

Article

A Multilevel Heterogeneous ADMM Algorithm for Elliptic Optimal Control Problems with L^1 -Control Cost

Xiaotong Chen ¹, Xiaoliang Song ², Zixuan Chen ^{3,*} and Lijun Xu ¹¹ School of Science, Dalian Maritime University, Dalian 116026, China² School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China³ College of Sciences, Northeastern University, Shenyang 110819, China

* Correspondence: chenxixuan@mail.neu.edu.cn

Abstract: In this paper, elliptic optimal control problems with L^1 -control cost and box constraints on the control are considered. To numerically solve the optimal control problems, we use the *First optimize, then discretize* approach. We focus on the inexact alternating direction method of multipliers (iADMM) and employ the standard piecewise linear finite element approach to discretize the subproblems in each iteration. However, in general, solving the subproblems is expensive, especially when the discretization is at a fine level. Motivated by the efficiency of the multigrid method for solving large-scale problems, we combine the multigrid strategy with the iADMM algorithm. Instead of fixing the mesh size before the computation process, we propose the strategy of gradually refining the grid. Moreover, to overcome the difficulty whereby the L^1 -norm does not have a decoupled form, we apply nodal quadrature formulas to approximately discretize the L^1 -norm and L^2 -norm. Based on these strategies, an efficient multilevel heterogeneous ADMM (mhADMM) algorithm is proposed. The total error of the mhADMM consists of two parts: the discretization error resulting from the finite-element discretization and the iteration error resulting from solving the discretized subproblems. Both errors can be regarded as the error of inexactly solving infinite-dimensional subproblems. Thus, the mhADMM can be regarded as the iADMM in function space. Furthermore, theoretical results on the global convergence, as well as the iteration complexity results $o(1/k)$ for the mhADMM, are given. Numerical results show the efficiency of the mhADMM algorithm.



Citation: Chen, X.; Song, X.; Chen, Z.; Xu, L. A Multilevel Heterogeneous ADMM Algorithm for Elliptic Optimal Control Problems with L^1 -Control Cost. *Mathematics* **2023**, *11*, 570. <https://doi.org/10.3390/math11030570>

Keywords: optimal control problems; ADMM; sparse regularization; multilevel

MSC: 49M41; 49M25; 65K10; 65M32

Academic Editor: Ioannis G. Tsoulos

Received: 31 December 2022

Revised: 18 January 2023

Accepted: 19 January 2023

Published: 21 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sparse optimal control problems are widespread in many areas, such as the placement of actuators [1], quantum spin systems [2], etc. [3,4]. It is well-known that adding the L^1 -norm penalty can lead to sparse optimal control problems, i.e., the infinite dimensional control variable is localized in its domain of action. In this paper, we consider the elliptic optimal control problem with L^1 -control cost and box constraints on the control:

$$\left\{ \begin{array}{l} \min_{(y,u) \in Y \times U} J(y,u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 + \beta \|u\|_{L^1(\Omega)} \\ \text{s.t.} \quad Ly = u + y_r \quad \text{in } \Omega, \\ y = 0 \quad \text{on } \partial\Omega, \\ u \in U_{ad} = \{v(x) | a \leq v(x) \leq b, \text{ a.e on } \Omega\} \subseteq U, \end{array} \right. \quad (1)$$

where $Y := H_0^1(\Omega)$, $U := L^2(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ ($n = 2, 3$) is a convex, open and bounded domain with $C^{1,1}$ - or polygonal boundary; the desired state $y_d \in H^1(\Omega)$ and the source term

$y_r \in H^1(\Omega)$ are given; parameters $\alpha, \beta > 0$, $-\infty < a < 0 < b < +\infty$; L is the uniformly elliptic differential operator given by

$$Ly := - \sum_{i,j=1}^n \partial_{x_j} (a_{ij} y_{x_i}) + c_0 y, \quad (2)$$

where $a_{ij}, c_0 \in L^\infty(\Omega)$, $c_0 \geq 0$, $a_{ij} = a_{ji}$ and there is a constant $\theta > 0$, such that

$$\sum_{i,j=1}^n a_{ij}(x) \xi_i \xi_j \geq \theta \|\xi\|^2, \quad \text{a.a. } x \in \Omega, \forall \xi \in \mathbb{R}^n. \quad (3)$$

Due to the L^1 -control cost, the objective function $J(y, u)$ is non-smooth. This means that the structure of the control significantly differs from that of the smooth one. To solve elliptic optimal control problems with L^1 -control cost, Stadler et al. proposed a semi-smooth Newton method to solve elliptic optimal control problems with L^1 -control cost [1]. Porcelli et al. proposed a semi-smooth Newton method with a robust preconditioner for different formulations of the Newton equation [5]. However, solving Newton equations is expensive, especially when the discretization is at a fine level. Thus, some efficient first-order algorithms have received much attention in recent years. In [6], Schindele and Borzi proposed an inexact accelerated proximal gradient (APG) method in function space to solve elliptic optimal control problems with L^1 -control cost. However, the efficiency of the APG method relies closely on the step length. The backtracking approach is applied to obtain the appropriate step length, but this greatly increases the computational cost. In [7], Song et al. proposed an inexact heterogeneous ADMM (ihADMM) algorithm. Different from the classical ADMM, two different weighted inner products are utilized to define the augmented Lagrangian function for two subproblems. Moreover, the ihADMM was applied to solve PDE-constrained optimization problems with L^2 -control cost [8] and elliptic optimal control problems with pointwise box constraints on the state [9]. For more applications of ADMM-type algorithms to solve PDE-constrained optimization problems, see [10–13]. Inspired by the efficiency of ADMM-type algorithms on PDE-constrained optimal control problems, we consider using ADMM-type algorithms to solve (1).

In classical finite-element-based algorithms, problems are always discretized and computed at a fixed grid level. As the mesh becomes finer and finer, the scale of the discretized problems will be larger and the computation cost will increase, causing a bottleneck. Thus, it is essential to develop new approaches to solve optimal control problems in an accurate and computationally efficient way. The multigrid method is a modern field of research, which started in the early 1970s. It is well-known that the multigrid method is the optimal solution to many discretized partial differential equations [14]. Different from the classical multigrid method, Deuffhard proposed a cascading multigrid method for elliptic partial differential equations, which solves problems from the coarse grid to the fine grid [15]. In [16], the multigrid method was applied to tackle infinite dimensional non-linear partial differential equations using Newton methods. Due to the efficiency of the multigrid method, it is a natural idea to construct the multigrid type numerical method for PDE-constrained optimization problems. The classical multigrid method for the optimal control problem is designed to solve the linear algebraic systems formulated on each step of the optimization algorithm; refer to [17] for an overview. In [18], Gong et al. proposed an adaptive multilevel correction method, which solves the optimal control problems from the coarsest mesh to the finest mesh. Chen et al. proposed a strategy of gradually refining the grid, and proposed a multilevel ADMM (mADMM) algorithm for PDE-constrained optimization problems with a L^2 -control cost in [19]. The authors proved the global convergence and the iteration complexity results $o(1/k)$ for the mADMM.

In this paper, we focus on the *first optimize, then discretize* [20] approach. This approach provides the freedom to discretize subproblems using different discretization schemes. Motivated by the success of the multigrid method, we combined the multigrid method and

the ihADMM algorithm to numerically solve (1). At the early stage of the whole iteration process of the algorithm, using a coarse grid will not make the precision worse, but will reduce the computational cost. Thus, we propose a strategy of gradually refining the grid. Specifically, we first introduced the iterative scheme of the iADMM in function space. Then, nodal quadrature formulas were used to approximately discretize the L^1 -norm and L^2 -norm to ensure that discretized subproblems have decoupled forms. Finally, appropriate methods, such as Krylov-based methods, are used to solve the discretized subproblems.

The main contribution of this paper can be summarized as follows:

- We propose an efficient multilevel heterogeneous ADMM (mADMM) algorithm for solving sparse optimal control problems with L^1 -control cost. Specifically, we first apply the iADMM in the function space as an outer optimization method. Then, the subproblems in each iteration are discretized by the finite element method. Instead of fixing the grid size before the computation process, we apply the strategy of gradually refining the grid. Moreover, we apply nodal quadrature formulas to approximately L^1 -norm and L^2 -norm to overcome the difficulty that L^1 -norm does not have a decoupled form and ensure z-subproblems have closed form solutions. Finally, we use appropriate methods to inexactly solve the subproblems in each iteration. The global convergence and the iteration complexity results for the mhADMM algorithm are also established.
- We present numerical experiments to illustrate the effectiveness of the mhADMM algorithm. In addition, we compare the performance of the mhADMM with two benchmark methods: the ihADMM and the classical ADMM. Compared to the two methods, the mhADMM has an evident advantage in terms of computational time. Moreover, the numerical results regarding iterations illustrate the mesh-independent performance of the mhADMM.

The paper is organized as follows. In Section 2, we first briefly review the iteration format of the inexact ADMM algorithm in function space, then introduce the finite element approximation. A strategy of gradually refining the grid is introduced and the mhADMM algorithm is proposed. Numerical computation of the subproblems is also discussed. In Section 3, we present the convergence results of the mhADMM algorithm. Numerical results are given in Section 4, and concluding remarks are drawn in Section 5.

2. An Multilevel Heterogeneous ADMM Algorithm

In this section, we propose an efficient convergent multilevel alternating direction method of multipliers (mhADMM). Moreover, we introduce the numerical computation of the subproblems in the mhADMM algorithm.

To apply ADMM to solve (1), we introduce an artificial variable z ; then, we can rewrite (1) in an equivalent form:

$$\left\{ \begin{array}{l} \min_{(y,u,z) \in Y \times U \times U} J(y, u, z) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 + \beta \|z\|_{L^1(\Omega)} \\ \text{s.t.} \quad Ly = u + y_r \quad \text{in } \Omega, \\ \quad \quad y = 0 \quad \quad \text{on } \partial\Omega, \\ \quad \quad u = z, \\ \quad \quad z \in U_{ad} = \{v(x) | a \leq v(x) \leq b, \text{ a.e on } \Omega\} \subseteq U. \end{array} \right. \tag{4}$$

For the existence and uniqueness of the solution of the elliptic PDE involved in (1), where L is defined by (2), the following proposition holds.

Proposition 1. For every $u \in L^2(\Omega)$ and $y_r \in H^1(\Omega)$, the elliptic PDE involved in (1):

$$\begin{array}{l} Ly = u + y_r \quad \text{in } \Omega, \\ y = 0 \quad \quad \text{on } \partial\Omega, \end{array} \tag{5}$$

has a unique weak solution $y = y(u) := S(u + y_r)$, where $S : L^2(\Omega) \rightarrow H_0^1(\Omega)$ denotes the solution operator. Moreover, S is a well-defined continuous linear injective operator. The adjoint operator $S^* : H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$ is also a continuous linear operator.

