Linearized version of the generalized alternating direction method of multipliers for three-block separable convex minimization problem*

Xueqing Zhang† · Jianwen Peng ‡

Abstract Recently, the generalized alternating direction method of multipliers (GADMM) proposed by Eckstein and Bertsekas has received wide attention, especially with respect to numerous applications. In this paper, we develop a new linearized version of generalized alternating direction method of multipliers (L-GADMM) for the linearly constrained separable convex programming whose objective functions are the sum of three convex functions without coupled variables. We give a sufficient condition to ensure the convergence of the L-GADMM for three-block separable convex minimization problem. Theoretically, we establish the worst-case $\mathcal{O}(1/t)$ convergence rate for the proposed L-GADMM in both ergodic and nonergodic senses under the sufficient condition. Moreover, we also show an example to prove its divergence of the proposed L-GADMM if the sufficient condition is lost and give some numerical results.

Keywords The linearized version of generalized alternating direction method of multipliers; Convergence; Three-block separable convex minimization problem; Matrix optimization problem.

1 Introduction

The alternating direction method of multipliers (ADMM) was originally proposed in [1] and it is now a benchmark for the following convex minimization model with two blocks of functions

*Research of this author was supported by the National Natural Science Foundation of China (11431004), the Chongqing Research Program of Basic Research and Frontier Technology (cstc2015jcyjBX0029), the Education Committee Project Foundation of Chongqing for "Bayu scholar" Distinguished Professor, Key Laboratory of optimization and control of Ministry of Education in China.

†School of Mathematical Sciences, Chongqing Normal University, Chongqing 400047, P.R.China. E-mail: zxqcqspb@163.com

‡Corresponding author. School of Mathematical Science, Chongqing Normal University, Chongqing, 400047, P.R. China; E-mail: jwpeng168@hotmail.com

and variables:

$$\begin{aligned} &\min \; f_1(x_1) + f_2(x_2) \\ &\text{s.t.} \quad A_1x_1 + A_2x_2 = b, \\ &\qquad\; x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2. \end{aligned} \tag{1.1}$$

where $A_i \in \Re^{p\times n_i}$ $(i = 1, 2), b \in \Re^p, \mathcal{X}_i \in \Re^{n_i}$ $(i = 1, 2)$ are closed convex sets; and $f_i : \Re^{n_i} \longrightarrow \Re$ $(i = 1, 2)$ are closed convex but not necessarily smooth function.

The iterative scheme of ADMM for (1.1) is

$$\begin{cases} x_1^{k+1} = \text{argmin}\{f_1(x_1) - x_1^\top A_1^\top \lambda^k + \dfrac{\rho}{2}\|A_1x_1 + A_2x_2^k - b\|^2 | x_1 \in \mathcal{X}_1\}, \\ x_2^{k+1} = \text{argmin}\{f_2(x_2) - x_2^\top A_2^\top \lambda^k + \dfrac{\rho}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2 | x_2 \in \mathcal{X}_2\}, \\ \lambda^{k+1} = \lambda^k - \rho(A_1x_1^{k+1} + A_2x_2^{k+1} - b) \end{cases} \tag{1.2}$$

where $\lambda \in \Re^p$ is the Lagrangian multiplier; $\rho > 0$ is a penalty parameter, and $\|\cdot\|$ is the Euclidean 2-norm. Furthermore, if we change the performance order of x_2 and λ of the ADMM (1.2), it becomes a new ADMM (ADMMnew) scheme:

$$\begin{cases} x_1^{k+1} = \text{argmin}\{f_1(x_1) - x_1^\top A_1^\top \lambda^k + \dfrac{\rho}{2}\|A_1x_1 + A_2x_2^k - b\|^2 | x_1 \in \mathcal{X}_1\}, \\ \lambda^{k+1} = \lambda^k - \rho(A_1x_1^{k+1} + A_2x_2^{k} - b) \\ x_2^{k+1} = \text{argmin}\{f_2(x_2) - x_2^\top A_2^\top \lambda^{k+1} + \dfrac{\rho}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2 | x_2 \in \mathcal{X}_2\}, \end{cases}$$

The iterative scheme of ADMM embeds a Gaussian-Seidel decomposition into each iteration of the augmented Lagrangian method (ALM) in [2, 3]; thus the functions f_1 and f_2 are treated individually and so easier subproblems could be generated. Recently, the ADMM has received wide attention from a broad spectrum of areas because of its easy implementation and impressive efficiency. We refer to [4, 5, 6] for excellent review papers for the applications and the history of ADMM. Gabay [7] pointed out that the ADMM was an application of the well-known Douglas-Rachford splitting method (DRSM) in [8] to the dual of (1.1); and as an application of the proximal point algorithm (PPA) in [9]. Eckstein and Bertsekas [10] proposed the following generalized ADMM (GADMM) scheme:

$$\begin{cases} x_1^{k+1} = \text{argmin}\{f_1(x_1) - x_1^\top A_1^\top \lambda^k + \dfrac{\rho}{2}\|A_1x_1 + A_2x_2^k - b\|^2 | x_1 \in \mathcal{X}_1\}, \\ x_2^{k+1} = \text{argmin}\{f_2(x_2) - x_2^\top A_2^\top \lambda^k + \dfrac{\rho}{2}\|\beta A_1x_1^{k+1} + (1-\beta)(b - A_2x_2^k) + A_2x_2 - b\|^2 | x_2 \in \mathcal{X}_2\}, \\ \lambda^{k+1} = \lambda^k - \rho\left(\beta A_1x_1^{k+1} + (1-\beta)(b - A_2x_2^k) + A_2x_2^{k+1} - b\right), \end{cases}$$

where the parameter $\beta \in (0, 2)$ is a relaxation factor. It is obvious that the above GADMM scheme reduces to the original ADMM scheme (1.2) when $\beta = 1$. There are some empirical studies of the acceleration performance of the GADMM, we thus refer to [11, 12, 13]. Recently, Fang et al. [14] proposed the following linearized versions of GADMM (L-GADMM) for (1.1):

$$\begin{cases} x_1^{k+1} = \text{argmin}\{f_1(x_1) - x_1^\top A_1^\top \lambda^k + \dfrac{\rho}{2}\|A_1x_1 + A_2x_2^k - b\|^2 + \dfrac{1}{2}\|x_1 - x_1^k\|^2_{G_1} | x_1 \in \mathcal{X}_1\}, \\ x_2^{k+1} = \text{argmin}\{f_2(x_2) - x_2^\top A_2^\top \lambda^k + \dfrac{\rho}{2}\|\beta A_1x_1^{k+1} + (1-\beta)(b - A_2x_2^k) + A_2x_2 - b\|^2 | x_2 \in \mathcal{X}_2\}, \\ \lambda^{k+1} = \lambda^k - \rho(\beta A_1x_1^{k+1} + (1-\beta)(b - A_2x_2^k) + A_2x_2^{k+1} - b), \end{cases} \tag{1.3}$$

and

$$
\begin{cases}
x_1^{k+1} = \text{argmin}\{f_1(x_1) - x_1^\top A_1^\top \lambda^k + \frac{\rho}{2}\|A_1x_1 + A_2x_2^k - b\|^2 + \frac{1}{2}\|x_1 - x_1^k\|_{G_1}^2 | x_1 \in \mathcal{X}_1\}, \\
x_2^{k+1} = \text{argmin}\{f_2(x_2) - x_2^\top A_2^\top \lambda^k + \frac{\rho}{2}\|\beta A_1x_1^{k+1} + (1-\beta)(b - A_2x_2^k) + A_2x_2 - b\|^2 \\
\qquad + \frac{1}{2}\|x_2 - x_2^k\|_{G_2}^2 | x_2 \in \mathcal{X}_2\}, \\
\lambda^{k+1} = \lambda^k - \rho(\beta A_1x_1^{k+1} + (1-\beta)(b - A_2x_2^k) + A_2x_2^{k+1} - b),
\end{cases}
\tag{1.4}
$$

where $G_i \in \Re^{n_i \times n_i}$ $(i = 1, 2)$ are symmetric positive definite matrices and $\|x_i\|_{G_i}$ is the G_i-norm of x_i which denotes the quantity $\sqrt{x_i^\top G_i x_i}$ for $i = 1, 2$. Obviously, if $\mathcal{X}_1 = \Re^{n_1}$ and $G_1 = \tau I_{n_1} - \rho A_1^\top A_1$ with the requirement $\tau > \rho\|A_1^\top A_1\|_2$, where $\|\cdot\|_2$ denotes the spectral norm of a matrix, then the x_1-subproblem in (1.3) or (1.4) reduces to estimating the resolvent operator of ∂f_1:

$$
x_1^{k+1} = (I + \frac{1}{\tau}\partial f_1)^{-1}(a) = \text{argmin}\{f_1(x_1) + \frac{\tau}{2}\|x_1 - a\|^2\},
$$

where $\partial(\cdot)$ denotes the subdifferential of a convex function and $a = \frac{1}{\tau}((\tau I_{n_1} - \rho A_1^\top A_1)x_1^k - \rho A_1^\top A_2 x_2^k + A_1^\top \lambda^k + \rho A_1^\top b))$. Therefore, the iterative scheme (1.3) includes the linearized version of ADMM as a special case with $G_1 = \tau I_{n_1} - \rho A_1^\top A_1$ and $\beta = 1$ (see [15, 16, 17]). And Fang et al. [14] also proved the worst-case $\mathcal{O}(1/t)$ convergence rate measured by the iteration complexity for the linearized versions of GADMM (1.3) and (1.4) in both the ergodic and a nonergodic senses.

Now, we consider the convex minimization model with linear constraints and an objective function which is sum of three functions without coupled variables

$$
\begin{array}{ll}
\min & f_1(x_1) + f_2(x_2) + f_3(x_3) \\
\text{s.t.} & A_1x_1 + A_2x_2 + A_3x_3 = b. \\
& x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, x_3 \in \mathcal{X}_3
\end{array}
\tag{1.5}
$$

where $A_i \in \Re^{p \times n_i}$ $(i = 1, 2, 3), b \in \Re^p, \mathcal{X}_i \in \Re^{n_i}$ $(i = 1, 2, 3)$ are closed convex sets; and $f_i : \Re^{n_i} \longrightarrow \Re$ $(i = 1, 2, 3)$ are closed convex but not necessarily smooth function. Throughout the paper, the solution set of (1.5) is assumed to be nonempty.

Recently, Chen et al. [18] extended the original ADMM (1.2) directly to (1.5) by taking the advantage of each f_i's properties individually and obtained the direct extension of ADMM (1.2) as follows:

$$
\begin{cases}
x_1^{k+1} = \text{argmin}\{L_\rho(x_1, x_2^k, x_3^k, \lambda^k) | x_1 \in \mathcal{X}_1\}, \\
x_2^{k+1} = \text{argmin}\{L_\rho(x_1^{k+1}, x_2, x_3^k, \lambda^k) | x_2 \in \mathcal{X}_2\}, \\
x_3^{k+1} = \text{argmin}\{L_\rho(x_1^{k+1}, x_2^{k+1}, x_3, \lambda^k) | x_3 \in \mathcal{X}_3\}, \\
\lambda^{k+1} = \lambda^k - \rho(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^{k+1} - b),
\end{cases}
\tag{1.6}
$$

where

$$L_\rho(x_1,x_2,x_3,\lambda)=f_1(x_1)+f_2(x_2)+f_3(x_3)-\lambda^\top(A_1x_1+A_2x_2+A_3x_3-b)$$
$$+\frac{\rho}{2}\|A_1x_1+A_2x_2+A_3x_3-b\|^2$$

Feng et al. [19] proposed the following linearized version of Alternating direction method of multipliers (L-ADMM) for (1.5):

$$\begin{cases} x_1^{k+1}=\operatorname{argmin}_{x_1} f_1(x_1)+\frac{\rho}{2\tau_1}\|x_1-(x_1^k-\tau_1A_1^\top(A_1x_1^k+A_2x_2^k+A_3x_3^k-b-\frac{1}{\rho}\lambda^k))\|^2,\\ x_2^{k+1}=\operatorname{argmin}_{x_2} f_2(x_2)+\frac{\rho}{2\tau_2}\|x_2-(x_2^k-\tau_2A_2^\top(A_1x_1^{k+1}+A_2x_2^k+A_3x_3^k-b-\frac{1}{\rho}\lambda^k))\|^2,\\ x_3^{k+1}=\operatorname{argmin}_{x_3} f_3(x_3)+\frac{\rho}{2\tau_3}\|x_3-(x_3^k-\tau_3A_3^\top(A_1x_1^{k+1}+A_2x_2^{k+1}+A_3x_3^k-b-\frac{1}{\rho}\lambda^k))\|^2,\\ \lambda^{k+1}=\lambda^k-\rho(A_1x_1^{k+1}+A_2x_2^{k+1}+A_3x_3^{k+1}-b), \end{cases} \tag{1.7}$$

where $\rho>0$ is a penalty parameter, and $\tau_i(i=1,2,3)$ is the step sizes for the above proximal gradient step. It is well known that ADMM (1.6) and L-ADMM (1.7) are not necessarily convergent without further assumptions (see [18, 19, 20]).

Inspired by the above ideas, we propose a linearized version of GADMM to (1.5) as follows:

$$\begin{cases} x_1^{k+1}=\operatorname{argmin}\{f_1(x_1)-x_1^\top A_1^\top\lambda^k+\frac{\rho}{2}\|A_1x_1+A_2x_2^k+A_3x_3^k-b\|^2+\frac{1}{2}\|x_1-x_1^k\|_{G_1}^2|x_1\in\mathcal{X}_1\},\\ x_2^{k+1}=\operatorname{argmin}\{f_2(x_2)-x_2^\top A_2^\top\lambda^k+\frac{\rho}{2}\|A_1x_1^{k+1}+A_2x_2-b+A_3x_3^k\|^2+\frac{1}{2}\|x_2-x_2^k\|_{G_2}^2|x_2\in\mathcal{X}_2\},\\ x_3^{k+1}=\operatorname{argmin}\{f_3(x_3)-x_3^\top A_3^\top\lambda^k+\frac{\rho}{2}\|\beta A_1x_1^{k+1}+(1-\beta)(b-A_3x_3^k)+\beta A_2x_2^{k+1}+A_3x_3-b\|^2\\ \qquad\qquad+\frac{1}{2}\|x_3-x_3^k\|_{G_3}^2|x_3\in\mathcal{X}_3\},\\ \lambda^{k+1}=\lambda^k-\rho(\beta A_1x_1^{k+1}+(1-\beta)(b-A_3x_3^k)+\beta A_2x_2^{k+1}+A_3x_3^{k+1}-b), \end{cases} \tag{1.8}$$

where $G_i\in\Re^{n_i\times n_i}$ $(i=1,2,3)$ are symmetric positive definite matrices and $\beta>0$ is a relaxation factor. For a matrix G, $G\succeq\mathbf{0}$ denotes that G is positive semidefinite and $G\succ\mathbf{0}$ denotes that G is positive definite. It is worthy noting that if $\beta=1$ and $G_i=\mathbf{0}$ for $i=1,2,3$, then the scheme (1.8) reduces to ADMM (1.6); if $\beta=1$, $G_i=\frac{\rho}{\tau_i}I_{n_i}-\rho A_i^\top A_i\succ\mathbf{0}$ ($i=1,2,3$), then the iterative scheme (1.8) reduces to L-ADMM (1.7).

