An inexact ADMM with linear convergence and its application to inverse coefficient problems*

Tiancan Feng† Wenyi Tian‡

Abstract

The alternating direction method of multipliers (ADMM) has been widely used for various separable convex optimization problems in different applications. When considering some complicated optimization problems arising in scientific computing areas, it is necessary to inexactly solve the subproblem in the ADMM to reduce vast computations. In this work, we propose and analyze a new inexact ADMM algorithm nested with some easily applicable inexactness criterion. Under the same conditions ensuring the linear convergence of the exactly-solved ADMM, we prove that the proposed inexactness criterion can guarantee the linear convergence of our inexact ADMM. Then we apply the algorithm to two nonlinear inverse problems in elliptic equations and present the specific implementation details. Numerical results for identifying discontinuous coefficients in elliptic equations are reported to demonstrate the feasibility and efficiency of the proposed inexact ADMM algorithm.

Keywords: convex optimization problems, inexact alternating direction method of multipliers, linear convergence rate, nonlinear inverse problems for PDEs, coefficient identification, total variation

AMS subject classifications: 65K10, 65N21, 90C25, 35R30

1 Introduction

Many practical problems can be attributed to solving the following regularized optimization problem

$$\begin{cases} \min\limits_{u,z} \ J(u) + R(z) \\ \text{s.t.} \ \ Au + Bz = b, \ u \in \mathcal{U}, \ z \in \mathcal{Z}, \end{cases} \tag{1.1}$$

where $\mathcal{U}$ and $\mathcal{Z}$ are two Hilbert spaces. The functional $J : \mathcal{U} \to \mathbb{R}$ is smooth and strongly convex, $R : \mathcal{Z} \to \mathbb{R}$ is a convex functional (not necessarily smooth), A, B are two non-zero bounded linear operators with $A : \mathcal{U} \to \mathcal{H}$, $B : \mathcal{Z} \to \mathcal{H}$, and $\mathcal{H}$ is also a Hilbert space. There are some special cases of the optimization problem (1.1). For instance, if $B = -I$ and $b = 0$, then (1.1) reduces to the problem $\min_u \ J(u) + R(Au)$; furthermore, if $A = I$, $B = -I$, $b = 0$ and $R(\cdot) = I_{\mathcal{C}}(\cdot)$ being an indicator function on $\mathcal{C} \subset \mathcal{U}$, then (1.1) becomes $\min_{u \in \mathcal{C}} \ J(u)$.

The alternating direction method of multipliers (ADMM), originally proposed by Glowinski and Marrocco [22], has gained prominence for its efficacy in solving complicated optimization problem (1.1), which arise in various areas, including inverse problems, signal processing, image reconstruction, control systems, statistics and so on. We refer to [5, 9, 13, 47, 51] for some applications, and [5, 16, 21] for comprehensive overviews of recent works. In order to apply the ADMM to the problem (1.1), it first introduces the augmented Lagrangian functional L_β of (1.1) as follows

$$L_\beta(u, z, \lambda) = J(u) + R(z) - (\lambda, Au + Bz - b) + \frac{\beta}{2}\|Au + Bz - b\|^2,$$

*This work was supported by the National Natural Science Foundation of China under grant 12071343 and basic research fund of Tianjin University under grant 2025XJ21-0010.

†Center for Applied Mathematics, KL-AAGDM, Tianjin University, Tianjin 300072, China (tiancan.feng@gmail.com).

‡Corresponding author. Center for Applied Mathematics and KL-AAGDM, Tianjin University, Tianjin 300072, China (twymath@gmail.com).

where λ is the Lagrange multiplier associated with the constraint $Au + Bz = b$, and $\beta > 0$ a penalty parameter. $(\cdot, \cdot)$ represents the inner product in the corresponding Hilbert spaces, and $\|\cdot\|$ is the norm induced by the inner product. Then we can solve problem (1.1) by the ADMM algorithm:

$$\begin{cases} u^{k+1} = \arg\min_{u \in \mathcal{U}} L_\beta(u, z^k, \lambda^k), \\ z^{k+1} = \arg\min_{z \in \mathcal{Z}} L_\beta(u^{k+1}, z, \lambda^k), \\ \lambda^{k+1} = \lambda^k - \beta(Au^{k+1} + Bz^{k+1} - b). \end{cases} \tag{1.2}$$

The ADMM (1.2) was regarded in [19] as an application of the Douglas-Rachford splitting method (DRSM) proposed in [12, 36] to the dual problem of (1.1), and was further explained in [15] as an application of the proximal point algorithm (PPA) [38, 40] from the maximal monotone operator perspective. Due to its extensive applications and subsequently a large amount of researches on the algorithm and its variants, there has been an abundance of works on the theoretical analysis on the global convergence and rate of convergence. The convergence of ADMM was well studied in earlier literature [19, 20] and the worst-case $O(1/K)$ convergence rate has been established in [32, 33, 39] as well, where K denotes the iteration counter. Under the assumption that both objective functions are strongly convex (one of them being quadratic), an accelerated ADMM (with a predictor-corrector-type acceleration step) was proposed in [25] and shown having an $O(1/K^2)$ convergence rate in the sense of primal and dual residuals. Whenever objective functions without strong convexity, [45] proposed a faster ADMM with a worst-case $O(1/K^2)$ convergence rate in ergodic sense for solving the two-block optimization problems with the constraint $Au = z$, where the Lagrangian penalty parameter was iteratively updated. More works on this topic can be referred to the review works [5, 21, 16] and the references therein.

Compared with the plentiful works devoted to the applications of the ADMM, the research on its linear convergence rate is relatively fewer. It is worth noting that the ADMM converges linearly at global sense [14] for linear programming, and [4] established a local linear convergence rate for convex quadratic programming under the condition that both the primal and dual optimization problems have unique solutions. While for more general case, the work [35] proved the linear convergence rate of the ADMM (more than two blocks) under a certain form that ensures each subproblem is strongly convex and the dual step size is sufficiently small. Motivated by the fact that a large amount of applications trace back to model (1.1) with at least one strongly convex function, [11] established the global and linear convergence of the generalized ADMM under the assumptions of strongly convex and Lipschitz continuous gradient on one of the two objective functions, and it summarized some linear convergence results for different scenarios along with certain conditions on A and B. An inexact ADMM was proposed in [27] for a separable convex optimization problem with multi-blocks in finite dimensional Euclidean spaces, where each subproblem was solved approximately via an accelerated gradient method. The algorithm has the ergodic convergence rate of $O(1/K)$ for convex problems and $O(1/K^2)$ under additionally strong convexity, and it is 2-step linearly convergent when an error-bound condition holds. Taking advantage of a specific feature of the ADMM with part of the perturbation being automatically zero, [37] proposed a partial error bound condition, which is weaker than those in the literatures (see e.g., [35, 49]), and derived the linear convergence rate of the ADMM. More conclusions about linear convergence of the ADMM can refer to [51].

1.1 Motivations and goals

The ADMM can be viewed as a splitting version of the classical augmented Lagrangian method (ALM) proposed in [34, 44], the latter minimizes $L_\beta(u, z, \lambda)$ with respect to u, z and then updates λ. Compared with the ALM for (1.1), the ADMM can decouple the constraint of the two variables u and z, then solve two unconstrained subproblems respectively, and it usually runs faster due to the relatively simple subproblems. However, for the optimization problem (1.1) in large scale dimensions or with complicated objective functions, at least one of the subproblems in the ADMM (1.2) is expensive and difficult to solve, such as the Lasso model [5, 52], the optimal control and inverse problems of PDEs [23, 24, 45]. Then it is necessary to consider an inexact version of the ADMM (1.2) by designing certain inexactness criterion which is applicable and easy to implement.

There are some earlier works on inexact ADMM algorithms which require a summable sequence of accuracy constants (see, e.g. [15, 16, 17, 31, 42, 50]). The summable requirement on the error sequences in terms of either absolute or relative errors often makes these inexact methods difficult to implement in practical applications, because manually choosing the parameters in the sequences is quite challenging, and inappropriate choices may ruin the numerical performance of the whole method. Recently, an elegant and easily implementable inexactness criterion $\|e_k(u^{k+1})\| \leq \sigma \|e_k(u^k)\|$ with $0 < \sigma < 1$ was proposed in [23, 52] as the stopping condition of some iterative algorithm for solving the first subproblem of (1.2), in which $e_k(u) := D_u L_\beta(u, z^k, \lambda^k)$. This inexact approach benefits from that the parameter σ measuring the relative error of $\|e_k(u^{k+1})\|$ is fixed by a specific formula as a constant through all the iterations, and makes the inexact ADMM fully automatic and practical for numerical implementations, which is significantly different from those in earlier works. It also has the characteristics of less computation and robustness to different settings, which are demonstrated in numerical experiments.

It should be mentioned that both the works [23, 52] were discussing the inexact ADMM for specific problems, i.e, a parabolic optimal control problem in [23] and the Lasso problem in [52]. Both the problems share the common feature that the objective function $J(\cdot)$ in (1.1) is in quadratic form, and the corresponding method was only theoretically proved to be of first-order convergence rate, which does not account for faster linear convergence rate observed in numerical results. These issues will be addressed in this work.

The first aim of this work is to analyze the convergence and convergence rate of an inexact version of the ADMM (1.2) with the criterion (2.3) for the generic problem (1.1). The second one is to design a new inexactness criterion for the ADMM (1.2) achieving linear convergence rate, and then to analyze it rigorously under the same assumptions on (1.1) that ensure the linear convergence of the exactly-solved ADMM (1.2). Moreover, we apply our proposed inexact ADMM to two nonlinear inverse coefficient problems of elliptic equations formulated in (4.1) and (4.2), to reduce the computational cost on the premise of ensuring linear convergence.

1.2 Organization

The rest of this paper is organized as follows. In section 2, we extend the inexact version of the ADMM (1.2) with the inexactness criterion (2.3) in [23, 52] to the generic problem (1.1), and obtain the global convergence and the worst-case $O(1/K)$ convergence rate measured by iteration complexity in both ergodic and non-ergodic senses. In section 3, a new inexactness criterion is proposed for the ADMM (1.2) guaranteeing the linear convergence rate under the assumption that $J(\cdot)$ is strongly convex and has a Lipschitz continuous gradient, then the global linear convergence rate of the new method is rigorously established. In section 4, we illustrate the feasibility of applying our method to the regularized nonlinear inverse coefficient problems (4.1) and (4.2) for elliptic equations, and specify some implementation details. Numerical results are reported in section 5 to validate the efficiency of the proposed method. Finally, some conclusions are drawn in section 6.

2 Inexact ADMM for the problem (1.1)

In this section, we consider the extension of the inexact ADMM algorithm, that is (1.2) with the inexactness criterion (2.3) proposed in [23, 52] for quadratic problems, to the general convex optimization problem (1.1) with the conditions in Assumption 1, and present some results on its global convergence and convergence rate.

Assumption 1. *The functional $J : \mathcal{U} \to \mathbb{R}$ is α-strongly convex with continuous first-order derivative $DJ(\cdot)$, i.e, there exists a constant $\alpha > 0$ such that*

$$\big(DJ(u_1) - DJ(u_2), u_1 - u_2\big) \geq \alpha \|u_1 - u_2\|^2, \ \ \forall\, u_1, u_2 \in \mathcal{U}. \tag{2.1}$$

The considered inexact ADMM reads: obtain the iterate $(u^{k+1}, z^{k+1}, \lambda^{k+1})$ by

$$\begin{cases} u^{k+1} \approx \arg\min_{u\in\mathcal{U}} L_\beta(u, z^k, \lambda^k), \\ z^{k+1} = \arg\min_{z\in\mathcal{Z}} L_\beta(u^{k+1}, z, \lambda^k), \\ \lambda^{k+1} = \lambda^k - \beta(Au^{k+1} + Bz^{k+1} - b), \end{cases} \tag{2.2}$$

such that u^{k+1} solved by some iterative algorithm (e.g., the conjugate gradient (CG) method) admits the following inexactness criterion

$$\|e_k(u^{k+1})\| \le \sigma \|e_k(u^k)\|, \tag{2.3}$$

where the constant σ satisfies

$$0 < \sigma < \frac{\sqrt{2\alpha}}{\sqrt{2\alpha} + \sqrt{\beta}\|A^*\|} \in (0, 1), \tag{2.4}$$

and the notation $e_k(u)$ denotes the derivative of $L_\beta(u, z^k, \lambda^k)$ with respective to u, that is

$$e_k(u) = D_u L_\beta(u, z^k, \lambda^k) = DJ(u) - A^*\lambda^k + \beta A^*(Au + Bz^k - b). \tag{2.5}$$

2.1 Convergence

To present the convergence analysis for the inexact ADMM algorithm, some notations are introduced as follows

$$w = \begin{pmatrix} u \\ z \\ \lambda \end{pmatrix}, \quad v = \begin{pmatrix} z \\ \lambda \end{pmatrix}, \quad F(w) = \begin{pmatrix} DJ(u) - A^*\lambda \\ -B^*\lambda \\ Au + Bz - b \end{pmatrix}, \quad H = \begin{pmatrix} \beta B^*B & 0 \\ 0 & \frac{1}{\beta} I \end{pmatrix},$$

and

$$W = \mathcal{U} \times \mathcal{Z} \times \mathcal{H}, \ \ V = \mathcal{Z} \times \mathcal{H},$$

where $DJ(u)$ stands for the first-order derivative of $J(\cdot)$ at $u \in \mathcal{U}$. We also define the norm

$$\|v\|_H := \sqrt{(v, Hv)} = \sqrt{\beta\|Bz\|^2 + \frac{1}{\beta}\|\lambda\|^2}, \ \ \forall\, v \in V.$$

It is easy to see that the problem (1.1) can be characterized by a variational inequality problem: find $w^* = (u^*, z^*, \lambda^*)^\top \in W$ such that

$$R(z) - R(z^*) + (w - w^*, F(w^*)) \ge 0, \ \ \forall\, w \in W. \tag{2.6}$$

We denote by W^* the solution set of (2.6), as analyzed in [32], W^* has the following reprentation.

