A Linearized Alternating Direction Method of Multipliers for a Special Three-Block Nonconvex Optimization Problem of Background/Foreground Extraction

CHUN ZHANG1,2, YANHONG YANG3, ZEYAN WANG4, AND YONGXIN CHEN5

1School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, China
2Department of Basic Teaching, Jiangsu Vocational Institute of Commerce, Nanjing 211101, China
3Department of Mathematics, College of Taizhou, Nanjing Normal University, Taizhou 225300, China
4Department of Basic Courses, The PLA Army Engineering University, Nanjing 211101, China
5School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, China

Corresponding author: Yanhong Yang (yzynj@163.com)

ABSTRACT

In this paper, we focus on the three-block nonconvex optimization problem of background/foreground extraction from a blurred and noisy surveillance video. The coefficient matrices of the equality constraints are nonidentity matrices. Regarding the separable structure of the objective function and linear constraints, the benchmark solver for the problem is the alternating direction method of multipliers (ADMM). The computational challenge is that there is no closed-form solution to the subproblem of ADMM since the objective function is not differentiable and the coefficient matrices of the equality constraints are not identity matrices. In this paper, we propose a linearized ADMM by choosing the proximal terms appropriately and add the dual step size to make the proposed algorithm more flexible. Under proper assumptions and the associated function satisfying the Kurdyka-Łojasiewicz property, we show that the proposed algorithm converges to a critical point of the given problem. We apply the proposed algorithm to the background/foreground extraction and the numerical results are used to demonstrate the effectiveness of the proposed algorithm.

INDEX TERMS

Alternating direction method of multipliers, global convergence, image processing, Kurdyka-Łojasiewicz property, linear constraint, nonconvex optimization.

I. INTRODUCTION

In this research, we study a special kind of three-block separable nonconvex optimization problem of background/foreground extraction [5], [7], [8], [14], [20], [44], [45] which is used to detect moving objects in blurred and noisy surveillance videos acquired from static cameras, as well as other fields [1], [10], [49]. In general, the problem can be summarized as the following problem:

$$\min_{x,y,z} f(x) + g(y) + h(z)$$

s.t. $Ax + By + z = b,$ (1)

The associate editor coordinating the review of this manuscript and approving it for publication was Radu-Emil Precup.

where $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous, $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is convex and continuous, and $h : \mathbb{R}^l \to \mathbb{R}$ is Lipschitz continuous differentiable with the modulus $L > 0$, and $A \in \mathbb{R}^{l \times m}$ and $B \in \mathbb{R}^{l \times n}$ are the matrices representing a regular blurring operator for the blurred data $b \in \mathbb{R}^l$.

In recent years, a large variety of algorithms [4], [26], [32], [46], [47] for solving problem (1) have been studied. Due to its separable structure and linear equality constraints, perhaps the first choice is the alternating direction method of multipliers (ADMM) proposed in [29], which has been well studied in two-block convex problems [16], [17], [19], [24], [25], [43], [54] and two-block nonconvex problems [31], [35], [53]. ADMM can be regarded as a splitting version of the
classical augmented Lagrangian method in [26], [46], which transforms the original high-dimensional problems into low-dimensional problems. The augmented Lagrangian function $L_\beta(\cdot)$ of (1) is

$$L_\beta(x, y, z, \lambda) := f(x) + g(y) + h(z) - \langle \lambda, Ax + By + z - b \rangle + \frac{\beta}{2} \|Ax + By + z - b\|^2,$$

where $\beta > 0$ is the penalty parameter and $\lambda \in \mathbb{R}^l$ is the Lagrange multiplier. The scheme of the extended ADMM for the three-block nonconvex problem (1) is:

$$\begin{align*}
x^{k+1} &= \arg \min \{L_\beta(x, y^k, z^k, \lambda^k) \mid x \in \mathbb{R}^m\}, \quad (3a) \\
y^{k+1} &= \arg \min \{L_\beta(x^{k+1}, y, z^k, \lambda^k) \mid y \in \mathbb{R}^n\}, \quad (3b) \\
z^{k+1} &= \lambda^k - \beta(Ax^{k+1} + By^k + z^k - b), \quad (3c) \\
\lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + By^k + z^k - b). \quad (3d)
\end{align*}$$

Note that without any assumptions, it is difficult to analyze the convergence of ADMM for three-block convex or nonconvex optimization problems. In particular, Chen et al. [11] constructed a counterexample to prove that the direct extension of the multi-block convex optimization problem does not necessarily converge, which has aroused the attention of the majority of scholars. Some scholars adopt strategies that correct the output of (3) to generate a new iteration or change the iterative scheme to guarantee the convergence for three-block convex optimization problems [9], [12], [13], [15], [21], [22], [28], [36], [38], [39], [51], and others have extended it to three-block nonconvex optimization problems [27], [30], [42], [50], [55].

Due to the special structure of the problem (1), that is, $f$ and $g$ are not differentiable and the coefficient matrices of the equality constraints are not identity matrices, the subproblems (3a) and (3b) do not have closed-form solutions. To address the two-block convex problems of the so-called sparse group least absolute shrinkage and selection operator (SGLASSO) and the fused least absolute shrinkage and selection operator (FLASSO), Li et al. [34] embed linearization technology into ADMM approach due to the simplicity of the resolvent operator of the subproblem, that is, the quadratic term of the subproblems without a closed-form solution is linearized and thus the resulting subproblem has a closed-form solution. The numerical experiments in [34] show that the linearization technique is very simple and effective at addressing such problems. Meanwhile, some scholars have performed similar works on linearizing the augmented term of augmented Lagrangian function [25], [37], [52], [56] for convex problems. The linearization technique is significant due to it easing the numerical implementation; therefore, it is popular in a wide range of applications, especially in ADMM for convex problems, such as linearizing the differentiable part of the objective function [34], [40], [41], [48], [52].

Based on the simplicity and efficiency of the linearization technique in numerical calculation, we propose a new ADMM by introducing a linearization technique for the three-block nonconvex problem (1). First, the subproblem (3a) is specified as

$$x^{k+1} \in \arg \min \{f(x) + \frac{\beta}{2} \|Ax + By + z - b\|^2 \mid x \in \mathbb{R}^m\}. \quad (4)$$

Linearizing the quadratic term $\frac{\beta}{2} \|Ax + By + z - b\|^2$ in (4) results in a simple format

$$x^{k+1} \in \arg \min \{L_\beta(x, y^k, z^k, \lambda^k) + q_k \mid x \in \mathbb{R}^m\}, \quad (5)$$

where

$$q_k = \frac{1}{2} \|x - x^k\|^2_{rI - \beta A^TA} + \frac{\beta}{2} \|Ax^k + By^k + z^k - b\|^2 + \langle \lambda^k, Ax^k + By^k + z^k - b \rangle - \frac{\beta}{2} \|Ax^k + By^k + z^k - b\|^2$$

and $rI > \beta A^TA$. Substituting (6) into (5), along with $\frac{\beta}{2} \|Ax + By + z - b\|^2 + \langle \lambda^k, Ax + By + z - b \rangle - \frac{\beta}{2} \|Ax + By + z - b\|^2$ being a constant, yields

$$x^{k+1} \in \arg \min \{L_\beta(x, y^k, z^k, \lambda^k) + \frac{1}{2} \|x - x^k\|^2_{rI - \beta A^TA} \mid x \in \mathbb{R}^m\}.$$

The linearization technique makes it easier to get the optimal solution of the subproblem, and it even gets the closed-form solution.

In the same way, by linearizing the quadratic term of (3b), we propose a more general proximal linearized version of ADMM (LADMM) as follows in Algorithm 1.

By adding appropriate proximal terms, the subproblem of the algorithm has a closed-form solution, which improves the computational effect and the convergence condition is

$$x^{k+1} \in \arg \min \{L_\beta(x, y^k, z^k, \lambda^k) + q^k \mid x \in \mathbb{R}^m\},$$

and

$$\beta > 0.$$
Algorithm 1 (LADMM)

Input Initial point \((x^{0T}, y^{0T}, z^{0T}, λ^{0T})\), dual step size parameter \(α ∈ (0, \frac{1 + √5}{2})\), penalty parameter \(β > β := \sqrt{\frac{1}{4} + \frac{2c}{α}L + \frac{1}{2}L_c}\), where \(c := \max\{\frac{α^2}{1 - α^2 + α}, 1\}\), \(rt > βA^TA\), and let \(sI > βB^TB\), \(k = 0\).

while a termination criterion is not met, do

Step 1 Set

\[
\begin{align*}
x^{k+1} &\in \arg\min \{L_β(x, y^{k}, z^{k}, λ^k) \} & (7a) \\
&+ \frac{1}{2}\|x - x^k\|^2_{A^TβA} | x ∈ \mathbb{R}^m], \quad \gamma^{k+1} ∈ \arg\min \{L_β(x^k+1, y^{k}, z^{k}) \} (7b) &+ \frac{1}{2}\|y - y^{k}\|^2_{βB^T} | y ∈ \mathbb{R}^n], \quad z^{k+1} ∈ \arg\min \{L_β(x^k+1, y^{k+1}, z) \} (7c) &+ \frac{1}{2}\|z - z^{k}\|^2_{x^k} | z ∈ \mathbb{R}^l], \quad λ^{k+1} := λ^k - αβ(Ax^{k+1} + By^{k+1} + z^{k+1} - b) \} (7d)
\end{align*}
\]

Step 2. Set \(k := k + 1\).

end while

Output \((x^{kT}, y^{kT}, z^{kT}, λ^{kT})\)

Definition 1: We say that \(x^*\) is a critical point of the augmented Lagrangian function \(L_β(\cdot)\) defined in (2), if it satisfies

\[
\begin{align*}
A^Tλ^* &∈ ∂f(x^*), \quad B^Tλ^* ∈ ∂g(y^*), \quad λ^* ∈ ∇h(z^*), \quad Ax^* + By^* + z^* - b = 0.
\end{align*}
\]

The set of critical points of \(L_β(\cdot)\) is denoted by crit \(L_β\).