We notice that the update order of (1.8) at each iteration is $x_1\to x_2\to x_3\to\lambda$ and it repeats cyclically. Equivalently, we can update the variables via the order $x_2\to x_3\to\lambda\to x_1$ and thus have the following iterative scheme

$$\begin{cases} x_2^{k+1}=\operatorname{argmin}\{f_2(x_2)-x_2^\top A_2^\top\lambda^k+\frac{\rho}{2}\|A_1x_1^k+A_2x_2+A_3x_3^k-b\|^2+\frac{1}{2}\|x_2-x_2^k\|_{G_2}^2|x_2\in\mathcal{X}_2\},\\ x_3^{k+1}=\operatorname{argmin}\{f_3(x_3)-x_3^\top A_3^\top\lambda^k+\frac{\rho}{2}\|\beta A_1x_1^k+(1-\beta)(b-A_3x_3^k)+\beta A_2x_2^{k+1}+A_3x_3-b\|^2\\ \qquad\qquad+\frac{1}{2}\|x_3-x_3^k\|_{G_3}^2|x_3\in\mathcal{X}_3\},\\ \lambda^{k+1}=\lambda^k-\rho(\beta A_1x_1^k+(1-\beta)(b-A_3x_3^k)+\beta A_2x_2^{k+1}+A_3x_3^{k+1}-b),\\ x_1^{k+1}=\operatorname{argmin}\{f_1(x_1)-x_1^\top A_1^\top\lambda^{k+1}+\frac{\rho}{2}\|A_1x_1+A_2x_2^{k+1}+A_3x_3^{k+1}-b\|^2+\frac{1}{2}\|x_1-x_1^k\|_{G_1}^2|x_1\in\mathcal{X}_1\}. \end{cases} \tag{1.9}$$

Note that the convergence analysis of the iterative scheme (1.9) is equivalent to that of (1.8) since the scheme (1.9) is exactly becomes the iterative scheme (1.8) if x_1^k is taken as x_1^{k+1} and x_1^{k+1} as x_1^{k+2}.

The rest of this paper is organized as follows. In Sect. 2, we summarize some preliminaries which are useful for further analysis. Then, we show the global convergence of the L-GADMM under a sufficient condition in Sect. 3. In Sect. 4, we derive the worst-case convergence rate for L-GADMM in both ergodic and nonergodic senses. In Sect. 5, we construct an example to show the divergence of the L-GADMM and report some numerical results on the basis of the analysis in Sect. 3. Finally, we draw some conclusions in Sect. 6.

2 Preliminaries

It is useful to characterize the model (1.5) by a variatioonal inequality. More specially, as well known in the literature (see [21, 22]), solving (1.5) is equivalent to solving the following variational inequality (VI) problem: Find $w^* \in \Omega$ such that

$$f(x) - f(x^*) + (w - w^*)^\top F(w^*) \geq 0, \qquad \forall w \in \Omega, \tag{2.1}$$

where

$$w = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix}, \qquad f(x) = f_1(x_1) + f_2(x_2) + f_3(x_3),$$

$$F(w) = \begin{pmatrix} -A_1^\top \lambda \\ -A_2^\top \lambda \\ -A_3^\top \lambda \\ A_1x_1 + A_2x_2 + A_3x_3 - b \end{pmatrix}, \quad \text{and} \quad \Omega = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \Re^p, \tag{2.2}$$

We denote by $VI(\Omega, F, f)$ the problem (2.1-2.2). It is easy to see that the mapping $F(w)$ defined in (2.2) is affine with a skew-symmetric matrix; it is thus monotone:

$$(w_1 - w_2)^\top (F(w_1) - F(w_2)) \geq 0, \qquad \forall \ \ w_1, w_2 \in \Omega.$$

Under the assumption that the solution set of (1.5) is nonempty, we denote the solution set of (2.1) by Ω^*.

Further more, a multifunction $\Psi : \mathcal{R}^n \to 2^{\mathcal{R}^n}$ means that $\Psi(x)$ is a set in $\mathcal{R}^n$. for a multifunction $\Psi : \mathcal{R}^n \to 2^{\mathcal{R}^n}$, we say that Ψ is monotone if for any $x_1, x_2 \in \mathcal{R}^n$

$$(x_1 - x_2)^\top (\xi - \zeta) \geq 0, \quad \forall \xi \in \Psi(x_1), \ \ \forall \ \ \zeta \in \Psi(x_2).$$

It is well known that ∂f is a monotone multifunction (see [23]) for any convex function f.

In the following, we define some auxiliary sequences for the convenience analysis of the L-GADMM for (1.5).

Specially, for the sequence $\{w^k\}$ generated by the L-GADMM (1.8), let

$$\tilde{w}^k = \begin{pmatrix} \tilde{x}_1^k \\ \tilde{x}_2^k \\ \tilde{x}_3^k \\ \tilde{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^k - \rho(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^k - b) \end{pmatrix}. \tag{2.3}$$

By the definition of λ^{k+1} in (1.8), we get

$$\lambda^k - \lambda^{k+1} = -\rho A_3 (x_3^k - x_3^{k+1}) + \rho\beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^k - b).$$

Moreover, because of the identities $\lambda^k - \tilde{\lambda}^k = \rho(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^k - b)$ and $x_3^{k+1} = \tilde{x}_3^k$, we know that

$$\lambda^k - \lambda^{k+1} = -\rho A_3 (x_3^k - \tilde{x}_3^k) + \beta(\lambda^k - \tilde{\lambda}^k). \tag{2.4}$$

Thus, we have

$$w^k - w^{k+1} = M_1 (w^k - \tilde{w}^k), \tag{2.5}$$

where M_1 is defined as

$$M_1 = \begin{pmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 \\ 0 & 0 & -\rho A_3 & \beta I_p \end{pmatrix}. \tag{2.6}$$

Similarly, we also define two auxiliary sequences for the sequence $\{w^k\}$ generated by the iterative scheme (1.9), let

$$\bar{w}^k = \begin{pmatrix} \bar{x}_1^k \\ \bar{x}_2^k \\ \bar{x}_3^k \\ \bar{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^k - \rho(A_1 x_1^k + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b) \end{pmatrix}. \tag{2.7}$$

By (1.9) and the above equation, we have

$$\lambda^k - \lambda^{k+1} = \rho(1-\beta) A_3 (\bar{x}_3^k - x_3^k) + \beta(\lambda^k - \bar{\lambda}^k). \tag{2.8}$$

Then, we have

$$w^k - w^{k+1} = M_2(w^k - \bar{w}^k), \tag{2.9}$$

where M_2 is defined as

$$M_2 = \begin{pmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 \\ 0 & 0 & -\rho(1-\beta)A_3 & \beta I_p \end{pmatrix}. \tag{2.10}$$

Furthermore, we give the following auxiliary sequences for the sequence $\{w^k\}$ generated by the iterative scheme (1.9), which is analogous to (2.7). let

$$\hat{w}^k = \begin{pmatrix} \hat{x}_1^k \\ \hat{x}_2^k \\ \hat{x}_3^k \\ \hat{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^k - \rho(A_1x_1^k + A_2x_2^{k+1} + A_3x_3^k - b) \end{pmatrix}. \tag{2.11}$$

Using (1.9) and the above equation, we can obtain

$$\lambda^k - \lambda^{k+1} = \rho A_3(\hat{x}_3^k - x_3^k) + \beta(\lambda^k - \hat{\lambda}^k). \tag{2.12}$$

And we have

$$w^k - w^{k+1} = M_1(w^k - \hat{w}^k), \tag{2.13}$$

where M_1 is defined in (2.6).

For notational simplicity, we define some matrices that will be used later.

$$Q_1 = \begin{pmatrix} G_1 & -\rho A_1^\top A_2 & 0 & 0 \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \rho A_3^\top A_3 & (1-\beta)A_3^\top \\ 0 & 0 & -A_3 & \frac{1}{\rho}I_p \end{pmatrix}, \tag{2.14}$$

$$Q_2 = \begin{pmatrix} G_1 + \rho A_1^\top A_1 & 0 & \rho(1-\beta)A_1^\top A_3 & -\beta A_1^\top \\ 0 & G_2 & -\rho A_2^\top A_3 & 0 \\ 0 & 0 & G_3 + \rho(1-\beta)A_3^\top A_3 & (1-\beta)A_3^\top \\ -A_1 & 0 & 0 & \frac{1}{\rho}I_p \end{pmatrix}, \tag{2.15}$$

$$Q_3 = \begin{pmatrix} G_1 + \rho A_1^\top A_1 & 0 & 2\rho A_1^\top A_3 & -\beta A_1^\top \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \rho A_3^\top A_3 & (1-\beta)A_3^\top \\ -A_1 & 0 & -A_3 & \frac{1}{\rho}I_p \end{pmatrix}, \tag{2.16}$$

$$\Gamma_1 = \begin{pmatrix} G_1 & 0 & 0 & 0 \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \frac{\rho}{\beta} A_3^\top A_3 & \frac{1-\beta}{\beta} A_3^\top \\ 0 & 0 & \frac{1-\beta}{\beta} A_3 & \frac{1}{\rho\beta} I_p \end{pmatrix}, \tag{2.17}$$

$$\Gamma_2 = \begin{pmatrix} G_1 + \rho A_1^\top A_1 & 0 & \rho(1-\beta) A_1^\top A_3 & -A_1^\top \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \frac{\rho}{\beta}(1-\beta) A_3^\top A_3 & \frac{1-\beta}{\beta} A_3^\top \\ -A_1 & 0 & \frac{1-\beta}{\beta} A_3 & \frac{1}{\rho\beta} I_p \end{pmatrix}, \tag{2.18}$$

$$\Gamma_3 = \begin{pmatrix} G_1 + \rho A_1^\top A_1 & 0 & 0 & -A_1^\top \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \frac{\rho}{\beta} A_3^\top A_3 & \frac{1-\beta}{\beta} A_3^\top \\ -A_1 & 0 & \frac{1-\beta}{\beta} A_3 & \frac{1}{\rho\beta} I_p \end{pmatrix}, \tag{2.19}$$

$$Y_1 = \begin{pmatrix} G_1 + \rho A_1^\top A_1 & 0 & 0 & -A_1^\top \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \rho\frac{(1-\beta)^2}{\beta} A_3^\top A_3 & \frac{1-\beta}{\beta} A_3^\top \\ -A_1 & 0 & \frac{1-\beta}{\beta} A_3 & \frac{1}{\rho\beta} I_p \end{pmatrix}. \tag{2.20}$$

Then it is easy to verify that if $A_1^\top A_2 = 0$, then we have

$$Q_1 = \Gamma_1 M_1, \tag{2.21}$$

if $A_2^\top A_3 = 0$, then we have

$$Q_2 = \Gamma_2 M_2, \tag{2.22}$$

and if $A_1^\top A_3 = 0$, then we get

$$Q_3 = \Gamma_3 M_1. \tag{2.23}$$

Note that the G_i (i=1, 2, 3) is assumed to be positive definite, then the positive definiteness of Γ_1 defined above is guaranteed when $\beta \in (0, 2)$. And both the positive definiteness of Y_1 and Γ_3 is guarateed by $G_3 - \rho(1-\beta)A_3^\top A_3 \succeq \mathbf{0}$ and $\beta \in (0, 1]$.

3 The global convergence of the L-GADMM scheme

In this section, we establish a global convergence for the L-GADMM for (1.5).

To derive the global convergence of the L-GADMM scheme, we firstly obtain the following lemma which characterizes the accuracy of the vector $\tilde{w}^k$ to a solution point of $VI(\Omega, F, f)$.

Lemma 3.1 Let the sequence $\{w^k\}$ be generated by the L-GADMM (1.8) and the associated sequence $\{\tilde{w}^k\}$ be defined in (2.3). Then we have

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(\tilde{w}^k) \geq (w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k), \quad \forall w \in \Omega, \tag{3.1}$$

where Q_1 is defined in (2.14), F is defined in (2.2).

Proof From the optimality condition of the x_1-subproblem of (1.8), we can obtain

$$f_1(x_1)-f_1(x_1^{k+1})+(x_1-x_1^{k+1})^\top(-A_1^\top\lambda^k+\rho A_1^\top(A_1x_1^{k+1}+A_2x_2^k+A_3x_3^k-b)+G_1(x_1^{k+1}-x_1^k)) \geq 0, \forall x_1 \in \mathcal{X}_1. \tag{3.2}$$

By $\tilde{x}_1^k$ and $\tilde{\lambda}^k$ defined in (2.3), (3.2) can be rewritten as

$$f_1(x_1) - f_1(\tilde{x}_1^k) + (x_1 - \tilde{x}_1^k)^\top(-A_1^\top\tilde{\lambda}^k + G_1(\tilde{x}_1^k - x_1^k) - \rho A_1^\top A_2(\tilde{x}_2^k - x_2^k)) \geq 0, \quad \forall x_1 \in \mathcal{X}_1. \tag{3.3}$$

By deriving the optimality condition for the x_2-subproblem in (1.8), then we have

$$\begin{aligned} f_2(x_2) - f_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top(-A_2^\top\lambda^k + \rho A_2^\top(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^k - b) \\ +G_2(x_2^{k+1} - x_2^k)) \geq 0, \quad \forall x_2 \in \mathcal{X}_2. \end{aligned}$$

It follows from (2.3) that

$$f_2(x_2) - f_2(\tilde{x}_2^k) + (x_2 - \tilde{x}_2^k)^\top \left(-A_2^\top\tilde{\lambda}^k + G_2(\tilde{x}_2^k - x_2^k)\right) \geq 0, \forall x_2 \in \mathcal{X}_2. \tag{3.4}$$

By deriving the optimality condition of the x_3-subproblem of (1.8), we get

$$\begin{aligned} f_3(x_3) - f_3(x_3^{k+1}) + (x_3 - x_3^{k+1})^\top(-A_3^\top\lambda^k + \rho A_3^\top(\beta A_1x_1^{k+1} + (1-\beta)(b - A_3x_3^k) + \beta A_2x_2^{k+1} \\ +A_3x_3^{k+1} - b) + G_3(x_3^{k+1} - x_3^k)) \geq 0, \quad \forall x_3 \in \mathcal{X}_3. \end{aligned}$$

By using $\tilde{x}^k$ and $\tilde{\lambda}^k$ defined in (2.3) and the equation defined in (2.4), the above inequality can be rewritten as

$$f_3(x_3)-f_3(\tilde{x}_3^k)+(x_3-\tilde{x}_3^k)^\top(-A_3^\top\tilde{\lambda}^k+(G_3+\rho A_3^\top A_3)(\tilde{x}_3^k-x_3^k)+(1-\beta)A_3^\top(\tilde{\lambda}^k-\lambda^k)) \geq 0, \quad \forall x_3 \in \mathcal{X}_3. \tag{3.5}$$

From (2.3), we have

$$A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3\tilde{x}_3^k - b - A_3(\tilde{x}_3^k - x_3^k) + \frac{1}{\rho}(\tilde{\lambda}^k - \lambda^k) = 0. \tag{3.6}$$

Combining (3.3), (3.4), (3.5) with (3.6), we get

$$f(x) - f(\tilde{x}^k) + \begin{pmatrix} x_1 - \tilde{x}_1^k \\ x_2 - \tilde{x}_2^k \\ x_3 - \tilde{x}_3^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} -A_1^\top\tilde{\lambda}^k \\ -A_2^\top\tilde{\lambda}^k \\ -A_3^\top\tilde{\lambda}^k \\ A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3\tilde{x}_3^k - b \end{pmatrix}$$

$$+ \begin{pmatrix} x_1 - \tilde{x}_1^k \\ x_2 - \tilde{x}_2^k \\ x_3 - \tilde{x}_3^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} G_1(\tilde{x}_1^k - x_k) - \rho A_1^\top A_2(\tilde{x}_2^k - x_2^k) \\ G_2(\tilde{x}_2^k - x_2^k) \\ (G_3 + \rho A_3^\top A_3)(\tilde{x}_3^k - x_3^k) + (1-\beta)A_3^\top(\tilde{\lambda}^k - \lambda^k) \\ -A_3(\tilde{x}_3^k - x_3^k) + \frac{1}{\rho}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \geq 0, \tag{3.7}$$

where $f(x) = f_1(x_1) + f_2(x_2) + f_3(x_3)$. Further more, the inequality in (3.7) can be rewritten as

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top (F(\tilde{w}^k) + Q_1(\tilde{w}^k - w^k)) \geq 0, \qquad \forall w \in \Omega, \tag{3.8}$$

where

$$F(\tilde{w}^k) = \begin{pmatrix} -A_1^\top \tilde{\lambda}^k \\ -A_2^\top \tilde{\lambda}^k \\ -A_3^\top \tilde{\lambda}^k \\ A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3\tilde{x}_3^k - b \end{pmatrix}.$$

Thus the assertion (3.1) is proved. □

According to (3.1), the accuracy of $\tilde{w}^k$ to a solution of $VI(\Omega, F, f)$ is measured by the quantity $(w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k)$. We will explore this term and express it in terms of some quadratic terms in the next lemma, which makes us conveniently estimate the accuracy of $\tilde{w}^k$ and the convergence rate for the scheme (1.8).