Theorem 2.1. *The solution set W^* is convex and it can be characterized as*

$$W^* = \bigcap_{w\in W} \{\hat{w} \in W : R(z) - R(\hat{z}) + (w - \hat{w}, F(w)) \ge 0\}.$$

It is easy to show that $(u^*, Bz^*, A^*\lambda^*)$ is unique by using the strong convexity of $J(\cdot)$ in (2.1) and the KKT conditions of (1.1). If we further assume that operator B is injective, then we can conclude the uniqueness of z^*. The sequence $\{w^{k+1}\} = \{(u^{k+1}, z^{k+1}, \lambda^{k+1})^\top\}$ generated by the inexact ADMM (2.2) satisfies the first-order optimality conditions as follows

$$\begin{cases} D_u L_\beta(u^{k+1}, z^k, \lambda^k) = e_k(u^{k+1}), \\ R(z) - R(z^{k+1}) - \big(z - z^{k+1}, B^*(\lambda^k - \beta(Au^{k+1} + Bz^{k+1} - b))\big) \ge 0, \ \ \forall\, z \in V, \\ \lambda^{k+1} = \lambda^k - \beta(Au^{k+1} + Bz^{k+1} - b). \end{cases} \tag{2.7}$$

To simplify the notation, we introduce an auxiliary variable $\bar{w}^k$ as

$$\bar{w}^k = \begin{pmatrix} \bar{u}^k \\ \bar{z}^k \\ \bar{\lambda}^k \end{pmatrix} = \begin{pmatrix} u^{k+1} \\ z^{k+1} \\ \lambda^k - \beta(Au^{k+1} + Bz^k - b) \end{pmatrix}.$$

A useful result in Lemma 2.2 can be easily derived from (2.7) and the notation $\bar{w}^k$ to characterize the difference between the points $\bar{w}^k$ and $w \in W$, which is analogous to Lemma 3.1 in [23].

Lemma 2.2. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ satisfy (2.7). Then for any $w \in W$, it holds that*

$$\begin{aligned} &R(z) - R(\bar{z}^k) + \big(w - \bar{w}^k, F(\bar{w}^k)\big) \\ &\geq \big(u - u^{k+1}, e_k(u^{k+1})\big) + \frac{1}{2}\big(\|v - v^{k+1}\|_H^2 - \|v - v^k\|_H^2 + \|v^k - v^{k+1}\|_H^2\big). \end{aligned}$$

We can derive an estimate of $e_k(u^{k+1})$ from (2.3) and (2.5) that

$$\begin{aligned} \|e_k(u^{k+1})\| \leq \sigma\|e_k(u^k)\| &= \sigma\|e_{k-1}(u^k) + A^*(\lambda^{k-1} - \lambda^k) - \beta A^* B(z^{k-1} - z^k)\| \\ &\leq \sigma\|e_{k-1}(u^k)\| + \sigma\|A^*\|\sqrt{\beta}\cdot\|v^{k-1} - v^k\|_H \qquad (2.8) \\ &\leq \sum_{i=0}^{k-1}\sigma^{k-i}\sqrt{\beta}\|A^*\|\cdot\|v^i - v^{i+1}\|_H + \sigma^k\|e_0(u^1)\|. \qquad (2.9) \end{aligned}$$

Then the upper bound property of the term $(u^{k+1} - u, e_k(u^{k+1}))$ is further obtained from (2.9) in Lemma 2.3 analogous to [23, Lemma 3.2], which will be used to prove the global convergence results eventually.

Lemma 2.3. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ be the sequence generated by the inexact ADMM (2.2) with the inexactness criterion (2.3). Then for any $\mu > 0$ and integer $K > 0$, we have*

$$\begin{aligned} \sum_{k=1}^{K}\big(u^{k+1} - u, e_k(u^{k+1})\big) \leq \frac{\mu}{2}\sum_{k=1}^{K}\frac{\sigma}{1-\sigma}\|u^{k+1} - u\|^2 + \frac{1}{2\mu}\sum_{k=1}^{K-1}\frac{\sigma}{1-\sigma}\beta\|A^*\|^2\|v^k - v^{k+1}\|_H^2 \\ + \frac{1}{2\mu}\frac{\sigma}{1-\sigma}\big(\|e_0(u^1)\| + \sqrt{\beta}\|A^*\|\cdot\|v^0 - v^1\|_H\big)^2, \ \ \forall\, u \in \mathcal{U}. \end{aligned}$$

With the above preparations, now we can establish the global convergence property of the inexact ADMM (2.2) with the inexactness criterion (2.3).

Theorem 2.4. *Let $w^* = (u^*, z^*, \lambda^*)^\top$ be a solution point of the variational inequality (2.6), and $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ the sequence generated by the inexact ADMM (2.2) with the inexactness criterion (2.3). Then we have the following assertions:*

(i) $\|u^k - u^*\| \to 0,\ \|Bz^k - Bz^*\| \to 0,\ \|A^*\lambda^k - A^*\lambda^*\| \to 0\ (k \to \infty)$,

(ii) $\|e_k(u^{k+1})\| \to 0,\ \|Bz^k - Bz^{k+1}\| \to 0,\ \|Au^{k+1} + Bz^{k+1} - b\| \to 0\ (k \to \infty)$,

(iii) there exists a subsequence $\{\lambda^{k_j}\} \subset \{\lambda^k\}$, such that λ^{k_j} weakly converges to λ^* in Hilbert space $\mathcal{H}$.

Proof. First, it yields from the definition of $F(\cdot)$ and the strong convexity of $J(\cdot)$ in (2.1) that

$$\big(w - \bar{w}^k, F(w) - F(\bar{w}^k)\big) = \big(u - \bar{u}^k, DJ(u) - DJ(\bar{u}^k)\big) \geq \alpha\|u - u^{k+1}\|^2, \ \ \forall\, w \in W. \qquad (2.10)$$

Then by using Lemmas 2.2 and 2.3, we obtain

$$0 \leq \sum_{k=1}^{K}\big\{R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(w^*)\big)\big\} \qquad (2.11)$$

$$
\begin{aligned}
&\leq \sum_{k=1}^{K}\left\{R(\bar{z}^k)-R(z^*)+\big(\bar{w}^k-w^*,F(\bar{w}^k)\big)+\big(\bar{w}^k-w^*,F(w^*)-F(\bar{w}^k)\big)\right\}\\
&\leq \frac{1}{2}\big(\|v^1-v^*\|_H^2-\|v^{K+1}-v^*\|_H^2\big)-\frac{1}{2}\sum_{k=1}^{K}\|v^k-v^{k+1}\|_H^2\\
&\quad+\sum_{k=1}^{K}\left\{\big(u^{k+1}-u^*,e_k(u^{k+1})\big)-\big(w^*-\bar{w}^k,F(w^*)-F(\bar{w}^k)\big)\right\}\\
&\leq \frac{1}{2}\big(\|v^1-v^*\|_H^2-\|v^{K+1}-v^*\|_H^2\big)-\frac{1}{2}\sum_{k=1}^{K}\|v^k-v^{k+1}\|_H^2-\alpha\sum_{k=1}^{K}\|u^{k+1}-u^*\|^2\\
&\quad+\frac{\mu}{2}\sum_{k=1}^{K}\frac{\sigma}{1-\sigma}\|u^{k+1}-u^*\|^2+\frac{1}{2\mu}\sum_{k=1}^{K-1}\frac{\sigma}{1-\sigma}\beta\|A^*\|^2\|v^k-v^{k+1}\|_H^2\\
&\quad+\frac{1}{2\mu}\frac{\sigma}{1-\sigma}\big(\|e_0(u^1)\|+\sqrt{\beta}\|A^*\|\cdot\|v^0-v^1\|_H\big)^2\\
&\leq \frac{1}{2}\big(\|v^1-v^*\|_H^2-\|v^{K+1}-v^*\|_H^2\big)+\Big(\frac{\mu}{2}\frac{\sigma}{1-\sigma}-\alpha\Big)\sum_{k=1}^{K}\|u^{k+1}-u^*\|^2\\
&\quad+\frac{1}{2}\Big(\frac{\sigma}{1-\sigma}\frac{\beta\|A^*\|^2}{\mu}-1\Big)\sum_{k=1}^{K-1}\|v^k-v^{k+1}\|_H^2-\frac{1}{2}\|v^K-v^{K+1}\|_H^2\\
&\quad+\frac{1}{2\mu}\frac{\sigma}{1-\sigma}\big(\|e_0(u^1)\|+\sqrt{\beta}\|A^*\|\cdot\|v^0-v^1\|_H\big)^2.
\end{aligned}
$$

The above estimate further implies that

$$
\Big(\alpha-\frac{\mu}{2}\frac{\sigma}{1-\sigma}\Big)\sum_{k=1}^{K}\|u^{k+1}-u^*\|^2+\frac{1}{2}\Big(1-\frac{\sigma}{1-\sigma}\frac{\beta\|A^*\|^2}{\mu}\Big)\sum_{k=1}^{K-1}\|v^k-v^{k+1}\|_H^2\leq C. \tag{2.12}
$$

It follows from (2.4) that

$$
2\alpha\cdot\frac{1-\sigma}{\sigma}>\beta\|A^*\|^2\frac{\sigma}{1-\sigma},
$$

then we can choose $\mu\in\big(\beta\|A^*\|^2\frac{\sigma}{1-\sigma},\ 2\alpha\cdot\frac{1-\sigma}{\sigma}\big)$ to ensure that

$$
\alpha-\frac{\mu}{2}\frac{\sigma}{1-\sigma}>0\ \text{ and }\ 1-\frac{\sigma}{1-\sigma}\frac{\beta\|A^*\|^2}{\mu}>0. \tag{2.13}
$$

Thus, the inequality (2.12) implies that

$$
\|u^{k+1}-u^*\|\to 0\ \text{ and }\ \|v^k-v^{k+1}\|_H\to 0\ (k\to\infty), \tag{2.14}
$$

which deduces

$$
\|B(z^k-z^{k+1})\|\to 0\ \text{ and }\ \|\lambda^k-\lambda^{k+1}\|\to 0\ (k\to\infty), \tag{2.15}
$$

and

$$
\|Au^{k+1}+Bz^{k+1}-b\|=\frac{1}{\beta}\|\lambda^k-\lambda^{k+1}\|\to 0\ (k\to\infty). \tag{2.16}
$$

From (2.14) and (2.16), we have

$$
\|B(z^{k+1}-z^*)\|=\|Au^{k+1}+Bz^{k+1}-b-(Au^*+Bz^*-b)-A(u^{k+1}-u^*)\|\to 0\ (k\to\infty). \tag{2.17}
$$

The convergence of $e_k(u^{k+1})$ can be easily obtained from (2.9) and (2.14) that

$$
\|e_k(u^{k+1})\|\to 0\ (k\to\infty). \tag{2.18}
$$

Together with the continuity property of $DJ(\cdot)$ in Assumption 1, we have from (2.5), (2.14), (2.16), (2.18) and the optimality condition $DJ(u^*) = A^*\lambda$ that

$$A^*\lambda^k - A^*\lambda^* = DJ(u^{k+1}) - DJ(u^*) + \beta A^*(Au^{k+1} - Au^k) + \beta A^*(Au^k + Bz^k - b) \\ - e_k(u^{k+1}) \to 0 \ \ (k \to \infty). \tag{2.19}$$

From Lemma 2.2, we have

$$\begin{aligned} 0 &\le R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(w^*)\big) \\ &= R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(\bar{w}^k)\big) - \big(\bar{w}^k - w^*, F(\bar{w}^k) - F(w^*)\big) \\ &\le \big(u^{k+1} - u^*, e_k(u^{k+1})\big) - \frac{1}{2}\big(\|v^{k+1} - v^*\|_H^2 - \|v^k - v^*\|_H^2 + \|v^k - v^{k+1}\|_H^2\big) \\ &\quad - \alpha\|u^{k+1} - u^*\|^2. \end{aligned}$$

Then according to Lemma 2.3 and (2.12), it yields

$$\begin{aligned} \|v^{k+1} - v^*\|_H^2 &\le \|v^k - v^*\|_H^2 + 2\big(u^{k+1} - u^*, e_k(u^{k+1})\big) \\ &\le \|v^{k-1} - v^*\|_H^2 + 2\big(u^k - u^*, e_k(u^k)\big) + 2\big(u^{k+1} - u^*, e_k(u^{k+1})\big) \\ &\le \|v^0 - v^*\|_H^2 + 2\sum_{i=0}^{k}\big(u^{i+1} - u^*, e_k(u^{i+1})\big) \\ &\le \|v^0 - v^*\|_H^2 + 2\big(u^1 - u^*, e_k(u^1)\big) \\ &\quad + \mu\sum_{k=1}^{K}\frac{\sigma}{1-\sigma}\|u^{k+1} - u^*\|^2 + \frac{1}{\mu}\sum_{k=1}^{K-1}\frac{\sigma}{1-\sigma}\beta\|A^*\|^2\|v^k - v^{k+1}\|_H^2 \\ &\quad + \frac{1}{\mu}\frac{\sigma}{1-\sigma}\big(\|e_0(u^1)\| + \sqrt{\beta}\|A^*\|\cdot\|v^0 - v^1\|_H\big)^2 \le C, \end{aligned} \tag{2.20}$$

which indicates that λ^k is bounded in Hilbert space $\mathcal{H}$, thus there exists a subsequence $\{\lambda^{k_j}\} \subset \{\lambda^k\}$ and $\hat{\lambda} \in \mathcal{H}$, such that λ^{k_j} weakly converges to $\hat{\lambda}$.

In addition, it follows from (2.7) that the subsequence $\{w^{k_j}\}$ satisfies the following first-order optimal conditions:

$$\begin{cases} e_k(u^{k_j+1}) + A^*\lambda^{k_j+1} + \beta A^*B(z^{k_j+1} - z^{k_j}) = DJ(u^{k_j+1}), & (2.21) \\ R(z) - R(z^{k_j+1}) - (\lambda^{k_j+1}, B(z - z^{k_j+1})) \ge 0, \ \ \forall\, z \in \mathcal{Z}. & (2.22) \end{cases}$$

Let $z = z^*$ in (2.22), and $z = z^{k_j+1}$ in $R(z) - R(z^*) - (\lambda^*, B(z - z^*)) \ge 0, \ \forall\, z \in \mathcal{Z}$, we have

$$-\|\lambda^*\|\cdot\|B(z^{k_j+1} - z^*)\| \le R(z^{k_j+1}) - R(z^*) \le \|\lambda^{k_j+1}\|\cdot\|B(z^{k_j+1} - z^*)\|.$$

From (2.17), and the boundness of λ^{k_j+1}, it can derive that

$$R(z^{k_j+1}) \to R(z^*) \ \ (j \to \infty). \tag{2.23}$$

Hence, together with (2.15), (2.18), (2.19), (2.21), the continuity of $DJ(\cdot)$, (2.17), (2.23), the boundeness of λ^{k_j+1} and the identity $(\lambda^{k_j+1}, Bz^{k_j+1}) = (\lambda^{k_j+1}, Bz^*) + (\lambda^{k_j+1}, Bz^{k_j+1} - Bz^*)$, we derive that

$$A^*\hat{\lambda} = DJ(u^*), \ \ B^*\hat{\lambda} \in \partial R(z^*), \ \ Au^* + Bz^* = b,$$

which concludes that $(u^*, z^*, \hat{\lambda})^\top$ is a KKT point of (1.1), with $\hat{\lambda} = \lambda^*$ and $\lambda^{k_j} \rightharpoonup \lambda^*$. The proof is completed. □

Remark 2.5. *If we further assume that the operator B is injective, then we can easily derive that*

$$z^k \to z^* \ (k \to \infty).$$

2.2 Convergence rate

In this subsection, we establish a worst-case $O(1/K)$ convergence rate in both ergodic and non-ergodic senses for the inexact ADMM (2.2) with the inexactness criterion (2.3), where K denotes the iteration counter.

