k}, \lambda^{k}\right)\right\}$ be the sequence generated by the symmetric ADMM (5), If at least one of the following statements holds

(i) $\liminf_{\|x\|\to+\infty} f(x)=+\infty$.

(ii) $\inf_{x} f(x)>-\infty$ and $\liminf_{\|y\|\to+\infty} g(y)=+\infty$.

Then we can conclude that the sequence $\left\{w^{k}:=\left(x^{k}, y^{k}, \lambda^{k}\right)\right\}$ is bounded.

Proof. In the first place, suppose that condition (i) holds. Based on Lemma 5, we arrive at

$$\mathcal{L}_{\beta}^{s}\left(x^{k}, y^{k}, \lambda^{k}\right) \le \mathcal{L}_{\beta}^{s}\left(x^{1}, y^{1}, \lambda^{1}\right).$$

By combining (4) with $\nabla g\left(y^{k+1}\right)=\lambda^{k+1}-\nabla_{y} H\left(x^{k+1}, y^{k+1}\right)$, we get

$$\begin{aligned}\mathcal{L}_{\beta}^{s}\left(x^{1}, y^{1}, \lambda^{1}\right) &\ge f\left(x^{k}\right)+g\left(y^{k}\right)+H\left(x^{k}, y^{k}\right)-\left\langle\lambda^{k}, Ax^{k}+y^{k}-b\right\rangle \\ &\quad+\frac{s\beta}{2}\left\|Ax^{k}+y^{k}-b\right\|^{2} \\ &= f\left(x^{k}\right)+g\left(y^{k}\right)+H\left(x^{k}, y^{k}\right) \\ &\quad-\left\langle\nabla g\left(y^{k}\right)+\nabla_{y} H\left(x^{k}, y^{k}\right), Ax^{k}+y^{k}-b\right\rangle+\frac{s\beta}{2}\left\|Ax^{k}+y^{k}-b\right\|^{2} \\ &\ge f\left(x^{k}\right)+\left(\frac{L_{g}}{2}+\frac{L_{h}}{2}+\frac{s\beta}{2}\right)\left\|Ax^{k}+y^{k}-b\right\|^{2}.\end{aligned}$$

Note that (i) implies that $\inf_{x} f(x)>-\infty$. Based on Assumption A, we can obtain $\frac{L_{g}}{2}+\frac{L_{h}}{2}+\frac{s\beta}{2}>0$. Thus, we can deduce that $\left\{x^{k}\right\}$ and $\left\{y^{k}\right\}$ are bounded. Hence, $\left\{\lambda^{k}\right\}$ is also bounded, and thus $\left\{w^{k}\right\}$ is bounded. This completes the proof. □

Theorem 11. (Convergence rate) Let the iterative sequence generated by the algorithm (5) be denoted as $\left\{w^k:\left(x^k,y^k,\lambda^k\right)\right\}$. Suppose that this sequence is bounded, that Assumption A holds and $\eta>0$, and $\left\{w^k:\left(x^k,y^k,\lambda^k\right)\right\}\to\left\{w^*=\left(x^*,y^*,\lambda^*\right)\right\}$. Assume that $\mathcal{L}_\beta^s(\cdot)$ possesses the KL property at $\left(x^*,y^*,\lambda^*\right)$, and that the corresponding function is given by $\varphi(s)=ct^{1-\theta}$, with $\theta\in[0,1)$, $c>0$. Then the following three statements hold

(i) If $\theta=0$, the sequence $\left\{w^k=\left(x^k,y^k,\lambda^k\right)\right\}$ converges in finitely many steps. This means we can find an index k with $w^k=w^*$.