Lemma 3.2 Let $A_1^\top A_2 = 0$, $\beta \in (0, 2)$, the sequence $\{w^k\}$ be generated by L-GADMM (1.8) and the associated sequence $\{\tilde{w}^k\}$ be defined in (2.3). Then we have

$$(w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k) \geq \frac{1}{2}(\|w - w^{k+1}\|_{\Gamma_1}^2 - \|w - w^k\|_{\Gamma_1}^2) + \frac{1}{2}\delta_1\|w^k - \tilde{w}^k\|^2, \ \forall w \in \Omega, \tag{3.9}$$

where Q_1 and Γ_1 are defined in (2.14), (2.17) and

$$\delta_1 = \min\{\lambda_{\max}(G_1), \lambda_{\max}(G_2), \lambda_{\max}(G_3), \frac{2-\beta}{\rho}\} > 0.$$

Proof Note that Γ_1 is positive definite and $Q_1 = \Gamma_1 M_1$ because of $A_1^\top A_2 = 0$ and $\beta \in (0, 2)$. It follows from (2.5) that

$$\begin{aligned}(w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k) &= (w - \tilde{w}^k)^\top \Gamma_1 M_1(w^k - \tilde{w}^k) \\ &= (w - \tilde{w}^k)^\top \Gamma_1(w^k - w^{k+1}).\end{aligned} \tag{3.10}$$

It is easy to verify the following identity

$$\begin{aligned}(w - \tilde{w}^k)^\top \Gamma_1(w^k - w^{k+1}) =& \frac{1}{2}(\|w - w^{k+1}\|_{\Gamma_1}^2 - \|w - w^k\|_{\Gamma_1}^2) \\ &+ \frac{1}{2}(\|w^k - \tilde{w}^k\|_{\Gamma_1}^2 - \|w^{k+1} - \tilde{w}^k\|_{\Gamma_1}^2).\end{aligned} \tag{3.11}$$

By (2.17), we have

$$
\begin{aligned}
&\|w^k - \tilde{w}^k\|_{\Gamma_1}^2 \\
=&\|x_1 - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 + \|x_3^k - \tilde{x_3}^k\|_{\frac{\rho}{\beta}A_3^\top A_3}^2 \\
&+ 2(x_3^k - \tilde{x}_3^k)^\top \frac{1-\beta}{\beta} A_3^\top (\lambda^k - \tilde{\lambda}^k) + \frac{1}{\rho\beta}\|\lambda^k - \tilde{\lambda}^k\|^2 \\
=&\|x_1 - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 + \frac{2-\beta}{\rho}\|\lambda^k - \tilde{\lambda}^k\|^2 \\
&+ \frac{1}{\rho\beta}\|\rho A_3(x_3^k - \tilde{x}_3^k) + (1-\beta)(\lambda^k - \tilde{\lambda}^k)\|^2.
\end{aligned} \tag{3.12}
$$

From (2.4), we get

$$
\lambda^{k+1} - \tilde{\lambda}^k = \rho A_3(x_3^k - \tilde{x}_3^k) + (1-\beta)(\lambda^k - \tilde{\lambda}^k). \tag{3.13}
$$

It follows from (2.3) that

$$
\|w^{k+1} - \tilde{w}^k\|_{\Gamma_1}^2 = \frac{1}{\rho\beta}\|\lambda^{k+1} - \tilde{\lambda}^k\|^2 = \frac{1}{\rho\beta}\|\rho A_3(x_3^k - \tilde{x}_3^k) + (1-\beta)(\lambda^k - \tilde{\lambda}^k)\|^2. \tag{3.14}
$$

By the above equation and the equation (3.12), we get

$$
\begin{aligned}
&\|w^k - \tilde{w}^k\|_{\Gamma_1}^2 - \|w^{k+1} - \tilde{w}^k\|_{\Gamma_1}^2 \\
&= \|x_1 - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 + \frac{2-\beta}{\rho}\|\lambda^k - \tilde{\lambda}^k\|^2.
\end{aligned} \tag{3.15}
$$

Note that $\frac{2-\beta}{\rho} > 0$ for any $\beta \in (0, 2)$; and G_i is a symmetric positive definite matrix (i=1, 2, 3). Let $\delta_1 = \min\{\lambda_{\max}(G_1), \lambda_{\max}(G_2), \lambda_{\max}(G_3), \frac{2-\beta}{\rho}\} > 0$, where $\lambda_{max}(G_i)$ denotes the largest eigenvalue of the positive definite matrix G_i, from (3.11) and (3.15), we have

$$
(w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k) \geq \frac{1}{2}(\|w - w^{k+1}\|_{\Gamma_1}^2 - \|w - w^k\|_{\Gamma_1}^2) + \frac{1}{2}\delta_1\|w^k - \tilde{w}^k\|^2. \tag{3.16}
$$

which implies the proof of (3.9). □

Theorem 3.1 Let $A_1^\top A_2 = 0$, $\beta \in (0, 2)$ and the sequence $\{w^k\}$ be generated by the L-GADMM (1.8), then

$$
\|w^{k+1} - w^*\|_{\Gamma_1}^2 \leq \|w^k - w^*\|_{\Gamma_1}^2 - \delta_1\|w^k - \tilde{w}^k\|^2, \forall w^* \in \Omega^*, \tag{3.17}
$$

where Γ_1 is defined in (2.17) and $\delta_1 = \min\{\lambda_{max}(G_1), \lambda_{max}(G_2), \lambda_{max}(G_3), \frac{2-\beta}{\rho}\} > 0$.

Proof By Lemma 3.1 and Lemma 3.2, we know that for any $w \in \Omega$,

$$
\begin{aligned}
&f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(\tilde{w}^k) \geq (w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k) \\
&\geq \frac{1}{2}(\|w - w^{k+1}\|_{\Gamma_1}^2 - \|w - w^k\|_{\Gamma_1}^2) + \frac{1}{2}\delta_1\|w^k - \tilde{w}^k\|^2.
\end{aligned} \tag{3.18}
$$

Set $w = w^* \in \Omega$ in the above inequality, we have

$$\begin{aligned} & f(x^*) - f(\tilde{x}^k) + (w^* - \tilde{w}^k)^\top F(\tilde{w}^k) \\ & \geq (w^* - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k) \\ & \geq \frac{1}{2}(\|w^* - w^{k+1}\|_{\Gamma_1}^2 - \|w^* - w^k\|_{\Gamma_1}^2) + \frac{1}{2}\delta_1\|w^k - \tilde{w}^k\|^2. \end{aligned} \tag{3.19}$$

On the other hand, by (3.1) and the monotonicity of F, we have

$$0 \leq f(\tilde{x}^k) - f(x^*) + (\tilde{w}^k - w^*)^\top F(\tilde{w}^k) \leq (\tilde{w}^k - w^*)^\top Q_1(w^k - \tilde{w}^k). \tag{3.20}$$

Consequently, (3.19) and the above inequality imply that

$$(\|w^k - w^*\|_{\Gamma_1}^2 - \|w^{k+1} - w^*\|_{\Gamma_1}^2) - \delta_1\|w^k - \tilde{w}^k\|^2 \geq 0. \tag{3.21}$$

The assertion of this theorem follows directly by (3.21). □

Lemma 3.3 Let the sequence $\{w^k\}$ be generated by L-GADMM (1.9) and the associated sequence $\{\bar{w}\}^k$ be defined in (2.7). Then we have

$$f(x) - f(\bar{x}^k) + (w - \bar{w}^k)^\top F(\bar{w}^k) \geq (w - \bar{w}^k)^\top Q_2(w^k - \bar{w}^k), \quad \forall w \in \Omega, \tag{3.22}$$

where Q_2 is defined in (2.15).

Proof The proof of this lemma is analogous to Lemma 1. By the optimality coditions of the x_1-subproblem for the scheme (1.9), we have

$$\begin{aligned} & f_1(x_1) - f_1(x_1^{k+1}) + (x_1 - x_1^{k+1})^\top\{-A_1^\top\lambda^{k+1} + \rho A_1^\top(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^{k+1} - b) \\ & \quad + G_1(x_1^{k+1} - x_1^k)\} \geq 0, \quad \forall x_1 \in \mathcal{X}_1. \end{aligned}$$

By (2.7), the above inequality can be rewritten as

$$\begin{aligned} & f_1(x_1) - f_1(\bar{x}_1^k) + (x_1 - \bar{x}_1^k)^\top\{-A_1^\top\bar{\lambda}^k + (G_1 + \rho A_1^\top A_1)(\bar{x}_1^k - x_1^k) + \rho(1-\beta)A_1^\top A_3(\bar{x}_3^k - x_3^k) \\ & \quad - \beta A_1^\top(\bar{\lambda}^k - \lambda^k)\} \geq 0, \quad \forall x_1 \in \mathcal{X}_1. \end{aligned} \tag{3.23}$$

By the optimality condition for the x_2-subproblem in (1.9), we get

$$\begin{aligned} & f_2(x_2) - f_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top\{-A_2^\top\lambda^k + \rho A_2^\top(A_1x_1^k + A_2x_2^{k+1} + A_3x_3^k - b) \\ & \quad + G_2(x_2^{k+1} - x_2^k)\} \geq 0, \quad \forall x_2 \in \mathcal{X}_2. \end{aligned}$$

From (2.7) and (2.8), we can obtain

$$f_2(x_2) - f_2(\bar{x}_2^k) + (x_2 - \bar{x}_2^k)^\top\{-A_2^\top\bar{\lambda}^k + G_2(\bar{x}_2^k - x_2^k) - \rho A_2^\top A_3(\bar{x}_3^k - x_3^k)\} \geq 0, \quad \forall x_2 \in \mathcal{X}_2. \tag{3.24}$$

By the optimality condition for the x_3-subproblem in (1.9) and (2.7) and (2.8), we obtain

$$f_3(x_3) - f_3(\bar{x}_3^k) + (x_3 - \bar{x}_3^k)^\top\{-A_3^\top\bar{\lambda}^k + (G_3 + \rho(1-\beta)A_3^\top A_3)(\bar{x}_3^k - x_3^k) + (1-\beta)A_3^\top(\bar{\lambda}^k - \lambda^k) - A_1(\bar{x}_3^k - x_3^k) + \frac{1}{\rho}(\bar{\lambda}^k - \lambda^k)\} \ge 0, \quad \forall x_3 \in \mathcal{X}_3. \tag{3.25}$$

It follows from (2.7) that

$$A_1\bar{x}_1^k + A_2\bar{x}_2^k + A_3\bar{x}_3^k - b - A_1(\bar{x}_1^k - x_1^k) + \frac{1}{\rho}(\bar{\lambda}^k - \lambda^k) = 0. \tag{3.26}$$

Using (3.23)-(3.26), we get

$$f(x) - f(\bar{x}^k) + \begin{pmatrix} x_1 - \bar{x}_1^k \\ x_2 - \bar{x}_2^k \\ x_3 - \bar{x}_3^k \\ \lambda - \bar{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} -A_1^\top\bar{\lambda}^k \\ -A_2^\top\bar{\lambda}^k \\ -A_3^\top\bar{\lambda}^k \\ A_1\bar{x}_1^k + A_2\bar{x}_2^k + A_3\bar{x}_3^k - b \end{pmatrix} + \begin{pmatrix} x_1 - \bar{x}_1^k \\ x_2 - \bar{x}_2^k \\ x_3 - \bar{x}_3^k \\ \lambda - \bar{\lambda}^k \end{pmatrix}^T \Phi(\bar{w}^k - w^k) \right\} \ge 0$$

where

$$\Phi(\bar{w}^k - w^k) = \begin{pmatrix} (G_1 + \rho A_1^\top A_1)(\bar{x}_1^k - x_1^k) + \rho(1-\beta)A_1^\top A_3(\bar{x}_3^k - x_3^k) - \beta A_1^\top(\bar{\lambda}^k - \lambda^k) \\ G_2(\bar{x}_2^k - x_2^k) - \rho A_2^\top A_3(\bar{x}_3^k - x_3^k) \\ (G_3 + \rho(1-\beta)A_3^\top A_3)(\bar{x}_3^k - x_3^k) + (1-\beta)A_3^\top(\bar{\lambda}^k - \lambda^k) \\ -A_1(\bar{x}_3^k - x_3^k) + \frac{1}{\rho}(\bar{\lambda}^k - \lambda^k) \end{pmatrix}.$$

By (2.2) and (2.15), the above inequality can be rewritten as

$$f(x) - f(\bar{x}^k) + (w - \bar{w}^k)^\top(F(\bar{w}^k) + Q_2(w^k - \bar{w}^k)) \ge 0, \qquad \forall w \in \Omega. \qquad \square$$

Lemma 3.4 Let $A_2^\top A_3 = 0$, $\beta \in (0,1]$, $G_i - \rho(1-\beta)A_i^\top A_i \succeq 0 (i = 1,3)$, the sequence $\{w^k\}$ be generated by the L-GADMM (1.9) and the associated sequence $\{\bar{w}^k\}$ be defined in (2.7). Then there exists $\delta_2 > 0$ such that

$$f(x) - f(\bar{x}^k) + (w - \bar{w}^k)^\top F(\bar{w}^k) \ge \frac{1}{2}(\|w^{k+1} - w\|_{Y_1}^2 - \|w^k - w\|_{Y_1}^2) + \frac{1}{2}\delta_2\|w^k - \bar{w}^k\|^2, \quad \forall w \in \Omega. \tag{3.27}$$

Where Y_1 is defined in (2.20).
Proof It follows from (2.9), (2.22), (3.22) and $A_2^\top A_3 = 0$ that

$$f(x) - f(\bar{x}^k) + (w - \bar{w}^k)^\top(F(\bar{w}^k) - (w - \bar{w}^k)^\top\Gamma_2(w^k - w^{k+1})) \ge 0, \quad \forall w \in \Omega, \tag{3.28}$$

Adding the term $(w - \bar{w}^k)Y_1(w^k - w^{k+1})$ to both sides of (3.28), we have

$$f(x) - f(\bar{x}^k) + (w - \bar{w}^k)^\top F(\bar{w}^k) - (w - w^{k+1})^\top\rho(1-\beta)\mathrm{T}A_3(x_3^k - x_3^{k+1}) \ge (w - \bar{w}^k)^\top Y_1(w^k - w^{k+1}), \tag{3.29}$$

where

$$\mathrm{T} = \begin{pmatrix} A_1^\top \\ A_2^\top \\ A_3^\top \\ 0 \end{pmatrix}.$$