Remark 2: We denote \(Ω^* \subseteq Ω\) as the set whose elements are the optimal solutions of the augmented Lagrangian function. Throughout the paper, we assume that \(Ω^*\) is nonempty.

Definition 3: Let \(f : \mathbb{R}^n → (-∞, +∞]\) be a proper and lower semicontinuous function. For the given \(α, β ∈ \mathbb{R}\), \(α ≤ β\), \([α ≤ f ≤ β]\) is called sublevel set, if

\[
[α ≤ f ≤ β] := \{x ∈ \mathbb{R}^n : α ≤ f(x) ≤ β\}.
\]

We define \([α < f < β]\) similarly. The level set of \(f\) is simply denoted by

\[
[f = α] := \{x ∈ \mathbb{R}^n : f(x) = α\}.
\]

Definition 4 ([3, 6]): Let \(f : \mathbb{R}^n → (-∞, +∞]\) be a proper and lower semicontinuous function.

(i) For a given \(x ∈ dom f\), the Fréchet subdifferential of \(f\) at \(x\), written by \(∂f(x)\), is the set of all vectors \(u ∈ \mathbb{R}^m\) which satisfy

\[
\lim inf_{y → x, y ≠ x} \frac{f(y) - f(x) - \langle u, y - x \rangle}{\|y - x\|} ≥ 0,
\]

and we set \(\partial f = \emptyset\) when \(x \notin dom f\).

(ii) The limiting-subdifferential, or simply the subdifferential, of \(f\) at \(x\), written by \(\partial f(x)\), is defined by

\[
\partial f(x) := \{u ∈ \mathbb{R}^n : \exists x_k → x, f(x_k) → f(x) and \ \partial f(x_k) \ni u_k → u \ as k → ∞\}.
\]

(iii) A point \(x^*\) is called (limiting-)critical point or stationary point of \(f\) if it satisfies \(0 ∈ \partial f(x^*)\), and the set of critical points of \(f\) is denoted by crit \(f\).

The Definition 4 means that the property \(\partial f(x) ⊆ \partial f(x^*)\) holds immediately, and the first set is closed and convex while the second one is closed. We also use the notation \(dom(\partial f) := \{x ∈ \mathbb{R}^n : \partial f \neq \emptyset\}\). Indeed, the subdifferential (9) reduces to the derivative of \(f\) denoted as \(∇f\) if \(f\) is continuously differentiable. Furthermore, if \(g\) is a continuously differentiable function, then \(∂(f + g) = ∂f + ∇g\).

We give some properties about the convexity of functions below.

Definition 5: Let \(C\) be a convex subset of \(\mathbb{R}^n\) and let \(f : C → \mathbb{R}\) be a function. Then \(f\) is convex over \(C\), if

\[
∀ x, y ∈ C, ∀ λ ∈ [0, 1], \quad f(λ x + (1 - λ)y) ≤ λf(x) + (1 - λ)f(y).
\]

Lemma 6: Let \(C\) be a convex subset of \(\mathbb{R}^n\) and \(f : C → \mathbb{R}\).

(i) \(f\) is convex over \(C\) if and only if there exists a vector \(u ∈ ∂f(x)\) such that

\[
f(z) ≥ f(x) + \langle u, z - x \rangle, \quad ∀ x, z ∈ C.
\]

Moreover, if \(f\) is differentiable over \(\mathbb{R}^n\), \(u = ∇f(x)\).

(ii) \(f\) is strictly convex over \(C\) if and only if the above inequality is strict whenever \(x \neq z\).

We give an important property of the smooth function and omit its proof (see e.g. [6]).

Lemma 7 ([6]): Let \(f : \mathbb{R}^n → \mathbb{R}\) be a Lipschitz differentiable function with the modulus \(L > 0\). Then for any \(x, y ∈ \mathbb{R}^n\), we have

\[
|f(x) - f(y) - \langle x - y, ∇f(y) \rangle| ≤ \frac{L}{2} \|x - y\|^2
\]

and

\[
\|∇f(x) - ∇f(y)\| ≤ \frac{L}{2} \|x - y\|^2.
\]
Next, we give the definition of Kurdyka-Łojasiewicz property which plays a central role in our further convergence analysis. Now we proceed to give the definition of Kurdyka-
Łojasiewicz (KL) property formally.

Definition 8 ([2], [6]): We say that a proper function \( f : \mathbb{R}^n \to (-\infty, +\infty] \) has Kurdyka-
Łojasiewicz property at \( x^* \in \text{dom}(\partial f) \) if there exist \( \eta \in (0, +\infty) \), a neighborhood \( U \) of \( x^* \) and a continuous and concave function \( \varphi : [0, \eta) \to \mathbb{R}_+ \) such that

(i) \( \varphi(0) = 0 \) and \( \varphi \) is continuously differentiable on \( (0, \eta) \) with \( \varphi' > 0 \).
(ii) for all \( x \in U \cap \{x \in \mathbb{R}^n : f(x) < f(x) < f(x) + \eta \} \), the following inequality holds

\[
\varphi'(f(x) - f(x^*))d(0, \partial f(x)) \geq 1.
\]

A proper lower semicontinuous \( f \) satisfies the KL property at each point of \( \text{dom}(\partial f) \) is called a KL function.

We denote \( \Phi_n \) be the set of functions which satisfy Definition 8. Now we give an important lemma which was established in [6] for the KL property, and it will be useful for further convergence analysis.

Lemma 9 ([6]): Let \( \Omega \) be a compact set and \( f : \mathbb{R}^n \to (-\infty, +\infty] \) be a proper and lower semicontinuous function. Assume that \( f \) is a constant on \( \Omega \) and satisfies that KL property at each point of \( \Omega \) Then there exist \( \varepsilon > 0 \), \( \eta > 0 \) and \( \varphi \in \Phi_n \) such that for all \( \bar{x} \in \Omega \) and all \( x \) in the following intersection

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

one has

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

III. CONVERGENCE ANALYSIS

Before proceeding the analysis, we first make the following assumptions.

Assumption 10: From LADMM I, A, B, \( \alpha, \beta, r, \) and \( s \) are satisfy

(i) The optimal set of (1) is \( \Omega^* \), and we assume \( \Omega^* \neq \emptyset \).
(ii) \( \alpha \in (0, \frac{1+\sqrt{3}}{2}) \).
(iii) \( \beta > \beta \) := \( \sqrt{\frac{1}{3} + \frac{2}{\alpha} L} + \frac{1}{2} L \), where \( c := \max\{\frac{\alpha^3}{1-\alpha^2+\alpha}, 1\} \).
(iv) \( M = \begin{bmatrix} rI - \beta A^T A & 0 & 0 \\ 0 & sI - \frac{\beta}{2} B^T B & 0 \\ 0 & 0 & \alpha \beta (\beta - L) - 2 \alpha L^2 \end{bmatrix} > 0 \).