(ii) If $\theta\in\left(0,\frac{1}{2}\right]$, then there is a constant $c_1>0$ and $\tau\in[0,1)$ so that

$$\left\|\left(x^k,y^k,\lambda^k\right)-\left(x^*,y^*,\lambda^*\right)\right\|\le c_1\tau^k.$$

(iii) If $\theta\in\left(\frac{1}{2},1\right)$, then there is a constant $c_2>0$ for which

$$\left\|\left(x^k,y^k,\lambda^k\right)-\left(x^*,y^*,\lambda^*\right)\right\|\le c_2k^{(\theta-1)/(2\theta-1)}.$$

Proof. For $\theta=0$, we have $\varphi(t)=ct$ and $\varphi'(t)=c$. Suppose, contrary to the claim, that $\left\{w^k=\left(x^k,y^k,\lambda^k\right)\right\}$ does not terminate in finitely many steps. Then, for large enough k, the KL property yields $cd\left(0,\partial\mathcal{L}_\beta^s\left(w^k\right)\right)\ge1$, which contradicts Lemma 7. Now let $\theta>0$ and set $\Delta_k=\sum_{i=k}^{+\infty}\left(\left\|x^{i+1}-x^i\right\|+\left\|y^{i+1}-y^i\right\|\right)$ for $k\ge0$. From (40) we deduce

$$\Delta_{\tilde{k}+1}\le\Delta_{\tilde{k}}-\Delta_{\tilde{k}+1}+\frac{2\delta}{\eta}\varphi\left(\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right).\tag{44}$$

Because $\mathcal{L}_\beta^s(\cdot)$ has the KL property at w^*, we have

$$\varphi'\left(\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)d\left(0,\partial\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)\right)\ge1.$$

This inequality can be rearranged as

$$\left(\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)^\theta\le c\cdot(1-\theta)d\left(0,\partial\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)\right).\tag{45}$$

Finally, applying Lemma 7 yields

$$d\left(0,\partial\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)\right)\le\delta\left(\left\|x^{\tilde{k}+1}-x^{\tilde{k}}\right\|+\left\|y^{\tilde{k}+1}-y^{\tilde{k}}\right\|\right)=\delta\left(\Delta_{\tilde{k}}-\Delta_{\tilde{k}+1}\right).\tag{46}$$

From (45) and (46), there exists $\gamma=\left[c(1-\theta)\delta\right]^{\frac{1-\theta}{\theta}}>0$ satisfying

$$\varphi\left(\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)=c\cdot\left(\mathcal{L}_\beta^s\left(w^{\tilde{k}+1}\right)-\mathcal{L}_\beta^s\left(w^*\right)\right)^{1-\theta}\le\gamma\left(\Delta_{\tilde{k}}-\Delta_{\tilde{k}+1}\right)^{(1-\theta)/\theta}.$$

Inserting this into (44) gives

$$\Delta_{\tilde{k}+1}\le\Delta_{\tilde{k}}-\Delta_{\tilde{k}+1}+\frac{2\delta}{\eta}\gamma\left(\Delta_{\tilde{k}}-\Delta_{\tilde{k}+1}\right)^{(1-\theta)/\theta}.\tag{47}$$

Now, by (47) and the results of Attouch and Bolte [25], we obtain

- Case 1: $\theta \in \left(0, \frac{1}{2}\right]$, then there exist $c_1 > 0$ and $\tau \in [0,1)$ such that

$$\left\|x^k - x^*\right\| + \left\|y^k - y^*\right\| \le c_1 \tau^k. \tag{48}$$

- Case 2: $\theta \in \left(\frac{1}{2}, 1\right)$, then there exists $c_2 > 0$ such that

$$\left\|x^k - x^*\right\| + \left\|y^k - y^*\right\| \le c_2 k^{\frac{\theta-1}{2\theta-1}}. \tag{49}$$

Next, applying (19) yields

$$\begin{aligned}\left\|\lambda^k - \lambda^*\right\| &\le \sqrt{2}\left[L_g^2 \left\|y^k - y^*\right\|^2 + L_h^2 \left\|y^k - y^*\right\|^2 + L_h^2 \left\|x^k - x^*\right\|^2\right]^{\frac{1}{2}} \\ &\le \sqrt{2}\left[L_g \left\|y^k - y^*\right\| + L_h^2 \left\|y^k - y^*\right\| + L_h \left\|x^k - x^*\right\|\right].\end{aligned} \tag{50}$$

Thus, (ii) and (iii) follow from (48)-(50).