We first introduce a criterion to measure the accuracy of the variational inequality (2.6). According to the criterion in [41] and Theorem 2.1, we say that $\hat{w} \in W$ is an approximate solution of problem (1.1) with an accuracy of $\varepsilon > 0$ if

$$R(\hat{z}) - R(z) + (\hat{w} - w, F(w)) \leq \varepsilon, \ \ \forall\, w \in \mathcal{D}(\hat{w}),$$

where $\mathcal{D}(\hat{w}) := \{w \in W : \|v - \hat{v}\|_H \leq 1\}$.

Then, applying the inequality (2.11) and the similar technique in the proof of Theorem 4.2 in [23], we can derive a worst-case $O(1/K)$ convergence rate in the ergodic sense for the inexact algorithm in the following theorem.

Theorem 2.6. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ be the sequence generated by the inexact ADMM (2.2) with the inexactness criterion (2.3). For any integer $K \geq 1$, we further define $\hat{w}_K = \frac{1}{K}\sum_{k=1}^{K} \bar{w}^k$. Then there exists $C(v) > 0$ such that*

$$R(\hat{z}_K) - R(z) + \big(\hat{w}_K - w, F(w)\big) \leq \frac{C(v)}{K}, \ \ \forall\, w \in \mathcal{D}(\hat{w}_K),$$

where

$$C(v) = \frac{1}{2}||v^1 - v||_H^2 + \frac{1}{2\mu}\frac{\sigma}{1-\sigma}\Big(||e_0(u^1)|| + \sqrt{\beta}||A^*|| \cdot ||v^0 - v^1||_H\Big)^2.$$

We confirm that $C(v)$ is bounded when $w \in \mathcal{D}(\hat{w}_K)$. Indeed, for any integer $k \geq 0$, it implies from (2.20) and (2.15) that $\|v^{k+1}\|_H \leq C$, $\|B(z^{k+1} - z^k)\| \leq C$, which implies $\|\hat{v}_K\|_H \leq C$ and

$$\|v^0 - v\|_H \leq \|v - \hat{v}_K\|_H + \|v^0 - \hat{v}_K\|_H \leq 1 + \|v^0\|_H + \|\hat{v}_K\|_H \leq C, \ \ \forall\, w \in \mathcal{D}(\hat{w}_K).$$

The optimality gap and feasibility error for the problem (1.1) are estimated as follows.

Theorem 2.7. *Let $w^* = (u^*, z^*, \lambda^*)^\top$ be a solution point to the variational inequality (2.6), and $\hat{w}_K$ be defined in Theorem 2.6. We have that*

$$|J(\hat{u}_K) + R(\hat{z}_K) - J(u^*) - R(z^*)| \leq \left(1 + \frac{\alpha}{2} \cdot \frac{1}{\alpha - \frac{\mu}{2}\frac{\sigma}{1-\sigma}}\right)\frac{C(v^*)}{K} + \frac{2C_\lambda\|\lambda^*\|}{\beta K}, \tag{2.24}$$

and

$$\|A\hat{u}_K + B\hat{z}_K - b\| \leq \frac{2C_\lambda}{\beta K}, \tag{2.25}$$

where C_λ is the upper bound of $\|\lambda^k\|$ (the boundedness of $\|\lambda^k\|$ is confirmed in the proof of Theorem 2.4).

Proof. It easily deduces from (2.11) that

$$\frac{1}{K}\sum_{k=1}^{K}\big\{R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(\bar{w}^k)\big)\big\} \leq \frac{C(v^*)}{K} + \frac{\mu}{2}\frac{\sigma}{1-\sigma}\frac{1}{K}\sum_{k=1}^{K}\|u^{k+1} - u^*\|. \tag{2.26}$$

By the strong convexity of $J(\cdot)$, we obtain the estimate for the left hand side of (2.26) that

$$\frac{1}{K}\sum_{k=1}^{K}\big\{R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(\bar{w}^k)\big)\big\}$$

$$
\begin{aligned}
&= \frac{1}{K}\sum_{k=1}^{K}\big\{R(\bar z^k)-R(z^*)+\big(\bar u^k-u^*,DJ(\bar u^k)\big)-\big(\bar u^k-u^*,A^*\bar\lambda^k\big)-\big(\bar z^k-z^*,B^*\bar\lambda^k\big)\\
&\qquad\qquad +\big(\bar\lambda^k-\lambda^*,A\bar u^k+B\bar z^k-b\big)\big\}\\
&\ge \frac{1}{K}\sum_{k=1}^{K}\big\{R(\bar z^k)-R(z^*)+J(\bar u^k)-J(u^*)+\frac{\alpha}{2}\|u^{k+1}-u^*\|^2-\big(\lambda^*,A\bar u^k+B\bar z^k-b\big)\big\}.
\end{aligned}
$$

Then it follows from (2.26), the convexity of $J(\cdot)$ and $R(\cdot)$ that

$$
\begin{aligned}
&J(\hat u_K)+R(\hat z_K)-J(u^*)-R(z^*)-\big(\lambda^*,A\hat u_K+B\hat z_K-b\big)\\
&\le \frac{C(v^*)}{K}+\Big(\frac{\mu}{2}\frac{\sigma}{1-\sigma}-\frac{\alpha}{2}\Big)\frac{1}{K}\sum_{k=1}^{K}\|u^{k+1}-u^*\|^2\\
&\le \frac{C(v^*)}{K}+\frac{\alpha}{2K}\sum_{k=1}^{K}\|u^{k+1}-u^*\|^2\\
&\le \frac{C(v^*)}{K}+\frac{\alpha}{2K}\frac{C(v^*)}{\alpha-\frac{\mu}{2}\frac{\sigma}{1-\sigma}}\\
&=\Big(1+\frac{\alpha}{2}\cdot\frac{1}{\alpha-\frac{\mu}{2}\frac{\sigma}{1-\sigma}}\Big)\frac{C(v^*)}{K},
\end{aligned}
$$

where $\alpha-\frac{\mu}{2}\frac{\sigma}{1-\sigma}>0$ in (2.13) is applied for the second inequality, and the third inequality is obtained from (2.12). Then (2.24) is obtained based on the fact that

$$
J(u)+R(z)-J(u^*)-R(z^*)-\big(\lambda^*,Au+Bz-b\big)\ge 0,\ \ \forall\, u\in\mathcal{U},\ z\in\mathcal{Z}.
$$

The boundedness of λ^k $(k=1,2,\ldots)$ confirmed in the proof of Theorem 2.4 implies that

$$
\begin{aligned}
\|A\hat u_K+B\hat z_K-b\| &= \Big\|\frac{1}{K}\sum_{k=1}^{K}(Au^{k+1}+Bz^{K+1}-b)\Big\|\le\Big\|\frac{1}{\beta K}\sum_{k=1}^{K}(\lambda^k-\lambda^{k+1})\Big\|\\
&\le\frac{1}{\beta K}\|(\lambda^1-\lambda^{K+1})\|\le\frac{2C_\lambda}{\beta K}.
\end{aligned}
$$

The proof is completed. □

It follows from (2.7) that the iterate $(u^{k+1},z^{k+1},\lambda^{k+1})^\top$ satisfies

$$
\begin{aligned}
&R(z)-R(z^{k+1})+\big(w-w^{k+1},F(w^{k+1})\big)\\
&+\left(w-w^{k+1},\begin{pmatrix}\beta A^*B(z^k-z^{k+1})-e_k(u^{k+1})\\ 0\\ \frac{1}{\beta}(\lambda^{k+1}-\lambda^k)\end{pmatrix}\right)\ge 0,
\end{aligned}
$$

for any $w\in W$. Then we can easily derive that $(u^{k+1},z^{k+1},\lambda^{k+1})^\top$ is a solution point of the algorithm if and only if

$$
\|v^k-v^{k+1}\|_H^2=0\ \text{ and }\ \|e_k(u^{k+1})\|^2=0.
$$

Hence, we can measure the accuracy of the iterate $(u^{k+1},z^{k+1},\lambda^{k+1})^\top$ by $\|v^k-v^{k+1}\|_H^2$ and $\|e_k(u^{k+1})\|^2$. The following theorem originates from the inequality (2.11) and the similar technique in the proof of Theorem 4.3 in [23], and it shows that the inexact method has a worst-case $O(1/K)$ convergence rate in non-ergodic sense.

Theorem 2.8. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ be the sequence generated by the inexact ADMM (2.2) with the inexactness criterion (2.3). Then, for any integer $K > 0$, there exist $C_1,\ C_2,\ C_3 > 0$, such that*

$$\min_{1\le k\le K} \{\|v^k - v^{k+1}\|_H^2\} \le \frac{C_1}{K} \quad \text{and} \quad \min_{1\le k\le K} \{\|e_k(u^{k+1})\|^2\} \le \frac{C_2}{K} + \frac{C_3}{K^2},$$

where

$$\begin{cases} C_1 = \dfrac{1}{\mu_0} C(v^*), \quad \mu_0 := \dfrac{1}{2}\Big(1 - \dfrac{\sigma}{1-\sigma} \dfrac{\beta||A^*||^2}{\mu}\Big), \\ C_2 = 2\Big(\dfrac{\sigma}{1-\sigma}\sqrt{\beta}||A^*||\Big)^2 \cdot \dfrac{1}{\mu_0} C(v^*), \\ C_3 = 2\Big(\dfrac{\sigma}{1-\sigma}\Big)^2 \cdot \big(||e_0(u^1)|| + \sqrt{\beta}||A^*|| \cdot ||v^0 - v^1||_H\big)^2. \end{cases}$$

Remark 2.9. *It is still open to prove the linear convergence of the inexact ADMM (2.2) with the inexactness criterion (2.3) as illustrated in numerical tests. Note that we only employ the third inequality (2.9) instead of the first one (2.3) to estimate the convergence and convergence rate results in Theorems 2.4, 2.6 and 2.8. Hence, strictly speaking, our discussion revolves around the inexact algorithm utilizing the inequality (2.9) as the inexactness criterion, which is looser than the original one (2.3). Such discrepancies could potentially explain why it is challenging to establish the linear convergence for the inexact ADMM (2.2) with the inexactness criterion (2.3).*

3 A new inexact ADMM with global linear convergence rate

In this section, we design a new inexactness criterion that can ensure the linear convergence of the corresponding inexact ADMM. Then the theoretical results are rigorously analyzed provided that $J(\cdot)$ also has a Lipschitz continuous gradient.

3.1 A new inexactness criterion for the ADMM (2.2)

It is intuitive that the new inexactness criterion should be linearly convergent as the resulting algorithm behaves. Then a linear convergence rate of $\{\|e_k(u^{k+1})\|^2\}$ is achieved by changing the inexactness criterion (2.3) into

$$\begin{cases} \|e_k(u^{k+1})\| \le \sigma \|e_k(u^k)\|, \quad k = 0, \cdots, K_0, \\ \|e_k(u^{k+1})\|^2 \le \varepsilon \bar{\sigma}^k \displaystyle\sum_{i=1}^{k} \|v^{i-1} - v^i\|_H^2, \quad k \ge K_0 + 1, \end{cases} \tag{3.1}$$

where $K_0 \ge 0$, $\varepsilon > 0$, $\bar{\sigma} \in (0, 1)$, and σ satisfies (2.4). In the numerical experiments in section 5, we will set K_0 as a relatively small positive integer to make use of the robustness of inexactness criterion (2.3) to obtain a good initial iteration, and then flexibly adjust $\bar{\sigma}$ when $k > K_0$ to control the desired convergence speed. Note that it only needs to calculate $\|v^{i-1} - v^i\|_H^2$ once in each iteration and accumulates to obtain the summation in (3.1).

For the new inexact ADMM (2.2) nested with the inexactness criterion (3.1), we derive the estimate in Lemma 3.1 similar to (2.12), which directly deduces the convergence of the new algorithm.

Lemma 3.1. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ be the sequence generated by the new inexact ADMM (2.2) nested with the inexactness criterion (3.1), and $w^* = (u^*, z^*, \lambda^*)^\top$ be the solution point of the variational inequality (2.6). Then for any integer $K > 0$, there exists $C > 0$, such that*

$$\sum_{k=1}^{K} \|u^{k+1} - u^*\|^2 + \sum_{k=1}^{K-1} \|v^k - v^{k+1}\|_H^2 \le C. \tag{3.2}$$

Proof. We only need to consider the case $k \geq K_0 + 1$. It derives from (3.1) that

$$
\begin{aligned}
&\sum_{k=K_0+1}^{K} \left(u^{k+1} - u^*, e_k(u^{k+1})\right) \leq \sum_{k=K_0+1}^{K} \|u^{k+1} - u^*\| \cdot \|e_k(u^{k+1})\| \\
&\leq \frac{\alpha}{2} \sum_{k=K_0+1}^{K} \|u^{k+1} - u^*\|^2 + \frac{1}{2\alpha} \sum_{k=K_0+1}^{K} \|e_k(u^{k+1})\|^2 \\
&\leq \frac{\alpha}{2} \sum_{k=K_0+1}^{K} \|u^{k+1} - u^*\|^2 + \frac{\varepsilon}{2\alpha} \sum_{k=K_0+1}^{K} \bar{\sigma}^k \sum_{i=1}^{k} \|v^{i-1} - v^i\|_H^2 \\
&\leq \frac{\alpha}{2} \sum_{k=K_0+1}^{K} \|u^{k+1} - u^*\|^2 + \frac{\varepsilon}{2\alpha} \sum_{k=K_0+1}^{K} \frac{\bar{\sigma}^k}{1-\bar{\sigma}} \|v^{k-1} - v^k\|_H^2 \\
&\quad + \frac{\varepsilon}{2\alpha} \frac{\bar{\sigma}^{K_0+1}}{1-\bar{\sigma}} \sum_{k=1}^{K_0} \|v^{k-1} - v^k\|_H^2.
\end{aligned}
$$