Proof. The weak formulation of (5) is given by the following:

$$\text{Find } y \in H_0^1(\Omega) : a(y, v) = (u + y_r, v)_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega),$$

where $a : V \times V \rightarrow \mathbb{R}$ is a bilinear form:

$$a(y, v) := \int_{\Omega} \left(\sum_{i,j=1}^n a_{ij} y_{x_i} v_{x_j} + c_0 y v \right) dx.$$

We know from the assumption (3) that $a(\cdot, \cdot)$ is symmetric, i.e., $a(y, v) = a(v, y)$ for all $y, v \in H_0^1(\Omega)$. Thus, $(y, v) := a(y, v)$ defines a new inner product on $H_0^1(\Omega)$. Then, the existence of a unique solution of (5) directly follows from the Riesz representation theorem. Moreover,

$$\theta \|y\|_{H_0^1(\Omega)}^2 \leq a(y, y) = (u + y_r, y)_{L^2(\Omega)} \leq \|u + y_r\|_{L^2(\Omega)} \|y\|_{L^2(\Omega)} \leq \|u + y_r\|_{L^2(\Omega)} \|y\|_{H_0^1(\Omega)},$$

where θ is a constant, depending only on Ω , and we use the equivalence of norms in $H_0^1(\Omega)$ in the first inequality. Then, we have

$$\|y\|_{H_0^1(\Omega)}^2 \leq \frac{1}{\theta} \|u + y_r\|_{L^2(\Omega)}.$$

Thus, the solution operator S is well-defined and called the control-to-state mapping, which is a continuous linear injective operator. Since H_0^1 is a Hilbert space, the adjoint operator $S^* : H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$ is also a continuous linear operator. \square

Since (4) is continuous and strongly convex, the solution of (4) exists and is unique. The following Karush–Kuhn–Tucker (KKT) conditions hold at the optimal solution of (4).

Theorem 1. (First-order optimality condition) (y^*, u^*, z^*) is the optimal solution of (4), if, and only if, adjoint state $p^* \in H_0^1(\Omega)$ and Lagrange multiplier $\lambda^* \in L^2(\Omega)$ exists, such that the following conditions hold in the weak sense:

$$\begin{aligned} y^* &= \mathcal{S}(u^* + y_r), \\ p^* &= \mathcal{S}^*(y^* - y_d), \\ \alpha u^* - p^* + \lambda^* &= 0, \\ u^* &= z^*, \\ z^* &\in U_{ad}, \\ \langle -\lambda^*, \tilde{z} - z^* \rangle_{L^2(\Omega)} + \beta \left(\|\tilde{z}\|_{L^1(\Omega)} - \|z^*\|_{L^1(\Omega)} \right) &\geq 0, \quad \forall \tilde{z} \in U_{ad}. \end{aligned}$$

Moreover, we have

$$u^* := \Pi_{U_{ad}} \left(\frac{1}{\alpha} \text{soft}(p^*, \beta) \right),$$

where the projection operator $\Pi_{U_{ad}}(\cdot)$ and the soft thresholding operator $\text{soft}(\cdot)$, respectively, are defined as follows:

$$\begin{aligned} \Pi_{U_{ad}}(v(x)) &:= \max\{a, \min\{v(x), b\}\}, \\ \text{soft}(v(x), \beta) &:= \text{sgn}(v(x)) \circ \max(|v(x)| - \beta, 0). \end{aligned}$$

Using the solution operator S , (4) can be equivalently rewritten as the following reduced form:

$$\begin{cases} \min_{(u,z) \in U \times U} & f(u) + g(z) \\ \text{s.t.} & u = z, \end{cases} \tag{6}$$

where

$$\begin{aligned} f(u) &:= \frac{1}{2} \|S(u + y_r) - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2, \\ g(z) &:= \beta \|z\|_{L^1(\Omega)} + \delta_{U_{ad}}(z), \end{aligned}$$

$\delta_{U_{ad}}$ denotes the indicator function of U_{ad} , i.e.,

$$\delta_{U_{ad}}(z) = \begin{cases} 0, & z \in U_{ad}, \\ \infty, & z \notin U_{ad}. \end{cases}$$

Notice that (6) is a two-block separable convex optimization problem with linear equality constraints; the ADMM algorithm can be used to solve (6). The augmented Lagrangian function of (6) is defined as follows:

$$L_\sigma(u, z, \lambda; \sigma) = f(u) + g(z) + \langle \lambda, u - z \rangle_{L^2(\Omega)} + \frac{\sigma}{2} \|u - z\|_{L^2(\Omega)}^2,$$

where $\lambda \in L^2(\Omega)$ denotes the Lagrangian multiplier; $\sigma > 0$ is the penalty parameter. Given the initial point $(u^0, z^0, \lambda^0) \in L^2(\Omega) \times \text{dom}(\delta_{U_{ad}}(\cdot)) \times L^2(\Omega)$, parameters $\sigma > 0$, $\tau \in \left(0, \frac{\sqrt{5} + 1}{2}\right)$, the iteration format of ADMM in function space is as follows:

$$\begin{cases} \text{Step 1 : } \bar{u}^{k+1} = \arg \min_u f(u) + \langle \bar{\lambda}^k, u - \bar{z}^k \rangle_{L^2(\Omega)} + \frac{\sigma}{2} \|u - \bar{z}^k\|_{L^2(\Omega)}^2, \\ \text{Step 2 : } \bar{z}^{k+1} = \arg \min_z g(z) + \langle \bar{\lambda}^k, \bar{u}^{k+1} - z \rangle_{L^2(\Omega)} + \frac{\sigma}{2} \|\bar{u}^{k+1} - z\|_{L^2(\Omega)}^2, \\ \text{Step 3 : } \bar{\lambda}^{k+1} = \bar{\lambda}^k + \tau \sigma (\bar{u}^{k+1} - \bar{z}^{k+1}). \end{cases} \tag{7}$$

However, computing the exact solution of each subproblem is usually expensive and unnecessary. In [7], Song et al. propose an inexact ADMM (iADMM) in function space for (6). Krylov-based methods are used to solve the subproblems, which are equivalent to large-scale linear systems. We show the iterative scheme of the iADMM algorithm in Algorithm 1.

It is easy to see that the z -subproblem has a closed-form solution

$$\begin{aligned} z^{k+1} &= \arg \min_z \frac{\sigma}{2} \|z - \frac{1}{\sigma} (\sigma u^{k+1} + \lambda^k)\|_{L^2(\Omega)}^2 + \beta \|z\|_{L^1(\Omega)} + \delta_{U_{ad}}(z) \\ &= \Pi_{U_{ad}} \left(\frac{1}{\sigma} \text{soft}(\sigma u^{k+1} + \lambda^k, \beta) \right). \end{aligned}$$

For the global convergence results and the iteration complexity for Algorithm 1, we have the following theorem.

Algorithm 1 Inexact ADMM (iADMM) algorithm for (6)

Input: Choose the initial point $(u^0, z^0, \lambda^0) \in L^2(\Omega) \times \text{dom}(\delta_{U_{ad}}(\cdot)) \times L^2(\Omega)$, parameters $\sigma > 0, \tau \in \left(0, \frac{\sqrt{5} + 1}{2}\right)$. Let $\{\epsilon_{k+1}\}_{k=0}^\infty$ be a sequence satisfying $\{\epsilon_{k+1}\}_{k=0}^\infty \subseteq [0, +\infty)$ and $\sum_{k=0}^\infty \epsilon_k < \infty$. Set $k = 0$.

Output: u^k, z^k, λ^k .

Step 1 Compute u^{k+1} as an approximation solution of

$$\min_u f(u) + \langle \lambda^k, u - z^k \rangle_{L^2(\Omega)} + \frac{\sigma}{2} \|u - z^k\|_{L^2(\Omega)}^2$$

such that the residual $\delta_u^{k+1} := \nabla f(u^{k+1}) + \lambda^k + \sigma(u^{k+1} - z^k)$ satisfies $\|\delta_u^{k+1}\|_{L^2(\Omega)} \leq \epsilon_{k+1}$.

Step 2 Compute z^{k+1} as follows:

$$z^{k+1} = \arg \min_z g(z) + \langle \lambda^k, u^{k+1} - z \rangle_{L^2(\Omega)} + \frac{\sigma}{2} \|u^{k+1} - z\|_{L^2(\Omega)}^2.$$

Step 3 Compute

$$\lambda^{k+1} = \lambda^k + \tau \sigma (u^{k+1} - z^{k+1}).$$

Step 4 If a termination criterion is met, stop; otherwise, set $k := k + 1$ and go to Step 1.

Theorem 2 ([7] Theorem 3). Let $(y^*, u^*, z^*, p^*, \lambda^*)$ be the KKT point of (4), the sequence $\{(u^k, z^k, \lambda^k)\}_{k=0}^\infty$ is generated by Algorithm 1 with the associated state $\{y^k\}_{k=0}^\infty$ and adjoint state $\{p^k\}_{k=0}^\infty$; then, we have

$$\begin{aligned} \lim_{k \rightarrow \infty} \{ \|u^k - u^*\|_{L^2(\Omega)} + \|z^k - z^*\|_{L^2(\Omega)} + \|\lambda^k - \lambda^*\|_{L^2(\Omega)} \} &= 0, \\ \lim_{k \rightarrow \infty} \{ \|y^k - y^*\|_{H_0^1(\Omega)} + \|p^k - p^*\|_{H_0^1(\Omega)} \} &= 0. \end{aligned}$$

Moreover, a constant C_0 only depends on the initial point (u^0, z^0, λ^0) and the optimal solution (u^*, z^*, λ^*) , such that, for $k \geq 1$,

$$\min_{1 \leq i \leq k} R(u^i, z^i, \lambda^i) \leq \frac{C_0}{k}, \quad \lim_{k \rightarrow \infty} (k \cdot \min_{1 \leq i \leq k} R(u^i, z^i, \lambda^i)) = 0,$$

where the function $R : (u, z, \lambda) \rightarrow [0, \infty)$ is defined as

$$R(u, z, \lambda) := \|\nabla f(u) + \lambda\|_{L^2(\Omega)}^2 + \text{dist}^2(0, -\lambda + \partial g(z)) + \|u - z\|_{L^2(\Omega)}^2.$$

2.1. The mhADMM Algorithm

To numerically solve (6), we consider the full discretization, in which both the state y and the control u are discretized by piecewise, linear, globally continuous finite elements. We introduce a family of regular and quasi-uniform triangulations $\{\mathcal{T}_h\}$ of $\bar{\Omega}$, i.e., $\bar{\Omega} = \bigcup_{T \in \mathcal{T}_h} \bar{T}$. With each element $T \in \mathcal{T}_h$, we define the diameter of the set T by $\rho_T := \text{diam } T$ and let σ_T denote the diameter of the largest ball contained in T . The grid size is defined by $h := \max_{T \in \mathcal{T}_h} \rho_T$. We suppose the following standard assumption holds (see [20]).

Assumption 1. (Regular and quasi-uniform triangulations) There are two positive constants κ and τ such that

$$\frac{\rho_T}{\sigma_T} \leq \kappa, \quad \frac{h}{\rho_T} \leq \tau$$

hold for all $T \in \mathcal{T}_h$ and all $h > 0$. Moreover, let us define $\bar{\Omega}_h = \bigcup_{T \in \mathcal{T}_h} \bar{T}$ and let $\Omega_h \subseteq \Omega$ and Γ_h denote its interior and boundary, respectively. In the case that Ω is a convex polyhedral domain, we have $\Omega = \Omega_h$. In the case that Ω is a domain with a $C^{1,1}$ -boundary Γ , we assume that $\bar{\Omega}_h$ is convex and all boundary vertices of $\bar{\Omega}_h$ are contained in Γ , such that

$$|\Omega \setminus \Omega_h| \leq c^* h^2,$$

where $|\cdot|$ denotes the measure of the set, and $c^* > 0$ is a constant.

