ESAIM: COCV 25 (2019) 5
https://doi.org/10.1051/cocv/2018019

**ESAIM: Control, Optimisation and Calculus of Variations**
www.esaim-cocv.org

# A FE-ADMM ALGORITHM FOR LAVRENTIEV-REGULARIZED STATE-CONSTRAINED ELLIPTIC CONTROL PROBLEM[☆]

Zixuan Chen, Xiaoliang Song, Xuping Zhang and Bo Yu[*]

**Abstract.** In this paper, Elliptic control problems with pointwise box constraints on the state is considered, where the corresponding Lagrange multipliers in general only represent regular Borel measure functions. To tackle this difficulty, the Lavrentiev regularization is employed to deal with the state constraints. To numerically discretize the resulted problem, full piecewise linear finite element discretization is employed. Estimation of the error produced by regularization and discretization is done. The error order of full discretization is not inferior to that of variational discretization because of the Lavrentiev-regularization. Taking the discretization error into account, algorithms of high precision do not make much sense. Utilizing efficient first-order algorithms to solve discretized problems to moderate accuracy is sufficient. Then a heterogeneous alternating direction method of multipliers (hADMM) is proposed. Different from the classical ADMM, our hADMM adopts two different weighted norms in two subproblems respectively. Additionally, to get more accurate solution, a two-phase strategy is presented, in which the primal-dual active set (PDAS) method is used as a postprocessor of the hADMM. Numerical results not only verify error estimates but also show the efficiency of the hADMM and the two-phase strategy.

**Mathematics Subject Classification.** 49J20, 49N05, 65G99, 68W15.

Received June 3, 2017. Accepted March 12, 2018.

## 1. Introduction

In this paper, we consider the following elliptic PDE-constrained optimal control problem with box constraints on the state

$$\begin{cases} \min\limits_{(y,u)\in Y\times U} & J(y,u)=\dfrac{1}{2}\|y-y_d\|^2_{L^2(\Omega)}+\dfrac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ \text{s.t.} & -\Delta y = u \quad \text{in } \Omega, \\ & y=0 \quad \text{on } \Gamma, \\ & a\le y(x)\le b \quad \text{a.e. on } \Omega, \end{cases} \tag{P}$$

[☆]This work is supported by the National Natural Science Foundation of China (Nos. 11571061, 11401075, 11701065) and the Fundamental Research Funds for the Central Universities (No. DUT16LK36).

*Keywords and phrases:* Optimal control, Pointwise state constraints, Lavrentiev regularization, Error estimates, Heterogeneous ADMM, Two-phase strategy.

School of Mathematical Sciences, Dalian University of Technology, Dalian, Liaoning 116025, China.

[*] Corresponding author: yubo@dlut.edu.cn

 

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 $\Gamma$; the desired state $y_d \in L^2(\Omega)$ is given; $a$, $b \in \mathbb{R}$ and $\alpha > 0$ are given parameters. Since the constraints in (P) denote closed convex set, (P) admits unique solution $(y^*, u^*)$. The solution operator $G$ of the elliptic equation in (P) mapping $u$ to $y$ is compact. To be more precise, $G = ES$, where $S : u \to y$ assigns $u \in L^2(\Omega)$ to the weak solution $y \in H_0^1(\Omega)$ and $E : H_0^1(\Omega) \to L^2(\Omega)$ is the compact embedding operator. We use $(\cdot, \cdot)$ to denote the inner product in $L^2(\Omega)$ and use $\|\cdot\|$ to denote the corresponding norm. Through this paper, let us suppose the following Slater condition for (P) holds.