Then by Lemma 2.2 and (2.10), we can obtain that

$$
\begin{aligned}
0 &\leq \sum_{k=K_0+1}^{K} \left\{R(\bar{z}^k) - R(z^*) + \left(\bar{w}^k - w^*, F(w^*)\right)\right\} \\
&\leq \sum_{k=K_0+1}^{K} \left\{R(\bar{z}^k) - R(z^*) + \left(\bar{w}^k - w^*, F(\bar{w}^k)\right) + \left(\bar{w}^k - w^*, F(w^*) - F(\bar{w}^k)\right)\right\} \\
&\leq \sum_{k=K_0+1}^{K} \left(u^{k+1} - u^*, e_k(u^{k+1})\right) + \frac{1}{2}\left(\|v^{K_0+1} - v^*\|_H^2 - \|v^{K+1} - v^*\|_H^2\right) \\
&\quad - \frac{1}{2} \sum_{k=K_0+1}^{K} \|v^k - v^{k+1}\|_H^2 - \sum_{k=K_0+1}^{K} \left(w^* - \bar{w}^k, F(w^*) - F(\bar{w}^k)\right) \\
&\leq -\frac{\alpha}{2} \sum_{k=K_0+1}^{K} \|u^{k+1} - u^*\|^2 + \frac{1}{2}\left(\|v^{K_0+1} - v^*\|_H^2 - \|v^{K+1} - v^*\|_H^2\right) - \frac{1}{2}\|v^K - v^{K+1}\|_H^2 \\
&\quad + \sum_{k=K_0+1}^{K-1} \left(\frac{\varepsilon}{2\alpha} \frac{\bar{\sigma}^{k+1}}{1-\bar{\sigma}} - \frac{1}{2}\right) \|v^k - v^{k+1}\|_H^2 + \frac{\varepsilon}{2\alpha} \frac{\bar{\sigma}^{K_0+1}}{1-\bar{\sigma}} \sum_{k=1}^{K_0+1} \|v^{k-1} - v^k\|_H^2.
\end{aligned}
$$

Thus, the above estimate implies that

$$
\frac{\alpha}{2} \sum_{k=K_0+1}^{K} \|u^{k+1} - u^*\|^2 + \frac{1}{2} \sum_{k=K_0+1}^{K-1} \left(1 - \frac{\varepsilon}{\alpha} \frac{\bar{\sigma}^{k+1}}{1-\bar{\sigma}}\right) \|v^k - v^{k+1}\|_H^2 \leq C, \tag{3.3}
$$

which can easily conclude the results (3.2) as $\bar{\sigma} \in (0, 1)$. □

From Lemma 3.1 and (3.1), it is easy to deduce that $\{\|e_k(u^{k+1})\|^2\}$ is linearly convergent, i.e., there exists $C > 0$, such that

$$
\|e_k(u^{k+1})\|^2 \leq C\bar{\sigma}^k, \quad k \geq K_0 + 1. \tag{3.4}
$$

In addition, we have $\|u^{k+1} - u^*\|^2 \to 0$ $(k \to \infty)$, then by the similar approach as in the proof of Theorem 2.4, analogous convergence conclusions still hold for the new inexact ADMM.

3.2 Analysis of linear convergence rate

Through out this subsection, we additionally make the following assumption for the optimization problem (1.1).

Assumption 2. *The functional $J(\cdot)$ in (1.1) has a Lipschitz continuous gradient around the solution point u^*, i.e., there exists a constant $\kappa > 0$, such that*

$$\|DJ(u) - DJ(u^*)\| \leq \kappa \|u - u^*\|, \ \ \forall\, u \in \mathcal{U}. \tag{3.5}$$

Assumption 3. *The operator A in (1.1) is surjective.*

It is noticed that the conditions in Assumptions 1-3 is analogous to those for the linear convergence of the exactly-solved ADMM (1.2) according to Table 1 in [11]. To obtain the linear convergence rate of $\{\|v^{k+1} - v^*\|_H^2\}$, we first establish a useful lemma.

Lemma 3.2. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ be the sequence generated by the new inexact ADMM (2.2) nested with the inexactness criterion (3.1), and $v^* = (z^*, \lambda^*)^\top$ satisfy the variational inequality (2.6). Then there exist $C_1,\ C_2 > 0$, such that*

$$\|v^k - v^*\|_H^2 - \|v^{k+1} - v^*\|_H^2 \geq C_1 \|v^{k+1} - v^*\|_H^2 - C_2 \|e_k(u^{k+1})\|^2. \tag{3.6}$$

Proof. From Lemma 2.2 and (2.10), we can obtain that

$$\begin{aligned}
0 &\leq R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(w^*)\big) \\
&= R(\bar{z}^k) - R(z^*) + \big(\bar{w}^k - w^*, F(\bar{w}^k)\big) - \big(\bar{w}^k - w^*, F(\bar{w}^k) - F(w^*)\big) \\
&\leq \big(u^{k+1} - u^*, e_k(u^{k+1})\big) - \frac{1}{2}\big(\|v^{k+1} - v^*\|_H^2 - \|v^k - v^*\|_H^2 + \|v^k - v^{k+1}\|_H^2\big) \\
&\quad - \alpha \|u^{k+1} - u^*\|^2,
\end{aligned} \tag{3.7}$$

which directly implies

$$\begin{aligned}
&\|v^k - v^*\|_H^2 - \|v^{k+1} - v^*\|_H^2 \\
&\geq \|v^k - v^{k+1}\|_H^2 + 2\alpha \|u^{k+1} - u^*\|^2 - 2\big(u^{k+1} - u^*, e_k(u^{k+1})\big) \\
&\geq \|v^k - v^{k+1}\|_H^2 + (2\alpha - \gamma)\|u^{k+1} - u^*\|^2 - \frac{1}{\gamma}\|e_k(u^{k+1})\|^2,
\end{aligned} \tag{3.8}$$

where $\gamma \in (0, 2\alpha)$.

On one hand, it holds that

$$\begin{aligned}
\|u^{k+1} - u^*\|^2 &\geq \frac{1}{\|A\|^2}\|A(u^{k+1} - u^*)\|^2 \\
&= \frac{1}{\|A\|^2}\|Au^{k+1} + Bz^{k+1} - b - B(z^{k+1} - z^*)\|^2 \\
&= \frac{1}{\|A\|^2}\Big\|\frac{1}{\beta}(\lambda^k - \lambda^{k+1}) - B(z^{k+1} - z^*)\Big\|^2 \\
&= \frac{1 - \mu_0}{\|A\|^2}\|B(z^{k+1} - z^*)\|^2 - \frac{1 - \mu_0}{\mu_0 \|A\|^2}\frac{1}{\beta^2}\|\lambda^k - \lambda^{k+1}\|^2,
\end{aligned} \tag{3.9}$$

where $\mu_0 \in (0,1)$. On the other hand, we deduce from (3.5) and (2.5) that

$$
\begin{aligned}
\|u^{k+1}-u^*\|^2 &\geq \frac{1}{\kappa^2}\|DJ(u^{k+1})-DJ(u^*)\|^2 \\
&= \frac{1}{\kappa^2}\|A^*(\lambda^{k+1}-\lambda^*)-\beta A^*B(z^k-z^{k+1})+e_k(u^{k+1})\|^2 \\
&\geq \frac{(1-\mu_1)\lambda_{\min}(AA^*)}{\kappa^2}\|\lambda^{k+1}-\lambda^*\|^2 \\
&\quad -\frac{(1-\mu_1)(1+\mu_2)\beta^2\|A^*\|^2}{\mu_1\kappa^2}\|B(z^k-z^{k+1})\|^2 \\
&\quad -\frac{(1-\mu_1)(1+\mu_2)}{\mu_1\mu_2\kappa^2}\|e_k(u^{k+1})\|^2,
\end{aligned}
\tag{3.10}
$$

where $\mu_1 \in (0,1)$ and $\mu_2 > 0$. Then, choosing $\mu_3 \in (0,1)$, it yields from (3.8), (3.9) and (3.10) that

$$
\begin{aligned}
&\|v^k-v^*\|_H^2-\|v^{k+1}-v^*\|_H^2 \\
&\geq \beta\|B(z^k-z^{k+1})\|^2+\frac{1}{\beta}\|\lambda^k-\lambda^{k+1}\|^2+(2\alpha-\gamma)\mu_3\|u^{k+1}-u^*\|^2 \\
&\quad +(2\alpha-\gamma)(1-\mu_3)\|u^{k+1}-u^*\|^2-\frac{1}{\gamma}\|e_k(u^{k+1})\|^2 \\
&\geq \left[\beta-(2\alpha-\gamma)\mu_3\frac{(1-\mu_1)(1+\mu_2)\beta^2\|A^*\|^2}{\mu_1\kappa^2}\right]\|B(z^k-z^{k+1})\|^2 \\
&\quad +\left[\frac{1}{\beta}-(2\alpha-\gamma)(1-\mu_3)\frac{(1-\mu_0)}{\mu_0\|A\|^2}\frac{1}{\beta^2}\right]\|\lambda^k-\lambda^{k+1}\|^2 \\
&\quad +(2\alpha-\gamma)(1-\mu_3)\frac{(1-\mu_0)}{\|A\|^2}\|B(z^{k+1}-z^*)\|^2 \\
&\quad +(2\alpha-\gamma)\mu_3\frac{(1-\mu_1)\lambda_{\min}(AA^*)}{\kappa^2}\|\lambda^{k+1}-\lambda^*\|^2 \\
&\quad -\left[\frac{1}{\gamma}+(2\alpha-\gamma)\mu_3\frac{(1-\mu_1)(1+\mu_2)}{\mu_1\mu_2\kappa^2}\right]\|e_k(u^{k+1})\|^2.
\end{aligned}
$$

We choose μ_1 and μ_0 satisfying

$$
\frac{(2\alpha-\gamma)(1+\mu_2)\mu_3\beta\|A^*\|^2}{(2\alpha-\gamma)(1+\mu_2)\mu_3\beta\|A^*\|^2+\kappa^2}\leq \mu_1<1 \ \text{ and } \ \frac{(2\alpha-\gamma)(1-\mu_3)}{(2\alpha-\gamma)(1-\mu_3)+\beta\|A\|^2}\leq \mu_0<1
$$

to insure that

$$
\beta-(2\alpha-\gamma)\mu_3\frac{(1-\mu_1)(1+\mu_2)\beta^2\|A^*\|^2}{\mu_1\kappa^2}\geq 0 \ \text{ and } \ \frac{1}{\beta}-(2\alpha-\gamma)(1-\mu_3)\frac{(1-\mu_0)}{\mu_0\|A\|^2}\frac{1}{\beta^2}\geq 0.
$$

Then it follows from (3.8) that

$$
\begin{aligned}
&\|v^k-v^*\|_H^2-\|v^{k+1}-v^*\|_H^2 \\
&\geq (2\alpha-\gamma)(1-\mu_3)\frac{(1-\mu_0)}{\|A\|^2}\|B(z^{k+1}-z^*)\|^2 \\
&\quad +(2\alpha-\gamma)\mu_3\frac{(1-\mu_1)\lambda_{\min}(AA^*)}{\kappa^2}\|\lambda^{k+1}-\lambda^*\|^2 \\
&\quad -\left[\frac{1}{\gamma}+(2\alpha-\gamma)\mu_3\frac{(1-\mu_1)(1+\mu_2)}{\mu_1\mu_2\kappa^2}\right]\|e_k(u^{k+1})\|^2.
\end{aligned}
\tag{3.11}
$$

Let

$$
\begin{cases}
C_1 := (2\alpha-\gamma)\min\left\{(1-\mu_3)\dfrac{(1-\mu_0)}{\beta\|A\|^2},\ \mu_3\dfrac{(1-\mu_1)\lambda_{\min}(AA^*)\beta}{\kappa^2}\right\}>0, \\
C_2 := \dfrac{1}{\gamma}+(2\alpha-\gamma)\mu_3\dfrac{(1-\mu_1)(1+\mu_2)}{\mu_1\mu_2\kappa^2},
\end{cases}
$$

then we can conclude that

$$\|v^k - v^*\|_H^2 - \|v^{k+1} - v^*\|_H^2 \geq C_1\|v^{k+1} - v^*\|_H^2 - C_2\|e_k(u^{k+1})\|^2.$$

The proof is completed. □

Remark 3.3. *Inspired by [48], for the problem (1.1) in finite dimensional Euclidean spaces, the condition in Assumption 3 can be reduced to $Im(B) \subset Im(A)$ for deriving the estimates (3.10) and (3.6), where $Im(\cdot)$ refers to the image of a matrix. In fact, taking $\lambda^0 + \beta b \in Im(A)$, we can deduce $\lambda^{k+1} = \lambda^0 - \sum_{i=1}^{k+1} \beta(Au^i + Bz^i - b) \in Im(A)$. Since $Im(A)$ is closed, then it obtains $\lambda^* \in Im(A)$ and $\lambda^{k+1} - \lambda^* \in Im(A)$. By the standard SVD $A = U\Sigma V^\top$ of A, there holds $\lambda^{k+1} - \lambda^* = U\eta$ and*

$$\begin{aligned}\|A^\top(\lambda^{k+1} - \lambda^*)\|^2 &= \|V\Sigma U^\top U\eta\|^2 \\ &= \|\Sigma\eta\|^2 \\ &\geq \lambda_+(AA^\top)\|\eta\|^2 \\ &= \lambda_+(AA^\top)\|\lambda^{k+1} - \lambda^*\|^2,\end{aligned}$$

where $\lambda_+(AA^\top)$ is the strictly positive minimum eigenvalue of $AA^\top$ (if $rank(A) = r$, then only the first r components of η are nonzero).

Note that (3.6) reduces to the standard contraction property of the sequence $\{v^k\}$ in [33] if the iterate u^{k+1} is computed exactly (i.e., $\|e_k(u^{k+1})\| = 0$). Now we are prepared to establish the global linear convergence rate of $\{\|v^{k+1} - v^*\|_H^2\}$.