Due to the homogeneous boundary condition of the state equation, we use

$$Y_h := \{y_h \in C(\bar{\Omega}) \mid y_h|_T \in \mathcal{P}_1, \forall T \in \mathcal{T}_h, y_h = 0 \text{ in } \bar{\Omega} \setminus \Omega_h\},$$

$$U_h := \{u_h \in C(\bar{\Omega}) \mid u_h|_T \in \mathcal{P}_1, \forall T \in \mathcal{T}_h, u_h = 0 \text{ in } \bar{\Omega} \setminus \Omega_h\},$$

as the discretized state space and the discretized control space, respectively. \mathcal{P}_1 denotes the space of polynomials of degree less than or equal to 1. For the given regular and quasi-uniform triangulation \mathcal{T}_h with nodes $\{x_i\}_{i=1}^{N_h}$, let $\{\phi_i(x)\}_{i=1}^{N_h}$ be a basis of Y_h, U_h , which satisfies the following properties:

$$\phi_i(x) \geq 0, \quad \|\phi_i(x)\|_\infty = 1 \quad \forall i = 1, \dots, N_h, \quad \sum_{i=1}^{N_h} \phi_i(x) = 1.$$

Then, $u_h \in U_h, y_h \in Y_h$ can be represented in the following forms, respectively,

$$y_h = \sum_{i=1}^{N_h} y_i \phi_i, \quad u_h = \sum_{i=1}^{N_h} u_i \phi_i,$$

where $y_i := y_h(x_i), u_i := u_h(x_i)$. Moreover, y_h can be expressed by $y_h(u) = S_h(u + y_r)$, where S_h denotes the discretized version of the solution operator S . Let $U_{ad,h}$ denotes the discretized feasible set, which is defined by

$$U_{ad,h} := U_h \cap U_{ad} = \left\{ z_h = \sum_{i=1}^{N_h} z_i \phi_i \mid a \leq z_i \leq b, \forall i = 1, \dots, N_h \right\} \subset U_{ad}.$$

To overcome the difficulty that the discretized L^1 -norm does not have a decoupled form, we choose the nodal quadrature formulas introduced in [21] to approximately discretized the L^1 -norm:

$$\|z_h\|_{L^1(\Omega_h)} = \sum_{i=1}^n |z_i| \int_{\Omega_h} \phi_i(x) dx.$$

Moreover, in order to obtain a closed form solution for the z -subproblem, a similar quadrature formulae introduced in [8] is also used to discretize the L^2 -norm in the z -subproblem:

$$\|z_h\|_{L^2_h(\Omega_h)} = \left(\sum_{i=1}^n (z_i)^2 \int_{\Omega_h} \phi_i(x) dx \right)^{\frac{1}{2}}.$$

For the given $y_r \in H^1(\Omega)$ and $u \in L^2(\Omega)$, the unique discretized state y_h associated with u is can be expressed by $y_h(u) = S_h(u + y_r)$, where S_h is the discretized version of the solution operator S . Then, we have the well-known error estimates:

Lemma 1 ([22], Theorem 4.4.6). *For a given $u \in L^2(\Omega)$, let y be the unique weak solution of the state Equation (5) and y_h be the unique discretized state. Then, there exists a constant $c > 0$ independent of h, u and y_r , such that*

$$\|y - y_h\|_{L^2(\Omega)} + h \|\nabla y - \nabla y_h\|_{L^2(\Omega)} \leq ch^2 (\|u\|_{L^2(\Omega)} + \|y_r\|_{L^2(\Omega)}).$$

In particular, this implies $\|S - S_h\|_{\mathcal{L}(L^2, L^2)} \leq ch^2$ and $\|S - S_h\|_{\mathcal{L}(L^2, H^1)} \leq ch$.

To project the solution obtained on the coarser grid to the finer grid, we introduce the definition of the node interpolation operator I_h .

Definition 1. For a given regular and quasi-uniform triangulation \mathcal{T}_h of Ω with nodes $\{x_i\}_{i=1}^{N_h}$, let $\{\phi_i(x)\}_{i=1}^{N_h}$ denotes the standard set of nodal basis functions. The interpolation operator I_h is defined as

$$I_h w(x) := \sum_{i=1}^{N_h} w(x_i) \phi_i(x) \text{ for any } w \in C^0(\Omega) \cap H^1(\Omega).$$

For the interpolation error estimate, we have the following Theorem 3.

Theorem 3 ([19] Theorem 2). For all $w \in C^0(\Omega) \cap H^1(\Omega)$, we have

$$\|w - I_h w\|_{L^2(\Omega)} \leq c_I h \|w\|_{H^1(\Omega)},$$

where c_I is a constant independent of h .

At the early stage of the whole process, computing on the coarser grid can reduce the computation cost without making the precision worse. While as the iteration process proceeds, the iteration precision is supposed to increase. In this case, it is necessary to use the finer grid at the late stage. Thus, we apply the strategy of gradually refining the grid. In the initial iteration, we obtained a solution on the coarse grid, then projected the obtained solution to the finer grid. For the convenience of representing subproblems on different grids, we define

$$\begin{aligned} f_{h_{k+1}}(u) &:= \frac{1}{2} \|S_{h_{k+1}}(u + I_{h_{k+1}} y_r) - I_{h_{k+1}} y_d\|_{L^2(\Omega_{h_{k+1}})}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega_{h_{k+1}})}^2, \\ g_{h_{k+1}}(z) &:= \beta \|z\|_{L_{h_{k+1}}^1(\Omega_{h_{k+1}})} + \delta_{U_{ad}, h_{k+1}}(z). \end{aligned}$$

Moreover, let

$$I_h y_r := \sum_{i=1}^{N_h} y_r^i \phi_i, \quad I_h y_d := \sum_{i=1}^{N_h} y_d^i \phi_i$$

denotes the L^2 -projection of y_r and y_d onto Y_h , respectively. Then, we show the iterative scheme of the multilevel heterogeneous ADMM alternating direction method of multipliers (mhADMM) in Algorithm 2.

It is easy to see that the z -subproblem has a closed-form solution:

$$\begin{aligned} z_{h_{k+1}}^{k+1} &= \arg \min_z \frac{\sigma}{2} \|z - \frac{1}{\sigma} (\sigma u_{h_{k+1}}^{k+1} + \lambda_{h_{k+1}}^k)\|_{L^2(\Omega_{h_{k+1}})}^2 + \beta \|z\|_{L_{h_{k+1}}^1(\Omega_{h_{k+1}})} + \delta_{U_{ad}, h_{k+1}}(z) \\ &= \Pi_{U_{ad}, h_{k+1}} \left(\frac{1}{\sigma} \text{soft}(\sigma u_{h_{k+1}}^{k+1} + \lambda_{h_{k+1}}^k, \beta) \right). \end{aligned}$$

Algorithm 2 Multilevel heterogeneous ADMM (mhADMM) algorithm for (6)

Input: Choose the initial point $(u_{h_1}^0, z_{h_1}^0, \lambda_{h_1}^0) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)$, parameters $\sigma > 0, \tau \in \left(0, \frac{\sqrt{5} + 1}{2}\right)$. Let $\{\epsilon_{k+1}\}_{k=0}^\infty$ be a sequence satisfying $\{\epsilon_{k+1}\}_{k=0}^\infty \subseteq [0, +\infty)$ and $\sum_{k=0}^\infty \epsilon_{k+1} < \infty$, mesh sizes $\{h_{k+1}\}_{k=0}^\infty$ of each level satisfying $\sum_{k=0}^\infty h_{k+1} < \infty$. Set $k = 0$.

Output: $u_{h_k}^k, z_{h_k}^k, \lambda_{h_k}^k$.

Step 1 Compute $u_{h_{k+1}}^{k+1}$ as an approximation solution of

$$\min_u f_{h_{k+1}}(u) + \langle \lambda_{h_{k+1}}^k, u - z_{h_{k+1}}^k \rangle_{L^2(\Omega_{h_{k+1}})} + \frac{\sigma}{2} \|u - z_{h_{k+1}}^k\|_{L^2(\Omega_{h_{k+1}})}^2$$

such that the residual $\delta_{u, h_{k+1}}^{k+1} := \nabla f_{h_{k+1}}(u_{h_{k+1}}^{k+1}) + \lambda_{h_{k+1}}^k + \sigma(u_{h_{k+1}}^{k+1} - z_{h_{k+1}}^k)$ satisfies $\|\delta_{u, h_{k+1}}^{k+1}\|_{L^2(\Omega_{h_{k+1}})} \leq \epsilon_{k+1}$.

Step 2 Compute $z_{h_{k+1}}^{k+1}$ as follows:

$$z_{h_{k+1}}^{k+1} = \arg \min_z g_{h_{k+1}}(z) + \langle \lambda_{h_{k+1}}^k, u_{h_{k+1}}^{k+1} - z \rangle_{L^2(\Omega_{h_{k+1}})} + \frac{\sigma}{2} \|u_{h_{k+1}}^{k+1} - z\|_{L^2(\Omega_{h_{k+1}})}^2.$$

Step 3 Compute

$$\lambda_{h_{k+1}}^{k+1} = \lambda_{h_{k+1}}^k + \tau \sigma (u_{h_{k+1}}^{k+1} - z_{h_{k+1}}^{k+1}).$$

Step 4 If a termination criterion is met, stop; otherwise, set $k := k + 1$ and go to Step 1.

2.2. Numerical Computation of the Subproblems in Algorithm 2

To rewrite the subproblems into matrix-vector forms, we define the following matrices

$$\begin{aligned} K_h &:= (a(\phi_i, \phi_j))_{i,j=1}^{N_h}, \\ M_h &:= \left(\int_{\Omega_h} \phi_i \phi_j dx \right)_{i,j=1}^{N_h}, \\ W_h &:= \text{diag} \left(\int_{\Omega_h} \phi_i(x) dx \right)_{i,j=1}^{N_h}, \end{aligned}$$

where K_h, M_h and W_h denote the finite element stiffness matrix, mass matrix and lump mass matrix, respectively.

For $u_h = \sum_{i=1}^{N_h} u_i \phi_i \in U_h, y_h = \sum_{i=1}^{N_h} y_i \phi_i \in Y_h$, let

$$\mathbf{u}_h = (u_1, \dots, u_{N_h}), \mathbf{y}_h = (y_1, \dots, y_{N_h})$$

be the relative coefficient vectors, respectively. For $I_h y_r = \sum_{i=1}^{N_h} y_r^i \phi_i, I_h y_d = \sum_{i=1}^{N_h} y_d^i \phi_i$, let

$$\mathbf{y}_{r,h} = (y_r^1, y_r^2, \dots, y_r^{N_h}), \mathbf{y}_{d,h} = (y_d^1, y_d^2, \dots, y_d^{N_h})$$

be the coefficient vectors, respectively. Let I_h denotes the vector version of the interpolation operator. Moreover, we define

$$f(\mathbf{u}) := \frac{1}{2} \|K_h^{-1} M_h(\mathbf{u} + \mathbf{y}_{r,h}) - \mathbf{y}_{d,h}\|_{M_h}^2 + \frac{\alpha}{2} \|\mathbf{u}\|_{M_h}^2,$$

$$g(\mathbf{z}) := \beta \|W_h \mathbf{z}\|_1^2 + \delta_{[a,b]^{N_h}}(\mathbf{z}).$$

Then, the matrix-vector form of Algorithm 2 is given in Algorithm 3.

Algorithm 3 Matrix-vector form of the mhADMM algorithm

Input: $(\mathbf{u}^0, \mathbf{z}^0, \boldsymbol{\lambda}^0) \in \mathbb{R}^{N_h} \times [a, b]^{N_h} \times \mathbb{R}^{N_h}$, parameters $\sigma > 0$, $\tau \in \left(0, \frac{\sqrt{5} + 1}{2}\right)$.