According to (2.7) and using $A_2^\top A_3 = 0$, (3.29) can be rewritten as

$$\begin{aligned} &(w - \bar{w}^k)^\top Y_1(w^k - w^{k+1}) - (x_3^k - \bar{x}_3^k)^\top (1-\beta) A_3^\top (\lambda^k - \bar{\lambda}^k) \\ &+ (x_3^k - \bar{x}_3^k)^\top \rho(1-\beta) A_3^\top A_1 (x_1^k - \bar{x}_1^k) \le (w - \bar{w}^k)^\top F(\bar{w}^k) + f(x) - f(\bar{x}^k) \end{aligned} \tag{3.30}$$

Note that the Y_1 is positive definite since $G_3 - \rho(1-\beta)A_3^\top A_3 \succeq \mathbf{0}$, it follows that

$$\begin{aligned} (w - \bar{w}^k)^\top Y_1 (w^k - w^{k+1}) =& \frac{1}{2}(\|w - w^{k+1}\|_{Y_1}^2 - \|w - w^k\|_{Y_1}^2) \\ &+ \frac{1}{2}(\|w^k - \bar{w}^k\|_{Y_1}^2 - \|w^{k+1} - \bar{w}^k\|_{Y_1}^2) \end{aligned} \tag{3.31}$$

By (3.30) and the above equation, we have

$$\begin{aligned} &\frac{1}{2}(\|w - w^{k+1}\|_{Y_1}^2 - \|w - w^k\|_{Y_1}^2) + \frac{1}{2}(\|w^k - \bar{w}^k\|_{Y_1}^2 - \|w^{k+1} - \bar{w}^k\|_{Y_1}^2) \\ &- (x_3^k - \bar{x}_3^k)^\top (1-\beta) A_3^\top (\lambda^k - \bar{\lambda}^k) + (x_3^k - \bar{x}_3^k)^\top \rho(1-\beta) A_3^\top A_1 (x_1^k - \bar{x}_1^k) \\ \le& (w - \bar{w}^k)^\top F(\bar{w}^k) + f(x) - f(\bar{x}^k). \end{aligned} \tag{3.32}$$

The remaining task is to prove that there exists $\delta_2 > 0$ such that

$$\begin{aligned} &\frac{1}{2}(\|w^k - \bar{w}^k\|_{Y_1}^2 - \|w^{k+1} - \bar{w}^k\|_{Y_1}^2) - (x_3^k - \bar{x}_3^k)^\top (1-\beta) A_3^\top (\lambda^k - \bar{\lambda}^k) \\ &\quad + (x_3^k - \bar{x}_3^k)^\top \rho(1-\beta) A_3^\top A_1 (x_3^k - \bar{x}_3^k) \ge \frac{1}{2}\delta_2 \|w^k - \bar{w}^k\|^2. \end{aligned} \tag{3.33}$$

By the definition of Y_1 and $A_2^\top A_3 = 0$, we have

$$\begin{aligned} &\|w^k - \bar{w}^k\|_{Y_1}^2 \\ =& \|x_1^k - \bar{x}_1^k\|_{(G_1 + \rho A_1^\top A_1)}^2 + \|x_2^k - \bar{x}_2^k\|_{G_2}^2 + \|x_3^k - \bar{x}_3^k\|_{(G_3 + \rho\frac{(1-\beta)^2 A_3^\top A_3}{\beta})}^2 \\ &+ \frac{1}{\rho\beta}\|\lambda^k - \bar{\lambda}^k\|^2 + 2(x_3^k - \bar{x}_3^k)^\top \frac{1-\beta}{\beta} A_3^\top (\lambda^k - \bar{\lambda}^k) - 2(x_1^k - \bar{x}_1^k)^\top A_1^\top (\lambda^k - \bar{\lambda}^k) \\ =& \|x_1^k - \bar{x}_1^k\|_{(G_1 + \rho A_1^\top A_1)}^2 + \|x_2^k - \bar{x}_2^k\|_{G_2}^2 + \|x_3^k - \bar{x}_3^k\|_{G_3}^2 \\ &+ (\frac{1}{\rho\beta} - \frac{(1-\beta)^2}{\rho\beta})\|\lambda^k - \bar{\lambda}^k\|^2 + 2(1-\beta)(x_3^k - \bar{x}_3^k)^\top A_3^\top (\lambda^k - \bar{\lambda}^k) \\ &- 2(x_1^k - \bar{x}_1^k)^\top A_1^\top (\lambda^k - \bar{\lambda}^k) + \frac{(1-\beta)^2}{\rho\beta}\|\rho A_3 (x_3^k - \bar{x}_3^k) + (\lambda^k - \bar{\lambda}^k)\|^2. \end{aligned} \tag{3.34}$$

On the other hand, by (2.7) and (2.8), we have

$$\lambda^{k+1} - \bar{\lambda}^k = \rho(1-\beta) A_3 (x_3^k - \bar{x}_3^k) + (1-\beta)(\lambda^k - \bar{\lambda}^k) \tag{3.35}$$

It follows from (2.7) and (3.35) that

$$\|w^{k+1} - \bar{w}^k\|_{Y_1}^2 = \frac{1}{\rho\beta}\|\lambda^{k+1} - \bar{\lambda}^k\|^2 = \frac{(1-\beta)^2}{\rho\beta}\|\rho A_3 (x_3^k - \bar{x}_3^k) + (\lambda^k - \bar{\lambda}^k)\|^2 \tag{3.36}$$

By (3.34) and (3.36), we have

$$\begin{aligned}
&\frac{1}{2}(\|w^k-\bar{w}^k\|_{Y_1}^2-\|w^{k+1}-\bar{w}^k\|_{Y_1}^2)\\
=&\frac{1}{2}\|x_1^k-\bar{x}_1^k\|_{(G_1+\rho A_1^\top A_1)}^2\\
&+\frac{1}{2}\|x_2^k-\bar{x}_2^k\|_{G_2}^2+\frac{1}{2}\|x_3^k-\bar{x}_3^k\|_{G_3}^2+\frac{1}{2}(\frac{1}{\rho\beta}-\frac{(1-\beta)^2}{\rho\beta})\|\lambda^k-\bar{\lambda}^k\|^2\\
&+(1-\beta)(x_3^k-\bar{x}_3^k)^\top A_3^\top(\lambda^k-\bar{\lambda}^k)-(x_1^k-\bar{x}_1^k)^\top A_1^\top(\lambda^k-\bar{\lambda}^k)
\end{aligned} \tag{3.37}$$

According to the above equation, we have

$$\begin{aligned}
&\frac{1}{2}(\|w^k-\bar{w}^k\|_{Y_1}^2-\|w^{k+1}-\bar{w}^k\|_{Y_1}^2)-(1-\beta)(x_3^k-\bar{x}_3^k)^\top A_3^\top(\lambda^k-\bar{\lambda}^k)\\
&+(x_3^k-\bar{x}_3^k)\rho(1-\beta)A_3^\top A_1(x_1^k-\bar{x}_1^k)\\
=&\frac{1}{2}\|x_1^k-\bar{x}_1^k\|_{(G_1+\rho A_1^\top A_1)}^2+\frac{1}{2}\|x_2^k-\bar{x}_2^k\|_{G_2}^2\\
&+\frac{1}{2}\|x_3^k-\bar{x}_3^k\|_{G_3}^2+\frac{1}{2}(\frac{1}{\rho\beta}-\frac{(1-\beta)^2}{\rho\beta})\|\lambda^k-\bar{\lambda}^k\|^2\\
&+\rho(1-\beta)(x_3^k-\bar{x}_3^k)^\top A_3^\top A_1(x_1^k-\bar{x}_1^k)-(x_1^k-\bar{x}_1^k)^\top A_1^\top(\lambda^k-\bar{\lambda}^k).
\end{aligned} \tag{3.38}$$

Using the Cauchy-Schwartz inequality, we can get

$$\begin{aligned}
&\rho(1-\beta)(x_3^k-\bar{x}_3^k)^\top A_3^\top A_1^\top(x_1^k-\bar{x}_1^k)-(x_1^k-\bar{x}_1^k)^\top A_1^\top(\lambda^k-\bar{\lambda}^k)\\
\geq&[-\frac{1}{2}\|x_1^k-\bar{x}_1^k\|_{\rho A_1^\top A_1}^2-\frac{1}{2}\|\lambda^k-\bar{\lambda}^k\|_{\frac{1}{\rho}I_p}^2]-[\frac{1}{2}\|x_3^k-\bar{x}_3^k\|_{\rho(1-\beta)A_3^\top A_3}^2\\
&+\frac{1}{2}\|x_1^k-\bar{x}_1^k\|_{\rho(1-\beta)A_1^\top A_1}^2]+\frac{1}{2}\rho(1-\beta)\|A_3(x_3^k-\bar{x}_3^k)+A_1(x_1^k-\bar{x}_1^k)\|^2
\end{aligned} \tag{3.39}$$

Substituting the above inequality into (3.38), we obtain

$$\begin{aligned}
&\frac{1}{2}(\|w^k-\bar{w}^k\|_{Y_1}^2-\|w^{k+1}-\bar{w}^k\|_{Y_1}^2)-(1-\beta)(x_3^k-\bar{x}_3^k)^\top A_3^\top(\lambda^k-\bar{\lambda}^k)\\
&+(x_3^k-\bar{x}_3^k)\rho(1-\beta)A_3^\top A_1^\top(x_1^k-\bar{x}_1^k)\\
\geq&\frac{1}{2}(x_1^k-\bar{x}_1^k)^\top[G_1-\rho(1-\beta)A_1^\top A_1](x_1^k-\bar{x}_1^k)+\frac{1}{2}\|x_2^k-\bar{x}_2^k\|_{G_2}^2\\
&+(x_3^k-\bar{x}_3^k)^\top[G_3-\rho(1-\beta)A_3^\top A_3](x_3^k-\bar{x}_3^k)+\frac{1}{2}(\lambda^k-\bar{\lambda}^k)^\top[\frac{1}{\rho\beta}\\
&-\frac{(1-\beta)^2}{\rho\beta}-\frac{1}{\rho}](\lambda^k-\bar{\lambda}^k)
\end{aligned} \tag{3.40}$$

Case (I) If $\beta\in(0,1)$ and $G_i-\rho(1-\beta)A_i^\top A_i\succ\mathbf{0}$ (i=1, 3), we set

$$\begin{aligned}
\delta_2^1=&\min\{\lambda_{\max}(G_1-\rho(1-\beta)A_1^\top A_1),\lambda_{\max}(G_2),\lambda_{\max}(G_3-\rho(1-\beta)A_3^\top A_3),\\
&(\frac{1}{\rho\beta}-\frac{(1-\beta)^2}{\rho\beta}-\frac{1}{\rho})\}>0.
\end{aligned}$$

Case (II) If $\beta\in(0,1)$, $G_1-\rho(1-\beta)A_1^\top A_1=\mathbf{0}$ and $G_3-\rho(1-\beta)A_3^\top A_3\succ\mathbf{0}$, we set

$$\delta_2^2=\min\{\lambda_{\max}(G_2),\lambda_{\max}(G_3-\rho(1-\beta)A_3^\top A_3),(\frac{1}{\rho\beta}-\frac{(1-\beta)^2}{\rho\beta}-\frac{1}{\rho})\}>0.$$

Case (III) If $\beta\in(0,1)$, $G_3-\rho(1-\beta)A_3^\top A_3=\mathbf{0}$ and $G_1-\rho(1-\beta)A_1^\top A_1\succ\mathbf{0}$, let

$$\delta_2^3=\min\{\lambda_{\max}(G_1-\rho(1-\beta)A_1^\top A_1),\lambda_{\max}(G_2),(\frac{1}{\rho\beta}-\frac{(1-\beta)^2}{\rho\beta}-\frac{1}{\rho})\}>0.$$

Case (IV) If $\beta \in (0,1)$ and $G_i - \rho(1-\beta)A_i^\top A_i = \mathbf{0}$ (i=1, 3), we set

$$\delta_2^4 = \min\{\lambda_{\max}(G_2), (\frac{1}{\rho\beta} - \frac{(1-\beta)^2}{\rho\beta} - \frac{1}{\rho})\} > 0.$$

Case (V) If $\beta = 1$ and $G_i - \rho(1-\beta)A_i^\top A_i \succ \mathbf{0}$ (i=1, 3), let

$$\delta_2^5 = \min\{\lambda_{\max}(G_1 - \rho(1-\beta)A_1^\top A_1), \lambda_{\max}(G_2), \lambda_{\max}(G_3 - \rho(1-\beta)A_3^\top A_3)\} > 0.$$

Case (VI) If $\beta = 1$, $G_3 - \rho(1-\beta)A_3^\top A_3 = \mathbf{0}$ and $G_1 - \rho(1-\beta)A_1^\top A_1 \succ \mathbf{0}$, we set

$$\delta_2^6 = \min\{\lambda_{\max}(G_1 - \rho(1-\beta)A_1^\top A_1), \lambda_{\max}(G_2)\} > 0.$$

Case (VII) If $\beta = 1$, $G_1 - \rho(1-\beta)A_1^\top A_1 = \mathbf{0}$ and $G_3 - \rho(1-\beta)A_3^\top A_3 \succ \mathbf{0}$, let

$$\delta_2^7 = \min\{\lambda_{\max}(G_2), \lambda_{\max}(G_3 - \rho(1-\beta)A_3^\top A_3)\} > 0.$$

Case (VIII) If $\beta = 1$ and $G_i - \rho(1-\beta)A_i^\top A_i = \mathbf{0}$ (i=1, 3), let

$$\delta_2^8 = \lambda_{\max}(G_2) > 0.$$

Let $\delta_2 = \min\{\delta_2^1, \delta_2^2, \delta_2^3, \delta_2^4, \delta_2^5, \delta_2^6, \delta_2^7, \delta_2^8\} > 0$, it follows from (3.40) that

$$\begin{aligned}&\frac{1}{2}(\|w^k - \bar{w}^k\|_{Y_1}^2 - \|w^{k+1} - \bar{w}^k\|_{Y_1}^2) - (1-\beta)(x_3^k - \bar{x}_3^k)^\top A_3^\top(\lambda^k - \bar{\lambda}^k)\\&+ (x_3^k - \bar{x}_3^k)\rho(1-\beta)A_3^\top A_1^\top(x_1^k - \bar{x}_1^k) \geq \frac{1}{2}\delta_2\|w^k - \bar{w}^k\|^2.\end{aligned} \tag{3.41}$$

That is, there exists $\delta_2 > 0$ such that (3.33) holds. The proof of this lemma follows directly by (3.32) and (3.33). □

Theorem 3.2 Let $A_2^\top A_3 = 0$, $\beta \in (0,1]$, $G_i - \rho(1-\beta)A_i^\top A_i \succeq \mathbf{0}$ $(i = 1, 3)$. Then there exists $\delta_2 > 0$ such that the sequence $\{w^k\}$ generated by L-GADMM (1.9) satisfies that for all $w^* \in \Omega^*$

$$\|w^{k+1} - w^*\|_{Y_1}^2 \leq \|w^k - w^*\|_{Y_1}^2 - \delta_2\|w^k - \bar{w}^k\|^2 \tag{3.42}$$

where Y_1 is defined in (2.20).