From the first order optimality condition of LADMM 1, we have

\[
0 \in \partial f(x^{k+1}) - A^T \lambda^{k+1} + (1 - \frac{1}{\alpha}) A^T (\lambda^{k+1} - \lambda^k) - \beta A^T (By^{k+1} - B^k) - \beta A^T (z^{k+1} - z^k) + (rI - \beta A^T A)(x^{k+1} - x^k),
\]

\[
0 \in \partial g(y^{k+1}) - B^T \lambda^{k+1} + (1 - \frac{1}{\alpha}) B^T (\lambda^{k+1} - \lambda^k) - \beta B^T (z^{k+1} - z^k) + (sI - \beta B^T B)(y^{k+1} - y^k),
\]

\[
0 = \nabla h(z^{k+1}) - \lambda^{k+1} + (1 - \frac{1}{\alpha}) (\lambda^{k+1} - \lambda^k),
\]

\[
0 = \lambda^{k+1} - \lambda^k + \alpha \beta (Ax^{k+1} + By^{k+1} + z^{k+1} - b).
\]

Lemma 11: Let \( \{\omega^k\} \) and \( \{\nu^k\} \) be the sequences generated by LADMM 1 and suppose that Assumption 10 holds. Then, we obtain

\[
L_\beta(x^{k+1}, y^{k+1}, z^{k+1}, \lambda^k) - L_\beta(x^{k+1}, y^{k+1}, z^{k+1}, \lambda^k) \geq \|x^{k+1} - x^k\|^2_{M_1} + \|y^{k+1} - y^k\|^2_{M_2} + \|z^{k+1} - z^k\|^2_{M_3} + \|\lambda^{k+1} - \lambda^k\|^2_{M_4},
\]

\[
M_1 = \frac{r \beta}{2} I - \frac{s \beta}{2} B^T B, M_2 = \frac{r \beta}{2} I - \frac{s \beta}{2} B^T B, M_3 = \frac{r \beta}{2} I - \frac{s \beta}{2} B^T B, M_4 = \frac{r \beta}{2} I - \frac{s \beta}{2} B^T B.
\]

Proof: From the definition of augmented Lagrangian function \( L_\beta(\cdot) \) in (2), we have

\[
L_\beta(x^{k+1} + y^{k+1} + z^{k+1} + \lambda^k) \leq L_\beta(x^{k+1}, y^{k+1}, z^{k+1}, \lambda^k) \leq \lambda^{k+1} - \lambda^k, \quad - \frac{1}{\alpha \beta} (\lambda^{k+1} - \lambda^k)
\]

\[
= - \frac{1}{\alpha \beta} \|\lambda^{k+1} - \lambda^k\|^2,
\]

where the second equality follows from (7g). Similarly,

\[
L_\beta(x^{k+1} + y^{k+1} + z^{k+1} + \lambda^k) = h(z^{k+1}) - h(z^{k+1}) + (\lambda^k, z^{k+1} - z^k)
\]

\[
+ \beta \|Ax^{k+1} + By^{k+1} + z^{k+1} - b\|^2
\]

\[
+ \beta \|Ax^{k+1} + By^{k+1} + z^{k+1} - b\|^2
\]

\[
\geq (z^k - z^{k+1}, \nabla h(z^{k+1})) \geq \|z^k - z^{k+1}\|^2
\]

\[
+ (\lambda^k, z^{k+1} - z^k) - \frac{1}{\alpha} \|z^{k+1} - z^k\|^2
\]

\[
- \beta (Ax^{k+1} + By^{k+1} + z^{k+1} - b)
\]

\[
\geq (z^k - z^{k+1}, \nabla h(z^{k+1})) \geq \|z^k - z^{k+1}\|^2
\]

\[
+ (\lambda^k, z^{k+1} - z^k) - \frac{1}{\alpha} \|z^{k+1} - z^k\|^2
\]

\[
- \beta (Ax^{k+1} + By^{k+1} + z^{k+1} - b)
\]

\[
+ \beta \|z^{k+1} - z^k\|^2
\]

\[
+ \beta \|z^{k+1} - z^k\|^2
\]

\[
+ \beta \|z^{k+1} - z^k\|^2
\]

\[
+ \beta \|z^{k+1} - z^k\|^2
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\langle \beta \xe - \beta k, Byk + Byk - Byk \rangle \\
= \|Byk + Byk - Byk\|^2 \\
\geq -\beta B^T (\xe - \xe) + (sI - \beta B^T B)(\ye - \ye)^2 \\
+ (\lambda, Byk + Byk - Byk) + \frac{\beta}{2} \|Byk + Byk - Byk\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\ye - \ye\|^2 \\
\leq \frac{1}{\alpha} \|\xe - \xe\|^2.
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\|\xe - \xe\|^2 \\
\leq L^2 \alpha (\xe - \xe) + \|Byk + Byk - Byk\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\xe - \xe\|^2 \\
\leq \frac{1}{\alpha} \|\xe - \xe\|^2.
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\langle \beta \xe - \beta \lambda, Byk + Byk - Byk \rangle \\
= \|Byk + Byk - Byk\|^2 \\
\geq -\beta B^T (\xe - \xe) + (sI - \beta B^T B)(\ye - \ye)^2 \\
+ (\lambda, Byk + Byk - Byk) + \frac{\beta}{2} \|Byk + Byk - Byk\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\ye - \ye\|^2 \\
\leq \frac{1}{\alpha} \|\xe - \xe\|^2.
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\|L^2 \alpha (\xe - \xe) + \|Byk + Byk - Byk\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\xe - \xe\|^2 \\
\leq \frac{1}{\alpha} \|\xe - \xe\|^2.
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\|\lambda k + \lambda k\|^2 \\
\geq \frac{\alpha}{\alpha - \alpha} \|\lambda k + \lambda k\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\xe - \xe\|^2 \\
\leq \frac{1}{\alpha} \|\xe - \xe\|^2.
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\langle \beta \xe - \beta \lambda, Byk + Byk - Byk \rangle \\
= \|Byk + Byk - Byk\|^2 \\
\geq -\beta B^T (\xe - \xe) + (sI - \beta B^T B)(\ye - \ye)^2 \\
+ (\lambda, Byk + Byk - Byk) + \frac{\beta}{2} \|Byk + Byk - Byk\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\ye - \ye\|^2 \\
\leq \frac{1}{\alpha} \|\xe - \xe\|^2.
\]
where the inequality follows from \((11)\). From \((7b)\), we get
\[
\langle \beta \xe - \beta \lambda, Byk + Byk - Byk \rangle \\
= \|Byk + Byk - Byk\|^2 \\
\geq -\beta B^T (\xe - \xe) + (sI - \beta B^T B)(\ye - \ye)^2 \\
+ (\lambda, Byk + Byk - Byk) + \frac{\beta}{2} \|Byk + Byk - Byk\|^2 \\
+ \frac{1}{\alpha} (\lambda - \lambda - Byk + Byk) + Byk + Byk - Byk \\
= \|\ye - \ye\|^2 \\
Then adding $-\alpha(\alpha - 1)\|\lambda^{k+1} - \lambda^k\|^2$ to both sides of the above inequality, simplifying the resulting and using the fact that

$$(1 - \alpha^2 + \alpha\alpha)\|\lambda^{k+1} - \lambda^k\|^2$$

$$\leq \alpha^3L^2\|z^{k+1} - z^k\|^2 + \alpha(\alpha - 1)\|\lambda^{k} - \lambda^{k-1}\|^2$$

$$\Rightarrow \|\lambda^{k+1} - \lambda^k\|^2,$$

i.e.,

$$\|\lambda^{k+1} - \lambda^k\|^2$$

$$\leq \frac{\alpha^3}{1 - \alpha^2 + \alpha}\|z^{k+1} - z^k\|^2$$

$$\frac{\alpha(\alpha - 1)}{1 - \alpha^2 + \alpha}I\|\lambda^{k} - \lambda^{k-1}\|^2$$

$$\|\lambda^{k+1} - \lambda^k\|^2$$

$$\leq \max\{\frac{\alpha^3}{1 - \alpha^2 + \alpha}, 1\}L^2\|z^{k+1} - z^k\|^2$$

$$+ \max\{\frac{\alpha(\alpha - 1)}{1 - \alpha^2 + \alpha}, 1\}I\|\lambda^{k} - \lambda^{k-1}\|^2$$

$$\|\lambda^{k+1} - \lambda^k\|^2$$

$$\leq \max\{\frac{\alpha^3}{1 - \alpha^2 + \alpha}, 1\}L^2\|z^{k+1} - z^k\|^2$$

$$+ \max\{\frac{\alpha^2(1 - \alpha)}{1 - \alpha^2 + \alpha}, 1\}I\|\lambda^{k} - \lambda^{k-1}\|^2$$

$$\|\lambda^{k+1} - \lambda^k\|^2.$$

where the first equality follows from (7g). Dividing the two sides of the above inequality by $-\alpha\beta$ and recall the definition of $\theta(\alpha)$, we get

$$\frac{1}{\alpha\beta}\|\lambda^{k+1} - \lambda^k\|^2$$

$$\geq -\frac{1}{\alpha\beta}\max\{\frac{\alpha^3}{1 - \alpha^2 + \alpha}, 1\}L^2\|z^{k+1} - z^k\|^2$$

$$- \frac{1}{\alpha\beta}\max\{\frac{\alpha^2(1 - \alpha)}{1 - \alpha^2 + \alpha}, 1\}I\|\lambda^{k} - \lambda^{k-1}\|^2$$

$$\|\lambda^{k+1} - \lambda^k\|^2$$

$$\geq -\frac{\alpha^3}{2}\|z^{k+1} - z^k\|^2 - \frac{\alpha^2(1 - \alpha)}{2}\|\lambda^{k} - \lambda^{k-1}\|^2.$$

where $c := \max\{\frac{\alpha^3}{1 - \alpha^2 + \alpha}, 1\}$ and $\theta(\alpha) := \max\{\frac{\alpha^2(1 - \alpha)}{1 - \alpha^2 + \alpha}, 1\}$, the proof is completed.

**Lemma 13:** Let $\{\omega^k\}$ be the sequence generated by (1) and suppose $0 < \alpha < 1 + \frac{1}{2\sqrt{2}}$. If there exists some positive constant $\delta$ and Assumption 10 holds, for $k \geq 1$, we have

$$m(\omega^k) - m(\omega^{k+1}) \geq \delta\|v^{k+1} - v^k\|^2. \quad (26)$$

Moreover, since $\delta > 0$, the sequence $\{m(\omega^k)\}$ is monotonically nonincreasing.