Let $\{\epsilon_{k+1}\}_{k=0}^\infty$ be a sequence satisfying $\{\epsilon_{k+1}\}_{k=0}^\infty \subseteq [0, +\infty)$ and $\sum_{k=0}^\infty \frac{\epsilon_{k+1}}{\|M_{h_{k+1}}\|_2} < \infty$,

mesh sizes $\{h_{k+1}\}_{k=0}^\infty$ of each iteration satisfy $\sum_{k=0}^\infty h_{k+1} < \infty$. Set $k = 0$.

Output: $\mathbf{u}_{h_k}^k, \mathbf{z}_{h_k}^k, \boldsymbol{\lambda}_{h_k}^k$.

Step 1: Compute $\mathbf{u}_{h_{k+1}}^{k+1}$ as an approximation solution of

$$\min_{\mathbf{u}} f(\mathbf{u}) + \langle M_{h_{k+1}} \boldsymbol{\lambda}_{h_{k+1}}^k, \mathbf{u} - \mathbf{z}_{h_{k+1}}^k \rangle + \frac{\sigma}{2} \|\mathbf{u} - \mathbf{z}_{h_{k+1}}^k\|_{M_{h_{k+1}}}^2$$

such that the residual $\boldsymbol{\delta}_{\mathbf{u}, h_{k+1}}^{k+1} := \nabla f(\mathbf{u}_{h_{k+1}}^{k+1}) + M_{h_{k+1}} \boldsymbol{\lambda}_{h_{k+1}}^k + \sigma M_{h_{k+1}} (\mathbf{u}_{h_{k+1}}^{k+1} - \mathbf{z}_{h_{k+1}}^k)$ satisfies $\|\boldsymbol{\delta}_{\mathbf{u}, h_{k+1}}^{k+1}\| \leq \frac{\epsilon_{k+1}}{\sqrt{\|M_{h_{k+1}}\|_2}}$.

Step 2: Compute $\mathbf{z}_{h_{k+1}}^{k+1}$ as follows:

$$\mathbf{z}_{h_{k+1}}^{k+1} = \arg \min_{\mathbf{z}} g(\mathbf{z}) + \langle M_h \boldsymbol{\lambda}_{h_{k+1}}^k, \mathbf{u}_{h_{k+1}}^{k+1} - \mathbf{z} \rangle + \frac{\sigma}{2} \|\mathbf{u}_{h_{k+1}}^{k+1} - \mathbf{z}\|_{W_h}^2.$$

Step 3: Compute

$$\boldsymbol{\lambda}_{h_{k+1}}^{k+1} = \mathbf{I}_{h_{k+1}} \boldsymbol{\lambda}_{h_k}^k + \tau \sigma (\mathbf{u}_{h_{k+1}}^{k+1} - \mathbf{z}_{h_{k+1}}^{k+1}).$$

Step 4: If a termination criterion is met, stop; otherwise, set $k := k + 1$ and go to Step 1.

The u -subproblem at the k th iteration is equivalent to the following linear system:

$$M_{h_k} K_{h_k}^{-1} M_{h_k} (K_{h_k}^{-1} M_{h_k} (\mathbf{u}_{h_k}^k + \mathbf{y}_{r,h_k}) - \mathbf{y}_{d,h_k}) + \alpha M_{h_k} \mathbf{u}_{h_k}^k + \boldsymbol{\lambda}_{h_k}^k + \sigma (\mathbf{u}_{h_k}^k - \mathbf{z}_{h_k}^k) = 0, \quad (8)$$

where $\mathbf{y}_{h_k}^k := K_{h_k}^{-1} M_{h_k} (\mathbf{u}_{h_k}^k + \mathbf{y}_{r,h_k})$, $\mathbf{p}_{h_k}^k := K_{h_k}^{-1} M_{h_k} (\mathbf{y}_{d,h_k} - \mathbf{y}_{h_k}^k)$ denote the discretized state and the discretized adjoint state, respectively. Then (8) can be rewritten as:

$$\begin{bmatrix} M_{h_k} & 0 & K_{h_k} \\ 0 & (\alpha + \sigma) M_{h_k} & -M_{h_k} \\ K_{h_k} & -M_{h_k} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}_{h_k}^k \\ \mathbf{u}_{h_k}^k \\ \mathbf{p}_{h_k}^k \end{bmatrix} = \begin{bmatrix} M_{h_k} \mathbf{y}_{d,h_k} \\ M_{h_k} (\sigma \mathbf{I}_{h_k} \mathbf{z}_{h_{k-1}}^{k-1} - \boldsymbol{\lambda}_{h_{k-1}}^{k-1}) \\ M_{h_k} \mathbf{y}_{r,h_k} \end{bmatrix}. \quad (9)$$

We know from (9) that $\mathbf{p}_{h_k}^k = (\alpha + \sigma) \mathbf{u}_{h_k}^k - \sigma \mathbf{I}_{h_k} \mathbf{z}_{h_{k-1}}^{k-1} + \mathbf{I}_{h_k} \boldsymbol{\lambda}_{h_{k-1}}^{k-1}$. By eliminating the variable $\mathbf{p}_{h_k}^k$, (9) can be rewritten in the following reduced form without any additional computational cost:

$$\begin{bmatrix} M_{h_k} & (\alpha + \sigma) K_{h_k} \\ -K_{h_k} & M_{h_k} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{h_k}^k \\ \mathbf{u}_{h_k}^k \end{bmatrix} = \begin{bmatrix} M_{h_k} \mathbf{y}_{d,h_k} + K_{h_k} (\sigma \mathbf{I}_{h_k} \mathbf{z}_{h_{k-1}}^{k-1} - \boldsymbol{\lambda}_{h_{k-1}}^{k-1}) \\ -M_{h_k} \mathbf{y}_{r,h_k} \end{bmatrix}. \quad (10)$$

The equation system (10) can be solved by the generalized minimal residual (GMRES) with the preconditioned variant of modified hermitian and skew-hermitian splitting (PMHSS) preconditioner [23,24].

For the z-subproblem, there is a closed-form solution

$$z_{h_{k+1}}^{k+1} = \Pi_{[a,b]^{N_{h_{k+1}}}} \left(\frac{1}{\sigma} \text{soft} \left(\sigma \mathbf{u}_{h_{k+1}}^{k+1} + W_h^{-1} M_h \lambda_{h_{k+1}}^k, \beta \right) \right).$$

3. Convergence Analysis

In this section, we establish the global convergence and the iteration complexity results in non-ergodic sense for the sequence generated by Algorithm 2. Before giving the convergence analysis, we first introduce the exact multi-level heterogeneous ADMM (mhADMM) algorithm. Each subproblem of the exact mhADMM algorithm is solved exactly. Given the initial point $(u^0, z^0, \lambda^0) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)$, parameters $\sigma > 0$, $\tau \in \left(0, \frac{\sqrt{5} + 1}{2}\right)$. The mesh sizes $\{h_{k+1}\}_{k=0}^\infty$ of each iteration satisfy $\sum_{k=0}^\infty h_{k+1} < \infty$. Then, each iteration of the exact mhADMM has three main steps:

$$\left\{ \begin{array}{l} \text{Step 1 : } \bar{u}_{h_{k+1}}^{k+1} = \arg \min_u f_{h_{k+1}}(u) + \langle \bar{\lambda}_{h_{k+1}}^k, u - \bar{z}_{h_{k+1}}^k \rangle_{L^2(\Omega_{h_{k+1}})} + \frac{\sigma}{2} \|u - \bar{z}_{h_{k+1}}^k\|_{L^2(\Omega_{h_{k+1}})}, \\ \text{Step 2 : } \bar{z}_{h_{k+1}}^{k+1} = \arg \min_z g_{h_{k+1}}(z) + \langle \bar{\lambda}_{h_{k+1}}^k, \bar{u}_{h_{k+1}}^{k+1} - z \rangle_{L^2(\Omega_{h_{k+1}})} + \frac{\sigma}{2} \|\bar{u}_{h_{k+1}}^{k+1} - z\|_{L^2_{h_{k+1}}(\Omega_{h_{k+1}})}^2, \\ \text{Step 3 : } \bar{\lambda}_{h_{k+1}}^{k+1} = \bar{\lambda}_{h_{k+1}}^k + \tau \sigma (\bar{u}_{h_{k+1}}^{k+1} - \bar{z}_{h_{k+1}}^{k+1}), \end{array} \right. \tag{11}$$

where $\bar{\lambda}_{h_{k+1}}^k := I_{h_{k+1}} \bar{\lambda}_{h_k}^k$, $\bar{z}_{h_{k+1}}^k := I_{h_{k+1}} z_{h_k}^k$. Then, we use the following lemma to measure the gap between the solution sequence obtained by the ADMM in function space and the exact mhADMM algorithm in finite dimensional space.

Lemma 2. Let the initial point be $(z^0, \lambda^0) \in H^1(\Omega) \times H^1(\Omega)$. Let $\{(\bar{u}^k, \bar{z}^k, \bar{\lambda}^k)\}_{k=0}^\infty$ defined in (7) be the sequence generated by the ADMM in function space and $\{(\bar{u}_{h_k}^k, \bar{z}_{h_k}^k, \bar{\lambda}_{h_k}^k)\}_{k=0}^\infty$ defined in (11) be the sequence generated by the exact mhADMM algorithm. Then, for all $k \geq 1$, we have

$$\begin{aligned} \|\bar{u}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} &\leq C_{u,k} h_k, \\ \|\bar{z}^k - \bar{z}_{h_k}^k\|_{L^2(\Omega_{h_k})} &\leq C_{z,k} h_k, \\ \|\bar{\lambda}^{k-1} - \bar{\lambda}_{h_k}^{k-1}\|_{L^2(\Omega_{h_k})} &\leq C_{\lambda,k} h_k, \end{aligned}$$

where $C_{u,k}, C_{z,k}, C_{\lambda,k}$ are constants independent of h_k and there is a constant C such that $C_{u,k} \leq C$ for any $k \geq 1$. Thus, we have

$$\sum_{k=1}^\infty \|\bar{u}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} \leq C \sum_{k=1}^\infty h_k.$$

Proof. We employ the mathematical induction to prove the conclusion. For $k = 1$, we know from Theorem 3 that

$$\begin{aligned} \|\bar{\lambda}^0 - \bar{\lambda}_{h_1}^0\|_{L^2(\Omega_{h_1})} &= \|\lambda^0 - I_{h_1} \lambda^0\|_{L^2(\Omega_{h_1})} \\ &\leq c_I h_1 \|\lambda^0\|_{H^1(\Omega_{h_1})} \\ &\leq C_{\lambda,1} h_1, \end{aligned}$$

where $C_{\lambda,1} := c_I \|\bar{\lambda}^0\|_{H^1(\Omega_{h_1})}$ is a constant independent of h .