Proof The proof of this result is analogous to Therorem 3.1. By Lemma 3.3 and Lemma 3.4, we get

$$\begin{aligned}&f(x) - f(\bar{x}^k) + (w - \bar{w}^k)^\top F(\bar{w}^k)\\&\geq (w - \bar{w}^k)^\top Q_2(w^k - \bar{w}^k)\\&\geq \frac{1}{2}(\|w - w^{k+1}\|_{Y_1}^2 - \|w - w^k\|_{Y_1}^2) + \frac{1}{2}\delta_2\|w^k - \bar{w}^k\|^2\end{aligned} \tag{3.43}$$

Set $w = w^* \in \Omega$ in the above equation, we have

$$\begin{aligned} & f(x^*) - f(\bar{x}^k) + (w^* - \bar{w}^k)^\top F(\bar{w}^k) \\ & \quad \geq (w^* - \bar{w}^k)^\top Q_2(w^k - \bar{w}^k) \\ & \quad \geq \frac{1}{2}(\|w^* - w^{k+1}\|_{Y_1}^2 - \|w^* - w^k\|_{Y_1}^2) + \frac{1}{2}\delta_2\|w^k - \bar{w}^k\|^2 \end{aligned} \tag{3.44}$$

On the other hand, by (3.1) and the monotonicity of F, we have

$$0 \leq f(\bar{x}^k) - f(x^*) + (\bar{w}^k - w^*)^\top F(\bar{w}^k) \leq (\bar{w}^k - w^*)^\top Q_2(w^k - \bar{w}^k) \tag{3.45}$$

Consequently, using (3.19) and the above equation, we obtain

$$(\|w^k - w^*\|_{Y_1}^2 - \|w^{k+1} - w^*\|_{Y_1}^2) - \delta_2\|w^k - \bar{w}^k\|^2 \geq 0 \tag{3.46}$$

The assertion (3.42) follows directly by (3.46). □

Lemma 3.5 Let the sequence $\{w^k\}$ be generated by the L-GADMM (1.9) and the associated sequence $\{\hat{w}\}^k$ be defined in (2.11). Then we have

$$f(x) - f(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k) \geq (w - \hat{w}^k)^\top Q_3(w^k - \hat{w}^k), \qquad \forall w \in \Omega, \tag{3.47}$$

where Q_3 is defined in (2.16).

Proof By the optimality conditions of the x_1-subproblem, x_2-subproblem and x_3-subproblem for the iterative scheme (1.9), then we can obtain

$$\begin{cases} f_1(x_1) - f_1(x_1^{k+1}) + (x_1 - x_1^{k+1})^\top(-A_1^\top\lambda^{k+1} + \rho A_1^\top(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^{k+1} - b) \\ \qquad + G_1(x_1^{k+1} - x_1^k)) \geq 0, \forall x_1 \in \mathcal{X}_1, \\ f_2(x_2) - f_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top(-A_2^\top\lambda^k + \rho A_2^\top(A_1x_1^k + A_2x_2^{k+1} + A_3x_3^k - b) \\ \qquad + G_2(x_2^{k+1} - x_2^k)) \geq 0, \quad \forall x_2 \in \mathcal{X}_2, \\ f_3(x_3) - f_3(x_3^{k+1}) + (x_3 - x_3^{k+1})^\top(-A_3^\top\lambda^k + \rho A_3^\top(\beta A_1x_1^k + (1-\beta)(b - A_3x_3^k) + \beta A_2x_2^{k+1} + A_3x_3^{k+1} - b) \\ \qquad + G_3(x_3^{k+1} - x_3^k)) \geq 0, \quad \forall x_3 \in \mathcal{X}_3, \\ \lambda^{k+1} = \lambda^k - \rho(\beta A_1x_1^k + (1-\beta)(b - A_3x_3^k) + \beta A_2x_2^{k+1} + A_3x_3^{k+1} - b). \end{cases}$$

It follows from (2.11) and (2.12) that

$$\begin{cases} f_1(x_1) - f_1(\hat{x}_1^k) + (x_1 - \hat{x}_1^k)^\top(-A_1^\top\hat{\lambda}^k - \beta A_1^\top(\hat{\lambda}^k - \lambda^k) + 2\rho A_1^\top A_3(\hat{x}_3^k) \\ \qquad + (G_1 + \rho A_1^\top A_1)(\hat{x}_1^k - x_1^k)) \geq 0, \forall x_1 \in \mathcal{X}_1, \\ f_2(x_2) - f_2(\hat{x}_2^k) + (x_2 - \hat{x}_2^k)^\top(-A_2^\top\hat{\lambda}^k + G_2(\hat{x}_2^k - x_2^k) \geq 0, \quad \forall x_2 \in \mathcal{X}_2, \\ f_3(x_3) - f_3(\hat{x}_3^k) + (x_3 - \hat{x}_3^k)^\top(-A_3^\top\hat{\lambda}^k + (1-\beta)A_3^\top(\hat{\lambda}^k - \lambda^k) \\ \qquad + (G_3 + \rho A_3^\top A_3)(\hat{x}_3^k - x_3^k)) \geq 0, \quad \forall x_3 \in \mathcal{X}_3, \\ A_1\hat{x}_1^k + A_2\hat{x}_2^k + A_3\hat{x}_3^k - b - A_1(\hat{x}_1^k - x_1^k) - A_3(\hat{x}_3^k - x_3^k) + \frac{1}{\rho}(\hat{\lambda}^k - \lambda^k) = 0. \end{cases} \tag{3.48}$$

According to (2.2) and (2.16), (3.50) can be written as

$$f(x) - f(\hat{x}^k) + (w - \hat{w}^k)^\top (F(\hat{w}^k) + Q_3(\hat{w}^k - w^k)) \geq 0, \qquad \forall w \in \Omega.$$

The assertion of (3.47) is proved by using the above inequality. $\square$

Lemma 3.6 Let $A_1^\top A_3 = 0$, $\beta \in (0, 1]$, the sequence $\{w^k\}$ be generated by L-GADMM (1.9) and the associated sequence $\{\hat{w}^k\}$ be defined in (2.11). Then there exists $\delta_3 > 0$ such that

$$f(x) - f(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k) \geq \frac{1}{2}(\|w^{k+1} - w\|_{\Gamma_3}^2 - \|w^k - w\|_{\Gamma_3}^2) + \frac{1}{2}\delta_3 \|w^k - \hat{w}^k\|^2, \quad \forall w \in \Omega. \tag{3.49}$$

Where Γ_3 is defined in (2.19).

Proof It is easy to see that Γ_3 is positive definite and $Q_3 = \Gamma_3 M_1$ since $A_1^\top A_3 = 0$ and $\beta \in (0, 1]$. It follows from (2.13) that

$$\begin{aligned} (w - \hat{w}^k)^\top Q_3(w^k - \hat{w}^k) &= (w - \hat{w}^k)^\top \Gamma_3 M_1 (w^k - \hat{w}^k) \\ &= (w - \hat{w}^k)^\top \Gamma_3 (w^k - w^{k+1}). \end{aligned} \tag{3.50}$$

By the positive definiteness of Γ_3, then we can obtain the following identity

$$\begin{aligned} (w - \hat{w}^k)^\top \Gamma_3 (w^k - w^{k+1}) =& \frac{1}{2}(\|w - w^{k+1}\|_{\Gamma_3}^2 - \|w - w^k\|_{\Gamma_3}^2) \\ &+ \frac{1}{2}(\|w^k - \hat{w}^k\|_{\Gamma_3}^2 - \|w^{k+1} - \hat{w}^k\|_{\Gamma_3}^2). \end{aligned} \tag{3.51}$$

Hence,

$$\begin{aligned} &\|w^k - \hat{w}^k\|_{\Gamma_3}^2 \\ =&\|x_1 - \hat{x}_1^k\|_{(G_1 + \rho A_1^\top A_1)}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \|x_3^k - \hat{x}_3^k\|_{G_3}^2 + \|x_3^k - \hat{x_3}^k\|_{\frac{\rho}{\beta} A_3^\top A_3} \\ &- 2(x_1^k - \hat{x}_1^k)^\top A_1^\top (\lambda^k - \hat{\lambda}^k) + 2(x_3^k - \hat{x}_3^k)\frac{1-\beta}{\beta} A_3^\top (\lambda^k - \hat{\lambda}^k) + \frac{1}{\rho\beta}\|\lambda^k - \hat{\lambda}^k\|_{I_p}^2 \\ =&\|x_1 - \hat{x}_1^k\|_{(G_1 + \rho A_1^\top A_1)}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \|x_3^k - \hat{x}_3^k\|_{G_3}^2 + \frac{2-\beta}{\rho}\|\lambda^k - \hat{\lambda}^k\|_{I_p}^2 \\ &+ \frac{1}{\rho\beta}\|\rho A_3(x_3^k - \tilde{x}_3^k) + (1-\beta)(\lambda^k - \hat{\lambda}^k)\|^2 - 2(x_1^k - \hat{x}_1^k) A_1^\top (\lambda^k - \hat{\lambda}^k). \end{aligned} \tag{3.52}$$

It follows from (2.12) that

$$\lambda^{k+1} - \hat{\lambda}^k = \rho A_3(x_3^k - \hat{x}_3^k) + (1-\beta)(\lambda^k - \hat{\lambda}^k). \tag{3.53}$$

By the definition of $\{\hat{w}^k\}$, we have

$$\|w^{k+1} - \hat{w}^k\|_{\Gamma_3}^2 = \frac{1}{\rho\beta}\|\lambda^{k+1} - \hat{\lambda}^k\|^2 = \frac{1}{\rho\beta}\|\rho A_3(x_3^k - \hat{x}_3^k) + (1-\beta)(\lambda^k - \hat{\lambda}^k)\|^2. \tag{3.54}$$

Using (3.52) and (3.54), we can obtain

$$
\begin{aligned}
&\|w^k - \hat{w}^k\|_{\Gamma_3}^2 - \|w^{k+1} - \hat{w}^k\|_{\Gamma_3}^2 \\
=&\|x_1 - \hat{x}_1^k\|_{(G_1+\rho A_1^\top A_1)}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \|x_3^k - \hat{x}_3^k\|_{G_3}^2 + \frac{2-\beta}{\rho}\|\lambda^k - \hat{\lambda}^k\|_{I_p}^2 \\
&- 2(x_1^k - \hat{x}_1^k)A_1^\top(\lambda^k - \hat{\lambda}^k) \\
\geq&\|x_1 - \hat{x}_1^k\|_{(G_1+\rho A_1^\top A_1)}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \|x_3^k - \hat{x}_3^k\|_{G_3}^2 + \frac{2-\beta}{\rho}\|\lambda^k - \hat{\lambda}^k\|_{I_p}^2 \\
&- \|x_1 - \hat{x}_1^k\|_{\rho A_1^\top A_1}^2 - \frac{1}{\rho}\|\lambda^k - \hat{\lambda}^k\|^2 \\
=&\|x_1 - \hat{x}_1^k\|_{G_1}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \|x_3^k - \hat{x}_3^k\|_{G_3}^2 + \frac{1-\beta}{\rho}\|\lambda^k - \hat{\lambda}^k\|_{I_p}^2.
\end{aligned}
\tag{3.55}
$$

Clearly, $\frac{1-\beta}{\rho} \geq 0$ for $\beta \in (0, 1]$. Thus by the positve definiteness of G_i (i=1, 2, 3), we have the following two cases.

Case (I) If $\beta \in (0, 1)$, we set

$$\delta_3^1 = \min\{\lambda_{\max}(G_1), \lambda_{\max}(G_2), \lambda_{\max}(G_3), \frac{1-\beta}{\rho}\}.$$

Case (II) If $\beta = 1$, we set

$$\delta_3^2 = \min\{\lambda_{\max}(G_1), \lambda_{\max}(G_2), \lambda_{\max}(G_3)\}.$$

Let $\delta_3 = \min\{\delta_3^1, \delta_3^2\} > 0$, it follows from (3.55) that

$$\|w^k - \hat{w}^k\|_{\Gamma_3}^2 - \|w^{k+1} - \hat{w}^k\|_{\Gamma_3}^2 \geq \delta_3\|w^k - \hat{w}^k\|^2,$$

Using the above inequality, (3.47), (3.50) and (3.51), the assertion (3.49) is proved. □

We show that the sequences $\{w^k\}$ generated by the scheme (1.9) are monotonous nonincreasing sequences in the following theorem. Its proof is analogous to that of Theorem 3.1, we thus omit it.

Theorem 3.3 Let $A_1^\top A_3 = 0$, $\beta \in (0, 1]$, the sequence $\{w^k\}$ be generated by L-GADMM (1.9) and Γ_3 be defined in (2.19). Then there exists $\delta_3 > 0$ such that for all $w^* \in \Omega^*$,

$$\|w^{k+1} - w^*\|_{\Gamma_3}^2 \leq \|w^k - w^*\|_{\Gamma_3}^2 - \delta_3\|w^k - \hat{w}^k\|^2. \tag{3.56}$$

Now we present the global convergence of L-GADMM for (1.5) as follows.

Theorem 3.4 Let $A_1^\top A_2 = 0$, $\beta \in (0, 2)$. Then the sequence $\{w^k\}$ generated by scheme (1.8) converges to an optimal solution to problem (1.1) from any starting point.

Proof By Theorem 3.1, we know that

$$\delta_1\|w^k - \tilde{w}^k\|^2 \leq \|w^k - w^*\|_{\Gamma_1}^2 - \|w^{k+1} - w^*\|_{\Gamma_1}^2 \tag{3.57}$$

where Γ_1 is defined in (2.17) and $\delta_1 > 0$.

The above inequality implies that the sequence $\{w^k\}$ generated by the L-GADMM (1.8) is contractive with respect to Ω^* on account of the positive definiteness of the matrix Γ_1. By applying the standard technique of contraction methods, it is easy to show

$$\|w^k - w^*\|_{\Gamma_1}^2 \text{ convergence, and } \|w^k - \tilde{w}^k\| \to 0. \tag{3.58}$$

Hence $\{w^k\}$ is bounded. Suppose that the subsequence $\{w^{k_j}\}$ converges to w^∞, then by taking the limits on both sides of (3.1), we obtain

$$f(x) - f(x^\infty) + (w - w^\infty)^\top F(w^\infty) \geq 0, \quad \forall w \in \Omega, \tag{3.59}$$

This means that $w^\infty \in \Omega^*$, and the sequence $\{w^k\}$ is Fejér monotone and $\|w^k - w^\infty\|_{\Gamma_1}^2$ convergence by (3.58). Then $\lim_{k\to\infty} w^k = w^\infty$. Therefore, we obtain the global convergence of L-GADMM (1.8). □

Theorem 3.5 If either one of the following two conditions holds:

(i) $A_2^\top A_3 = 0$, $\beta \in (0, 1]$, $G_i - \rho(1-\beta)A_i^\top A_i \succeq \mathbf{0}$ $(i = 1, 3)$,

(ii) $A_1^\top A_3 = 0$, $\beta \in (0, 1]$,

then the sequence $\{w^k\}$ generated by L-GADMM scheme (1.9) convergenes to an optimal solution to problem (1.1) from any starting point.

Proof Suppose Condition (i) holds. Then the inequality (3.42) implies that the sequence $\{w^k\}$ generated by the L-GADMM (1.9) is contractive with respect to Ω^* on account of the positive definiteness of the matrix Y_1. By the standard technique of contraction methods, it is easy to see

$$\|w^k - w^*\|_{Y_1}^2 \text{ convergence, and } \|w^k - \bar{w}^k\| \to 0. \tag{3.60}$$

Hence $\{w^k\}$ is bounded. And we assume that the subsequence $\{w^{k_j}\}$ converges to w^∞, then by taking the limits on both sides of (3.22), we obtain

$$f(x) - f(x^\infty) + (w - w^\infty)^\top F(w^\infty) \geq 0, \quad \forall w \in \Omega, \tag{3.61}$$

And we can obtain that $w^\infty \in \Omega^*$. Obviously, the sequence $\{w^k\}$ is Fejér monotone and $\|w^k - w^\infty\|_{Y_1}^2$ convergence by (3.64), then $\lim_{k\to\infty} w^k = w^\infty$. Similarly, due to the Lemma 3.5 and Theorem 3.3, we can obtain the above conclusion under Condition (ii) and its proof is omitted. □

4 The worst case $\mathcal{O}(1/t)$ convergence rate of L-GADMM

In this section, we establish a worst-case convergence rate for the L-GADMM in both ergodic and nonergodic senses.