Theorem 3.4. *Let $\{w^k\} = \{(u^k, z^k, \lambda^k)^\top\}$ be the sequence generated by the new inexact ADMM (2.2) nested with the inexactness criterion (3.1), and $v^* = (z^*, \lambda^*)^\top$ satisfy the variational inequality (2.6). Then there exists $C > 0$, such that*

$$\|v^{k+1} - v^*\|_H^2 \leq C\hat{\sigma}^{k+1}, \quad \hat{\sigma} \in (0,1), \quad k \geq K_0 + 1. \tag{3.12}$$

Proof. Let $\sigma_1 := 1/(1 + C_1)$. By using (3.4) and (3.6), we obtain that

$$\begin{aligned}\|v^{k+1} - v^*\|_H^2 &\leq \sigma_1\|v^k - v^*\|_H^2 + C_2\sigma_1\|e_k(u^{k+1})\|^2 \\ &\leq \sigma_1\big(\sigma_1\|v^{k-1} - v^*\|_H^2 + C\sigma_1\bar{\sigma}^{k-1}\big) + C\sigma_1\bar{\sigma}^k \\ &= \sigma_1^2\|v^{k-1} - v^*\|_H^2 + C\sigma_1^2\bar{\sigma}^{k-1} + C\sigma_1\bar{\sigma}^k \\ &\leq \sigma_1^{k-K_0}\|v^{K_0+1} - v^*\|_H^2 + C\sigma_1 \sum_{i=K_0+1}^{k} \sigma_1^{k-i}\bar{\sigma}^i \\ &= \sigma_1^{k-K_0}\|v^{K_0+1} - v^*\|_H^2 + C\sigma_1^{k+1} \sum_{i=K_0+1}^{k} \left(\frac{\bar{\sigma}}{\sigma_1}\right)^i,\end{aligned}$$

then it implies

$$\|v^{k+1} - v^*\|_H^2 \leq \begin{cases} C\sigma_1^{k+1}, & \bar{\sigma} < \sigma_1, \\ C\bar{\sigma}^{k+1}, & \bar{\sigma} > \sigma_1, \\ C(k - K_0 + 1)\bar{\sigma}^{k+1}, & \bar{\sigma} = \sigma_1. \end{cases} \tag{3.13}$$

Therefore, the result (3.12) holds with some $\hat{\sigma} \in [\max\{\bar{\sigma}, \sigma_1\}, 1)$. □

Remark 3.5. *It follows from (3.8) that*

$$(2\alpha - \gamma)\|u^{k+1} - u^*\|^2 \leq \|v^k - v^*\|_H^2 + \frac{1}{\gamma}\|e_k(u^{k+1})\|^2,$$

then we can easily obtain from (3.4) and (3.12) that $\{\|u^{k+1} - u^*\|^2\}$ has a global linear convergence rate. The estimate (3.12) can guarantee the linear convergence of $\|Bz^k - Bz^*\|^2$. If we further assume that B is injective, then the linear convergence of $\|z^k - z^*\|^2$ can be obtained.

4 An implementable inexact ADMM framework for identifying coefficients in elliptic equations

In this section, we consider to apply the inexact ADMM for solving two nonlinear inverse coefficient problems with L^2-TV regularization for elliptic equations. In the following, $(\cdot,\cdot)$ and $\|\cdot\|$ represent the L^2-inner product and the L^2-norm, respectively.

Inverse diffusion coefficient problem. The diffusion coefficient u in an elliptic equation is identified by the following L^2-TV regularized optimization problem

$$\begin{cases} \min\limits_{u} \ \dfrac{1}{2}\displaystyle\int_\Omega u|\nabla y - \nabla y_\delta|^2 dx + \dfrac{\alpha}{2}\|u\|^2 + \tau\|u\|_{TV}, \\ \text{s.t. } -\nabla\cdot(u\nabla y) = f, \ \ (u,y)\in K\times H_0^1(\Omega), \end{cases} \tag{4.1}$$

where Ω is a bounded convex polyhedral domain in $\mathbb{R}^d$ $(d \le 2)$, $\nabla y_\delta \in (L^2(\Omega))^d$ is the observation data, and

$$K := \{u \in L^\infty \cap BV(\Omega) : 0 < a_0 \le u \le a_1, \text{ a.e. in } \Omega\},$$

is the admissible set. $BV(\Omega)$ endowed with the norm $\|u\|_{BV} := \|u\|_{L^1(\Omega)} + \|u\|_{TV}$ is a Banach space, and $\|u\|_{TV}$ is the total variation (TV) norm defined in [2] as

$$\|u\|_{TV} := \sup\Big\{\int_\Omega u \text{ div } \varphi \, dx : \varphi \in C_c^1(\Omega;\mathbb{R}^d), \ \|\varphi\|_\infty \le 1\Big\},$$

where $C_c^1(\Omega;\mathbb{R}^d)$ is the set of once continuously differentiable functions with compact support in Ω, and $\|\varphi\|_\infty = \sup_{x\in\Omega}(\sum_{i=1}^d |\varphi_i(x)|^2)^{1/2}$. In addition, $\tau > 0$ is a penalty parameter of the TV term.

Inverse potential coefficient problem. The potential coefficient u in an elliptic equation is recovered by the L^2-TV regularized optimization problem as follows

$$\begin{cases} \min\limits_{u} \ \dfrac{1}{2}\displaystyle\int_\Omega |\nabla y - \nabla y_\delta|^2 dx + \dfrac{1}{2}\displaystyle\int_\Omega u|y - y_\delta|^2 dx + \dfrac{\alpha}{2}\|u\|^2 + \tau\|u\|_{TV}, \\ \text{s.t. } -\Delta y + uy = f, \ \ (u,y)\in K\times H_0^1(\Omega), \end{cases} \tag{4.2}$$

where $y_\delta \in (L^2(\Omega))^d$ is the observation data.

The two inverse coefficient problems of elliptic equations with only TV regularization have been discussed in [10, 28, 29, 54], and the TV regularization is popular for identifying discontinuous or oscillated coefficients (see, e.g. [1, 6, 7, 10, 26]). As we know that $BV(\Omega) \hookrightarrow L^2(\Omega)$ when $d = 1, 2$ (see [2]), then $L^2(\Omega)\cap BV(\Omega)$ and $BV(\Omega)$ are equivalent Banach spaces [8]. Then adding the L^2 regularization term is reasonable in (4.1) and (4.2), which also makes the functionals strongly convex and keeps favorable characters of the TV regularization. In the rest of this section, we mainly focus on the application of the new inexact ADMM in section 3 to the two optimization problems (4.1) and (4.2).

4.1 Conceptual application of the ADMM

We first discretize the problems (4.1) and (4.2) in the piecewise linear finite element space V_h and obtain the following discrete problem:

$$\min_{u_h\in K_h} J(u_h) + \tau\|\nabla u_h\|_{L^1(\Omega)}, \tag{4.3}$$

where $K_h = V_h \cap K$, $J(\cdot) = J_1(\cdot)$ or $J_2(\cdot)$ depends on the context, and

$$J_1(u_h) := \int_\Omega u_h|\nabla y_h(u_h) - \nabla y_\delta|^2 dx + \frac{\alpha}{2}\|u_h\|^2, \tag{4.4}$$

$$J_2(u_h) := \frac{1}{2}\int_\Omega |\nabla y_h(u_h) - \nabla y_\delta|^2 dx + \frac{1}{2}\int_\Omega u_h|y_h(u_h) - y_\delta|^2 dx + \frac{\alpha}{2}\|u_h\|^2, \tag{4.5}$$

where $y_h(u_h)$ in $J_1(\cdot)$ and $J_2(\cdot)$ are respectively the solutions of the following variational forms

$$\int_\Omega u_h \nabla y_h(u_h) \cdot \nabla \phi_h dx = (f, \phi_h) \tag{4.6}$$

and

$$\int_\Omega \nabla y_h(u_h) \cdot \nabla \phi_h dx + \int_\Omega u_h y_h(u_h) \phi_h dx = (f, \phi_h) \tag{4.7}$$

for any $\phi_h \in V_h \cap H_0^1(\Omega)$.

We next show that the two regularized inverse problems (4.1) and (4.2) satisfy the assumptions in the previous sections.

Theorem 4.1 ([30])**.** *The functionals $J_1(\cdot)$ and $J_2(\cdot)$ are α-strongly convex on the convex set K_h, and both the problems* (4.1) *and* (4.2) *exist unique solutions.*

The next two theorems indicate that the two regularized inverse problems (4.1) and (4.2) satisfy Assumption 2.

Theorem 4.2. *The functional $J_1(\cdot)$ has a Lipschitz continuous gradient around the solution point $u_h^* \in K_h$, i.e., there exists $\kappa > 0$ such that*

$$\|DJ_1(u_h) - DJ_1(u_h^*)\| \le \kappa \|u_h - u_h^*\|, \ \ \forall\, u_h \in V_h. \tag{4.8}$$

Proof. For any $u_h \in K_h$, we have from [28] that the gradient of the functional $J_1(\cdot)$ satisfies

$$\big(DJ_1(u_h), v_h\big) = -\frac{1}{2}\int_\Omega \big(|\nabla y_h(u_h)|^2 - |\nabla y_\delta|^2\big) v_h dx + \alpha \int_\Omega u_h v_h dx, \ \ \forall\, v_h \in V_h. \tag{4.9}$$

Then it obtains that

$$\begin{aligned}
&\big(DJ_1(u_h) - DJ_1(u_h^*), u_h - u_h^*\big) \\
&= -\frac{1}{2}\int_\Omega (u_h - u_h^*)\big(|\nabla y_h(u_h)|^2 - |\nabla y_h(u_h^*)|^2\big) dx + \alpha \|u_h - u_h^*\|^2 \\
&\le \frac{1}{2}\|u_h - u_h^*\|_{L^\infty(\Omega)} \cdot \big(\|\nabla y_h(u_h)\| + \|\nabla y_h(u_h^*)\|\big) \cdot \|\nabla y_h(u_h) - \nabla y_h(u_h^*)\| \\
&\quad + \alpha \|u_h - u_h^*\|^2.
\end{aligned} \tag{4.10}$$

From the variational form (4.6), it follows that

$$\begin{aligned}
0 &= (u_h \nabla y_h(u_h), \nabla \phi_h) - (u_h^* \nabla y_h(u_h^*), \nabla \phi_h) \\
&= (u_h^* \nabla y_h(u_h) - u_h^* \nabla y_h(u_h^*), \nabla \phi_h) + (u_h \nabla y_h(u_h), \nabla \phi_h) - (u_h^* \nabla y_h(u_h), \nabla \phi_h),
\end{aligned}$$

holds for any $\phi_h \in V_h \cap H_0^1(\Omega)$, then

$$\int_\Omega u_h^* \big(\nabla y_h(u_h) - \nabla y_h(u_h^*)\big) \cdot \nabla \phi_h \, dx = -\int_\Omega (u_h - u_h^*) \nabla y_h(u_h) \cdot \nabla \phi_h \, dx.$$

Taking $\phi_h = y_h(u_h) - y_h(u_h^*)$, we have

$$\begin{aligned}
&\int_\Omega u_h^* \big(\nabla y_h(u_h) - \nabla y_h(u_h^*)\big) \cdot \big(\nabla y_h(u_h) - \nabla y_h(u_h^*)\big) \, dx \\
&= -\int_\Omega (u_h - u_h^*) \nabla y_h(u_h) \cdot \big(\nabla y_h(u_h) - \nabla y_h(u_h^*)\big) \, dx,
\end{aligned}$$

which implies that

$$a_0 \|\nabla y_h(u_h) - \nabla y_h(u_h^*)\|^2 \le \|u_h - u_h^*\|_{L^\infty(\Omega)} \cdot \|\nabla y_h(u_h)\| \cdot \|\nabla y_h(u_h) - \nabla y_h(u_h^*)\|,$$

and then

$$\|\nabla y_h(u_h) - \nabla y_h(u_h^*)\| \leq \frac{1}{a_0} \|\nabla y_h(u_h)\| \cdot \|u_h - u_h^*\|_{L^\infty(\Omega)}. \tag{4.11}$$

Now by using (4.10), (4.11), and the equivalence property of norms in finite dimensional space V_h, there exists $C > 0$, that

$$\begin{aligned} &\big(DJ_1(u_h) - DJ_1(u_h^*), u_h - u_h^*\big) \\ &\leq \frac{1}{a_0} \big(\|\nabla y_h(u_h)\| + \|\nabla y_h(u_h^*)\|\big) \cdot \|\nabla y_h(u_h)\| \cdot \|u_h - u_h^*\|_{L^\infty(\Omega)}^2 + \alpha \|u_h - u_h^*\|^2 \\ &\leq (C + \alpha) \cdot \|u_h - u_h^*\|^2. \end{aligned} \tag{4.12}$$

Therefore, the estimate (4.8) is obtained from (4.12) and [3, Corollary 18.14]. □

The gradient of functional $J_2(\cdot)$ is as follows [29]

$$\big(DJ_2(u_h), v_h\big) = -\frac{1}{2} \int_\Omega \big(|y_h(u_h)|^2 - |y_\delta|^2\big) v_h dx + \alpha \int_\Omega u_h v_h dx, \quad \forall\, u_h,\ v_h \in V_h, \tag{4.13}$$

which obtains that

$$\begin{aligned} &\big(DJ_2(u_h) - DJ_2(v_h), u_h - v_h\big) \\ &= \int_\Omega \Big(-\frac{1}{2}|y_h(u_h)|^2 + \frac{1}{2}|y_\delta|^2 + \frac{1}{2}|y_h(v_h)|^2 - \frac{1}{2}|y_\delta|^2 \Big) (u_h - v_h) dx + \alpha \|u_h - v_h\|^2. \end{aligned}$$

Similar to the derivation of (4.10)-(4.12), we can also deduce that the functional $J_2(\cdot)$ also has a property similar to $J_1(\cdot)$.

Theorem 4.3. *The functional $J_2(\cdot)$ has a Lipschitz continuous gradient around the solution point $u_h^* \in K_h$, i.e., there exists $\kappa > 0$ such that*

$$\|DJ_2(u_h) - DJ_2(u_h^*)\| \leq \kappa \|u_h - u_h^*\|, \quad \forall\, u_h \in V_h.$$

To apply the ADMM algorithm for solving the problems (4.1) and (4.2), we reformulate (4.3) into an equivalent form

$$\begin{cases} \min\limits_{(u_h, z_h) \in V_h \times K_h} & J(u_h) + \tau \|\nabla z_h\|_{L^1(\Omega)} \\ \quad \text{s.t.} & u_h = z_h. \end{cases} \tag{4.14}$$

The problem (4.14) is in the form of (1.1) with $A = I$, $B = -I$ and $b = 0$, and it satisfies Assumptions 1-3, which are verified in Theorems 4.1-4.3. The corresponding ADMM for solving problem (4.3) is as follows

$$\begin{cases} u_h^{k+1} \approx \arg \min\limits_{u_h \in V_h} \Big\{ J(u_h) - (\lambda_h^k, u_h - z_h^k) + \dfrac{\beta}{2} \|u_h - z_h^k\|^2 \Big\}, \\ z_h^{k+1} = \arg \min\limits_{z_h \in K_h} \Big\{ \tau \|\nabla z_h\|_{L^1(\Omega)} - (\lambda_h^k, u_h^{k+1} - z_h) + \dfrac{\beta}{2} \|u_h^{k+1} - z_h\|^2 \Big\}, \\ \lambda_h^{k+1} = \lambda_h^k - \beta(u_h^{k+1} - z_h^{k+1}). \end{cases} \tag{4.15}$$

4.2 A nonlinear CG method for the u_h-subproblem

We reformulate the first subproblem of (4.15) in Euclidean space for the convenience of calculation, that is

$$\min_{\boldsymbol{u} \in \mathbb{R}^N} L_k(\boldsymbol{u}) := J(\boldsymbol{u}) - (\boldsymbol{\lambda}^k, \boldsymbol{u} - \boldsymbol{z}^k)_M + \frac{\beta}{2} \|\boldsymbol{u} - \boldsymbol{z}^k\|_M^2,$$

where M denotes the mass matrix as $M_{i,j} = (\phi_h^j, \phi_h^i)$, and $\boldsymbol{u} := (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N)^\top$ is the coefficients of $u_h = \sum_{i=1}^N \boldsymbol{u}_i \phi_h^i$, where $\{\phi_h^i\}_{i=1}^N$ are the finite element basis functions in V_h. The same setting is also applied to z_h and λ_h, with the coefficients $\boldsymbol{z}$, $\boldsymbol{\lambda} \in \mathbb{R}^N$, respectively.