For u -subproblems, \bar{u}^1 and $\bar{u}_{h_1}^1$ satisfy the following optimality conditions, respectively,

$$\begin{aligned} S^*[S(\bar{u}^1 + y_r) - y_d] + \alpha\bar{u}^1 + \lambda^0 + \sigma(\bar{u}^1 - z^0) &= 0, \\ S_{h_1}^*[S_{h_1}(\bar{u}_{h_1}^1 + I_{h_1}y_r) - I_{h_1}y_d] + \alpha\bar{u}_{h_1}^1 + I_{h_1}\lambda^0 + \sigma(\bar{u}_{h_1}^1 - I_{h_1}z^0) &= 0. \end{aligned}$$

By subtracting the two equalities above, we have

$$\begin{aligned} & [-(\alpha + \sigma)I - S_{h_1}^*S_{h_1}](\bar{u}^1 - \bar{u}_{h_1}^1) \\ &= S^*S\bar{u}^1 - S_{h_1}^*S_{h_1}\bar{u}_{h_1}^1 + S^*Sy_r - S_{h_1}^*S_{h_1}I_{h_1}y_r - S^*y_d + S_{h_1}^*I_{h_1}y_d + \lambda^0 - I_{h_1}\lambda^0 + \sigma I_{h_1}z^0 - \sigma z^0. \end{aligned}$$

Then, we know from Lemma 2 in [19] that $\|\bar{u}^1 - \bar{u}_{h_1}^1\|_{L^2(\Omega_{h_1})} \leq C_{u,1}h_1$, where $C_{u,1}$ is a constant independent of h_1 .

For z -subproblems, \bar{z}^1 and $\bar{z}_{h_1}^1$ satisfy

$$\begin{aligned} \bar{z}^1 &= \Pi_{U_{ad}}\left(\frac{1}{\sigma}\text{soft}(\sigma\bar{u}^1 + \lambda^0, \beta)\right), \\ \bar{z}_{h_1}^1 &= \Pi_{U_{ad},h_1}\left(\frac{1}{\sigma}\text{soft}(\sigma\bar{u}_{h_1}^1 + \lambda_{h_1}^0, \beta)\right), \end{aligned}$$

respectively. Then, we know from the projection operator Π and the soft thresholding operator $\text{soft}(\cdot)$ are nonexpansive, such that

$$\begin{aligned} \|\bar{z}^1 - \bar{z}_{h_1}^1\|_{L^2(\Omega_{h_1})} &= \|\Pi_{U_{ad}}\left(\frac{1}{\sigma}\text{soft}(\sigma\bar{u}^1 + \lambda^0, \beta)\right) - \Pi_{U_{ad},h_1}\left(\frac{1}{\sigma}\text{soft}(\sigma\bar{u}_{h_1}^1 + \lambda_{h_1}^0, \beta)\right)\|_{L^2(\Omega_{h_1})} \\ &\leq \frac{1}{\sigma}\|\sigma\bar{u}^1 - \sigma\bar{u}_{h_1}^1 + \lambda^0 - \lambda_{h_1}^0\|_{L^2(\Omega_{h_1})} \\ &\leq \|\bar{u}^1 - \bar{u}_{h_1}^1\|_{L^2(\Omega_{h_1})} + \frac{1}{\sigma}\|\lambda^0 - \lambda_{h_1}^0\|_{L^2(\Omega_{h_1})} \\ &\leq C_{u,1}h_1 + \frac{1}{\sigma}C_{\lambda,1}h_1 \\ &= C_{z,1}h_1, \end{aligned}$$

where $C_{z,1} := C_{u,1} + \frac{1}{\sigma}C_{\lambda,1}$. Hence, the statement is true for $k = 1$.

For $k > 1$, we assume the statement is true for $\forall j \leq k$. Then, for $j = k + 1$, we have

$$\begin{aligned} \|\bar{\lambda}^k - \bar{\lambda}_{h_{k+1}}^k\|_{L^2(\Omega_{h_{k+1}})} &= \|\bar{\lambda}^k - \bar{\lambda}_{h_k}^k + \bar{\lambda}_{h_k}^k - I_{h_{k+1}}\bar{\lambda}_{h_k}^k\|_{L^2(\Omega_{h_{k+1}})} \\ &\leq \|\bar{\lambda}^k - \bar{\lambda}_{h_k}^k\|_{L^2(\Omega_{h_k})} + \|\bar{\lambda}_{h_k}^k - I_{h_{k+1}}\bar{\lambda}_{h_k}^k\|_{L^2(\Omega_{h_{k+1}})} \\ &\leq \|\bar{\lambda}^{k-1} - \bar{\lambda}_{h_k}^{k-1} + \tau\sigma(\bar{u}^k - \bar{u}_{h_k}^k) - \tau\sigma(\bar{z}^k - \bar{z}_{h_k}^k)\|_{L^2(\Omega_{h_k})} \\ &\quad + \|\bar{\lambda}_{h_k}^k - I_{h_{k+1}}\bar{\lambda}_{h_k}^k\|_{L^2(\Omega_{h_{k+1}})} \\ &\leq \|\bar{\lambda}^{k-1} - \bar{\lambda}_{h_k}^{k-1}\|_{L^2(\Omega_{h_k})} + \tau\sigma(\|\bar{u}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} + \|\bar{z}^k - \bar{z}_{h_k}^k\|_{L^2(\Omega_{h_k})}) \\ &\quad + \|\bar{\lambda}_{h_k}^k - I_{h_{k+1}}\bar{\lambda}_{h_k}^k\|_{L^2(\Omega_{h_{k+1}})} \\ &\leq (C_{\lambda,k} + \tau\sigma C_{u,k} + \tau\sigma C_{z,k})h_k + c_I h_{k+1} \|\bar{\lambda}_{h_k}^k\|_{H^1(\Omega_{h_{k+1}})} \\ &\leq C_{\lambda,k+1}h_{k+1}, \end{aligned}$$

where $C_{\lambda,k+1} := C_{k+1}(C_{\lambda,k}h_k + \tau\sigma C_{u,k} + \tau\sigma C_{z,k}) + c_I \|\bar{\lambda}_{h_k}^k\|_{H^1(\Omega_{h_{k+1}})}$ is a constant independent of h_{k+1} . In the last equality, we use the property $\sum_{k=0}^{\infty} h_k < \infty$; thus, there exists a constant C_{k+1} such that $h_k < C_{k+1}h_{k+1}$.

For u -subproblems, \bar{u}^{k+1} and $\bar{u}_{h_{k+1}}^{k+1}$ satisfy the following optimality conditions respectively,

$$S^*[S(\bar{u}^k + y_r) - y_d] + \alpha \bar{u}^{k+1} + \bar{\lambda}^k + \sigma(\bar{u}^{k+1} - \bar{z}^k) = 0,$$

$$S_{h_{k+1}}^*[S_{h_{k+1}}(\bar{u}_{h_{k+1}}^{k+1} + I_{h_{k+1}}y_r) - I_{h_{k+1}}y_d] + \alpha \bar{u}_{h_{k+1}}^{k+1} + \bar{\lambda}_{h_{k+1}}^k + \sigma(\bar{u}_{h_{k+1}}^{k+1} - \bar{z}_{h_{k+1}}^k) = 0.$$

By subtracting the two equalities above, we have

$$- [(\alpha + \sigma)I + S_{h_{k+1}}^* S_{h_{k+1}}] (\bar{u}^{k+1} - \bar{u}_{h_{k+1}}^{k+1})$$

$$= S^* S \bar{u}^{k+1} - S_{h_{k+1}}^* S_{h_{k+1}} \bar{u}_{h_{k+1}}^{k+1} + S^* S y_r - S_{h_{k+1}}^* S_{h_{k+1}} I_{h_{k+1}} y_r - S^* y_d + S_{h_{k+1}}^* I_{h_{k+1}} y_d$$

$$+ \bar{\lambda}^k - \bar{\lambda}_{h_{k+1}}^k - \sigma(\bar{z}^k - \bar{z}_{h_{k+1}}^k).$$

Then, we know from Lemma 2 in [19] that

$$\|\bar{u}^{k+1} - \bar{u}_{h_{k+1}}^{k+1}\|_{L^2(\Omega_{h_{k+1}})} \leq C_{u,k+1} h_{k+1},$$

where $C_{u,k+1}$ is a constant independent of h_{k+1} .

For z -subproblems, \bar{z}^{k+1} and $\bar{z}_{h_{k+1}}^{k+1}$ satisfy

$$\bar{z}^{k+1} = \Pi_{U_{ad}} \left(\frac{1}{\sigma} \text{soft}(\sigma \bar{u}^{k+1} + \lambda^k, \beta) \right),$$

$$\bar{z}_{h_{k+1}}^{k+1} = \Pi_{U_{ad}, h_1} \left(\frac{1}{\sigma} \text{soft}(\sigma \bar{u}_{h_{k+1}}^{k+1} + \lambda_{h_{k+1}}^k, \beta) \right),$$

respectively. We know the projection operator Π and the soft thresholding operator $\text{soft}(\cdot)$ are nonexpansive, such that

$$\|\bar{z}^{k+1} - \bar{z}_{h_{k+1}}^{k+1}\|_{L^2(\Omega_{h_{k+1}})} \leq \frac{1}{\sigma} \|\sigma \bar{u}^{k+1} - \sigma \bar{u}_{h_{k+1}}^{k+1} + \lambda^k - \lambda_{h_{k+1}}^k\|_{L^2(\Omega_{h_{k+1}})}$$

$$\leq \|\bar{u}^{k+1} - \bar{u}_{h_{k+1}}^{k+1}\|_{L^2(\Omega_{h_{k+1}})} + \frac{1}{\sigma} \|\lambda^k - \lambda_{h_{k+1}}^k\|_{L^2(\Omega_{h_{k+1}})}$$

$$\leq C_{u,k+1} h_{k+1} + \frac{1}{\sigma} C_{\lambda,k+1} h_{k+1}$$

$$= C_{z,k+1} h_{k+1},$$

where $C_{z,k+1} := C_{u,k+1} + \frac{1}{\sigma} C_{\lambda,k+1}$. Hence, the statement is true for $j = k + 1$. We complete the whole proof of Lemma 2. \square

Similarly, we have the following lemma. Lemma 3 shows the gap between the sequence (u^k, z^k, λ^k) generated by Algorithm 1 and the sequence $(u_{h_k}^k, z_{h_k}^k, \lambda_{h_k}^k)$ generated by Algorithm 2.

Lemma 3. *Let the initial point be $(u^0, z^0, \lambda^0) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)$. Let $\{(u^k, z^k, \lambda^k)\}_{k=0}^\infty$ be the sequence generated by Algorithm 1 and $\{(u_{h_k}^k, z_{h_k}^k, \lambda_{h_k}^k)\}_{k=0}^\infty$ be the sequence generated by Algorithm 2. Then, for all $k \geq 1$, we have*

$$\|u^k - u_{h_k}^k\|_{L^2(\Omega_{h_k})} \leq \hat{C}_{u,k}(h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}),$$

$$\|z^k - z_{h_k}^k\|_{L^2(\Omega_{h_k})} \leq \hat{C}_{z,k}(h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}),$$

$$\|\lambda^{k-1} - \lambda_{h_k}^{k-1}\|_{L^2(\Omega_{h_k})} \leq \hat{C}_{\lambda,k}(h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}),$$

where $\hat{C}_{u,k}, \hat{C}_{z,k}, \hat{C}_{\lambda,k}$ are constants independent of h_k , and there exists a constant \hat{C} such that $\hat{C}_{u,k} \leq \hat{C}$ for any $k \geq 1$. Thus we have

$$\sum_{k=1}^{\infty} \|u^k - u_{h_k}^k\|_{L^2(\Omega)} \leq \hat{C} \sum_{k=1}^{\infty} (h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}).$$

Proof. We employ the mathematical induction to prove the conclusion. The proof is similar to Lemma 2. We do not discuss this in detail here. \square

The total error of utilizing numerical methods to solve PDE-constrained optimal control problem consists of two parts: the discretization error and the iteration error. These two kinds of errors can be regarded as the error of inexactly solving infinite-dimensional subproblems. Thus, the mhADMM algorithm can be regarded as the iADMM algorithm in function space. Inspired by the results of Theorem 2, we have the following convergence results.