4.1 A worst case $\mathcal{O}(1/t)$ convergence rate for L-GADMM in the ergodic sense

By the above analysis, we will show the worst case $\mathcal{O}(1/t)$ convergence rate of L-GADMM in a ergodic sense.

Theorem 4.1 Let $A_1^\top A_2 = 0$, $\beta \in (0,2)$ and $\{w^k\}$ be the sequence generated by the L-GADMM (1.8), Γ_1 be given by (2.18). For any integer $t > 0$, let $w_t = \frac{1}{t+1}\sum_{k=0}^{t} \tilde{w}^k$, then we have $w_t \in \Omega$ and

$$f(x_t) - f(x) + (w_t - w)^\top F(w) \leq \frac{1}{2(t+1)} \|w - w^0\|_{\Gamma_1}^2, \quad \forall w \in \Omega. \tag{4.1}$$

Proof It is easy to see that $\tilde{w}^k \in \Omega$ for all $k \geq 0$. It follows the convexity of $\mathcal{X}_1$, $\mathcal{X}_2$ with $\mathcal{X}_3$ that $w_t \in \Omega$. Because of the monotonicity of the function, we have

$$(w - \tilde{w}^k)^\top F(w) \geq (w - \tilde{w}^k)^\top F(\tilde{w}^k), \tag{4.2}$$

By (3.1) and (3.9), we have

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(\tilde{w}) + \frac{1}{2}\|w^k - w\|_{\Gamma_1}^2 \geq \frac{1}{2}\|w^{k+1} - w\|_{\Gamma_1}^2 + \frac{1}{2}\delta_1 \|w^k - \tilde{w}^k\|^2,$$

Due to the above inequality and (4.2), we have for all $w \in \Omega$,

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(w) + \frac{1}{2}\|w^k - w\|_{\Gamma_1}^2 \geq \frac{1}{2}\|w^{k+1} - w\|_{\Gamma_1}^2 + \frac{1}{2}\delta_1 \|w^k - \tilde{w}^k\|^2.$$

Hence, we obtain

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(w) + \frac{1}{2}\|w^k - w\|_{\Gamma_1}^2 \geq \frac{1}{2}\|w^{k+1} - w\|_{\Gamma_1}^2, \quad \forall w \in \Omega. \tag{4.3}$$

Summing the above inequality over $k = 0, 1, ..., t$, we have

$$(t+1)f(x) - \sum_{k=0}^{t} f(\tilde{x}^k) + ((t+1)w - \sum_{k=0}^{t} \tilde{w}^k)^\top F(w) + \frac{1}{2}\|w - w^0\|_{\Gamma_1}^2 \geq 0, \quad \forall w \in \Omega. \tag{4.4}$$

According to the definition of w_t and the convexity of $f(x)$, (4.4) can be rewritten as

$$f(x_t) - f(x) + (w_t - w)^\top F(w) \leq \frac{1}{2(t+1)} \|w - w^0\|_{\Gamma_1}^2. \tag{4.5}$$

The assertion of this theorem thus follows directly. □

Let an compact set $E \subset \Omega$ and define

$$e = \sup\{\|w - w^0\|_{\Gamma_1} | w \in E\}, \tag{4.6}$$

where $w^0 = (x_1^0, x_2^0, x_3^0, \lambda^0)$ is the initial point. After t iterations of L-GADMM (1.8), we can find a $w_t \in \Omega$ such that

$$\sup_{w\in E}\{f(x_t) - f(x) + (w_t - w)^\top F(w)\} \le \frac{e^2}{2t}. \tag{4.7}$$

We thus established a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.8) in the ergodic sense.

Similarly, by Lemma 3.4 and Lemma 3.6, respectively, we can obtain the following convergence rate results:

Theorem 4.2 Let $A_2^\top A_3 = 0$, $\beta \in (0, 1]$, $G_i - \rho(1 - \beta)A_i^\top A_i \succeq \mathbf{0}$ $(i = 1, 3)$, $\{w^k\}$ be the sequence generated by the L-GADMM (1.9) and Y_1 be given by (2.20). For any integer $t > 0$, let $w_t^1 = \frac{1}{t+1}\sum_{k=0}^t \bar{w}^k$, we have $w_t^1 \in \Omega$ and

$$f(x_t) - f(x) + (w_t^1 - w)^\top F(w) \le \frac{1}{2(t+1)}\|w - w^0\|_{Y_1}^2, \quad \forall w \in \Omega,$$

where Y_1 is defined in (2.20).

Theorem 4.3 Let $A_1^\top A_3 = 0$, $\beta \in (0, 1]$, $\{w^k\}$ be the sequence generated by the L-GADMM (1.9) and Γ_3 be given by (2.19). For any integer $t > 0$, let $w_t^2 = \frac{1}{t+1}\sum_{k=0}^t \hat{w}^k$, then we have $w_t^2 \in \Omega$ and

$$f(x_t) - f(x) + (w_t^2 - w)^\top F(w) \le \frac{1}{2(t+1)}\|w - w^0\|_{\Gamma_3}^2, \quad \forall w \in \Omega,$$

where Γ_3 is defined in (2.19).

Moreover, based on the above two theorems, we can also establish a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.9) in the ergodic sense, which is analogous to the analysis of (4.7).

4.2 A worst case $\mathcal{O}(1/t)$ convergence rate for L-GADMM in a nonergodic sense

Next, we prove a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM in a nonergodic sense. We first have to mention that the term $\|w^k - w^{k+1}\|_{\Gamma_1}^2$ can be used to measure the accuracy of an iterate.

Lemma 4.1 Let $A_1^\top A_2 = 0$, $\beta \in (0,2)$. For a given sequence $\{w^k\}$, let $\{w^{k+1}\}$ be generated by the L-GADMM (1.8). Then $\tilde{w}^k$ defined in (2.3) is a solution to (2.1) if $\|w^k - w^{k+1}\|_{\Gamma_1}^2 = 0$.

Proof By Lemma 3.1, we have

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(\tilde{w}^k) \geq (w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k), \quad \forall w \in \Omega, \tag{4.8}$$

By (2.5) and (2.21), we can obtain

$$f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(\tilde{w}^k) \geq (w - \tilde{w}^k)^\top \Gamma_1(w^k - w^{k+1}), \quad \forall w \in \Omega, \tag{4.9}$$

Clearly, Γ_1 is positive definite, the right-hand side of (4.9) vanishes if $\|w^{k+1} - w^k\|_{\Gamma_1}^2 = 0$, since we can obtain $\Gamma_1(w^{k+1} - w^k) = 0$ whenever $\|w^{k+1} - w^k\|_{\Gamma_1}^2 = 0$. The assertion is proved. □

Now, we are ready to show a worst-case $\mathcal{O}(1/t)$ convergence rate for the iterative scheme (1.8) in a nonergodic sense. First, we prove the following two lemmas.

Lemma 4.2 Let the sequence $\{w^k\}$ be generated by the L-GADMM (1.8) with $\beta \in (0,2)$ and the associated $\{\tilde{w}^k\}$ be defined in (2.3); the matrix Q_1 be defined in (2.14). Then, we have

$$(\tilde{w}^k - \tilde{w}^{k+1})^\top Q_1 \left[(w^k - w^{k+1}) - (\tilde{w}^k - \tilde{w}^{k+1}) \right] \geq 0. \tag{4.10}$$

Proof By setting $w = \tilde{w}^{k+1}$ in (3.1), we get

$$f(\tilde{x}^{k+1}) - f(\tilde{x}^k) + (\tilde{w}^{k+1} - \tilde{w}^k)^\top F(\tilde{w}^k) \geq (\tilde{w}^{k+1} - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k). \tag{4.11}$$

It is easy to see that (3.1) is also true for $k := k+1$, that is

$$f(x) - f(\tilde{x}^{k+1}) + (w - \tilde{w}^{k+1})^\top F(\tilde{w}^{k+1}) \geq (w - \tilde{w}^{k+1})^\top Q_1(w^{k+1} - \tilde{w}^{k+1}), \quad \forall w \in \Omega,$$

By setting $w = \tilde{w}^k$ in the above inequality, we have

$$f(\tilde{x}^k) - f(\tilde{x}^{k+1}) + (\tilde{w}^k - \tilde{w}^{k+1})^\top F(\tilde{w}^{k+1}) \geq (\tilde{w}^k - \tilde{w}^{k+1})^\top Q_1(w^{k+1} - \tilde{w}^{k+1}). \tag{4.12}$$

Adding (4.11) and (4.12) and using the monotonicity of F, the assertion (4.10) is proved immediately. □

Lemma 4.3 Let $A_1^\top A_2 = 0$, $\beta \in (0,2)$, the sequence $\{w^k\}$ be generated by the L-GADMM (1.8) and $\{\tilde{w}^k\}$ be defined in (2.3); the matrices M_1, Γ_1 and Q_1 be defined in (2.6), (2.17) and (2.14), respectively. Then we have

$$\begin{aligned}&(w^k-\tilde{w}^k)^\top M_1^\top \Gamma_1 M_1 \left[(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\right]\\ &\geq \frac{1}{2}\|(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\|^2_{(Q_1^\top+Q_1)}.\end{aligned} \tag{4.13}$$

Proof Adding the equation

$$\begin{aligned}&\left[(w^k-w^{k+1})-(\tilde{w}^k-\tilde{w}^{k+1})\right]^\top Q_1\left[(w^k-w^{k+1})-(\tilde{w}^k-\tilde{w}^{k+1})\right]\\ &=\frac{1}{2}\|(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\|^2_{(Q_1^\top+Q_1)}\end{aligned}$$

to both sides of (4.10), we have

$$\begin{aligned}&(w^k-w^{k+1})^\top Q_1\left[(w^k-w^{k+1})-(\tilde{w}^k-\tilde{w}^{k+1})\right]\\ &\geq\frac{1}{2}\|(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\|^2_{(Q_1^\top+Q_1)}\end{aligned} \tag{4.14}$$

It follows from $A_1^\top A_2=0$, (2.5), (2.21) and (4.14) that the assertion (4.13) holds. □

Theorem 4.4 Let $A_1^\top A_2=0$, $\beta\in(0,2)$, the sequence $\{w^k\}$ be generated by the L-GADMM (1.8) and the matrix Γ_1 be defined in (2.17). Then, we have

$$\|w^{k+1}-w^{k+2}\|^2_{\Gamma_1}\leq\|w^k-w^{k+1}\|^2_{\Gamma_1}. \tag{4.15}$$

Proof By setting $c=M_1(w^k-\tilde{w}^k)$ and $d=M_1(w^{k+1}-\tilde{w}^{k+1})$, we get

$$\|c\|^2_{\Gamma_1}-\|d\|^2_{\Gamma_1}=2c^\top\Gamma_1(c-d)-\|c-d\|^2_{\Gamma_1}.$$

That is,

$$\begin{aligned}&\|M_1(w^k-\tilde{w}^k)\|^2_{\Gamma_1}-\left\|M_1(w^{k+1}-\tilde{w}^{k+1})\right\|^2_{\Gamma_1}\\ &=2(w^k-\tilde{w}^k)M_1^\top\Gamma_1M_1\left[(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\right]\\ &\quad-\left\|M_1[(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})]\right\|^2_{\Gamma_1}.\end{aligned} \tag{4.16}$$

By (4.13) and (4.16), we have

$$\begin{aligned}&\|M_1(w^k-\tilde{w}^k)\|^2_{\Gamma_1}-\left\|M_1(w^{k+1}-\tilde{w}^{k+1})\right\|^2_{\Gamma_1}\\ &\geq\|(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\|^2_{(Q_1^\top+Q_1)}\\ &\quad-\left\|M_1[(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})]\right\|^2_{\Gamma_1}\\ &=\|(w^k-\tilde{w}^k)-(w^{k+1}-\tilde{w}^{k+1})\|^2_{(Q_1^\top+Q_1)-M_1^\top\Gamma_1M_1}.\end{aligned} \tag{4.17}$$

It follows from (2.21) that

$$
\begin{aligned}
&(Q_1^\top + Q_1) - M_1^\top \Gamma_1 M_1 \\
=&(Q_1^\top + Q_1) - M_1^\top Q_1 \\
=&\begin{pmatrix} 2G_1 & 0 & 0 & 0 \\ 0 & 2G_2 & 0 & 0 \\ 0 & 0 & 2(G_3 + \rho A_3^\top A_3) & -\beta A_3^\top \\ 0 & 0 & -\beta A_3 & \frac{2}{\rho} I_p \end{pmatrix} \\
&-\begin{pmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 \\ 0 & 0 & -\rho A_3 & \beta I_p \end{pmatrix}^\top \begin{pmatrix} G_1 & 0 & 0 & 0 \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 + \rho A_3^\top A_3 & (1-\beta) A_3^\top \\ 0 & 0 & -A_3 & \frac{1}{\rho} I_p \end{pmatrix} \\
=&\begin{pmatrix} G_1 & 0 & 0 & 0 \\ 0 & G_2 & 0 & 0 \\ 0 & 0 & G_3 & 0 \\ 0 & 0 & 0 & \frac{2-\beta}{\rho} I_p \end{pmatrix}.
\end{aligned}
$$

Using the positive definiteness of G_i (i=1, 2, 3), it is easy to see that $(Q_1^\top + Q_1) - M_1^\top \Gamma_1 M_1$ is positive definite when $\beta \in (0, 2)$. Then, we have

$$\|M_1(w^k - \tilde{w}^k)\|_{\Gamma_1}^2 - \left\| M_1(w^{k+1} - \tilde{w}^{k+1}) \right\|_{\Gamma_1}^2 \geq 0.$$

By $w^k - w^{k+1} = M_1(w^k - \tilde{w}^k)$, the assertion (4.15) follows immediately. □

In order to establish a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.8) in a nonergodic sense, we need the following two lemmas.