Conceptually, we can utilize nonlinear CG algorithms (such as FR-PR type, see e.g. [43]) to solve the minimum of $L_k(\boldsymbol{u})$ with the new inexactness criterion (3.1) in the form of

$$\|\boldsymbol{g}^{k_m+1}\|_2^2 \le \begin{cases} \sigma^2\|\boldsymbol{g}^k\|_2^2, & k \le K_0, \\ \varepsilon\bar{\sigma}^k \sum_{i=1}^{k} \Big[\beta\|M(\boldsymbol{z}^{i-1}-\boldsymbol{z}^i)\|_2^2 + \frac{1}{\beta}\|M(\boldsymbol{\lambda}^{i-1}-\boldsymbol{\lambda}^i)\|_2^2\Big], & k \ge K_0+1, \end{cases} \tag{4.16}$$

where k_m be the counter of the internal iteration, $\boldsymbol{g}^k := DL_k(\boldsymbol{u}^k)$, and $\boldsymbol{g}^{k_m+1} := DL_k(\boldsymbol{u}^{k_m+1})$. We set $\boldsymbol{u}^{k+1} = \boldsymbol{u}^{k_m+1}$ if (4.16) is satisfied.

However, two crucial steps should be further addressed, including the computations of the gradient $DJ(u)$ and the step sizes, which are discussed in the next subsections.

4.2.1 Computation of $DJ(u)$

The gradient of $J(\boldsymbol{u})$, denoted as $J'(\boldsymbol{u}^{k_m})$ at $\boldsymbol{u}^{k_m}$, is in the form of

$$\big(J'(\boldsymbol{u}^{k_m})\big)_i = \big(DJ(u_h^{k_m}), \phi_h^i\big), \quad i = 1, \cdots, N.$$

- **Case** $J(u_h) = J_1(u_h)$ **in** (4.4). We first compute $\boldsymbol{y}^{k_m}$ from (4.6) via solving

$$A_1^{k_m}\boldsymbol{y}^{k_m} = \boldsymbol{f}, \text{ with } (A_1^{k_m})_{i,j} = (u_h^{k_m}\nabla\phi_h^j, \nabla\phi_h^i), \tag{4.17}$$

  where $\boldsymbol{f}_i = (f, \phi_h^i)$. Then we substitute $y_h(u_h^{k_m}) = \sum_{j=1}^N \boldsymbol{y}_j^{k_m}\phi_h^j$ into (4.9) and compute $J_1'(\boldsymbol{u}^{k_m})$ by

$$\big(J_1'(\boldsymbol{u}^{k_m})\big)_i = \Big(-\frac{1}{2}|\nabla y_h(u_h^{k_m})|^2 + \frac{1}{2}|\nabla y_\delta|^2 + \alpha u_h^{k_m}, \phi_h^i\Big).$$

- **Case** $J(u_h) = J_2(u_h)$ **in** (4.5). Similarly, we can evaluate $J_2'(\boldsymbol{u}^{k_m})$ from (4.13) by

$$\big(\hat{J}_2'(\boldsymbol{u}^{k_m})\big)_i = \Big(-\frac{1}{2}|y_h(u_h^{k_m})|^2 + \frac{1}{2}|y_\delta|^2 + \alpha u_h^{k_m}, \phi_h^i\Big),$$

  where $y_h(u_h^{k_m})$ satisfies the variational form (4.7), and its coefficients $\boldsymbol{y}^{k_m}$ is computed by solving

$$A_2^{k_m}\boldsymbol{y}^{k_m} = \boldsymbol{f}, \text{ with } (A_2^{k_m})_{i,j} = (\nabla\phi_h^j, \nabla\phi_h^i) + (u_h^{k_m}\phi_h^j, \phi_h^i). \tag{4.18}$$

4.2.2 Computation of the step size ρ^{k_m}

The other crucial step in the nonlinear CG method is to obtain the step size ρ^{k_m} satisfying

$$L_k(\boldsymbol{u}^{k_m} - \rho^{k_m}\boldsymbol{w}^{k_m}) \le L_k(\boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m}), \quad \forall\, \rho \in \mathbb{R}, \tag{4.19}$$

where $\boldsymbol{w}^{k_m}$ refers to the conjugate direction. As mentioned in [46], it is computationally demanding to implement the line search which requires multiple computations of the objective function values and/or the corresponding gradients. To reduce the computational cost, instead of applying some known line search strategies (see e.g., [43]), we choose an approximate approach similar as in [24]. It first approximates the scale function $H_{k_m}(\rho)$ defined as

$$H_{k_m}(\rho) := L_k(\boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m}),$$

by linear expansion of the solution mappings determined by (4.6) and (4.7), then obtains an approximate step size by an explicit formula. However, it seems that the approach by linear expansion in [24] is not suitable for our problems. To address this issue, we need to first reformulate the functionals in (4.4) and (4.5), then approximate the solution mappings by quadratic expansions.

(1) Case $J(u_h) = J_1(u_h)$ in (4.4). We have from (4.6) the identity (see e.g. [46]) as follows

$$\begin{aligned}&\frac{1}{2}\int_\Omega u_h|\nabla y_h(u_h) - \nabla y_\delta|^2 dx \\ &= \frac{1}{2}\int_\Omega (u_h|\nabla y_\delta|^2 - 2fy_\delta)dx - \frac{1}{2}\int_\Omega (u_h|\nabla y_h(u_h)|^2 - 2fy_h(u_h))dx,\end{aligned}$$

then the functional $J_1(u_h)$ in (4.4) can be rewritten as

$$J_1(u_h) = \frac{1}{2}\int_\Omega (u_h|\nabla y_\delta|^2 - 2fy_\delta)dx + \frac{1}{2}\int_\Omega fy_h(u_h)dx + \frac{\alpha}{2}\|u_h\|^2. \tag{4.20}$$

Let $y_h(u_h) = S_1(\boldsymbol{u})$ with $S_1(\cdot)$ being the solution mapping defined by (4.6), the quadratic approximation of $\rho \to S_1(\boldsymbol{u}^{k_m} - \rho \boldsymbol{w}^{k_m})$ at $\rho = 0$ is

$$S_1(\boldsymbol{u}^{k_m}) - \rho S_1'(\boldsymbol{u}^{k_m})\boldsymbol{w}^{k_m} + \frac{\rho^2}{2} S_1''(\boldsymbol{u}^{k_m})(\boldsymbol{w}^{k_m})^2. \tag{4.21}$$

We define

$$p_h^{k_m} := S_1'(\boldsymbol{u}^{k_m})\boldsymbol{w}^{k_m} \text{ and } q_h^{k_m} := S_1''(\boldsymbol{u}^{k_m})(\boldsymbol{w}^{k_m})^2,$$

then $p_h^{k_m}$ and $q_h^{k_m}$ satisfy

$$(u_h^{k_m}\nabla p_h^{k_m}, \nabla \phi_h) = -(w_h^{k_m}\nabla y_h^{k_m}, \nabla \phi_h), \ \ \forall\, \phi_h \in V_h \cap H_0^1(\Omega), \tag{4.22}$$

$$(u_h^{k_m}\nabla q_h^{k_m}, \nabla \phi_h) = -2(w_h^{k_m}\nabla p_h^{k_m}, \nabla \phi_h), \ \ \forall\, \phi_h \in V_h \cap H_0^1(\Omega). \tag{4.23}$$

Replacing $y_h(u_h^{k_m})$ in (4.20) by (4.21), we thus obtain an approximation of $H_{k_m}(\rho)$ by $Q_{k_m}(\rho)$ as follows

$$\begin{aligned}Q_{k_m}(\rho) =& \frac{1}{2}\int_\Omega u_h^{k_m}|\nabla y_\delta|^2 dx - \frac{\rho}{2}\int_\Omega w_h^{k_m}|\nabla y_\delta|^2 dx - \int_\Omega fy_\delta dx + \frac{1}{2}\int_\Omega fy_h^{k_m} dx \\ &- \frac{\rho}{2}\int_\Omega fp_h^{k_m} dx + \frac{\rho^2}{4}\int_\Omega fq_h^{k_m} dx + \frac{\alpha}{2}\|\boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m}\|_M^2 \\ &- (\boldsymbol{\lambda}^k, \boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m} - \boldsymbol{z}^k)_M + \frac{\beta}{2}\|\boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m} - \boldsymbol{z}^k\|_M^2.\end{aligned}$$

Let $Q_{k_m}'(\rho^{k_m}) = 0$, we get

$$\rho^{k_m} = \frac{(\boldsymbol{w}^{k_m})^\top \boldsymbol{g}^{k_m}}{(\alpha+\beta)\|\boldsymbol{w}^{k_m}\|_M^2 + \frac{1}{2}\int_\Omega fq_h^{k_m} dx}. \tag{4.24}$$

Together with (4.6), (4.22) and (4.23), $\int_\Omega fq_h^{k_m} dx$ in (4.24) can be calculated directly from $p_h^{k_m}$ without solving $q_h^{k_m}$ in (4.23), and

$$\int_\Omega fq_h^{k_m} dx = (u_h^{k_m}\nabla y_h^{k_m}, \nabla q_h^{k_m}) = -2(w_h^{k_m}\nabla p_h^{k_m}, \nabla y_h^{k_m}) = 2(u_h^{k_m}\nabla p_h^{k_m}, \nabla p_h^{k_m}),$$

where $p_h^{k_m}$ sloves (4.22).

(2) Case $J(u_h) = J_2(u_h)$ in (4.5). We have from (4.7) that $J_2(u_h)$ can be reformulated as

$$J_2(u_h) = \frac{1}{2}\int_\Omega (|\nabla y_\delta|^2 + u_h|y_\delta|^2 - 2fy_\delta)dx + \frac{1}{2}\int_\Omega fy_h(u_h)dx + \frac{\alpha}{2}\|u_h\|^2. \tag{4.25}$$

We also define the solution mapping of (4.7) as $y_h(u_h^{k_m}) = S_2(\boldsymbol{u}^{k_m})$, and substitute the quadratic expansion of $\rho \to S_2(\boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m})$ at $\rho = 0$ in the form of

$$S_2(\boldsymbol{u}^{k_m}) - \rho S_2'(\boldsymbol{u}^{k_m})\boldsymbol{w}^{k_m} + \frac{\rho^2}{2} S_2''(\boldsymbol{u}^{k_m})(\boldsymbol{w}^{k_m})^2$$

into (4.25), which leads to an approximate function $Q_{k_m}(\rho)$ of $H_{k_m}(\rho)$ as follows

$$\begin{aligned}Q_{k_m}(\rho) = &\frac{1}{2}\int_\Omega u_h^{k_m}|\nabla y_\delta|^2 dx - \frac{\rho}{2}\int_\Omega w_h^{k_m}|\nabla y_\delta|^2 dx - \int_\Omega f y_\delta dx + \frac{1}{2}\int_\Omega f y_h^{k_m} dx \\ &+ \frac{1}{2}\int_\Omega |\nabla y_\delta|^2 dx - \frac{\rho}{2}\int_\Omega f p_h^{k_m} dx + \frac{\rho^2}{4}\int_\Omega f q_h^{k_m} dx + \frac{\alpha}{2}\|\boldsymbol{u}^{k_m} - \rho \boldsymbol{w}^{k_m}\|^2 \\ &- (\boldsymbol{\lambda}^k, \boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m} - \boldsymbol{z}^k)_M + \frac{\beta}{2}\|\boldsymbol{u}^{k_m} - \rho\boldsymbol{w}^{k_m} - \boldsymbol{z}^k\|_M^2,\end{aligned}$$

where

$$p_h^{k_m} = S_2'(\boldsymbol{u}^{k_m})\boldsymbol{w}^{k_m} \text{ and } q_h^{k_m} = S_2''(\boldsymbol{u}^{k_m})(\boldsymbol{w}^{k_m})^2,$$

with $p_h^{k_m}$ and $q_h^{k_m}$ satisfying

$$(\nabla p_h^{k_m}, \nabla \phi_h) + (u_h^{k_m} p_h^{k_m}, \phi_h) = -(w_h^{k_m} y_h^{k_m}, \phi_h), \ \ \forall \, \phi_h \in V_h \cap H_0^1(\Omega), \tag{4.26}$$

$$(\nabla q_h^{k_m}, \nabla \phi_h) + (u_h^{k_m} q_h^{k_m}, \phi_h) = -2(w_h^{k_m} p_h^{k_m}, \phi_h), \ \ \forall \, \phi_h \in V_h \cap H_0^1(\Omega). \tag{4.27}$$

Then the root of $Q'_{k_m}(\rho^{k_m}) = 0$ is given by

$$\rho^{k_m} = \frac{(\boldsymbol{w}^{k_m})^\top \boldsymbol{g}^{k_m}}{(\alpha+\beta)\|\boldsymbol{w}^{k_m}\|_M^2 + \frac{1}{2}\int_\Omega f q_h^{k_m} dx},$$

where $\int_\Omega f q_h^{k_m} dx$ can be evaluated without solving $q_h^{k_m}$ in (4.27) by

$$\int_\Omega f q_h^{k_m} dx = 2(\nabla p_h^{k_m}, \nabla p_h^{k_m}) + 2(u_h^{k_m} p_h^{k_m}, p_h^{k_m}),$$

which is reformulated by using (4.7), (4.26) and (4.27).

4.3 An inexact ADMM numerical approach for (4.1) and (4.2)

The z_h-subproblem in (4.15) can be solved in V_h firstly by an deep convolutional neural network (CNN) solver in [46, Subroutine 4] which originates from [53], and then project it onto K_h.

With all of the discussions in previous subsections, we are ready to present an inexact ADMM numerical approach for solving the regularized inverse problems (4.1) and (4.2). The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 An Inexact ADMM numerical approach for (4.1) and (4.2).