Theorem 4. Let $(y^*, u^*, z^*, p^*, \lambda^*)$ be the KKT point of (1), $(u_{h_k}^k, z_{h_k}^k, \lambda_{h_k}^k)_{k=0}^{\infty}$ be the sequence generated by Algorithm 2 with the associated state $\{y_{h_k}^k\}_{k=0}^{\infty}$ and the adjoint state $\{p_{h_k}^k\}_{k=0}^{\infty}$. Then we have

$$\begin{aligned} \lim_{k \rightarrow \infty} \{ \|u_{h_k}^k - u^*\|_{L^2(\Omega_{h_k})} + \|z_{h_k}^k - z^*\|_{L^2(\Omega_{h_k})} + \|\lambda_{h_k}^k - \lambda^*\|_{L^2(\Omega_{h_k})} \} &= 0, \\ \lim_{k \rightarrow \infty} \{ \|y_{h_k}^k - y^*\|_{H_0^1(\Omega_{h_k})} + \|p_{h_k}^k - p^*\|_{H_0^1(\Omega_{h_k})} \} &= 0. \end{aligned}$$

Moreover, there exists a constant \bar{C} that only depends on the initial point (u^0, z^0, λ^0) and the optimal solution (u^*, z^*, λ^*) such that, for $k \geq 1$,

$$\min_{1 \leq i \leq k} R_{h_i}(u_{h_i}^i, z_{h_i}^i, \lambda_{h_i}^i) \leq \frac{\bar{C}}{k}, \quad \lim_{k \rightarrow \infty} (k \cdot \min_{1 \leq i \leq k} R_{h_i}(u_{h_i}^i, z_{h_i}^i, \lambda_{h_i}^i)) = 0,$$

where $R_{h_i} : (u_{h_i}^i, z_{h_i}^i, \lambda_{h_i}^i) \rightarrow [0, \infty)$ is defined as

$$R_h(u_{h_i}^i, z_{h_i}^i, \lambda_{h_i}^i) := \|\nabla f_h(u_{h_i}^i) + \lambda_{h_i}^{i-1}\|_{L^2(\Omega_{h_i})}^2 + \text{dist}^2(0, -\lambda_{h_i}^{i-1} + \partial g_h(z_{h_i}^i)) + \|u_{h_i}^i - z_{h_i}^i\|_{L^2(\Omega_{h_i})}^2.$$

Proof. Note that $(u_{h_k}^k, z_{h_k}^k, \lambda_{h_k}^k)$ can be regarded as the inexact solution obtained by Algorithm 1. The total error δ_u^k consists of two parts, the discretization error from gradually refining the grid and the iteration error from the inexactly solving the subproblems. Then, we know from the optimality conditions of the u -subproblem in Algorithm 1 that

$$S^*[S(u_{h_k}^k + y_r) - y_d] + \alpha u_{h_k}^k + \lambda_{h_{k-1}}^{k-1} + \sigma(u_{h_k}^k - z_{h_{k-1}}^{k-1}) = \delta_u^k. \tag{12}$$

Moreover, we know from the optimality condition of the u -subproblem in ADMM in function space, the optimality conditions of the u -subproblem in Algorithm 2 and the exact multi-level ADMM that

$$S^*[S(\bar{u}^k + y_r) - y_d] + \alpha \bar{u}^k + \bar{\lambda}^{k-1} + \sigma(\bar{u}^k - \bar{z}^{k-1}) = 0, \tag{13}$$

$$S_{h_k}^*[S_{h_k}(u_{h_k}^k + I_{h_k}y_r) - I_{h_k}y_d] + \alpha u_{h_k}^k + \lambda_{h_k}^{k-1} + \sigma(u_{h_k}^k - z_{h_k}^{k-1}) = \delta_{u,h_k}^k, \tag{14}$$

$$S_{h_k}^*[S_{h_k}(\bar{u}_{h_k}^k + I_{h_k}y_r) - I_{h_k}y_d] + \alpha \bar{u}_{h_k}^k + \bar{\lambda}_{h_k}^{k-1} + \sigma(\bar{u}_{h_k}^k - \bar{z}_{h_k}^{k-1}) = 0. \tag{15}$$

Then, we know from (12)–(15) that

$$\begin{aligned}
 \delta_u^k &= \delta_u^k - \delta_{u,h_k}^k + \delta_{u,h_k}^k \\
 &= \delta_{u,h_k}^k + S^* S(u_{h_k}^k - \bar{u}^k) + (\alpha + \sigma)(u_{h_k}^k - \bar{u}^k) + (\lambda^{k-1} - \bar{\lambda}^{k-1}) - \sigma(z^{k-1} - \bar{z}^{k-1}) \\
 &\quad + S_{h_k}^* S_{h_k}(\bar{u}_{h_k}^k - u_{h_k}^k) + (\alpha + \sigma)(\bar{u}_{h_k}^k - u_{h_k}^k) + (\bar{\lambda}_{h_k}^{k-1} - \lambda_{h_k}^{k-1}) - \sigma(\bar{z}_{h_k}^{k-1} - z_{h_k}^{k-1}) \\
 &= \delta_{u,h_k}^k + \underbrace{[(\alpha + \sigma)I + S^* S](\bar{u}_{h_k}^k - u_{h_k}^k)}_{I_1} + \underbrace{(\lambda^{k-1} - \lambda_{h_k}^{k-1}) - \sigma(z^{k-1} - z_{h_k}^{k-1})}_{I_2} \\
 &\quad + \underbrace{(\bar{\lambda}_{h_k}^{k-1} - \bar{\lambda}^{k-1}) - \sigma(\bar{z}_{h_k}^{k-1} - \bar{z}^{k-1})}_{I_4} + \underbrace{(S^* S - S_{h_k}^* S_{h_k})(u_{h_k}^k - \bar{u}_{h_k}^k)}_{I_6}.
 \end{aligned} \tag{16}$$

For the term I_1 and I_4 , we know from Lemma 2 that

$$\|I_1\|_{L^2(\Omega_{h_k})} \leq (\alpha + \sigma + \|S^*\| \|S\|) C_{u,k} h_k, \tag{17}$$

$$\|I_4\|_{L^2(\Omega_{h_k})} \leq C_{\lambda,k} h_k. \tag{18}$$

For the term I_5 , we know from Theorem 3 and Lemma 2 that

$$\begin{aligned}
 \|I_5\|_{L^2(\Omega_{h_k})} &= \sigma \|\bar{z}_{h_k}^{k-1} - \bar{z}_{h_{k-1}}^{k-1} + \bar{z}_{h_{k-1}}^{k-1} - \bar{z}^{k-1}\|_{L^2(\Omega_{h_k})} \\
 &\leq \sigma \|\bar{z}_{h_k}^{k-1} - \bar{z}_{h_{k-1}}^{k-1}\|_{L^2(\Omega_{h_k})} + \sigma \|\bar{z}_{h_{k-1}}^{k-1} - \bar{z}^{k-1}\|_{L^2(\Omega_{h_k})} \\
 &\leq \sigma c_I h_k \|\bar{z}_{h_{k-1}}^{k-1}\|_{H^1(\Omega_{h_k})} + \sigma C_{z,k-1} h_{k-1} \\
 &\leq \sigma c_I h_k \|\bar{z}_{h_{k-1}}^{k-1}\|_{H^1(\Omega_{h_k})} + \sigma C_{z,k-1} C_k h_k \\
 &\leq c_5 h_k,
 \end{aligned} \tag{19}$$

where $c_5 := \sigma c_I \|\bar{z}_{h_{k-1}}^{k-1}\|_{H^1(\Omega_{h_k})} + \sigma C_{z,k-1} C_k$ is a constant, $h_{k-1} \leq C_k h_k$.

For the term I_2 , we know from Lemma 3 that

$$\|I_2\|_{L^2(\Omega_{h_k})} \leq \hat{C}_{\lambda,k}(h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}). \tag{20}$$

For the term I_3 , we know from Theorem 3 and Lemma 2 that

$$\|I_3\|_{L^2(\Omega_{h_k})} \leq c_3(h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}), \tag{21}$$

where c_3 is a constant.

Finally, for the term I_6 , we make use of the decomposition

$$\begin{aligned}
 \|u_{h_k}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} &= \|u_{h_k}^k - u^k + u^k - u^* + u^* - \bar{u}^k + \bar{u}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} \\
 &\leq \|u_{h_k}^k - u^k\|_{L^2(\Omega_{h_k})} + \|u^k - u^*\|_{L^2(\Omega_{h_k})} + \|u^* - \bar{u}^k\|_{L^2(\Omega_{h_k})} + \|\bar{u}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} \\
 &\leq \hat{C}_{u,k}(h_k + \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})}) + C^*,
 \end{aligned} \tag{22}$$

where $\hat{C}_{u,k}, C^*$ are constants in dependent of h_k . In the last equality, we used Lemma 2, Lemma 3, the convergence property of ADMM in function space and the inexact ADMM in function space. Then, we know from Proposition 1 and Lemma 1 that

$$\begin{aligned}
 \|I_6\|_{L^2(\Omega_{h_k})} &= \|(S^* S - S^* S_{h_k} + S^* S_{h_k} - S_{h_k}^* S_{h_k})(u_{h_k}^k - \bar{u}_{h_k}^k)\|_{L^2(\Omega_{h_k})} \\
 &\leq \|(S^* S - S^* S_{h_k})(u_{h_k}^k - \bar{u}_{h_k}^k)\|_{L^2(\Omega_{h_k})} + \|(S^* S_{h_k} - S_{h_k}^* S_{h_k})(u_{h_k}^k - \bar{u}_{h_k}^k)\|_{L^2(\Omega_{h_k})} \\
 &\leq \|S^*\| \|S - S_{h_k}\| \|u_{h_k}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} + \|S^* - S_{h_k}^*\| \|S_{h_k}\| \|u_{h_k}^k - \bar{u}_{h_k}^k\|_{L^2(\Omega_{h_k})} \\
 &\leq c_6 h_k,
 \end{aligned} \tag{23}$$

where c_6 is a constant.

Then, we know from (16)–(23) that there are constants C_1^* and C_2^* such that

$$\sum_{k=1}^{\infty} \|\delta_u^k\|_{L^2(\Omega)} \leq C_1^* \sum_{k=1}^{\infty} \|\delta_{u,h_k}^k\|_{L^2(\Omega_{h_k})} + C_2^* \sum_{k=1}^{\infty} h_k.$$

Moreover, the mesh sizes $\{h_{k+1}\}_{k=0}^{\infty}$ of each mhADMM iteration satisfy $\sum_{k=0}^{\infty} h_{k+1} < \infty$, the residuals of each mhADMM iteration satisfy $\sum_{k=0}^{\infty} \|\delta_{u,h_{k+1}}^{k+1}\|_{L^2(\Omega_{h_{k+1}})} \leq \sum_{k=0}^{\infty} \epsilon_{k+1} < \infty$, thus we have

$$\sum_{k=1}^{\infty} \|\delta_u^k\|_{L^2(\Omega)} < \infty.$$

Then, we know from Theorem 2 that the global convergence and the iteration complexity results $o(1/k)$ for Algorithm 2 are guaranteed. \square

4. Numerical Experiments

In this section, we illustrate the numerical performance of the mhADMM algorithm for the elliptic PDE-constrained optimization problems with L^1 -control cost. For our numerical experiment, we used MATLAB R2021b with the FEM package iFEM [25] on a Thinkpad laptop with 2.8 GHz Intel Core i7 processor with 16GB of RAM.