**Proof:** From the definition of $m(-)$ (21) and (16), we get

$$m(x^k, y^k, z^k, \lambda^k) = m(x^{k+1}, y^{k+1}, z^{k+1}, \lambda^{k+1})$$

$$L_\beta(x^k, y^k, z^k, \lambda^k) + \theta(\alpha)\beta\|Ax^k + By^k + z^k - b\|^2$$

$$\geq \|x^k - x^k\|^2_M + \|y^k - y^k\|^2_M + \|z^k - z^k\|^2_M$$

$$+ \delta\|x^k - x^k\|^2_M + \|y^k - y^k\|^2_M + \|z^k - z^k\|^2_M$$

$$\|\lambda^{k+1} - \lambda^k\|^2,$$

where $M_\varepsilon = M_M - \frac{\varepsilon^2}{2\alpha\beta}I$. Therefore, from Assumption 10 (iii), there exists positive constant $\delta$, such that

$$m(x^k, y^k, z^k, \lambda^k) - m(x^{k+1}, y^{k+1}, z^{k+1}, \lambda^{k+1})$$

$$\geq \delta\|x^k - x^k\|^2_M + \|y^k - y^k\|^2_M + \|z^k - z^k\|^2_M.$$

Further, the sequence $\{m(\omega^k)\}$ is monotonically nonincreasing. The proof is completed.

Before getting the convergence, we prove the boundedness of the sequence $\{\omega^k\}$ generated by (1).

**Lemma 14:** Let $\{\omega^k\}$ be the sequence generated by LADMM (1) and suppose that Assumption 10 holds. Assume that

$$\hat{h} = \inf_z (h(z) - \max\{\alpha, \frac{1}{2}\alpha\beta\|\nabla h(z)\|^2\}) \geq -\infty,$$

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

and

$$\liminf_{\|y\| \to \infty} g(y) = +\infty.$$  \quad (27)

Thus, the sequence $\{\omega^k\}$ is bounded.

**Proof:** Before prove the boundedness of the sequence $\{\omega^k\}$, we have derive an upper bound of $\|\lambda^k\|^2$. Form (15c), we have

$$\nabla h(z^k) - \lambda^k + (1 - \alpha)(\lambda^k - \lambda^{k-1}) = 0,$$
Thus, combining (31) and (32), we get

\[ \alpha \lambda^k = \alpha \nabla h(z^k) + (1 - \alpha)(\lambda^{k-1} - \lambda^k). \]  

(30)

Now, we prove the boundedness of \( \|\lambda^k\| \) in two respects:

(i) For \( 0 < \alpha \leq 1 \),

\[ \|\alpha \lambda^k\|^2 \leq \alpha \|\nabla h(z^k)\|^2 + (1 - \alpha)\|\lambda^k - \lambda^{k-1}\|^2 \]

\[ = \alpha \|\nabla h(z^k)\|^2 + (1 - \alpha)\alpha^2 \beta^2 \|Ax^k + By^k + z^k - b\|^2, \]

where the inequality follows from the convexity of \( \|\cdot\|^2 \), (10) and the equality follows from (7g), i.e.,

\[ \|\lambda^k\|^2 \leq \frac{1}{\alpha} \|\nabla h(z^k)\|^2 + (1 - \alpha)\beta^2 \|Ax_k + By^k + z^k - b\|^2. \]  

(31)

(ii) For \( 1 < \alpha \leq \frac{1 + \sqrt{3}}{2} \), from (30), we obtain

\[ \lambda^k = \frac{1}{\alpha}(\alpha \nabla h(z^k)) + (1 - \frac{1}{\alpha})(\lambda^k - \lambda^{k-1}). \]

Using the convexity of \( \|\cdot\|^2 \), (10), we have

\[ \|\lambda^k\|^2 \leq \frac{1}{\alpha} \alpha \|\nabla h(z^k)\|^2 + (1 - \frac{1}{\alpha})\alpha \|\lambda^k - \lambda^{k-1}\|^2 \]

\[ = \frac{1}{\alpha} \alpha \|\nabla h(z^k)\|^2 + (1 - \frac{1}{\alpha})\alpha \alpha^2 \beta^2 \|Ax^k + By^k + z^k - b\|^2 \]

\[ = \alpha \|\nabla h(z^k)\|^2 + (1 - \alpha)\alpha \beta^2 \|Ax^k + By^k + z^k - b\|^2, \]  

(32)

where the first equality follows from (7g).

Thus, combining (31) and (32), we get

\[ \|\lambda^k\|^2 \leq \max\{\frac{1}{\alpha} \|\nabla h(z^k)\|^2 \]

\[ + \max\{1 - \alpha, (\alpha - 1)\alpha\} \beta^2 \|Ax^k + By^k + z^k - b\|^2, \]

i.e.,

\[ - \frac{1}{2\beta} \|\lambda^k\|^2 \]

\[ \geq - \max\{\frac{1}{\alpha} \|\nabla h(z^k)\|^2 \]

\[ + \max\{1 - \alpha, (\alpha - 1)\alpha\} \beta^2 \|Ax^k + By^k + z^k - b\|^2, \]

\[ \geq - \max\{\frac{1}{\alpha} \|\nabla h(z^k)\|^2 \]

\[ + \max\{1 - \alpha, (\alpha - 1)\alpha\} \beta^2 \|Ax^k + By^k + z^k - b\|^2, \]

(33)

which implies that the boundedness of \( \{\lambda^k\} \) can be deduced from the boundedness of \( \{z^k\} \).

From (26), we have

\[ m(x^1, y^1, z^1, \lambda) \geq m(x^k, y^k, z^k, \lambda^k) \]

\[ = f(x^k) + g(y^k) + h(z^k) - (\lambda^k, Ax^k + By^k + z^k - b) \]

\[ + \frac{\beta}{2} \|Ax^k + By^k + z^k - b\|^2 \]

\[ + \theta(\alpha)\beta \|Ax^k + By^k + z^k - b\|^2 \]

\[ = f(x^k) + g(y^k) + h(z^k) + \frac{\beta}{2} \|Ax^k + By^k + z^k - b - \frac{1}{\beta} \lambda^k\|^2 \]

\[ - \frac{1}{2\beta} \|\lambda^k\|^2 + \theta(\alpha)\beta \|Ax^k + By^k + z^k - b\|^2 \]

\[ \geq f(x^k) + g(y^k) + h(z^k) + \frac{\beta}{2} \|Ax^k + By^k + z^k - b - \frac{1}{\beta} \lambda^k\|^2 \]

\[ + \theta(\alpha)\beta \|Ax^k + By^k + z^k - b\|^2 \]

\[ - \max\{\frac{1}{\alpha} \|\nabla h(z^k)\|^2 \]

\[ - \max\{1 - \alpha, (\alpha - 1)\alpha\} \beta \|Ax^k + By^k + z^k - b\|^2 \]

\[ = f(x^k) + g(y^k) + h(z^k) - \max\{\frac{1}{\alpha} \|\nabla h(z^k)\|^2 \]

\[ + \frac{\beta}{2} \|Ax^k + By^k + z^k - b - \frac{1}{\beta} \lambda^k\|^2 \]

\[ + (\theta(\alpha) - \frac{1}{2} \max\{1 - \alpha, (\alpha - 1)\alpha\}) \beta \|Ax^k + By^k + z^k - b\|^2 \]

\[ = f(x^k) + g(y^k) + h(z^k) + \frac{\beta}{2} \|Ax^k + By^k + z^k - b - \frac{1}{\beta} \lambda^k\|^2 \]

\[ + (\theta(\alpha) - \frac{1}{2} \max\{1 - \alpha, (\alpha - 1)\alpha\}) \beta \|Ax^k + By^k + z^k - b\|^2, \]

where the inequality follows from (33) and the last equality follows from (27). Since \( \alpha \in (0, \frac{1 + \sqrt{3}}{2}) \) and \( \beta > \frac{1}{\beta} \), we have

\[ \theta(\alpha) - \frac{1}{2} \max\{1 - \alpha, (\alpha - 1)\alpha\} \]

\[ = \begin{cases} \frac{1}{2}(1 - \alpha), & \alpha \in (0, 1), \\ 0, & \alpha = 1, \\ \frac{\alpha(\alpha - 1)(\alpha^2 + 1 - \alpha)}{2(1 - \alpha^2)}, & \alpha \in (1, \frac{1 + \sqrt{3}}{2}) \end{cases} \]

Consequently, \( \theta(\alpha) - \frac{1}{2} \max\{1 - \alpha, (\alpha - 1)\alpha\} \geq 0 \). From (28) and (29), we have \( \liminf_{n \rightarrow +\infty} f(x) = +\infty \), \( \liminf_{n \rightarrow +\infty} g(y) = +\infty \). Together with (27), we have \( \{z^k\}, \{y^k\}, \{Ax^k + By^k + z^k - b\}, \{Ax^k + By^k + z^k - b - \frac{1}{\beta} \lambda^k\} \) are bounded. And hence, we have \( \{\lambda^k\} \) and \( \{z^k\} \) are bounded. The proof is completed.