Lemma 4.4 Let $\{x_3^k\}$ be the sequence generated by the L-GADMM (1.8) with $\beta \in (0, 2)$. Then, we have

$$(x_3^k - x_3^{k+1})^\top A_3^\top (\lambda^k - \lambda^{k+1}) \geq \frac{1}{2}\|x_3^k - x_3^{k+1}\|_{G_3}^2 - \frac{1}{2}\|x_3^{k-1} - x_3^k\|_{G_3}^2. \tag{4.18}$$

Proof By the optimality condition of the x_3-subproblem in lemma 3.1 and (1.8), there exists $\zeta_1 \in \partial f_3(x_3^{k+1})$ such that

$$(x_3 - x_3^{k+1})^\top \left[\zeta_1 - A_3^\top \lambda^{k+1} + G_3(x_3^{k+1} - x_3^k)\right] \geq 0, \quad \forall\ x_3 \in \mathcal{X}_3, \tag{4.19}$$

where $\partial f_3(x_3)$ is a subdifferential of $f_3(x_3)$. Setting $x_3 = x_3^k$ in (4.19), then we have

$$(x_3^k - x_3^{k+1})^\top \left[\zeta_1 - A_3^\top \lambda^{k+1} + G_3(x_3^{k+1} - x_3^k)\right] \geq 0. \tag{4.20}$$

Moreover, setting $k := k+1$ in (4.19), there exists $\zeta_2 \in \partial f_3(x_3^k)$ such that

$$(x_3 - x_3^k)^\top \left[\zeta_2 - A_3^\top \lambda^k + G_3(x_3^k - x_3^{k-1})\right] \geq 0.$$

Similarly, by $x_3 = x_3^{k+1}$ in the above inequality, we have

$$(x_3^{k+1} - x_3^k)^\top \left[\zeta_2 - A_3^\top \lambda^k + G_3(x_3^k - x_3^{k-1})\right] \geq 0. \tag{4.21}$$

Adding (4.20) to (4.21) and using the monotonicity of the $\partial f_i(x_i)$, we get

$$\begin{aligned}(x_3^k - x_3^{k+1})^\top A_3^\top (\lambda^k - \lambda^{k+1}) &\geq (x_3^{k+1} - x_3^k)^\top G_3(x_3^{k+1} - x_3^k + x_3^{k-1} - x_3^k) \\ &= \|x_3^{k+1} - x_3^k\|_{G_3}^2 + (x_3^{k+1} - x_3^k)^\top G_3(x_3^{k-1} - x_3^k)\end{aligned}$$

It follows from the inequality

$$(x_3^{k+1} - x_3^k)^\top G_3(x_3^{k-1} - x_3^k) \geq -\frac{1}{2}\|x_3^k - x_3^{k+1}\|_{G_3}^2 - \frac{1}{2}\|x_3^{k-1} - x_3^k\|_{G_3}^2$$

that the assertion (4.18) is proved. □

Lemma 4.5 The sequence $\{w^k\}$ generated by the L-GADMM (1.8) with $\beta \in (0,2)$ and the associated $\{\tilde{w}^k\}$ be defined in (2.3), then there exists $0 < \varsigma_\beta \leq 1$ such that

$$\begin{aligned}&\|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2 \\ &\geq \varsigma_\beta \left(\|x_1^k - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 + \frac{\beta}{\rho}\|\lambda^k - \tilde{\lambda}^k\|^2\right).\end{aligned} \tag{4.22}$$

Proof. Using the definition of Q_1, M_1 and Γ_1, we have

$$\begin{aligned}&\|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2 \\ =&\|x_1^k - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 + \frac{2-\beta}{\rho}\|\lambda^k - \tilde{\lambda}^k\|^2 \\ \geq&\min\{\frac{2-\beta}{\beta}, 1\}(\|x_1^k - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 + \frac{\beta}{\rho}\|\lambda^k - \tilde{\lambda}^k\|^2).\end{aligned} \tag{4.23}$$

Let $\varsigma_\beta = \min\left\{\frac{2-\beta}{\beta}, 1\right\}$, then by (4.23), we know that the assertion (4.22) holds. □

In the next lemma, we refine the bound of $(w - \tilde{w}^k)^\top Q_1(w^k - \tilde{w}^k)$ in (4.8). The refined bound consists of the terms $\|w - w^{k+1}\|_{\Gamma_1}^2$ recursively, which is favorable for establishing a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.8) in a nonergodic sense. It is easy to see that the proof is similar to the Lemma 13 in [14]. We thus omit its proof.

Lemma 4.6 The sequence $\{w^k\}$ generated by the L-GADMM (1.8) with $\beta \in (0,2)$. Then, $\tilde{w}^k \in \Omega$ and

$$\begin{aligned}&f(x) - f(\tilde{x}^k) + (w - \tilde{w}^k)^\top F(w) \\ \geq&\frac{1}{2}(\|w - w^{k+1}\|_{\Gamma_1}^2 - \|w - w^k\|_{\Gamma_1}^2) + \frac{1}{2}\|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2, \quad \forall\, w \in \Omega.\end{aligned} \tag{4.24}$$

Finally, we establish a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.8) in a nonergodic sense.

Theorem 4.5 Let $A_1^\top A_2 = 0$, $\beta \in (0,2)$ and the sequence $\{w^k\}$ be generated by the iterative scheme L-GADMM (1.8) with $\beta \in (0,2)$. Then we have

$$\|w^t - w^{t+1}\|_{\Gamma_1}^2 \le \frac{1}{t}\left(\frac{1}{\varsigma_\beta}\|w^0 - w^*\|_{\Gamma_1}^2 + \|x_3^0 - x_3^1\|_{\Gamma_1}^2\right). \tag{4.25}$$

Proof Setting $w = w^*$ in (4.24), we can obtain

$$\begin{aligned} & f(x^*) - f(\tilde{x}^k) + (w^* - \tilde{w}^k)^\top F(w^*) \\ \ge & \frac{1}{2}(\|w^* - w^{k+1}\|_{\Gamma_1}^2 - \|w^* - w^k\|_{\Gamma_1}^2) + \frac{1}{2}\|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2. \end{aligned}$$

By (2.1) and the above inequality, we have

$$\|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2 \le \|w^* - w^k\|_{\Gamma_1}^2 - \|w^* - w^{k+1}\|_{\Gamma_1}^2.$$

Based on the analysis of the Theorem 4.2, by using the positive definition of G_i (i=1, 2, 3), it is easy to see that $(Q_1^\top + Q_1) - M_1^\top \Gamma_1 M_1$ is positive definite. Thus, we have

$$\sum_{k=0}^{\infty} \|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2 \le \|w^0 - w^*\|_{\Gamma_1}^2. \tag{4.26}$$

Furthermore, recall the definition of Γ_1 in (2.13), we have

$$\begin{aligned} \|w^k - w^{k+1}\|_{\Gamma_1}^2 =& \|x_1^k - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 \\ & + \frac{1}{\rho\beta}(\|\rho A_3(x_3^k - x_3^{k+1})\|^2 + \|\lambda^k - \lambda^{k+1}\|^2 + 2(1-\beta)\rho(x_3^k - x_3^{k+1})^\top A_3^\top(\lambda^k - \lambda^{k+1})) \\ =& \|x_1^k - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 \\ & + \frac{1}{\rho\beta}\|\rho A_3(x_3^k - x_3^{k+1}) + (\lambda^k - \lambda^{k+1})\|^2 - 2(x_3^k - x_3^{k+1})^\top A_3^\top(\lambda^k - \lambda^{k+1}). \end{aligned}$$

Moreover, by (2.4), we have

$$\begin{aligned} \|w^k - w^{k+1}\|_{\Gamma_1}^2 =& \|x_1^k - \tilde{x}_1^k\|_{G_1}^2 + \|x_2^k - \tilde{x}_2^k\|_{G_2}^2 + \|x_3^k - \tilde{x}_3^k\|_{G_3}^2 \\ & + \frac{\beta}{\rho}\|\lambda^k - \tilde{\lambda}^k\|^2 - 2(x_3^k - x_3^{k+1})^\top A_3^\top(\lambda^k - \lambda^{k+1}). \end{aligned} \tag{4.27}$$

It follows from (4.18), (4.22), (4.26) and (4.27) that

$$\begin{aligned} \sum_{k=1}^{t} \|w^k - w^{k+1}\|_{\Gamma_1}^2 \le & \frac{1}{\varsigma_\beta}\sum_{k=1}^{t}\|w^k - \tilde{w}^k\|_{Q_1^\top + Q_1 - M_1^\top \Gamma_1 M_1}^2 \\ & + \sum_{k=1}^{t}\left(\|x_3^{k-1} - x_3^k\|_{G_3}^2 - \|x_3^k - x_3^{k+1}\|_{G_3}^2\right) \\ \le & \frac{1}{\varsigma_\beta}\|w^0 - w^*\|_{\Gamma_1}^2 + \|x_3^0 - x_3^1\|_{G_3}^2. \end{aligned} \tag{4.28}$$

Thanks to the theorem 4.2, the sequence $\{\|w^k - w^{k+1}\|^2_{\Gamma_1}\}$ is non-increasing. Thus, we get

$$\begin{aligned} t\|w^t - w^{t+1}\|^2_{\Gamma_1} &\le \sum_{k=1}^{t} \|w^k - w^{k+1}\|^2_{\Gamma_1} \\ &\le \frac{1}{\varsigma_\beta}\|w^0 - w^*\|^2_{\Gamma_1} + \|x_3^0 - x_3^1\|^2_{\Gamma_1}. \end{aligned} \tag{4.29}$$

Using the above inequality, the assertion (4.25) is proved. □

The conclusion in Theorem 4.5 demonstrates a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.8) in a nonergodic sense. Similarly, we can easily obtain that a worst-case $\mathcal{O}(1/t)$ convergence rate for the L-GADMM (1.9) in a nonergodic sense holds when $A_1^\top A_3 = 0$, $\beta \in (0, 1]$.

5 Numerical experiments

In section 3, we have studied that if some conditions hold, then the L-GADMM scheme is convergent. In the following, We first give an example to show the divergence of the scheme (1.8) if the proposed conditions are lost. Then, we report some numerical result under the above some conditions.

All experiments are implemented in MATLAB R2010b on a hp-notebook with an Intel Core i5-3340M CPU at 2.70 GHz and 8 GB memory.

5.1 An example showing the divergence of the scheme (1.8)

In this subsection, we give an example to show the divergence of the scheme (1.8). Specifically, we consider the following linear homogeneous equation with three variables:

$$A_1x_1 + A_2x_2 + A_3x_3 = 0 \tag{5.1}$$

where $A_i \in \Re^3 (i = 1, 2, 3)$ are all column vectors and $x_i \in \Re (i = 1, 2, 3)$. If the matrix $[A_1, A_2, A_3]$ is assumed to be nonsingular, the unique solution of (5.1) will be thus $x_1 = x_2 = x_3 = 0$. Clearly, (5.1) is a special case of (1.5) where the objective fuction is null, b is the all-zero vector in $\Re^3$, and the corresponding optimal Lagrange multipliers are all 0.

Now, we show the iterative scheme when we solve the (5.1) by L-GADMM (1.8) with $\rho = 1$,

$b = 0$, which can be represented as the following matrix recursion:

$$\begin{cases} -A_1^\top \lambda^k + A_1^\top (A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^k) + G_1(x_1^{k+1} - x_1^k) = 0, \\ -A_2^\top \lambda^k + A_2^\top (A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^k) + G_2(x_2^{k+1} - x_2^k) = 0, \\ -A_3^\top \lambda^k + A_3^\top (\beta A_1 x_1^{k+1} + (1-\beta)(-A_3 x_3^k) + \beta A_2 x_2^{k+1} + A_3 x_3^{k+1}) \\ \quad + G_3(x_3^{k+1} - x_3^k) = 0, \\ (\beta A_1 x_1^{k+1} + (1-\beta)(-A_3 x_3^k) + \beta A_2 x_2^{k+1} + A_3 x_3^{k+1}) + \lambda^{k+1} - \lambda^k = 0. \end{cases} \tag{5.2}$$

Furthermore we can recast the scheme (5.2) as

$$\begin{cases} (G_1 + A_1^\top A_1) x_1^{k+1} = G_1 x_1^k + A_1^\top A_2 x_2^k + A_1^\top A_3 x_3^k + A_1^\top \lambda^k, \\ \alpha A_2^\top A_1 x_1^{k+1} + (G_2 + A_2^\top A_2) x_2^{k+1} = G_2 x_2^k - A_2^\top A_3 x_3^k + A_2^\top \lambda^k, \\ \beta A_3^\top A_1 x_1^{k+1} + \beta A_3^\top A_2 x_2^{k+1} + (G_3 + A_3^\top A_3) x_3^{k+1} = (G_3 + (1-\beta) A_3^\top A_3) x_3^k + A_3^\top \lambda^k, \\ \beta A_1 x_1^{k+1} + \beta A_2 x_2^{k+1} + A_3 x_3^{k+1} + \lambda^{k+1} = (1-\beta) A_3 x_3^k + \lambda^k. \end{cases} \tag{5.3}$$

And then, we obtain a reformulation of (5.3)

$$\begin{pmatrix} G_1 + A_1^\top A_1 & 0 & 0 & 0_{1\times 3} \\ A_2^\top A_1 & G_2 + A_2^\top A_2 & 0 & 0_{1\times 3} \\ \beta A_3^\top A_1 & \beta A_3^\top A_2 & G_3 + A_3^\top A_3 & 0_{1\times 3} \\ \beta A_1 & \beta A_2 & A_3 & I_{3\times 3} \end{pmatrix} \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^{k+1} \end{pmatrix}$$

$$= \begin{pmatrix} G_1 & A_1^\top A_2 & A_1^\top A_3 & A_1^\top \\ 0 & G_2 & -A_2^\top A_3 & A_2^\top \\ 0 & 0 & G_3 + (1-\beta) A_3^\top A_3 & A_3^\top \\ 0 & 0 & (1-\beta) A_3 & I_{3\times 3} \end{pmatrix} \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \lambda^k \end{pmatrix}. \tag{5.4}$$

Let

$$P_1 = \begin{pmatrix} G_1 + A_1^\top A_1 & 0 & 0 & 0_{1\times 3} \\ A_2^\top A_1 & G_2 + A_2^\top A_2 & 0 & 0_{1\times 3} \\ \beta A_3^\top A_1 & \beta A_3^\top A_2 & G_3 + A_3^\top A_3 & 0_{1\times 3} \\ \beta A_1 & \beta A_2 & A_3 & I_{3\times 3} \end{pmatrix}, \tag{5.5}$$

and

$$P_2 = \begin{pmatrix} G_1 & A_1^\top A_2 & A_1^\top A_3 & A_1^\top \\ 0 & G_2 & -A_2^\top A_3 & A_2^\top \\ 0 & 0 & G_3 + (1-\beta) A_3^\top A_3 & A_3^\top \\ 0 & 0 & (1-\beta) A_3 & I_{3\times 3} \end{pmatrix}. \tag{5.6}$$

Then the (5.4) can be rewritten as the fixed matrix mappings:

$$\begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \Upsilon \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \lambda^k \end{pmatrix} = \cdots = \Upsilon^{k+1} \begin{pmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ \lambda^0 \end{pmatrix}, \tag{5.7}$$

where

$$\Upsilon = P_1^{-1} P_2. \tag{5.8}$$

Therefore, the L-GADMM (1.8) is convergent for any starting point when the matrix mapping is a contraction or $\rho(\Upsilon) < 1$ (the spectral radius of Υ is denoted by $\rho(\Upsilon)$). Thus, we could seek for a A such that $\rho(\Upsilon) > 1$ to construct a divergent example.