In the mhADMM algorithm, the accuracy of a numerical solution is measured by the following residual. Let ϵ be a given accuracy tolerance, we terminate the algorithm when $\eta < \epsilon$, where $\eta := \max\{\eta_1, \eta_2, \eta_3, \eta_4, \eta_5\}$, where

$$\eta_1 := \frac{\|K_h \mathbf{y} - M_h \mathbf{u} - M_h \mathbf{y}_r\|}{1 + \|M_h \mathbf{y}_r\|}, \eta_2 := \frac{\|M_h(\mathbf{u} - \mathbf{z})\|}{1 + \|\mathbf{u}\|}, \eta_3 := \frac{\|M_h(\mathbf{y} - \mathbf{y}_d) + K_h \mathbf{p}\|}{1 + \|M_h \mathbf{y}_d\|},$$

$$\eta_4 := \frac{\|\alpha M_h \mathbf{u} - M_h \mathbf{p} + M_h \boldsymbol{\lambda}\|}{1 + \|\mathbf{u}\|}, \eta_5 := \frac{\|\mathbf{z} - \Pi_{[a,b]^{N_h}}(\text{soft}(W_h^{-1} M_h \boldsymbol{\lambda}, \beta))\|}{1 + \|\mathbf{z}\|}.$$

To present the finite element error estimates' results, we introduce the experimental order of convergence, a brief EOC defined by

$$\text{EOC} := \frac{\log E(h_1) - \log E(h_2)}{\log h_1 - \log h_2},$$

where $h_1, h_2 > 0, h_1 \neq h_2$ denotes different grid sizes, E denotes the positive error functional

$$E(h) := \|u - u_h\|_{L^2(\Omega)}.$$

We note that if $E(h) = \mathcal{O}(h^\gamma)$ holds, then $\text{EOC} \approx \gamma$.

As shown in Section 2.1, instead of using the standard piecewise linear and continuous finite elements, nodal quadrature formulas are used to approximately discretize the L^1 -norm and L^2 -norm in Algorithm 2. In both examples, the mhADMM algorithm, the ihADMM algorithm and the classical ADMM algorithm are employed to obtain numerical solutions of different grid sizes. For both numerical examples and all algorithms, we chose $(u^0, z^0, \lambda^0) = (0, 0, 0)$ as the initial values. The penalty parameter σ was chosen as $\sigma = \alpha$. For the step length τ , we chose $\tau = 1.618$. We terminate the algorithms when the residual $\eta < 10^{-6}$ with the maximum number of iterations set to 500.

In numerical experiments, we show the numerical results for different final mesh sizes. In Tables 1 and 2, h denotes the final mesh size, '#dofs' denotes the dimension of the control variable on each grid level, and 'iter' represents the times of iteration. To guarantee the sequence $\{\epsilon_{k+1}\}_{k=0}^{\infty} \subseteq [0, +\infty)$ satisfies $\sum_{k=0}^{\infty} \frac{\epsilon_{k+1}}{\sqrt{\|M_{h_{k+1}}\|_2^2}} < \infty$, and the mesh

sizes $\{h_k\}_{k=0}^\infty \subseteq [0, +\infty)$ of each iteration satisfy $\sum_{k=0}^\infty h_{k+1} \leq \infty$, we choose $\epsilon_{k+1} = \frac{C}{(k+1)^2}$, where C is a constant and $h_k = \frac{\sqrt{2}}{2^{k+3}}, k \in \mathbb{Z}, k \geq 1$ in both examples. Moreover, we would like to point out that, in the iteration of mhADMM algorithm, once the mesh size h_k reaches the final mesh size h , we continue the iteration in the final mesh until the stopping criterion above is satisfied.

Before providing examples, we first introduce the following algorithm, which can help us formulate sparse optimal control problems.

According to the first-order optimality condition given in Theorem 1, it is easy to see that Algorithm 4 provides a construction strategy for problems with known optimal solutions (y^*, u^*) .

Algorithm 4 Construct the optimal control problem

Step 1 Choose $y^* \in H_0^1(\Omega)$ and $p^* \in H_0^1(\Omega)$ arbitrarily.

Step 2

$$u^* := \Pi_{U_{ad}} \left(\frac{1}{\alpha} \text{soft}(p^*, \beta) \right) = \begin{cases} \min \left\{ \frac{p^* - \beta}{\alpha}, b \right\}, & \text{on } x \in \Omega : p^*(x) > \beta, \\ \max \left\{ \frac{p^* + \beta}{\alpha}, a \right\}, & \text{on } x \in \Omega : p^*(x) < -\beta \\ 0, & \text{elsewhere.} \end{cases}$$

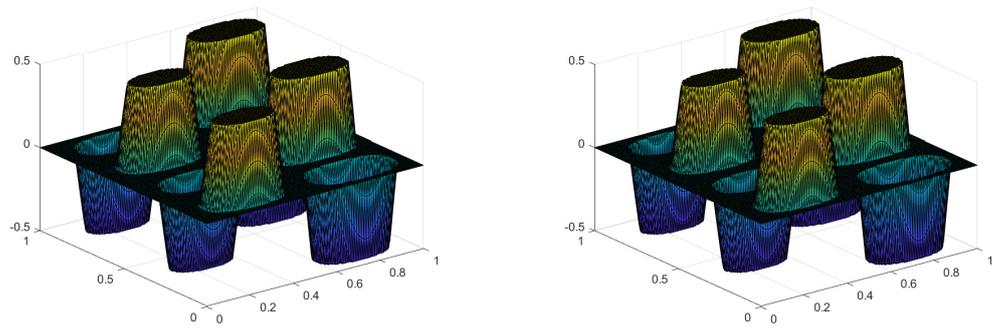
Step 3 Set $y_r = S^{-1}y^* - u^*$ and $y_d = y^* - (S^*)^{-1}p^*$.

Example 1. Consider

$$\begin{cases} \min_{(y,u) \in H_0^1(\Omega) \times L^2(\Omega)} & J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 + \beta \|u\|_{L^1(\Omega)} \\ \text{s.t.} & -\Delta y = u + y_r \quad \text{in } \Omega, \\ & y = 0 \quad \text{on } \partial\Omega, \\ & u \in U_{ad} = \{v(x) | a \leq v(x) \leq b, \text{ a.e on } \Omega\}, \end{cases}$$

where $\Omega = (0, 1)^2$, the parameters $\alpha = 0.5, \beta = 0.5, a = -0.5, b = 0.5$. As this is a constructed problem, we set $y^* = \sin(\pi x_1) \sin(\pi x_2)$ and $p^* = 2\beta \sin(2\pi x_1) \exp(0.5x_1) \sin(4\pi x_2)$. Then, through Algorithm 4, we can obtain the optimal control solution $u^* = \Pi_{U_{ad}} \left(\frac{1}{\alpha} \text{soft}(p^*, \beta) \right)$, the source term $y_r = S^{-1}y^* - u^*$ and the desired state $y_d = y^* - (S^*)^{-1}p^*$. Thus, we construct the example for which we know the exact solution.

We then test the mhADMM, the ihADMM and the classical ADMM for Example 1. The exact optimal control u and an example for the numerical optimal control obtained by mhADMM on the grid with $h = \frac{\sqrt{2}}{2^7}$ are shown in Figure 1.



(a) exact optimal control u

(b) numerical optimal control u_h

Figure 1. (a) The exact optimal control on the grid with $h = \frac{\sqrt{2}}{2^7}$ for Example 1. (b) The numerical optimal control obtained by the mhADMM on the grid with $h = \frac{\sqrt{2}}{2^7}$ for Example 1.

In Table 1, we report the dimension of the control variable on each grid level, the error E of the control u , the EOC, the residual η , the computational time and the number of iterations obtained by the mhADMM, the ihADMM and the classical ADMM. As can be seen from the fourth and fifth column of Table 1, a high-subfigures are correct.precision solution does not improve the accuracy of the discretization error and the EOC. The computational time on the seventh, eighth and ninth columns show that the mhADMM is much faster than the ihADMM and the classical ADMM, especially when the discretization is at a fine level. The mhADMM algorithm can significantly reduce the computational cost and make the algorithm faster. This is mainly because the mhADMM adopts the strategy of gradually refining the grid, while the ihADMM and the classical ADMM compute the problem on a fixed grid size, which is their computational bottleneck. Moreover, the seventh column illustrates the mesh-independent performance of mhADMM; that is, the number of iteration of the mhADMM is independent of the discretization level. Above all, we can see that the mhADMM is much more efficient than the ihADMM and the classical ADMM.

Example 2. Consider

$$\left\{ \begin{array}{l} \min_{(y,u) \in H_0^1(\Omega) \times L^2(\Omega)} J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 + \beta \|u\|_{L^1(\Omega)} \\ \text{s.t.} \quad -\Delta y = u \quad \text{in } \Omega, \\ \quad \quad y = 0 \quad \text{on } \partial\Omega, \\ \quad \quad u \in U_{ad} = \{v(x) | a \leq v(x) \leq b, \text{ a.e on } \Omega\}, \end{array} \right.$$

where $\Omega = (0, 1)^2$, the parameters $\alpha = 10^{-4}$, $\beta = 10^{-3}$, $a = -10$, $b = 10$. The exact sparse solution of this problem is not known in advance. Instead, we use the numerical solutions computed on the grid with the grid size $h = \frac{\sqrt{2}}{2^9}$ as reference solutions.

As an example, Figure 2 presents the numerical optimal control for Example 2 on the grid with $h = \frac{\sqrt{2}}{2^7}$.

Table 2 shows the dimension of the control variable at each grid level, the error E of the control u , the EOC, the residual η , the computational time and the number of iterations for three methods. Similar to Example 1, the mhADMM still outperforms the ihADMM and the classical ADMM in terms of the computational time. The fourth and fifth column of Table 2 clearly show that a high-precision solution does not improve the accuracy of the discretization error and the EOC. Furthermore, the numerical results in the seventh column also illustrate the mesh-independent performance of our mhADMM algorithm. These results demonstrate that the mhADMM is more efficient than the ihADMM and the classical ADMM.

Table 1. The convergence behavior of the mhADMM, the ihADMM and the classical ADMM for Example 1.