4. Conclusion

In this paper, we propose a symmetric alternating direction method of multipliers with two different relaxation factors for minimizing the sum of two non-separable nonconvex functions. For problems where the objective function contains a coupling term, *i.e.*, the case where f and g are non-separable, research remains limited in both convex and non-convex settings. We review the development of symmetric ADMM for solving non-separable problems and find that although many existing works have proposed symmetric ADMM variants incorporating techniques such as Bregman distances, inertial terms, regularization terms, or linearization, they often introduce only one relaxation factor. Inspired by this, we introduce two different relaxation factors and apply the algorithm to non-separable nonconvex problems, thereby refining the basic form of symmetric ADMM for solving non-separable problems. This makes the parameter range of the algorithm broader, allowing it to be adapted to more practical problems by adjusting the parameters. It also provides fundamental theoretical support for further integration with other techniques. Finally, based on the Kurdyka-Łojasiewicz (KL) property, we prove that the sequence generated by the algorithm converges to a stationary point of the problem and further analyze its finite-step convergence, linear convergence, and sublinear convergence.

Acknowledgements

Sincere thanks to the members of JAMP for their professional performance, and special thanks to managing editor *Hellen XU* for a rare attitude of high quality.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Shang, F., Xu, T., Liu, Y., Liu, H., Shen, L. and Gong, M. (2021) Differentially Private

ADMM Algorithms for Machine Learning. *IEEE Transactions on Information Forensics and Security*, **16**, 4733-4745. https://doi.org/10.1109/tifs.2021.3113768

[2] Chan, S.H. (2019) Performance Analysis of Plug-And-Play ADMM: A Graph Signal Processing Perspective. *IEEE Transactions on Computational Imaging*, **5**, 274-286. https://doi.org/10.1109/tci.2019.2892123

[3] Xu, H., Caramanis, C. and Mannor, S. (2013) Outlier-robust PCA: The High-Dimensional Case. *IEEE Transactions on Information Theory*, **59**, 546-572. https://doi.org/10.1109/tit.2012.2212415

[4] Glowinski, R. and Marroco, A. (1975) Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *Revue française d automatique, informatique, recherche opérationnelle. Analyse numérique*, **9**, 41-76. https://doi.org/10.1051/m2an/197509r200411

[5] Douglas, J. and Rachford, H.H. (1956) On the Numerical Solution of Heat Conduction Problems in Two and Three Space Variables. *Transactions of the American Mathematical Society*, **82**, 421-439. https://doi.org/10.1090/s0002-9947-1956-0084194-4

[6] He, B., Liu, H., Wang, Z. and Yuan, X. (2014) A Strictly Contractive Peaceman—Rachford Splitting Method for Convex Programming. *SIAM Journal on Optimization*, **24**, 1011-1040. https://doi.org/10.1137/13090849x

[7] He, B., Ma, F. and Yuan, X. (2016) Convergence Study on the Symmetric Version of ADMM with Larger Step Sizes. *SIAM Journal on Imaging Sciences*, **9**, 1467-1501. https://doi.org/10.1137/15m1044448

[8] Jia, Z., Gao, X., Cai, X. and Han, D. (2021) The Convergence Rate Analysis of the Symmetric ADMM for the Nonconvex Separable Optimization Problems. *Journal of Industrial & Management Optimization*, **17**, 1943-1971. https://doi.org/10.3934/jimo.2020053

[9] Liu, P., Jian, J., He, B. and Jiang, X. (2022) Convergence of Bregman Peaceman-Rachford Splitting Method for Nonconvex Nonseparable Optimization. *Journal of the Operations Research Society of China*, **11**, 707-733. https://doi.org/10.1007/s40305-022-00411-x

[10] Gao, X. and Zhang, S. (2016) First-Order Algorithms for Convex Optimization with Nonseparable Objective and Coupled Constraints. *Journal of the Operations Research Society of China*, **5**, 131-159. https://doi.org/10.1007/s40305-016-0131-5

[11] Chen, C., Li, M., Liu, X. and Ye, Y. (2017) Extended ADMM and BCD for Nonseparable Convex Minimization Models with Quadratic Coupling Terms: Convergence Analysis and Insights. *Mathematical Programming*, **173**, 37-77. https://doi.org/10.1007/s10107-017-1205-9