**Assumption 1.1.** There exists a $\widehat{u} \in L^2(\Omega)$ such that

$$a < (S\widehat{u})(x) < b \qquad \forall x \in \overline{\Omega}.$$

**Remark 1.2.** Our considerations can also carry over to uniformly elliptic operators

$$\mathscr{A} y = -\sum_{i,j=1}^{n} \partial_{x_j}(a_{ij} y_{x_j}) + c_0 y, \qquad a_{ij}, c_0 \in L^\infty,\ c_0 \geq 0,\ a_{ij} = a_{ji}$$

and there is a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \theta \|\xi\|^2 \qquad \text{for almost all } \xi \in \mathbb{R}^n.$$

Boundary condition can also expand to

$$\partial_n y = 0 \qquad \text{on } \Gamma.$$

Optimal control problems with state constraints and their numerical realization have been studied extensively recently (see [3, 17, 30]). Since the Lagrange multiplier associated to (P) in general only represents a regular Borel measure (see Casas [8] or Alibert and Raymond [1]) because of the presence of the pointwise state constraints, the complementarity condition in the optimality conditions cannot be written into a pointwise form. Hence, nonsmooth pointwise reformulations, which are needed in semismooth Newton methods, are not possible. To overcome this difficulty, there are two common approaches, Moreau-Yosida regularization and Lavrentiev regularization. Moreau-Yosida regularization [19, 20, 25, 27] is to convert the state constraint into a penalty term. As in [34], the authors showed that a semismooth Newton method applied to the Moreau-Yosida regularization of (P) leads to a $3 * 3$ block saddle point linear system, whose coefficient matrix is symmetric and indefinite. While in our paper, we focus on the Lavrentiev regularization, whose idea is to replace the state constraint by control-state mixed constraint. We can see from Section 4 that only a $2 * 2$ block saddle point system has to be solved in each iteration by applying our hADMM, which is based on the inherent structure of the problem.

The Lavrentiev regularized problem has the form of a control-constrained elliptic optimal control problem. As we know, since projection has to be carried out to get the control in each iteration in variational discretization [22], which means mesh refinement for the control, the error order of the control of variational discretization is generally higher than that of full discretization. However, our error analysis indicates that because of the employment of the Lavrentiev-regularization, the error order of the control of full discretization is not inferior to that of variational discretization, which is the most important reason prompting us to use Lavrentiev regularization. The Lavrentiev regularized problem associated to (P) is:

$$\left\{\begin{array}{ll}\min\limits_{(y,u)\in Y\times U} & J(y,u)=\frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)}+\frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}\\ \text{s.t.} & -\Delta y=u \quad \text{in } \Omega,\\ & y=0 \quad \text{on } \Gamma,\\ & a\leq \lambda u+y\leq b \quad \text{a.e. on } \Omega,\end{array}\right. \tag{$\mathrm{P}_\lambda$}$$

where $\lambda>0$ denotes the regularization parameter. Since the constraints in ($\mathrm{P}_\lambda$) denote closed convex set, ($\mathrm{P}_\lambda$) admits unique solution $(\overline{y}_\lambda,\overline{u}_\lambda)$. In [32], the authors prove the convergence of $(\overline{y}_\lambda,\overline{u}_\lambda)\to(y^*,u^*)$ in $L^2(\Omega)$ for $\lambda\to 0$. Also, they show that the Lagrange multiplier associated to the mixed control-state constraint in ($\mathrm{P}_\lambda$) is an $L^2$-function for every $\lambda>0$. In addition, [23] proves the weak convergence of the adjoint states in $L^2$ for $\lambda$ tending to zero and the weak-$*$ convergence of the multipliers in $C(\overline{\Omega})^*$ to their counterparts of problem (P) for $\lambda\downarrow 0$. Without loss of generality, we assume that $0<\lambda<1$. We know from [36] that for the error resulted from Lavrentiev-regularization, the following estimate holds

$$\|u^*-\overline{u}_\lambda\|\leq c\sqrt{\lambda}, \tag{1.1}$$

where $c$ is a constant independent of $\lambda$. If we introduce an artificial variable $v=y+\lambda u$, ($\mathrm{P}_\lambda$) can be transformed into a pure control constrained optimal control problem:

$$\left\{\begin{array}{ll}\min & J(y,v)=\frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)}+\frac{\alpha}{2\lambda^2}\|v-y\|^2_{L^2(\Omega)}\\ \text{s.t.} & -\Delta y+\frac{1}{\lambda}y=\frac{1}{\lambda}v \quad \text{in } \Omega,\\ & y=0 \quad \text{on } \Gamma,\\ & a\leq v\leq b \quad \text{a.e. on } \Omega.\end{array}\right. \tag{$\widetilde{\mathrm{P}}_\lambda$}$$

Since ($\widetilde{\mathrm{P}}_\lambda$) is a pure control-constrained problem, it admits a unique Lagrange multiplier in $L^2(\Omega)$ associated to the inequality constraint.