Input $(\boldsymbol{u}^0, \boldsymbol{z}^0, \boldsymbol{\lambda}^0)$, $\beta > 0$, $\varepsilon > 0$, σ satisfies (2.4), $\bar{\sigma} \in (0,1)$, and $K_0 \geq 0$.
while $k \geq 0$ **do**
  $(\boldsymbol{u}^k, \boldsymbol{z}^k, \boldsymbol{\lambda}^k) \to (\boldsymbol{u}^{k+1}, \boldsymbol{z}^{k+1}, \boldsymbol{\lambda}^{k+1})$ via:
  Compute $\boldsymbol{u}^{k+1}$ by the nonlinear CG method with the inexactness criterion (4.16) in subsection 4.2;
  Compute $\tilde{\boldsymbol{z}}^{k+1}$ by the CNN solver (Subroutine 4 in [46]), and obtain $\boldsymbol{z}^{k+1} = \mathcal{P}_{K_h}(\tilde{\boldsymbol{z}}^{k+1}) = \max\{a_0, \min\{\tilde{\boldsymbol{z}}^{k+1}, a_1\}\}$;
  Update the Lagrange multiplier by $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k - \beta(\boldsymbol{u}^{k+1} - \boldsymbol{z}^{k+1})$.
end while

5 Numerical results

In this section, we report some numerical results to verify the efficiency of our new inexact ADMM numerical approach applied to the regularized inverse problems (4.1) and (4.2) for recovering the discontinuous diffusion and potential coefficients, respectively. The following numerical experiments are performed on a laptop equipped with Windows 11 operation system, Intel(R) Core(TM) i7-8750H CPU (2.20GHz), and 16 GB of memory.

For the z_h-subproblem, we use the same network architecture and training dataset as [53] (see https://github.com/cszn/DnCNN) to train our networks (total 25 CNN networks $\mathcal{C}_\zeta$ with $\zeta = 1, 2, \cdots, 25$), and the settings of the deep CNN solver are the same as in [46]. We also follow the approach in [46] to adjust the parameters (i.e. $\beta\zeta \sim \delta$) satisfying the Morozov's discrepancy principle (see, e.g. [18]).

Example 5.1. *Let $a_0 = 0.1$, $a_1 = 5$, and the exact diffusion coefficient $u^\dagger(x_1, x_2)$ over $\Omega = (0,1)^2$ as*

$$u^\dagger(x_1, x_2) = \begin{cases} 1, & x_2 \in [0, 0.5], \\ 2, & x_2 \in (0.5, 1], \end{cases} \tag{5.1}$$

with which the noisy observation data ∇y_δ in (4.1) is constructed as $\nabla y_\delta(x) = \nabla y_h + \delta\|\nabla y_h\| rand(x)$ for the regularized inverse problem (4.1) for recovering the diffusion coefficient, where y_h is the finite element solution to (4.6) with $f = 10$, $rand(x)$ is a uniformly distributed vector-valued function in $[-1, 1]$, and $\delta > 0$ is the noise level.

Table 1: Parameters for Example 5.1.

δ	β	ζ	$(\beta\zeta)/\delta$	α
0.01	0.1	9	90	10^{-4}
0.05	0.5	9	90	10^{-4}
0.1	0.6	15	90	10^{-4}

The parameters δ, β, α and ζ are chosen as in Table 1 for performing the numerical computations. The initial values for Algorithm 1 are set as $\boldsymbol{u}^0 = \boldsymbol{z}^0 = \mathbf{1}$, and $\boldsymbol{\lambda}^0 = \mathbf{0}$. The inexactness criterion (4.16) in Algorithm 1 is denoted as $\mathsf{criterion}_{new}$, which is the specific form of (3.1) for the regularized inverse problem (4.14), and the inexactness criterion in (2.3) for the problem (4.14) is named as $\mathsf{criterion}_{old}$. In (3.1), we set $\sigma = 0.999\sqrt{2\alpha}/(\sqrt{2\alpha} + \sqrt{\beta})$, and $K_0 = 5$. In Table 2, the relative L^2 errors, inner CG iteration numbers and total computational times of the first 20 iterations by Algorithm 1 are reported for the comparison with the inexact ADMM under the original inexactness criterion (2.3). In Figure 1, we plot the curves of L^2 errors $\|u_h^k - \hat{u}_h\|$ with $h = 2^{-8}$ and $\delta = 0.01$ by the inexact ADMM with two inexactness criteria, which verifies the theoretical linear convergence result derived in section 3. The reference solution $\hat{u}_h$ is approximated by the 40-th outer iterate of exactly-solved ADMM (the internal stop condition is $\|e_k(u^{k+1})\| \leq 10^{-10}$), which helps to eliminate the influence of the finite element and regularization errors. In addition, the numerical solution u_h^{20} and the error $u_h^{20} - u_h^\dagger$ after 20 outer iterations with $h = 2^{-8}$ are displayed in Figure 2.

Example 5.2. *Let $\Omega = (0,1)^2$, $f = 10$, $a_0 = 0.1$, and $a_1 = 5$. The noisy observation data y_δ in the regularized inverse problem (4.2) for recovering the potential coefficient is set as $y_\delta(x) = y_h + \delta\|y_h\| rand(x)$, where $rand(x)$ is a uniformly distributed vector-valued function in $[-1, 1]$, $\delta > 0$ is the noise level, and y_h is the finite element solution of (4.7) with the precise coefficient $u^\dagger(x_1, x_2)$ as*

$$u^\dagger(x_1, x_2) = 1 + I_{\Omega_1}, \quad \Omega_1 = \{(x_1, x_2) \mid (x_1 - 0.5)^2 + (x_2 - 0.5)^2 \leq 1/8\}, \tag{5.2}$$

where I_{Ω_1} denotes the characteristic function over Ω_1.

The parameters δ, β, α and ζ are set in Table 3 for Example 5.2, and the initial values of the inexact ADMM is chosen as $\boldsymbol{u}^0 = \boldsymbol{z}^0 = \mathbf{1}$, and $\boldsymbol{\lambda}^0 = \mathbf{0}$. In (3.1), σ and K_0 are chosen the same ones as in the previous example. The relative L^2 errors, inner CG iteration numbers and total computational times of the first 20 iterations by Algorithm 1 as well as the algorithm with the inexactness criterion (2.3) are

Table 2: Numerical comparison of algorithms for Example 5.1.

Mesh(h)	δ	Criterion	$\|u_h^{20} - u_h^\dagger\|/\|u_h^\dagger\|$	CG Nos	Time(s)	$\bar{\sigma}/\varepsilon$
2^{-5}	0.01	criterion$_{old}$	8.7383×10^{-3}	349	3.86	$\sim$
		criterion$_{new}$	8.6635×10^{-3}	171	2.08	$0.59/10^{-2}$
	0.05	criterion$_{old}$	1.5085×10^{-2}	213	2.54	$\sim$
		criterion$_{new}$	1.5086×10^{-2}	180	2.08	$0.52/10^{-2}$
	0.1	criterion$_{old}$	4.4639×10^{-2}	205	2.39	$\sim$
		criterion$_{new}$	4.4636×10^{-2}	167	1.97	$0.52/10^{-2}$
2^{-6}	0.01	criterion$_{old}$	6.5952×10^{-3}	386	12.47	$\sim$
		criterion$_{new}$	6.5952×10^{-3}	260	8.67	$0.62/10^{-2}$
	0.05	criterion$_{old}$	1.0930×10^{-2}	216	7.47	$\sim$
		criterion$_{new}$	1.0932×10^{-2}	208	7.15	$0.52/10^{-2}$
	0.1	criterion$_{old}$	4.0888×10^{-2}	209	7.34	$\sim$
		criterion$_{new}$	4.0881×10^{-2}	190	6.78	$0.52/10^{-2}$
2^{-7}	0.01	criterion$_{old}$	5.3821×10^{-3}	412	52.17	$\sim$
		criterion$_{new}$	5.3821×10^{-3}	316	40.52	$0.59/10^{-2}$
	0.05	criterion$_{old}$	1.6478×10^{-2}	202	26.55	$\sim$
		criterion$_{new}$	1.6469×10^{-2}	183	24.42	$0.59/10^{-2}$
	0.1	criterion$_{old}$	4.0566×10^{-2}	192	25.85	$\sim$
		criterion$_{new}$	4.0565×10^{-2}	164	22.32	$0.59/10^{-2}$
2^{-8}	0.01	criterion$_{old}$	5.1069×10^{-3}	383	219.31	$\sim$
		criterion$_{new}$	5.1021×10^{-3}	387	222.13	$0.54/10^{-2}$
	0.05	criterion$_{old}$	3.0223×10^{-2}	199	118.90	$\sim$
		criterion$_{new}$	3.0204×10^{-2}	188	114.67	$0.57/10^{-2}$
	0.1	criterion$_{old}$	4.5675×10^{-2}	185	112.17	$\sim$
		criterion$_{new}$	4.5679×10^{-2}	177	107.93	$0.57/10^{-2}$

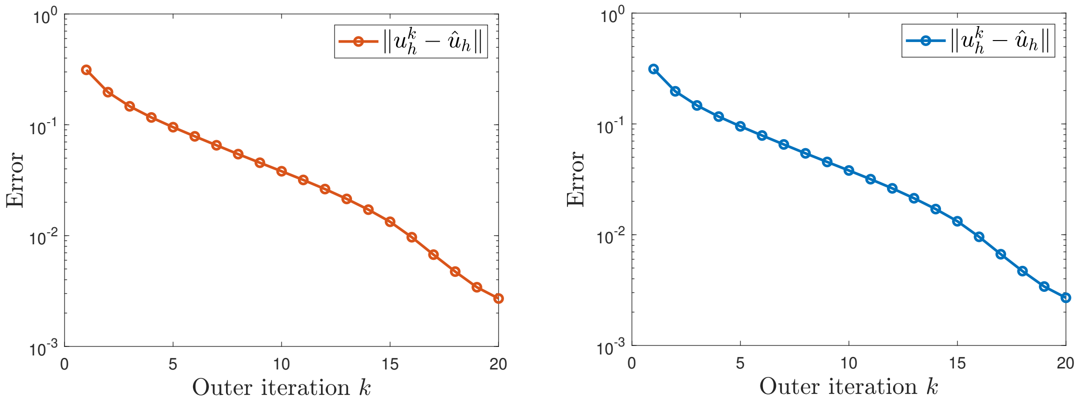

Figure 1: L^2 errors by the inexact ADMM equipped with criterion$_{new}$ (left) and criterion$_{old}$ (right) with respect to outer ADMM iterations for Example 5.1 with $\delta = 0.01$ and $h = 2^{-8}$.

reported in Table 4. In Figure 3, the plots of L^2 errors $\|u_h^k - \hat{u}_h\|$ by the inexact ADMM with two different inexactness criteria are shown for the case $h = 2^{-8}$ and $\delta = 0.001$. Similar to the previous example, we also use the 40-th outer iterate of exactly-solved ADMM as the reference solution $\hat{u}_h$ for comparison. In Figure 4, it presents the graphs of the numerical solutions u_h and the error $u_h - u_h^\dagger$ after 20 outer iterations with $h = 2^{-8}$.

From the numerical results of these examples, we can observe that the convergence speed and com-

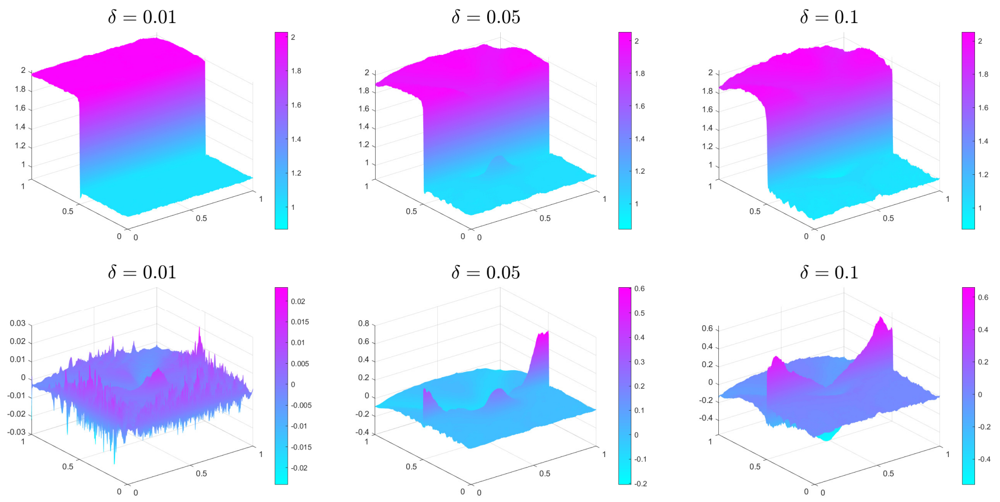

Figure 2: Numerical solutions u_h and errors $u_h - u_h^\dagger$ after 20 iterations for Example 5.1 with $h = 2^{-8}$. Columns 1-3 are the recovered results for $\delta = 0.01$, 0.05, and 0.1, respectively.

Table 3: Parameters for Example 5.2.

δ	β	ζ	$(\beta\zeta)/\delta$	α
0.001	0.001	9	9	10^{-6}
0.01	0.0075	12	9	10^{-6}
0.1	0.05	18	9	10^{-6}

putational cost of our method in Algorithm 1 are basically similar as those of the inexact ADMM with the inexactness criterion (2.3) proposed in [23]. Our method in Algorithm 1 has the theoretical guarantee of linear convergence established in section 3, while the theoretical linear convergence of the inexact ADMM with (2.3) is still an open problem.

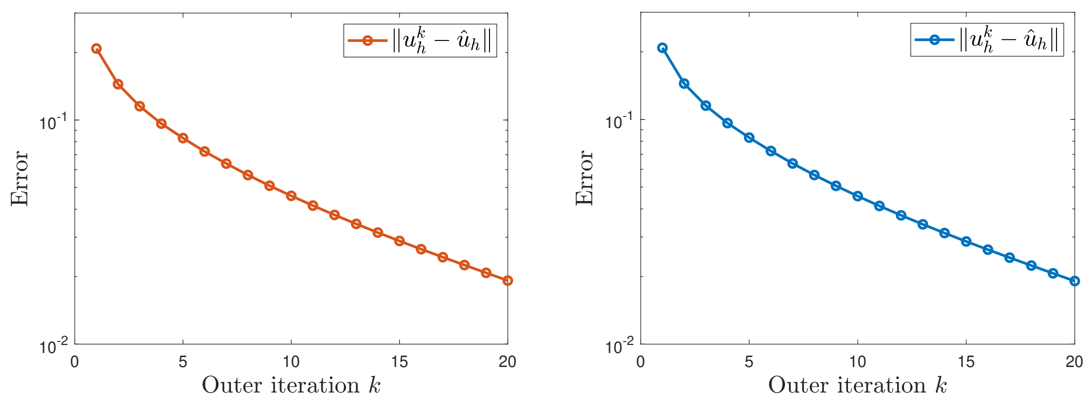

Figure 3: L^2 errors by the inexact ADMM equipped with criterion$_{new}$ (left) and criterion$_{old}$ (right) with respect to outer ADMM iterations for Example 5.2 with $\delta = 0.001$ and $h = 2^{-8}$.