We prove the bound of \( \|\lambda^{k+1} - \lambda^k\| \) below.

Lemma 15: Let \( \{\alpha^k\} \) be the sequence generated by LADMM (1), there exists positive constant \( \delta_1 \), such that

\[ \|\lambda^{k+1} - \lambda^k\| \leq \delta_1 \|A^{k+1} - I\|^2. \]  

(34)

Proof: From (30) and through some simple calculations, we have

\[ \lambda^{k+1} - \lambda^k \]

\[ = \alpha(\nabla h(z^{k+1}) - \nabla h(z^k)) + (1 - \alpha)(\lambda^k - \lambda^{k-1}). \]  

(35)

We prove the conclusion (34) from two aspects:

(i) For \( 0 < \alpha \leq 1 \), it follows from trigonometric inequality that

\[ \|\lambda^{k+1} - \lambda^k\| \]

\[ \leq \|\alpha \nabla h(z^{k+1}) - \nabla h(z^k)\| + \|(1 - \alpha)(\lambda^k - \lambda^{k-1})\| \]

\[ = \alpha \|\nabla h(z^{k+1}) - \nabla h(z^k)\| + (1 - \alpha)\|\lambda^k - \lambda^{k-1}\| \]

\[ \leq \alpha L \|z^{k+1} - z^k\| + (1 - \alpha)\|\lambda^k - \lambda^{k-1}\|, \]
where the second inequality follows from \( h(z) \) is Lipschitz continuous differentiable with the modulus \( L \).
Add \((\alpha - 1)\|\lambda^{k+1} - \lambda^k\|\) to both sides of the above inequality, we get
\[
\alpha \|\lambda^{k+1} - \lambda^k\| \\
\leq \alpha L \|z^{k+1} - z^k\| \\
+ (1 - \alpha) (\|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\|),
\]
i.e.,
\[
\|\lambda^{k+1} - \lambda^k\| \\
\leq L \|z^{k+1} - z^k\| \\
+ \frac{1}{\alpha} (\|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\|).
\]
(36)

(ii) For \( 1 < \alpha < \frac{1 + \sqrt{5}}{2} \), from (35), we have
\[
\|\lambda^{k+1} - \lambda^k\| \\
= \|\alpha (\nabla h(z^{k+1}) - \nabla h(z^k)) + (1 - \alpha) (\lambda^k - \lambda^{k-1})\| \\
\leq \alpha \|\nabla h(z^{k+1}) - \nabla h(z^k)\| + (1 - \alpha) \|\lambda^k - \lambda^{k-1}\| \\
\leq \alpha L \|z^{k+1} - z^k\| + (1 - \alpha) \|\lambda^k - \lambda^{k-1}\|,
\]
where the first inequality follows from trigonometric inequality and the second inequality follows from \( h(z) \) is Lipschitz continuous differentiable with the modulus \( L \).
Adding \((1 - \alpha)\|\lambda^{k+1} - \lambda^k\|\) to both sides of the above inequality, we get
\[
(2 - \alpha) \|\lambda^{k+1} - \lambda^k\| \\
\leq \alpha L \|z^{k+1} - z^k\| \\
+ (\alpha - 1) (\|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\|),
\]
i.e.,
\[
\|\lambda^{k+1} - \lambda^k\| \\
\leq \frac{\alpha}{2 - \alpha} L \|z^{k+1} - z^k\| \\
+ \frac{\alpha - 1}{2 - \alpha} (\|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\|).
\]
(37)

Combining (36) and (37), we have
\[
\|\lambda^{k+1} - \lambda^k\| \\
\leq \max \left\{ \frac{\alpha}{2 - \alpha}, \frac{\alpha - 1}{2 - \alpha} \right\} \|\lambda^k - \lambda^{k-1}\| \\
+ \frac{\alpha - 1}{2 - \alpha} (\|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\|).
\]
(38)

For \( \alpha \in (0, \frac{1 + \sqrt{5}}{2}) \), we always have \( \max \left\{ \frac{\alpha}{2 - \alpha}, \frac{\alpha - 1}{2 - \alpha} \right\} < 1 \) and \( \max \left\{ \frac{\alpha}{2 - \alpha}, \frac{\alpha - 1}{2 - \alpha} \right\} < +\infty \).
Next, we prove the upper bound of \( \|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\| \)
\[
\|\lambda^k - \lambda^{k-1}\| - \|\lambda^{k+1} - \lambda^k\| \\
= \alpha \beta (\|Ax^k + By^k + z^k - b\| - \|Ax^{k+1} + By^{k+1} + z^{k+1} - b\|) \\
\leq \alpha \beta (\|Ax^k + By^k + z^k - b\| - |Ax^{k+1} + By^{k+1} + z^{k+1} - b|).
\]
(39)

Substituting (39) into (38), there exists a positive constant \( \delta_1 \), such that
\[
\|\lambda^{k+1} - \lambda^k\| \leq \delta_1 \|v^{k+1} - v^k\|.
\]
The proof is completed.

Lemma 6: Let \( \{\omega^k\} \) be the sequence generated by LADMM (1) which is assumed to be bounded and suppose Assumption 10 holds. There exists positive integer \( k \), we define
\[
\begin{align*}
N_x^{k+1} &:= -(1 + \frac{2\theta(\alpha)}{\alpha})A^T (\lambda^{k+1} - \lambda^k) \\
&\quad + (1 - \beta A^T (B y^{k+1} - B y^k) \\
&\quad + \beta A^T (z^{k+1} - z^k) \\
&\quad - (r I - \beta A) (\lambda^{k+1} - \lambda^k), \\
N_y^{k+1} &:= -(1 + \frac{2\theta(\alpha)}{\alpha})B^T (\lambda^{k+1} - \lambda^k) \\
&\quad + \beta B^T (z^{k+1} - z^k) \\
&\quad - (s I - \beta B) (\lambda^{k+1} - y^k), \\
N_z^{k+1} &:= -(1 + \frac{2\theta(\alpha)}{\alpha}) (\lambda^{k+1} - \lambda^k), \\
N_\lambda^{k+1} &:= \frac{1}{\alpha} (\lambda^{k+1} - \lambda^k).
\end{align*}
\]
(40)

Then, there exists a positive constant \( \delta_2 \), such that
\[
d(0, \partial L_B(\omega^{k+1})) \\
\leq \| (N_x^{k+1}, N_y^{k+1}, N_z^{k+1}, N_\lambda^{k+1}) \| \\
\leq \|v^{k+1} - v^k\|.
\]
(41)
Proof: From (2), we obtain
\[
\begin{align*}
\partial \alpha m(\omega^{k+1}) &= \partial f(\lambda^{k+1}) - A^T \lambda^{k+1} + (1 + 2\theta(\alpha)) \beta A^T (Ax^{k+1} + By^{k+1} + z^{k+1} - b) \\
\partial \alpha m(\omega^{k+1}) &= \partial g(\lambda^{k+1} - B^T \lambda^{k+1} + (1 + 2\theta(\alpha)) \beta B^T (Ax^{k+1} + By^{k+1} + z^{k+1} - b) \\
\partial \alpha m(\omega^{k+1}) &= \partial h(z^{k+1} - \lambda^{k+1} + (1 + 2\theta(\alpha)) \beta (Ax^{k+1} + By^{k+1} + z^{k+1} - b) \\
\partial \alpha m(\omega^{k+1}) &= -(Ax^{k+1} + By^{k+1} + z^{k+1} - b) \\
\end{align*}
\]
Together with (15), yields
\[
(N_x^{k+1}, N_y^{k+1}, N_z^{k+1}, N_\lambda^{k+1}) \in \partial m(\omega^{k+1}).
\]

Together with (34) and (40), there exists a constant \( \delta_2 > 0 \), such that
\[
\| (N_x^{k+1}, N_y^{k+1}, N_z^{k+1}, N_\lambda^{k+1}) \| \\
\leq \|N_x^{k+1}\| + \|N_y^{k+1}\| + \|N_z^{k+1}\| + \|N_\lambda^{k+1}\| \\
\leq \delta_2 \|v^{k+1} - v^k\|.
\]
The proof is completed.
Lemma 17: Let \( \{\omega^k\} \) be the sequence generated by LADMM (1) which is assumed to be bounded. Then,
\[
\sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\|^2 < +\infty. \tag{42}
\]

Proof: Since \( \{\omega^k\} \) is bounded, there exists a subsequence \( \{\omega^{k_j}\} \) such that \( \omega^{k_j} \to \omega^* \) as \( j \to +\infty \). Due to the lower semicontinuous of \( m(\cdot) \) and hence
\[
m(\omega^*) \leq \lim \inf_{j \to +\infty} m(\omega^{k_j}). \tag{43}
\]
Consequently, \( m(\omega^{k_j}) \) is below bounded.