[12] Guo, K. and Han, D.R. and Wu, T.T. (2018) Convergence Analysis for Optimization Problems with Nonseparable Nonconvex Objective and Linear Constraints. *Pacific Journal of Optimization*, **14**, 489-506. https://webofscience.clarivate.cn/wos/woscc/full-record/WOS:000461411400008

[13] Guo, K. and Wang, X. (2018) Convergence of Generalized Alternating Direction Method of Multipliers for Nonseparable Nonconvex Objective with Linear Constraints. *Journal of Mathematical Research with Applications*, **38**, 18. https://d.wanfangdata.com.cn/periodical/sxyjypl201805010

[14] Liu, Q., Shen, X. and Gu, Y. (2019) Linearized ADMM for Nonconvex Nonsmooth Optimization with Convergence Analysis. *IEEE Access*, **7**, 76131-76144. https://doi.org/10.1109/access.2019.2914461

[15] Wu, Z., Li, M., Wang, D.Z.W. and Han, D. (2017) A Symmetric Alternating Direction

Method of Multipliers for Separable Nonconvex Minimization Problems. *Asia-Pacific Journal of Operational Research*, **34**, Article ID: 1750030. https://doi.org/10.1142/s0217595917500300

[16] Dang, Y.Z. and Cui, T.T. (2023) Convergence Analysis of Linear Symmetric Proximal ADMM for Nonconvex Nonsmooth Nonseparable Optimization. *Journal of Systems Science and Mathematical Sciences*, **43**, 2949-2969. (In Chinese) http://dx.chinadoi.cn/10.12341/jssms23295

[17] Dang, Y. and Liu, K. (2026) A Two-Step Inertial Bregman Symmetric Admm-Type Algorithm with Kl-Property for Nonconvex Nonsmooth Nonseparable Optimization Problems with Application. *Journal of Computational and Applied Mathematics*, **472**, Article ID: 116815. https://doi.org/10.1016/j.cam.2025.116815

[18] Mei, L. and Ke, G. (2026) Convergence Analysis of the Symmetric Alternating Direction Method of Multipliers for Two-Block Separable Nonconvex Optimization Problems with Linear Constraints. *Pacific Journal of Optimization*. https://doi.org/10.61208/pjo-2026-007

[19] Beck, A. (2017) First-Order Methods in Optimization. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611974997

[20] Guo, K., Han, D.R. and Wu, T.T. (2016) Convergence of Alternating Direction Method for Minimizing Sum of Two Nonconvex Functions with Linear Constraints. *International Journal of Computer Mathematics*, **94**, 1653-1669. https://doi.org/10.1080/00207160.2016.1227432

[21] Nesterov, Y. (2013) Introductory Lectures on Convex Optimization: A Basic Course, Springer Science and Business Media.

[22] Attouch, H., Bolte, J., Redont, P. and Soubeyran, A. (2010) Proximal Alternating Minimization and Projection Methods for Nonconvex Problems: An Approach Based on the Kurdyka-Łojasiewicz Inequality. *Mathematics of Operations Research*, **35**, 438-457. https://doi.org/10.1287/moor.1100.0449

[23] Bolte, J., Sabach, S. and Teboulle, M. (2013) Proximal Alternating Linearized Minimization for Nonconvex and Nonsmooth Problems. *Mathematical Programming*, **146**, 459-494. https://doi.org/10.1007/s10107-013-0701-9

[24] Attouch, H., Bolte, J. and Svaiter, B.F. (2011) Convergence of Descent Methods for Semi-Algebraic and Tame Problems: Proximal Algorithms, Forward-Backward Splitting, and Regularized Gauss-Seidel Methods. *Mathematical Programming*, **137**, 91-129. https://doi.org/10.1007/s10107-011-0484-9

[25] Attouch, H. and Bolte, J. (2007) On the Convergence of the Proximal Algorithm for Nonsmooth Functions Involving Analytic Features. *Mathematical Programming*, **116**, 5-16. https://doi.org/10.1007/s10107-007-0133-5