6 Conclusions

In this paper, we discuss inexact ADMM algorithms for a general two block convex optimization problem arising from inverse problems, where the first functional is strongly convex with continuous gradient. We

Table 4: Numerical comparison of algorithms for Example 5.2.

Mesh(h)	δ	Criterion	$\lVert u_h^{20} - u_h^{\dagger}\rVert / \lVert u_h^{\dagger}\rVert$	CG Nos	Time(s)	$\bar{\sigma}/\varepsilon$
2^{-5}	0.001	criterion$_{old}$	5.3451×10^{-2}	141	2.34	$\sim$
		criterion$_{new}$	5.3396×10^{-2}	110	1.56	$0.60/10^{-4}$
	0.01	criterion$_{old}$	1.1834×10^{-1}	100	1.44	$\sim$
		criterion$_{new}$	1.1780×10^{-1}	73	1.13	$0.70/10^{-4}$
	0.1	criterion$_{old}$	1.8683×10^{-1}	83	1.32	$\sim$
		criterion$_{new}$	1.8586×10^{-1}	55	1.02	$0.70/10^{-4}$
2^{-6}	0.001	criterion$_{old}$	6.2381×10^{-2}	132	5.40	$\sim$
		criterion$_{new}$	6.2372×10^{-2}	109	4.92	$0.60/10^{-4}$
	0.01	criterion$_{old}$	9.9454×10^{-2}	79	3.86	$\sim$
		criterion$_{new}$	9.9405×10^{-2}	62	3.22	$0.70/10^{-4}$
	0.1	criterion$_{old}$	1.7438×10^{-1}	74	3.72	$\sim$
		criterion$_{new}$	1.7327×10^{-1}	51	2.90	$0.70/10^{-4}$
2^{-7}	0.001	criterion$_{old}$	6.1630×10^{-2}	102	16.64	$\sim$
		criterion$_{new}$	6.1667×10^{-2}	100	16.52	$0.57/10^{-4}$
	0.01	criterion$_{old}$	1.0708×10^{-1}	76	13.36	$\sim$
		criterion$_{new}$	1.0682×10^{-1}	58	10.90	$0.70/10^{-4}$
	0.1	criterion$_{old}$	1.6494×10^{-1}	62	11.82	$\sim$
		criterion$_{new}$	1.6472×10^{-1}	49	9.70	$0.70/10^{-4}$
2^{-8}	0.001	criterion$_{old}$	6.0424×10^{-2}	87	69.31	$\sim$
		criterion$_{new}$	6.0475×10^{-2}	86	68.14	$0.57/10^{-4}$
	0.01	criterion$_{old}$	9.9142×10^{-2}	73	58.34	$\sim$
		criterion$_{new}$	9.9084×10^{-2}	57	48.58	$0.70/10^{-4}$
	0.1	criterion$_{old}$	1.4906×10^{-1}	61	53.21	$\sim$
		criterion$_{new}$	1.4877×10^{-1}	48	43.00	$0.70/10^{-4}$

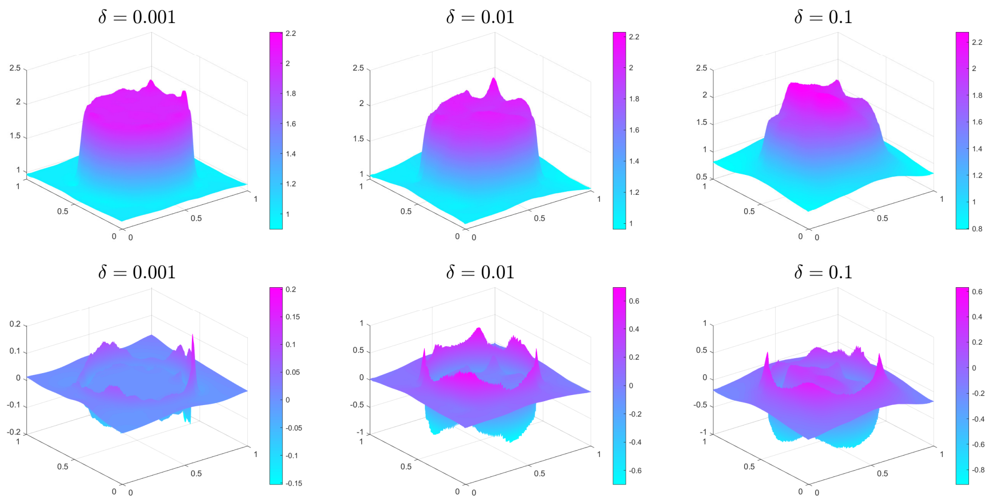

Figure 4: Numerical solutions u_h and errors $u_h - u_h^{\dagger}$ after 20 iterations for Example 5.2 with $h = 2^{-8}$. Columns 1-3 are the recovered results for $\delta = 0.001, 0.01$, and 0.1, respectively.

first prove some strongly convergence properties for the inexact ADMM with the inexactness criterion (2.3) proposed in [23, 52], and also obtain a worst-case $O(1/K)$ convergence rate of the algorithm in both ergodic and non-ergodic senses. However, it seems difficult to prove the linear convergence of the inexact ADMM with (2.3). To address this issue, we propose a new inexactness criterion such that $\{\|e_k(u^{k+1})\|^2\}$ converges linearly to 0. Then we further prove that the resulting inexact ADMM has a linear convergence rate. After the theoretical part, we consider the application of the proposed inexact ADMM to two nonlinear inverse problems of elliptic equations and present the detailed implementations. Numerical results illustrate the efficiency and theoretical results of our algorithm.

Acknowledgments

The authors are very grateful to the anonymous referees for careful reading of the original manuscript and their valuable suggestions.

Data Availability Enquiries about data availability should be directed to the authors.

Declarations

Competing interests The authors have not disclosed any competing interests.

References

[1] Acar, R., Vogel, C.R.: Analysis of bounded variation penalty methods for ill-posed problems. Inverse Problems **10**(6), 1217–1229 (1994).

[2] Attouch, H., Buttazzo, G., Michaille, G.: Variational analysis in Sobolev and BV spaces: Applications to PDEs and Optimization. SIAM/MPS, Philadelphia, PA (2006).

[3] Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces. Springer, New York (2011).

[4] Boley, D.: Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs. SIAM J. Optim. **23**(4), 2183–2207 (2013).

[5] 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. Learning **3**(1), 1–122 (2010).

[6] Casas, E., Kunisch, K., Pola, C.: Regularization by functions of bounded variation and applications to image enhancement. Appl. Math. Optim. **40**(2), 229–257 (1999).

[7] Chan, T.F., Tai, X.: Level set and total variation regularization for elliptic inverse problems with discontinuous coefficients. J. Comput. Phys. **193**(1), 40–66 (2004).

[8] Chavent, G., Kunisch, K.: Regularization of linear least squares problems by total bounded variation. ESAIM Control Optim. Calc. Var. **2**, 359–376 (1997).

[9] Chen, C., Ng, M.K., Zhao, X.: Alternating direction method of multipliers for nonlinear image restoration problems. IEEE Trans. Image Process. **24**(1), 33–43 (2015).

[10] Chen, Z., Zou, J.: An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems. SIAM J. Control Optim. **37**(3), 892–910 (1999).

[11] Deng, W., Yin, W.: On the global and linear convergence of the generalized alternating direction method of multipliers. J. Sci. Comput. **66**(3), 889–916 (2016).

[12] Douglas, J., Jr., Rachford, H.H., Jr.: On the numerical solution of heat conduction problems in two and three space variables. Trans. Amer. Math. Soc. **82**(2), 421–439 (1956).

[13] Du, Y., Zhang, W., Zhang, T.: ADMM based distributed state estimation for integrated energy system. CSEE Journal of Power and Energy Systems **5**(2), 275–283 (2019).

[14] Eckstein, J., Bertsekas, D.P.: An alternating direction method for linear programming. Tech report (1990).

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

[16] Eckstein, J., Yao, W.: Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives. Pac. J. Optim. **11**(4), 619–644 (2015).

[17] Eckstein, J., Yao, W.: Relative-error approximate versions of Douglas-Rachford splitting and special cases of the ADMM. Math. Program. **170**(2, Ser. A), 417–444 (2018).

[18] Engl, H.W., Hanke, M., Neubauer, A.: Regularization of inverse problems. Kluwer Academic Publishers Group, Dordrecht (1996).

[19] Fortin, M., Glowinski, R.: Augmented Lagrangian methods: Applications to the numerical solution of boundary value problems. North-Holland Publishing Co., Amsterdam (1983).

[20] Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Comput. Math. Appl. **2**(1), 17–40 (1976).

[21] Glowinski, R.: On alternating direction methods of multipliers: A historical perspective. Fitzgibbon, W., Kuznetsov, Y.A., Neittaanmäki, P., Pironneau, O. (Editors), Modeling, Simulation and Optimization for Science and Technology, (59–82). Springer Netherlands (2014).

[22] Glowinski, R., Marrocco, A.: Sur l'approximation par éléments finis et la résolution par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. R.A.I.R.O. Anal. Numér. **9**(R-2), 41–76 (1975).

[23] Glowinski, R., Song, Y., Yuan, X., Yue, H.: Application of the alternating direction method of multipliers to control constrained parabolic optimal control problems and beyond. Ann. Appl. Math. **38**(2), 115–158 (2022).

[24] Glowinski, R., Song, Y., Yuan, X., Yue, H.: Bilinear optimal control of an advection-reaction-diffusion system. SIAM Rev. **64**(2), 392–421 (2022).

[25] Goldstein, T., O'Donoghue, B., Setzer, S., Baraniuk, R.: Fast alternating direction optimization methods. SIAM J. Imaging Sci. **7**(3), 1588–1623 (2014).

[26] Gutman, S.: Identification of discontinuous parameters in flow equations. SIAM J. Control Optim. **28**(5), 1049–1060 (1990).

[27] Hager, W.W., Zhang, H.: Convergence rates for an inexact admm applied to separable convex optimization. Comput. Optim. Appl. **77**(3), 729–754 (2020).

[28] Hào, D.N., Quyen, T.N.T.: Convergence rates for Tikhonov regularization of coefficient identification problems in Laplace-type equations. Inverse Problems **26**(12), 125014 (23pp) (2010).

[29] Hào, D.N., Quyen, T.N.T.: Convergence rates for total variation regularization of coefficient identification problems in elliptic equations I. Inverse Problems **27**(7), 075008 (2011).

[30] Hào, D.N., Quyen, T.N.T.: Convergence rates for total variation regularization of coefficient identification problems in elliptic equations II. J. Math. Anal. Appl. **388**(1), 593–616 (2012).

[31] He, B.S., Liao, L.Z., Han, D., Yang, H.: A new inexact alternating directions method for monotone variational inequalities. Math. Program. **92**(1), 103–118 (2002).

[32] He, B.S., Yuan, X.M.: On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method. SIAM J. Numer. Anal. **50**(2), 700–709 (2012).

[33] He, B.S., Yuan, X.M.: On non-ergodic convergence rate of Douglas-Rachford alternating direction method of multipliers. Numer. Math. **130**(3), 567–577 (2015).

[34] Hestenes, M.R.: Multiplier and gradient methods. J. Optim. Theory Appl. **4**(5), 303–320 (1969).

[35] Hong, M., Luo, Z.Q.: On the linear convergence of the alternating direction method of multipliers. Math. Program. **162**, 165–199 (2017).

[36] Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(6), 964–979 (1979).

[37] Liu, Y., Yuan, X., Zeng, S., Zhang, J.: Partial error bound conditions and the linear convergence rate of the alternating direction method of multipliers. SIAM J. Numer. Anal. **56**(4), 2095–2123 (2018).

[38] Martinet, B.: Régularisation dinéquations variationnelles par approximations successives. Rev. Franç. Inform. Rech. Opér. **4**, 154–158 (1970).

[39] Monteiro, R.D.C., Svaiter, B.F.: Iteration-complexity of block-decomposition algorithms and the alternating direction method of multipliers. SIAM J. Optim. **23**(1), 475–507 (2013).

[40] Moreau, J.J.: Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. France **93**, 273–299 (1965).

[41] Nesterov, Y.: Gradient methods for minimizing composite functions. Math. Program. **140**(1), 125–161 (2013).

[42] Ng, M.K., Wang, F., Yuan, X.M.: Inexact alternating direction methods for image recovery. SIAM J. Sci. Comput. **33**(4), 1643–1668 (2011).

[43] Nocedal, J., Wright, S.J.: Numerical optimization. Springer, New York, 2nd edition (2006).

[44] Powell, M.J.D.: A method for nonlinear constraints in minimization problems. Fletcher, R. (Editor), Optimization, (283–298). Academic Press, New York (1969).

[45] Tian, W.Y., Yuan, X.M.: An alternating direction method of multipliers with a worst-case $O(1/n^2)$ convergence rate. Math. Comp. **88**(318), 1685–1713 (2019).

[46] Tian, W.Y., Yuan, X.M., Yue, H.: An ADMM-Newton-CNN numerical approach to a TV model for identifying discontinuous diffusion coefficients in elliptic equations: convex case with gradient observations. Inverse Problems **37**(8), 085004 (2021).

[47] Wang, X.F., Yuan, X.M.: The linearized alternating direction method of multipliers for Dantzig selector. SIAM J. Sci. Comput. **34**(5), A2792–A2811 (2012).

[48] Wang, Y., Yin, W., Zeng, J.: Global convergence of ADMM in nonconvex nonsmooth optimization. J. Sci. Comput. **78**, 29–63 (2019).

[49] Yang, W.H., Han, D.: Linear convergence of the alternating direction method of multipliers for a class of convex optimization problems. SIAM J. Numer. Anal. **54**(2), 625–640 (2016).

[50] Yuan, X.: The improvement with relative errors of He et al.'s inexact alternating direction method for monotone variational inequalities. Math. Comput. Modelling **42**(11-12), 1225–1236 (2005).

[51] Yuan, X., Zeng, S., Zhang, J.: Discerning the linear convergence of ADMM for structured convex optimization through the lens of variational analysis. J. Mach. Learn. Res. **21**(83), 1–75 (2020).

[52] Yue, H., Yang, Q., Wang, X., Yuan, X.: Implementing the alternating direction method of multipliers for big datasets: A case study of least absolute shrinkage and selection operator. SIAM J. Sci. Comput. **40**(5), A3121–A3156 (2018).

[53] Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. Image Process. **26**(7), 3142–3155 (2017).

[54] Zou, J.: Numerical methods for elliptic inverse problems. Int. J. Comput. Math. **70**(2), 211–232 (1998).