Specifically, we assume that the matrices A_1, A_2, A_3 in (5.1) are non-orthogonal and let $G_i = 2A_i^\top A_i$ (i=1,2,3). We consider the matrix A which is proposed by Chen et al. in [18]:

$$A = (A_1, A_2, A_3) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}. \tag{5.9}$$

Let $\beta = 4.5$ and by using this matrix A, (5.4) can be specified as

$$\begin{pmatrix} 9.0000 & 0 & 0 & 0 & 0 & 0 \\ 4.0000 & 18.0000 & 0 & 0 & 0 & 0 \\ 22.5000 & 31.5000 & 27.0000 & 0 & 0 & 0 \\ 4.5000 & 4.5000 & 1.0000 & 1.0000 & 0 & 0 \\ 4.5000 & 4.5000 & 2.0000 & 0 & 1.0000 & 0 \\ 4.5000 & 9.0000 & 2.0000 & 0 & 0 & 1.0000 \end{pmatrix} \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda_1^{k+1} \\ \lambda_2^{k+1} \\ \lambda_3^{k+1} \end{pmatrix}$$

$$= \begin{pmatrix} 6.0000 & 4.0000 & 5.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0 & 12.0000 & -7.0000 & 1.0000 & 1.0000 & 2.0000 \\ 0 & 0 & -13.5000 & 1.0000 & 2.0000 & 2.0000 \\ 0 & 0 & -3.5000 & 1.0000 & 0 & 0 \\ 0 & 0 & -7.0000 & 0 & 1.0000 & 0 \\ 0 & 0 & -7.0000 & 0 & 0 & 1.0000 \end{pmatrix} \begin{pmatrix} x_1^{k} \\ x_2^{k} \\ x_3^{k} \\ \lambda_1^{k} \\ \lambda_2^{k} \\ \lambda_3^{k} \end{pmatrix}.$$

Note with the specification in (5.9), the matrices P_1 in (5.5) and P_2 in (5.6) reduce to

$$P_1 = \begin{pmatrix} 9.0000 & 0 & 0 & 0 & 0 & 0 \\ 4.0000 & 18.0000 & 0 & 0 & 0 & 0 \\ 22.5000 & 31.5000 & 27.0000 & 0 & 0 & 0 \\ 4.5000 & 4.5000 & 1.0000 & 1.0000 & 0 & 0 \\ 4.5000 & 4.5000 & 2.0000 & 0 & 1.0000 & 0 \\ 4.5000 & 9.0000 & 2.0000 & 0 & 0 & 1.0000 \end{pmatrix},$$

$$
\text{and} \quad P_2 = \begin{pmatrix} 6.0000 & 4.0000 & 5.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0 & 12.0000 & -7.0000 & 1.0000 & 1.0000 & 2.0000 \\ 0 & 0 & -13.5000 & 1.0000 & 2.0000 & 2.0000 \\ 0 & 0 & -3.5000 & 1.0000 & 0 & 0 \\ 0 & 0 & -7.0000 & 0 & 1.0000 & 0 \\ 0 & 0 & -7.0000 & 0 & 0 & 1.0000 \end{pmatrix}.
$$

Thus we have

$$
\Upsilon = P_1^{-1} P_2 = \begin{pmatrix} 0.6667 & 0.4444 & 0.5556 & 0.1111 & 0.1111 & 0.1111 \\ -0.1481 & 0.5679 & -0.5123 & 0.0309 & 0.0309 & 0.0864 \\ -0.3827 & -1.0329 & -0.3652 & -0.0916 & -0.0545 & -0.1193 \\ -1.9506 & -3.5226 & -3.3292 & 0.4527 & -0.5844 & -0.7695 \\ -1.5679 & -2.4897 & -6.4640 & -0.4558 & 0.4702 & -0.6502 \\ -0.9012 & -5.0453 & -4.1584 & -0.5947 & -0.6687 & -0.0391 \end{pmatrix}. \tag{5.10}
$$

It follows from the matrix decomposition that

$$
\rho(\Upsilon) = 1.1582 > 1,
$$

from which we can construct a divergent sequence $\{(x_1^k, x_2^k, x_3^k, \lambda_1^k, \lambda_2^k, \lambda_3^k)\}$ starting from certain initial points.

5.2 Some numerical results

In this subsection, we report some numerical result for the proposed method by calibrating the correlation matrices.

We consider to solve the following matrix optimization problem

$$
\begin{array}{ll} \min & \frac{1}{2}\|X_1 - C\|^2 + \frac{1}{2}\|X_2 - C\|^2 + \frac{1}{2}\|X_3 - C\|^2 \\ \text{s.t.} & X_1 - X_2 = \mathbf{0}. \\ & X_2 - X_3 = \mathbf{0}. \\ & X_1 \in S_+^n, X_2 \in S_+^n, X_3 \in S_B. \end{array}
$$

where

$$
S_+^n = \{H \in R^{n\times n} | H^\top = H, H \succeq \mathbf{0}\}, \quad S_B = \{H \in R^{n\times n} | H^\top = H, H_L \preceq H \preceq H_U\}.
$$

The above problem could be converted to the following equivalent form:

$$
\begin{array}{ll} \min & \frac{1}{2}\|X_1 - C\|^2 + \frac{1}{2}\|X_2 - C\|^2 + \frac{1}{2}\|X_3 - C\|^2 \\ \text{s.t.} & A_1 X_1 + A_2 X_2 + A_3 X_3 = \mathbf{0}, \\ & X_1 \in S_+^n, X_2 \in S_+^n, X_3 \in S_B, \end{array} \tag{5.11}
$$

where $A_1 = \begin{pmatrix} I \\ \mathbf{0} \end{pmatrix}$, $A_2 = \begin{pmatrix} -I \\ I \end{pmatrix}$, $A_3 = \begin{pmatrix} \mathbf{0} \\ -I \end{pmatrix}$. It is easy to see that the L-GADMM could be applicable and we obtain the following subproblem for model (5.14):

$$\begin{cases} X_2^{k+1} = \text{argmin}\{\frac{1}{2}\|X_2 - C\|^2 - X_2^\top A_2^\top \lambda^k + \frac{\rho}{2}\|A_1 X_1^k + A_2 X_2 + A_3 X_3^k - b\|^2 \\ \qquad + \frac{1}{2}\|X_2 - X_2^k\|_{G_2}^2 | X_2 \in \mathcal{X}_2\}, \\ X_3^{k+1} = \text{argmin}\{\frac{1}{2}\|X_3 - C\|^2 - X_3^\top A_3^\top \lambda^k + \frac{\rho}{2}\|\beta A_1 X_1^k + (1-\beta)(b - A_3 X_3^k) + \beta A_2 X_2^{k+1} + A_3 X_3 - b\|^2 \\ \qquad + \frac{1}{2}\|X_3 - X_3^k\|_{G_3}^2 | X_3 \in \mathcal{X}_3\}, \\ \lambda^{k+1} = \lambda^k - \rho(\beta A_1 X_1^k + (1-\beta)(b - A_3 X_3^k) + \beta A_2 X_2^{k+1} + A_3 X_3^{k+1} - b), \\ X_1^{k+1} = \text{argmin}\{\frac{1}{2}\|X_1 - C\|^2 - X_1^\top A_1^\top \lambda^{k+1} + \frac{\rho}{2}\|A_1 X_1 + A_2 X_2^{k+1} + A_3 X_3^{k+1} - b\|^2 \\ \qquad + \frac{1}{2}\|X_1 - X_1^k\|_{G_1}^2 | X_1 \in \mathcal{X}_1\}. \end{cases} \tag{5.12}$$

Here, we let $C = \text{rand}(n, n)$, $C = (C + C) - \text{ones}(n, n) + eye(n)$, $HU = \text{ones}(n, n) * 0.1$ and $HL = -HU$. Furthermore, we set $n = 100$, $\beta = 0.9$ and $G_1 = G_2 = G_3 = 5I$, then we apply the L-GADMM (1.9) to solve the problem (5.14) and show a numeric comparison by our algorithm and ADMM with Gaussian back substitution (ADMM-G) [24]. And the stopping criterion is written as

$$\max\left\{ \frac{\|X_1^{k+1} - X_1^k\|}{\|X_1^1 - X_1^0\|}, \frac{\|X_2^{k+1} - X_2^k\|}{\|X_2^1 - X_2^0\|}, \frac{\|X_3^{k+1} - X_3^k\|}{\|X_3^1 - X_3^0\|}, \frac{\|\lambda^{k+1} - \lambda^k\|}{\|\lambda^1 - \lambda^0\|} \right\} < \text{tol},$$

where $\text{tol} = 10^{-6}$. In Fig.5.1, we plot the evolutions of the objective function value and number

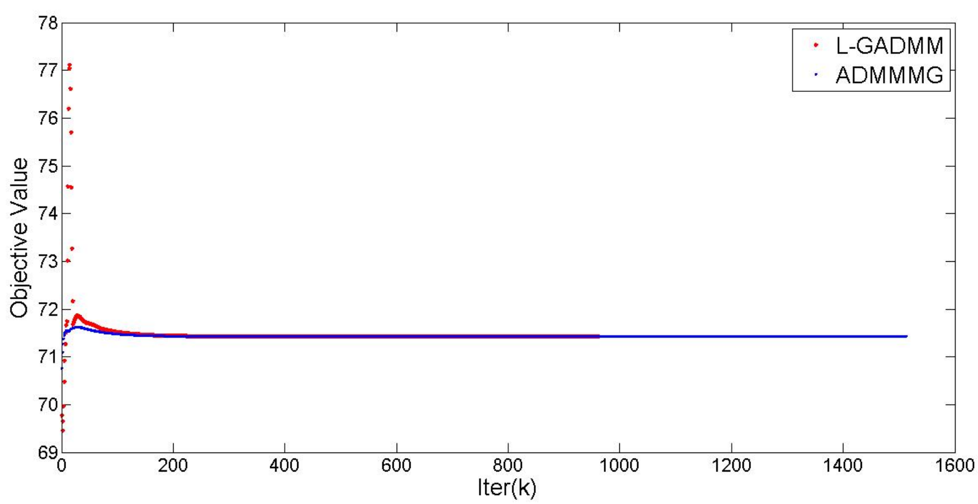

Figure 5.1: evolutions of the objective function value and number of iterations

of iterations for the above two algorithm. In fact, L-GADMM is not so different from ADMM-G in the averaged numerical performance. But the figure shows that L-GADMM requires a much less iteration than the ADMM-G when the stopping criterion is achieved. On the other hand, we show some analysis of the Error of the above two iterative schemes in Figure 5.2. From Fig.5.2, the accuracy of L-GADMM is higher than ADMM-G when the stopping criterion is achieved and the algorithm (1.4) has better convergence speed than ADMM-G.

It follows from the above numerical results that our proposed algorithm is efficient.

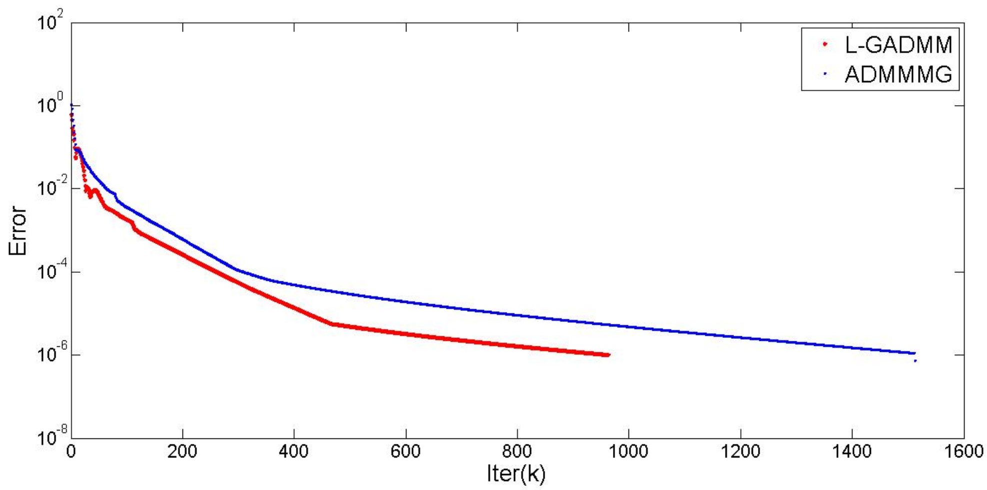

Figure 5.2: evolutions of the error and the number of iterations

Rmark 5.1 In our experiment, the numerical performance of our proposed method is well if $A_1^\top A_3 = 0$ and $\beta \in (0, 1]$.

6 Conclusions

We have discuss by an example that the linearized version of GADMM is not necessarily convergent for solving a convex minimization model with linear constraints and a separable objective with three function blocks. In this paper, we first investigate the convergence of the linearized version of GADMM when the orthogonality of any two coefficient matrices in the model is satisfied and the relaxation factor $\beta \in (0, 1]$ or $\beta \in (0, 2)$. Then, we also establish the worst-case $\mathcal{O}(1/t)$ convergence rate when $A_2^\top A_3 = 0$, $\beta \in (0, 1]$ or $A_1^\top A_3 = 0$, $\beta \in (0, 1]$ or $A_1^\top A_2 = 0$, $\beta \in (0, 2)$. Finally, we construct a counter example to show the divergence of the proposed L-GADMM when the above sufficient condition is lost; moreover, we also report some numerical results when $A_1^\top A_3 = 0$ and $\beta \in (0, 1]$.

References

1 D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximations, Comput. Math. Appl. 2, 17-40 (1976).

2 Hestenes, M.R.: Multiplier and gradient methods. J. Optim. Theory Appl. 4, 302-320 (1969)

3 Powell, M.J.D.: A method for nonlinear constraints in minimization problems. In: Fletcher, R. (ed.) Optimization. Academic Press (1969)

4 Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3, 1-122 (2011)

5 Eckstein, J., Yao, W.: Augmented Lagrangian and alternating direction methods for convex optimization: a tutorial and some illustrative computational results. RUTCOR Research Report RRR 32-2012 (2012)

6 Glowinski, R.: On alternating directon methods of multipliers: a historical perspective. Modeling, Simulation and Optimization for Science and Technology, 34, 59-82 (2014)

7 Gabay, D.: Applications of the method of multipliers to variational inequalities, Augmented Lagrange Methods: applications to the solution of boundary-valued problems. Fortin, M. Glowinski, R. eds. North Holland pp. 299-331 (1983)

8 Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operator. SIAM J. Numer. Anal. 16, 964-979 (1979)

9 Martinet, B.: Regularisation, dinquations variationelles par approximations succesives. Rev. Francaise dInform. Recherche Oper. 4, 154-159 (1970)

10 Eckstein, J., Bertsekas, D.: On theDouglas-Rachford splittingmethod and the proximal point algorithm for maximal monotone operators. Math. Program. 55, 293-318 (1992)

11 Bertsekas, D.P.: Constrained optimization and Lagrange multiplier methods. Academic Press, New York (1982)

12 Cai, X., Gu, G., He, B., Yuan, X.: A proximal point algorithm revisit on alternating direction method of multipliers. Sci. China Math. 56(10), 2179-2186 (2013)

13 Eckstein, J.: Parallel alternating direction multiplier decomposition of convex programs. J. Optim. Theory Appli. 80(1), 39-62 (1994)

14 Ethan X,Fang, Bingsheng He, Han Liu, Xiaoming Yuan, Generalized alternating direction method of multipliers: new theoretical insights and applications. Math. Prog. Comp. 7, 149-187 (2015)

15 Wang, X., Yuan, X.: The linearized alternating direction method of multipliers for Dantzig Selector. SIAM J. Sci. Comput. 34, 2782-2811 (2012)

16 Yang, J., Yuan, X.: Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization. Math. Comput. 82, 301-329 (2013)

17 Zhang, X.Q., Burger, M., Osher, S.: A unified primal-dual algorithm framework based on Bregman iteration. J. Sci. Comput. 6, 20-46 (2010)

18 Chen, C.H., He, B.S., Ye, Y.Y., Yuan, X.M.: The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. Math. Program. 155, 57-79 (2016)

19 Jun-Kai Feng, Hai-Bin Zhang, Cao-Zong Cheng, Hui-Min Pei, Convergence Analysis of L-ADMM for Multi-block Linear-Constrained Separable Convex Minimization Problem. J. Oper. Res. Soc. China 3, 563-579 (2015)

20 Xingju Cai, Deren Han, Xiaoming 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. 66, 39-73 (2017)

21 He, B., Yang, H.: Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities. Oper. Res. Let. 23, 151-161 (1998)

22 He, B., Yuan, X.: On the O(1/n) convergence rate of Douglas-Rachford alternating direction method. SIAM J. Numer. Anal. 50, 700-709 (2012)

23 R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

24 B. S. He, M. Tao, X. M. Yuan, Alternating direction method with Gaussian back substitution for separable convex programming, SIAM J. Optim. Vol. 22, No. 2, pp. 313-340 (2012)