On the other hand, since \( x^{k+1} \) is the optimal solution of (7b), we have
\[
m(x^{k+1}, y^k, z^k, \lambda^k) \leq m(x^*, y^*, z^*, \lambda^k). \tag{44}
\]
It follows from (44) and the continuous of \( m(\cdot) \) with respect to \( y, z \) and \( \lambda \), we get
\[
\lim sup_{j \to +\infty} m(x^{k_j+1}, y^{k_j+1}, z^{k_j+1}, \lambda^{k_j+1}) = \lim sup_{j \to +\infty} m(x^{k_j+1}, y^{k_j}, z^{k_j}, \lambda^{k_j}) \leq \lim sup_{j \to +\infty} m(x^*, y^{k_j}, z^{k_j}, \lambda^{k_j}) \leq m(x^*, y^*, z^*, \lambda^*).
\]
Together with (43), we have
\[
\lim_{j \to +\infty} m(\omega^{k_j+1}) = m(\omega^*).
\]
Together with the monotonically nonincreasing of \( m(\cdot) \), we have \( m(\omega^k) \geq m(\omega^*) \) and
\[
\lim_{k \to +\infty} m(\omega^k) = m(\omega^*). \tag{45}
\]
From (26), we have
\[
\delta \|v^{k+1} - v^k\|^2 \leq m(\omega^k) - m(\omega^{k+1}).
\]
Summing up the above inequality from \( k = 0 \) to \( +\infty \), we get
\[
\sum_{k=0}^{+\infty} \delta \|v^{k+1} - v^k\|^2 \leq \sum_{k=0}^{+\infty} (m(\omega^k) - m(\omega^{k+1})) \leq m(\omega^0) - (m(\omega^*)) < +\infty.
\]
Since \( \delta > 0 \), yields
\[
\sum_{k=0}^{+\infty} \|v^{k+1} - v^k\|^2 \leq +\infty.
\]
Thus,
\[
\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 \leq +\infty, \quad \sum_{k=0}^{+\infty} \|y^{k+1} - y^k\|^2 \leq +\infty, \quad \sum_{k=0}^{+\infty} \|z^{k+1} - z^k\|^2 \leq +\infty.
\]
and
\[
\sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\|^2 \leq +\infty.
\]
Together with (34), we obtain
\[
\sum_{k=0}^{+\infty} \|\lambda^{k+1} - \lambda^k\|^2 \leq +\infty.
\]
Therefore, \( \sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\|^2 < +\infty \). The proof is completed. \(\square\)

Lemma 18: Let \( \{\omega^k\} \) be the sequence generated by LADMM (1) which is assumed to be bounded and denote the limit point as \( \Omega^0 \). Then, we get the assertions below:

(i) \( \Omega^0 \) is a nonempty compact set and
\[
d(\omega^k, \Omega^0) \to 0, \quad k \to +\infty.
\]
(ii) \( \emptyset \neq \Omega^0 \subseteq \text{crit} \, L_\beta \).
(iii) \( m(\cdot) \) is finite and constant on \( \Omega^0 \), i.e.,
\[
\inf_{k \in N} m(\omega^k) = \lim_{k \to +\infty} m(\omega^k).
\]

Proof: (i) \( \Omega^0 \) is nonempty and \( d(\omega^k, \Omega^0) \to 0, \quad k \to +\infty \) are obviously. The proof of \( \Omega^0 \) is a compact set can be found in [6, lemma 5 (iii)].
(ii) Let \( \omega^* \in \Omega^0 \) and \( \{\omega^k\} \) be defined as Lemma 17. From (42), we have
\[
\lim_{k \to +\infty} \|\omega^{k+1} - \omega^k\| = 0.
\]
On the other hand, from (41), we have
\[
\lim_{k \to +\infty} (N_{x^k}^k, N_{y^k}^k, N^k_z, N^k) = 0.
\]
From the continuity of \( \nabla h \) and the closeness of \( \partial f \) and \( \partial g \), we have \( 0 \in \partial d(m(\omega^*)) \), which implies \( \omega^* \) satisfies (8) and it is the critical point of \( m(\cdot) \).
(iii) From (45), \( m(\cdot) \) is a constant on \( \Omega^0 \).

The proof is completed. \(\square\)

Theorem 19: Let \( \{\omega^k\} \) be the sequence generated by LADMM 1 which is assumed to be bounded and suppose that \( m(\cdot) \) is a KL function. Then, the sequence \( \{\omega^k\} \) has finite length, that is,
\[
\sum_{k=0}^{+\infty} \|\omega^{k+1} - \omega^k\| < +\infty, \tag{46}
\]
and sequence \( \{\omega^k\} \) globally converges to a critical point of \( m(\cdot) \).

Proof: We prove the assertion (46) from two cases.

(i) For the continuous of \( m(\cdot) \) and (45), if there exists a positive integer \( k_1 \), such that
\[
m(x^{k_1}) = m(\omega^*).
\]
Since the monotonicity of \( m(\cdot) \) and rearranging terms of (26), for any \( k > k_1 \), we have
\[
\delta \|v^{k+1} - v^k\|^2 \leq m(\omega^*) - m(\omega^{k+1}) \leq m(\omega^{k+1}) - m(\omega^*) = 0.
\]
Thus, \( v^{k+1} = v^k \), for any \( k > k_1 \). Together with (34), we have \( o^{k+1} = o^k \). Thus, the assertion (46) holds.

(ii) From Lemma 18, we have \( d(o^k, \Omega^0) \rightarrow 0 \), for \( k \rightarrow +\infty \), we have for any \( \varepsilon > 0 \), there exists a nonnegative integer \( k' \), such that \( d(o^k, \Omega^0) < \varepsilon \), for any \( k > k' \). From (45), for any \( \eta > 0 \), there exists a nonnegative integer \( k_3 \), such that \( m(o^k) < m(o^\ast) + \eta \), where \( o^\ast \in \Omega^0 \). Therefore, for all \( \varepsilon, \eta > 0 \), when \( k > k_4 := \max\{k', k_3\} \), we have

\[
d(o^k, \Omega^0) < \varepsilon \quad \text{and} \quad m(o^k) < m(o^k) < m(o^\ast) + \eta.
\]

Since \( \Omega^0 \) is a nonempty compact set and \( m(\cdot) \) is a constant on \( \Omega^0 \), for \( k > k_4 \), and together \( \Omega := \Omega^0 \) in (14), we get

\[
\varphi'(m(o^k) - m(o^\ast)) d(0, \partial m(o^k)) \geq 1.
\]

For convenience, we denote

\[
\Delta_{p,q} := \varphi(m(o^p) - m(o^\ast)) - \varphi(m(o^q) - m(o^\ast)).
\]

Since the concavity of \( \varphi \), we have

\[
\Delta_{k,k+1} \geq \varphi'(m(o^k) - m(o^\ast)) [(m(o^k) - m(o^\ast))
- (m(o^{k+1}) - m(o^\ast))]
= \varphi'(m(o^k) - m(o^\ast))(m(o^k) - m(o^{k+1})). (47)
\]

Together with (41) and \( \varphi(m(o^k) - m(o^\ast)) > 0 \), rearranging terms of (47), we obtain

\[
m(o^k) - m(o^{k+1}) \leq \frac{\Delta_{k,k+1}}{\varphi'(m(o^k) - m(o^\ast))} \leq \frac{\delta_2}{\delta} \| v^k - v^{k-1} \| \| \Delta_{k,k+1} \|.
\]

Together with (26), we get

\[
\delta \| v^{k+1} - v^k \| \leq m(o^k) - m(o^{k+1}) \leq \frac{\delta_2}{\delta} \| v^k - v^{k-1} \| \| \Delta_{k,k+1} \|.
\]

for \( k > k_4 \), i.e.,

\[
\| v^{k+1} - v^k \| \leq \frac{\sqrt{\delta_2}}{\delta} \| \Delta_{k,k+1} \| v^k - v^{k-1} \| \| \Delta_{k,k+1} \| \leq \frac{\delta_2}{\delta} \| v^k - v^{k-1} \| \| \Delta_{k,k+1} \|,
\]

where the second inequality follows from \( 2\sqrt{ab} \leq a + b \), for any \( a, b > 0 \).

After some simple calculation and summing up the above inequality from \( k = k_4 + 1 \) to \( k_5 \), yields

\[
\sum_{k=k_4+1}^{k_5} \| v^{k+1} - v^k \| \leq \| v^{k_4+1} - v^{k_4} \| - \| v^{k_5+1} - v^{k_5} \| + \frac{\delta_2}{\delta} \| \Delta_{k_4+1,k_5} \|.
\]

From KL property (14), we have \( \varphi(m(o^{k_5}) - m(o^\ast)) > 0 \) and let \( k_5 \rightarrow +\infty \), thus,

\[
\sum_{k=k_4+1}^{+\infty} \| v^{k+1} - v^k \| \leq \| v^{k_4+1} - v^{k_4} \| + \frac{\delta_2}{\delta} \| \varphi(m(o^{k_4+1}) - m(o^\ast)) \|
\]

which means

\[
\sum_{k=0}^{+\infty} \| v^{k+1} - v^k \| \leq +\infty.
\]

Thus,

\[
\sum_{k=0}^{+\infty} \| x^{k+1} - x^k \| \leq +\infty, \sum_{k=0}^{+\infty} \| y^{k+1} - y^k \| \leq +\infty,
\]

and

\[
\sum_{k=0}^{+\infty} \| z^{k+1} - z^k \| \leq +\infty.
\]

Furthermore, it follows from (34) that

\[
\sum_{k=0}^{+\infty} \| \chi^{k+1} - \chi^k \| \leq +\infty.
\]

Consequently,

\[
\sum_{k=0}^{+\infty} \| o^{k+1} - o^k \| \leq +\infty.
\]

Thus, \( \{ o^{k} \} \) is a Cauchy sequence [6] and convergent.