Case	h	#dofs	E	EOC	Index	mhADMM	ihADMM	ADMM
1	$\frac{\sqrt{2}}{2^4}$	225	9.66×10^{-2}	0.99	residual η	8.93×10^{-7}	6.69×10^{-7}	6.46×10^{-7}
					time (s)	0.12	0.13	0.13
					#iter	20	16	18
2	$\frac{\sqrt{2}}{2^5}$	961	4.46×10^{-2}	1.05	residual η	9.44×10^{-7}	6.60×10^{-7}	8.63×10^{-7}
					time (s)	0.40	0.42	0.53
					#iter	20	18	24
3	$\frac{\sqrt{2}}{2^6}$	3969	1.49×10^{-2}	1.23	residual η	3.30×10^{-7}	7.37×10^{-7}	9.40×10^{-7}
					time (s)	2.16	2.67	8.51
					#iter	22	21	49
4	$\frac{\sqrt{2}}{2^7}$	16,129	4.92×10^{-3}	1.32	residual η	5.57×10^{-7}	5.54×10^{-7}	8.87×10^{-7}
					time (s)	14.86	18.35	262.56
					#iter	21	23	120
5	$\frac{\sqrt{2}}{2^8}$	65,025	1.65×10^{-3}	1.37	residual η	7.02×10^{-7}	4.61×10^{-7}	9.97×10^{-7}
					time (s)	114.33	200.43	7576.14
					#iter	20	25	257
6	$\frac{\sqrt{2}}{2^9}$	261,121	5.83×10^{-4}	1.39	residual η	6.91×10^{-7}	2.82×10^{-7}	1.17×10^{-5}
					time (s)	2457.43	3850.03	279,881.28
					#iter	20	27	500

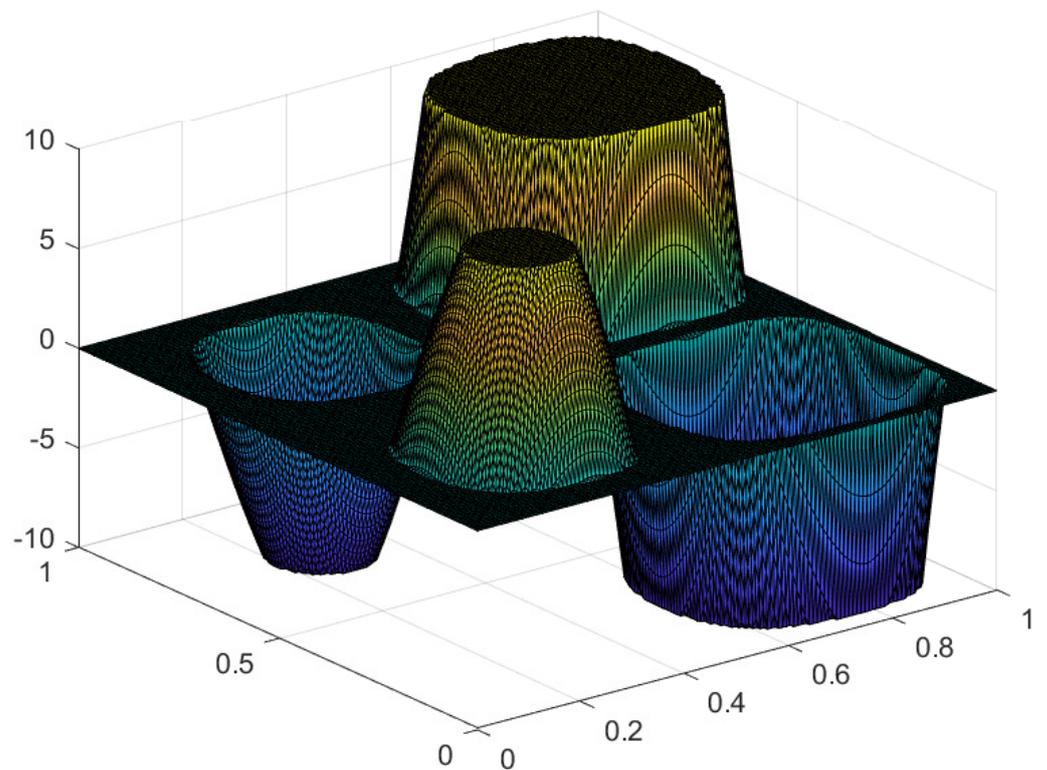


Figure 2. The numerical optimal control for Example 2 on the grid with $h = \frac{\sqrt{2}}{2^7}$.

Table 2. The convergence behavior of the mhADMM, the ihADMM and the classical ADMM for Example 2.

Case	h	#dofs	E	EOC	Index	mhADMM	ihADMM	ADMM
1	$\frac{\sqrt{2}}{2^4}$	225	9.69×10^{-1}	1.19	residual η	7.12×10^{-7}	7.12×10^{-7}	8.04×10^{-7}
					time (s)	0.25	0.30	0.33
					#iter	22	22	35
2	$\frac{\sqrt{2}}{2^5}$	961	5.72×10^{-1}	0.97	residual η	7.71×10^{-7}	7.77×10^{-7}	9.37×10^{-7}
					time (s)	0.95	1.10	1.66
					#iter	19	19	88
3	$\frac{\sqrt{2}}{2^6}$	3969	1.64×10^{-1}	1.25	residual η	6.03×10^{-7}	6.02×10^{-7}	7.63×10^{-7}
					time (s)	5.03	5.57	35.73
					#iter	20	20	198
4	$\frac{\sqrt{2}}{2^7}$	16,129	4.44×10^{-2}	1.41	residual η	8.48×10^{-7}	8.25×10^{-7}	8.10×10^{-7}
					time (s)	28.13	37.70	851.48
					#iter	20	20	454
5	$\frac{\sqrt{2}}{2^8}$	65,025	1.72×10^{-2}	1.40	residual η	7.15×10^{-7}	7.74×10^{-7}	1.66×10^{-5}
					time (s)	134.46	196.07	8659.08
					#iter	21	21	500
6	$\frac{\sqrt{2}}{2^9}$	261,121	-	-	residual η	9.09×10^{-7}	9.01×10^{-7}	6.17×10^{-5}
					time (s)	1052.59	1972.46	151,623.99
					#iter	21	21	500

5. Conclusions

In this paper, we propose a new, efficient, multilevel, heterogeneous ADMM (mhADMM) algorithm for solving sparse elliptic PDE-constrained optimal control problems with L^1 -control cost and box constraints on the control. Specifically, the inexact ADMM is first applied in the function space. Then, we propose the strategy of gradually refining the grid and employ the standard piecewise linear finite element to discretize the related subproblems appearing in each iteration of the inexact ADMM algorithm. Moreover, nodal quadrature formulas are utilized to approximately discretize the L^1 -norm and L^2 -norm to overcome the difficulty that the L^1 -norm does not have a decoupled form. Finally, subproblems are solved by appropriate numerical methods. Theoretical results regarding the global convergence and iteration complexity are presented. In our numerical experiments, we show that the proposed mhADMM is superior to the ihADMM and the classical ADMM in terms of the efficiency.

Author Contributions: Conceptualization, X.C. and Z.C.; methodology, X.S.; software, X.C.; validation, X.C. and L.X.; formal analysis, X.C. and X.S.; writing—original draft preparation, X.C.; writing—review and editing, L.X. and Z.C.; visualization, X.C.; supervision, X.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 11971092), the National Natural Science Foundation of China (No. 42274166), the China Postdoctoral Science Foundation of China (No. 2020M670717), the Fundamental Research Funds for the Central Universities (No. N2105019), the Fundamental Research Funds for the Central Universities (No. 3132022200) and the Fundamental Research Funds for the Central Universities.

Data Availability Statement: The data presented in this study are available on request from the first author and corresponding author.

Acknowledgments: We will thank the reviewers for taking time off their busy schedule to review this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stadler, G. Elliptic optimal control problems with L^1 -control cost and applications for the placement of control devices. *Comp. Optim. Appl.* **2009**, *44*, 159–181. [\[CrossRef\]](#)
2. Ciarrella, G.; Borzi, A. A LONE code for the sparse control of quantum systems. *Comput. Phys. Commun.* **2016**, *200*, 312–323. [\[CrossRef\]](#)
3. Garcke, H.; Lam, K.F.; Signori, A. Sparse optimal control of a phase field tumor model with mechanical effects. *SIAM J. Control Optim.* **2021**, *59*, 1555–1580. [\[CrossRef\]](#)
4. Casas, E.; Tröltzsch, F. Sparse optimal control for a semilinear heat equation with mixed control-state constraints-regularity of Lagrange multipliers. *ESAIM Control Optim. Calc. Var.* **2021**, *27*, 2. [\[CrossRef\]](#)
5. Porcelli, M.; Simoncini, V.; Stoll, M. Preconditioning PDE-constrained optimization with L^1 - sparsity and control constraints. *Comput. Math. Appl.* **2017**, *74*, 1059–1075. [\[CrossRef\]](#)
6. Schindele, A.; Borzi, A. Proximal methods for elliptic optimal control problems with sparsity cost functional. *Appl. Math.* **2016**, *7*, 967–992. [\[CrossRef\]](#)
7. Song, X.; Yu, B.; Wang, Y.; Zhang, X. An FE-inexact heterogeneous ADMM for elliptic optimal control problems with L^1 -control cost. *J. Syst. Sci. Complex.* **2018**, *31*, 1659–1697. [\[CrossRef\]](#)
8. Song, X.; Yu, B. A two-phase strategy for control constrained elliptic optimal control problem. *Numer. Linear Algebra Appl.* **2018**, *25*, e2138. [\[CrossRef\]](#)
9. Chen, Z.; Song, X.; Zhang, X.; Yu, B. A FE-ADMM algorithm for Lavrentiev-regularized state-constrained elliptic control problem. *ESAIM Control Optim. Calc. Var.* **2019**, *25*, 5. [\[CrossRef\]](#)
10. Zhang, K.; Li, J.; Song, Y.; Wang, X. An alternating direction method of multipliers for elliptic equation constrained optimization problem. *Sci. China Math.* **2017**, *60*, 361–378. [\[CrossRef\]](#)
11. Li, J.; Wang, X.; Zhang, K. An efficient alternating direction method of multipliers for optimal control problems constrained by random Helmholtz equations. *Numer. Algorithms* **2018**, *78*, 161–191. [\[CrossRef\]](#)
12. Glowinski, R.; Song, Y.; Yuan, X. An ADMM numerical approach to linear parabolic state constrained optimal control problems. *Numer. Math.* **2020**, *144*, 931–966. [\[CrossRef\]](#)
13. 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.* **2022**, *38*, 115–158. [\[CrossRef\]](#)
14. Shaidurov, V.V. *Multigrid Methods for Finite Elements*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995. [\[CrossRef\]](#)
15. Bornemann, F.A.; Deuffhard P. The cascadic multigrid method for elliptic problems. *Numer. Math.* **1996**, *75*, 135–152.
16. Deuffhard, P. *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*; Springer: Berlin/Heidelberg, Germany, 2011. [\[CrossRef\]](#)
17. Borzi, A.; Schulz, V. Multigrid methods for PDE optimization. *SIAM Rev.* **2009**, *51*, 361–395.
18. Gong, W.; Xie, H.; Yan, N. Adaptive multilevel correction method for finite element approximations of elliptic optimal control problems. *J. Sci. Comput.* **2017**, *72*, 820–841. [\[CrossRef\]](#)
19. Chen, X.; Song, X.; Chen, Z.; Yu, B. A multilevel ADMM algorithm for elliptic PDE-constrained optimization problems. *Comp. Appl. Math.* **2020**, *39*, 331. [\[CrossRef\]](#)
20. Hinze, M.; Pinnau, R.; Ulbrich, M.; Ulbrich, S. *Optimization with PDE Constraints*; Springer: Berlin/Heidelberg, Germany, 2009. [\[CrossRef\]](#)
21. Wachsmuth, G.; Wachsmuth, D. Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM Control Optim. Calc. Var.* **2011**, *17*, 858–886.
22. Ciarlet, P.G. *The Finite Element Method for Elliptic Problems*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2002. [\[CrossRef\]](#)
23. Bai, Z.; Benzi, M.; Chen, F.; Wang, Z. Preconditioned MHSS iteration methods for a class of block two-by-two linear systems with applications to distributed control problems. *IMA J. Numer. Anal.* **2013**, *33*, 343–369.
24. Cao, S.; Wang, Z. PMHSS iteration method and preconditioners for Stokes control PDE-constrained optimization problems. *Numer. Algorithms* **2021**, *87*, 365–380.
25. Chen, L. *iFEM: An Integrated Finite Element Methods Package in MATLAB*; Technical Report; University of California at Irvine: Irvine, CA, USA, 2008. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.