The proof is completed. \( \square \)

We show the convergence rate of (1) in the following theorem, which is similar to [6], [31], [53]. We omit its proof.

Theorem 20: Let \( \{ o^{k} \} \) be the sequence generated by LADMM 1 and it converges to \( o^\ast \). Assume that \( L(\cdot) \) has the KL property at \( (x^\ast, y^\ast, z^\ast, \lambda^\ast) \) with \( \varphi(s) = cs^{1-\theta}, \theta \in [0, 1) \) and \( c > 0 \).

(i) if \( \theta = 0 \), then the sequence \( \{ o^{k} \} \) converges in a finite number of steps;

(ii) if \( \theta \in (0, \frac{1}{2}] \), then there exist \( c > 0 \) and \( \sigma \in [0, 1) \) such that \( \| o^k - o^\ast \| \leq c o^k \);

(iii) if \( \theta \in (\frac{1}{2}, 1) \), then there exists \( c > 0 \) such that \( \| o^k - o^\ast \| \leq c k^{(\theta-1)/(2\theta-1)} \).

IV. NUMERICAL RESULTS

In this section, all experiments are run in MATLAB R2014a on a 64-bit PC with an Intel Core i3-4030 CPU (1.90 GHz) and 4 gb of RAM equipped with the Windows 4.6 OS. We test LADMM 1 on solving the following nonconvex feasibility problem of background/foreground extraction from a blurred and noisy surveillance video:

\[
\min_{x,y,z} f(x) + \frac{1}{2} \| z \|^2
\]

s.t. \( A(x + y) + z = b, y \in \Omega \),

\[
(48)
\]
where $A$ is a linear map and $\Omega = \{ y \in \mathbb{R}^{m \times l} \mid \| y \|_\infty \leq 1, y_1 = y_2 = \cdots = y_m \}$. The model (48) corresponds to (1) with $g(y) = P_{Q_2}(y) := \arg \min_{y \in \Omega} \{ \| y - v \|_F \mid \forall v \in \mathbb{R}^{m \times n} \}$ and $A = B$, and hence $r = s$. We apply LADM 1 to solve the problem (48). This yields

$$
\begin{align*}
x^k+1 \in \arg \min \{ f(x) - (\lambda T)^T Ax & + \frac{\beta}{2} \| A(x + y^k) + z^k - b \| + \frac{1}{2} \| x - x^k \|_F^2 \} + \frac{1}{r} - \beta A^T A, \\
y^k+1 \in \arg \min \{ \lambda T \lambda A^T & (Ax^k+1 + Ay + z^k - b) + y^k \}, \\
z^k+1 = \frac{1}{1+\beta} & (\lambda - (Ax^k+1 + Ay + z^k - b)), \\
\lambda^k+1 = \lambda^k - \alpha \beta (Ax^k+1 + Ay + z^k - b),
\end{align*}
$$

(49)

To better illustrate the effectiveness of LADM 1, we compare it with the algorithm (ADMMy) proposed by Yang et al. [55]. Its iterative scheme for solving the problem (48) is given as follows:

$$
\begin{align*}
x^k+1 \in \arg \min \{ L_0(x, y^k, z^k, \lambda^k) & \mid x \in \mathbb{R}^n \}, \\
y^k+1 \in \arg \min \{ L_0(x^k+1, y, z, \lambda^k) & \mid y \in \mathbb{R}^m \}, \\
z^k+1 \in \arg \min \{ L_0(x^k+1, y^k+1, z, \lambda^k) & \mid z \in \mathbb{R}^l \}, \\
\lambda^k+1 = \lambda^k - \alpha \beta (Ax^k+1 + Ay + z^k - b),
\end{align*}
$$

(50)

where $\alpha \in (0, 1+\sqrt{3})$ is the double step size. We update $\beta$ using the following heuristics: for the given $\beta$ in Assumption 10, initialize $n_\beta = 0$ and $\beta = 0.65 \beta$; compute

$p_n = \| y^k \|_F + \| z^k \|_F,$
$q_n = \| y^k - y^{k-1} \|_F + \| z^k - z^{k-1} \|_F,$

and then if $p_n > 0.99 p_{n-1}$, we increase $n_\beta$ by 1. Then, we replace $\beta$ by $1.1 \beta$ whenever $\beta \leq 1.1 \beta$ and the sequence is either $n_\beta \geq 0.3$ or $p_n > 10^{10}$. For $\lambda_{\text{max}}(A^T A) = 0.07960$, we set $r = 0.1$. The initialization we consider is

$\gamma^0 = P_{Q_2}(b), \quad x^0 = 0, \quad z^0 = Ay, \quad \lambda^0 = A^T (b - Az).$

In our experiments, we choose the following three choices of sparse regularizer function $f(x)$ as suggested in [55]:

(i) bridge regularizer: $f(x) = \mu \| x \|_p^p$ for $0 < p < 1$;

(ii) fraction regularizer: $f(x) = \mu \sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{|x_{ij}|}$ for $\xi > 0$;

(iii) logistic regularizer: $f(x) = \mu \sum_{i=1}^{m} \sum_{j=1}^{n} \log(1 + \xi |x_{ij}|)$ for $\xi > 0$.

Our test problems are shown in TABLE 1, where we choose 3 sparse regularizers, 10 choices of $\mu$, and 6 choice of $p$ or $\xi$. Since the sensitivity of parameters values is different for different problems, we choose the best parameter values for our numerical experiment comparison.

|| Parameter Selection of Sparse Regulator Function. |
|---|---|
| $\lambda T$ | $\mu$ |
| 1, 1 ± 0.5, 1 ± 0.1 | 1, 1 ± 0.5, 1 ± 0.1 |

TABLE 1. Parameter selection of sparse regulator function.

All the algorithms use the following stopping criteria. For some positive constants $\text{ERR}_1$ and $\text{ERR}_2$, we check if

$$
\frac{\| y^k - y^{k-1} \|_F + \| z^k - z^{k-1} \|_F}{\| x^k \|_F + \| \lambda^k \|_F + 1} < \text{ERR}_1
$$

holds. Then, we further check

$$
\frac{\| y^k - y^{k-1} \|_F + \| z^k - z^{k-1} \|_F}{\| y^k \|_F + \| \lambda^k \|_F + 1} < \text{ERR}_2.
$$

We choose $\text{ERR}_1 = 5 \times 10^{-4}$ and $\text{ERR}_2 = 5 \times 10^{-3}$ as in [55]. The results of the numerical experiments satisfy the stopping criterion, that is, $\| 0^k - 0^{k-1} \|_F$ is less than a certain value, which guarantees the convergence of LADM 1 and ADMMy (49), such as in Theorem 19.

We choose the ‘Hall’ video that contains 200 1444 × 176 frames (from airport 2001 to airport 2200) following Li et al. [33]. We choose two frames (airport 60 and airport 180) of the video as the test images, as shown in the first line of FIGURE 1. The images on the second line of FIGURE 1 are the blurred images, which are assessed in a numerical test.

We denote ‘iter’, ‘time’, and ‘F-measure’ as the number of iterations, the computing time, and the measure of the accuracy of the separation results, as in [55], respectively. The ‘F-measure’ is close to the maximum value of 1, including that the foreground is completely restored. We report the
TABLE 4. Numerical results of LADMM with bridge regularizer.

| $\alpha$ | $\xi$ | $\mu$ | $\text{Iter}$ | $\text{Time}$ | $\text{F-measure}$ |
|--------|------|------|---------|--------|-----------------|
| 0.1    | 0.5  | 0.0  | 50      | 7.49   | 0.7790          |
| 0.1    | 0.5  | 1.0  | 50      | 9.23   | 0.8643          |
| 0.5    | 1.0  | 1.0  | 20      | 25.96  | 0.8971          |

Figure 2. The images restored by LADMM and ADMM with (a) $\xi = 0.1$, $\mu = 1e - 2$, and (b) $\xi = 1$, $\mu = 1e - 3$.

Figure 3. The images restored by LADMM and ADMM with (a) $\xi = 0.1$, $\mu = 1e - 2$, and (b) $\xi = 1$, $\mu = 1e - 3$.

Figure 4. The images restored by LADMM and ADMM with (a) $p = 1$, $\mu = 1e - 2$, (b) $p = 0.2$, $\mu = 1e - 4$.

Figure 5. The sensitivity of LADMM to $\alpha$. For fraction regularizer, the effect is better when $\alpha = 0.8$. For bridge regularizer function, the effect is good when $\alpha = 1.6$.

V. CONCLUSION

This paper shows that the background/foreground extraction with a blurred and noisy surveillance video can be efficiently solved by the linearized alternating direction method of multipliers. The introduction of linearization technology makes it easy to solve the subproblem. Under the powerful Kurdyka-Łojasiewicz property and some assumptions on the correlation function, we prove that the sequence generated by the proposed algorithm converges to a critical point of
the augmented Lagrangian function. We also report some preliminary numerical results to indicate the feasibility and effectiveness of the linearization strategy.

REFERENCES

[1] B. H. Abed-Alguni, “Island-based cuckoo search with highly disruptive polynomial mutation,” Int. J. Artif. Intell., vol. 17, no. 1, pp. 57–82, 2019.
[2] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, “Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the kurdyka–Łojasiewicz inequality,” Math. Oper. Res., vol. 35, no. 2, pp. 438–457, May 2010, doi: 10.1287/moor.1090.0449.
[3] H. Attouch, J. Bolte, and B. F. Svaiter, “Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods,” Math. Program., vol. 137, nos. 1–2, pp. 91–129, Feb. 2013.
[4] A. Azadeh, M. Hosseinabadi Farahani, H. Eivazv, S. Nazari-Shirkoubi, and G. Asadipour, “A hybrid meta-heuristic algorithm for optimization of crew scheduling,” Appl. Soft Comput., vol. 13, no. 1, pp. 158–164, 2013.
[5] W. Bian and X. Chen, “Linearly constrained non-Lipschitz optimization for image restoration,” SIAM J. Imag. Sci., vol. 8, no. 4, pp. 2294–2322, Jan. 2015.
[6] J. Bolte, S. Sabach, and M. Teboulle, “Proximal alternating linearized minimization for nonconvex and nonsmooth problems,” Math. Program., vol. 146, nos. 1–2, pp. 459–494, Aug. 2014.
[7] T. Bouwmans, “Recent advanced statistical background modeling for foreground Detection–A systematic survey,” Recent Patents Comput. Sci., vol. 4, no. 3, pp. 147–176, Sep. 2011.
[8] T. Bouwmans, “Traditional and recent approaches in background modeling for foreground detection: An overview,” Comput. Sci. Rev., vol. 11–12, pp. 31–66, May 2014.
[9] X. Cai, D. Han, and X. Yuan, “On the convergence of the direct extension of ADMM for three-block separable convex minimization models with one strongly convex function,” Comput. Optim. Appl., vol. 66, no. 1, pp. 39–73, Jan. 2017, doi: 10.1007/s10589-016-9860-y.
[10] E. Candès, X. Li, Y. Ma, and J. Wright, Robust Principal Component Analysis. New York, NY, USA: Springer-Verlag, 2011.
[11] C. Chen, B. He, Y. Ye, and X. Yuan, “The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent,” Math. Program., vol. 155, nos. 1–2, pp. 57–79, Jan. 2016, doi: 10.1007/s10107-014-0826-5.
[12] C. Chen, M. Li, X. Liu, and Y. Ye, “Extended ADMM and BCD for nonseparable convex minimization models with quadratic coupling terms: Convergence analysis and insights,” Math. Program., vol. 173, nos. 1–2, pp. 37–77, Jan. 2019, doi: 10.1007/s10107-017-1205-9.
[13] C. Chen, Y. Shen, and Y. You, “On the convergence analysis of the alternating direction method of multipliers with three blocks,” Abstract Appl. Anal., vol. 2013, pp. 1–7, Oct. 2013.
[14] X. Chen and W. Zhou, “Smoothing nonlinear conjugate gradient method for image restoration using nonsmooth nonconvex minimization,” SIAM J. Imag. Sci., vol. 3, no. 4, pp. 765–790, Jan. 2010.
[15] W. Deng, M.-J. Lai, Z. Peng, and W. Yin, “Parallel multi-block ADMM with o(1/k) convergence,” J. Sci. Comput., vol. 71, no. 2, pp. 712–736, May 2017.
[16] W. Deng and W. Yin, “On the global and linear convergence of the generalized alternating directionmethod of multipliers,” J. Sci. Comput., vol. 66, pp. 889–916, Mar. 2016, doi: 10.1007/s10915-015-0048-x.
[17] J. Eckstein and M. C. Ferris, “Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control,” INFORMS J. Comput., vol. 10, no. 2, pp. 218–235, May 1998.
[18] J. Friedman, T. Hastie, and R. Tibshirani, “A note on the group lasso and a sparse group lasso,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 2862–2869.
[19] D. Han and X. Yuan, “A note on the alternating direction method of multipliers,” J. Optim. Theory Appl., vol. 155, no. 1, pp. 227–238, Oct. 2012, doi: 10.1007/s10957-012-0003-z.
[43] R. D. C. Monteiro and B. F. Svaiter, “Iteration-complexity of block-decomposition algorithms and the alternating direction method of multipliers,” SIAM J. Optim., vol. 23, no. 1, pp. 475–507, Jan. 2013, doi: 10.1137/110849468.

[44] M. Nikolova, Energy Minimization Methods. O. Scherzer, Ed. New York, NY, USA: Springer, 2011.

[45] M. Nikolova, M. K. Ng, S. Zhang, and W.-K. Ching, “Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization,” SIAM J. Imag. Sci., vol. 1, no. 1, pp. 2–25, Jan. 2008.

[46] M. Powell, “A method for nonlinear constraints in minimization problems,” in Optimization, R. Fletcher, Ed. New York, NY, USA: Academic, 1969, pp. 283–298.

[47] C. Purcaru, R.-E. Precup, D. Iercan, L.-O. Fedorovici, R.-C. David, and F. Dragan, “Optimal robot path planning using gravitational search algorithm,” Int. J. Artif. Intell., vol. 10, no. 13, pp. 1–20, 2013.

[48] L. Qiao, B. Zhang, X. Lu, and J. Su, “Adaptive linearized alternating direction method of multipliers for non-convex composite regularized optimization problems,” Tsinghua Sci. Technol., vol. 22, pp. 328–341, Jun. 2017, doi: 10.23919/TST.2017.7914204.

[49] G. Rigatos, P. Siano, D. Selisteanu, and R. Precup, “Nonlinear optimal control of oxygen and carbon dioxide levels in blood,” Intell. Ind. Syst., vol. 3, pp. 61–75, Jun. 2017, doi: 10.1007/s40903-016-0060-y.

[50] Y. Wang, W. Yin, and J. Zeng, “Global convergence of ADMM in nonconvex nonsmooth optimization,” J. Sci. Comput., vol. 78, no. 1, pp. 29–63, Jan. 2019, doi: 10.1007/s10915-018-0757-z.

[51] D. Sun, K. Toh, and L. Yang, “A convergent 3-block semi-proximal alternating direction method of multipliers for conic programming with 4-block constraints,” SIAM J. Optim., vol. 25, no. 2, pp. 882–915, 2015, doi: 10.1137/140964357.

[52] Z. Wu, X. Cai, and D. Han, “Linearized block-wise alternating direction method of multipliers for multiple-block convex programming,” J. Ind. Manage. Optim., vol. 14, no. 3, pp. 833–855, 2018, doi: 10.3934/jimo.2017078.

[53] Z. Wu, M. Li, D. Wang, and D. Han, “A symmetric alternating direction method of multipliers for separable nonconvex minimization problems,” Asia Pacific J. Oper. Res., vol. 34, no. 6, pp. 531–547, 2017, doi: 10.1142/S0217595917500300.

[54] W. Yang and D. Han, “Linear convergence of the alternating direction method of multipliers for a class of convex optimization problems,” SIAM J. Numer. Anal., vol. 54, no. 2, pp. 625–640, 2016.

[55] L. Yang, T. Pong, and X. Chen, “Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction,” SIAM J. Imag. Sci., vol. 10, no. 1, pp. 74–110, 2017, doi: 10.1137/15M1027528.

[56] X. Yu, “Application of linearized alternating direction multiplier method in dictionary learning,” J. Appl. Math. Phys., vol. 7, no. 1, pp. 138–147, 2019, doi: 10.4236/jamp.2019.71012.

CHUN ZHANG was born in Jiangsu, China, in 1980. She received the M.S. degree from the School of Mathematical Sciences, Nanjing Normal University, Nanjing, China, in 2007, where she is currently pursuing the Ph.D. degree in computational mathematics. Her research interests include nonconvex optimization problems and image processing.

YANHONG YANG was born in Jiangsu, China, in 1982. She received the Ph.D. degree in computational mathematics and mechanical engineering from Nanjing Normal University, Nanjing, in 2018. Her research interests include structure preserving algorithms and image processing.

ZEYAN WANG was born in Anhui, China, in 1972. He received the M.S. degree from the Institute of Communication Engineering, The PLA Army Engineering University, China, in 1999. His research interests include optimization theory and its applications in algorithms.

YONGXIN CHEN was born in Jiangsu, China, in 1996. He is currently pursuing the M.S. degree with the School of Mathematical Sciences, Nanjing Normal University, China. His research interests include sparse optimization and nonconvex optimization.