To numerically solve the regularized problems, we use the *First discretize, then optimize* approach. With respect to the discrete methods, the variational discretization has been applied in dealing with ($\mathrm{P}_\lambda$) in [23], where the authors give the following error estimates.

$$\|u^*-\overline{u}_{\lambda,h}\|\leq C\left(\sqrt{\lambda}+\frac{1}{\lambda^2}\left(h^2+\frac{1}{\lambda}h^3+\frac{1}{\lambda^2}h^4\right)\right) \tag{1.2}$$

and

$$\|u^*-\overline{u}_{\lambda,h}\|\leq C\left(\sqrt{\lambda}+\max\{h|\log(h)|,h^{2-\frac{n}{2}}\}\right), \tag{1.3}$$

where $n=2,3$ denotes the space dimension and $C$ is a positive constant independent of the finite element grid size $h$ and regularization parameter $\lambda$.

Although the variational discretization avoids explicit discretization of the controls, compared with full discretization method, it is not convenient to be numerically implemented. Hence, in this paper, we use the full discretization method, in which both the state and control are discretized by piecewise linear functions. The remarkable advantage of full discretization is that it can transform the problem into a finite dimensional problem with a good structure, which is convenient to be implemented numerically. More importantly, we extend the results of [23] to the full discretization case, which results in the following two error estimates.

$$\|u^* - \overline{u}_{\lambda,h}\| \leq C\left(\sqrt{\lambda} + \frac{1}{\sqrt{\lambda}}h + \frac{1}{\lambda^2}\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right)\right) \tag{1.4}$$

and

$$\|u^* - \overline{u}_{\lambda,h}\| \leq C\left(\sqrt{\lambda} + \max\{h|\log(h)|, h^{2-\frac{n}{2}}\}\right), \tag{1.5}$$

where $n = 2, 3$ denotes the space dimension and $C$ is a positive constant independent of $\lambda$ and $h$.

Although at first glance, the precision of (1.2) is higher than (1.4) from the view of $h$, actually it depends on the matching relation between $\lambda$ and $h$. For example, we take $h = 2^{-9}$, which is small enough in general. Meanwhile, we take $\lambda = 10^{-4}$, where $\lambda$ often has to be smaller in practice. In this case, $\frac{h^2}{\lambda^2} = \frac{h}{\lambda^{\frac{3}{2}}} \cdot \frac{h}{\sqrt{\lambda}}$ is bigger than $\frac{h}{\sqrt{\lambda}}$. In addition, the second error estimate (1.5) is the same as (1.3). So it does not mean that the error order of full discretization is inferior to that of variational discretization because of the effect of $\lambda$, i.e. the employment of Lavrentiev regularization, especially when $\lambda$ is very small.

An algorithm called the primal-dual active set method (PDAS) has been used in solving the Lavrentiev-regularized state constrained elliptic control problems in [31], which was proved to be a special semismooth Newton method in [21]. Benefiting from the local superlinear convergence rate, semismooth Newton method is a prior choice for solving nonsmooth optimization problem. The total error of utilizing numerical methods to solve PDE constrained problem consists of two parts: discretization error and the error of algorithm for discretized problem. The discretization error order of piecewise linear finite element method in terms of mesh size is $O(h)$, which accounts for the main part. Taking the precision of discretization error into account, using fast first-order algorithms to get a solution of moderate precision is sufficient, rather than necessarily using algorithms of high precision. In addition, it is seen in Section 4 that in general we have to solve a $4*4$ block equation system in each iteration, which makes the calculation very large, especially when the finite element grid size $h$ is very small. In [35], the authors give a method to transform the $4*4$ block equation system to a $2*2$ block one, however, it brings additional computation for the inverse of the mass matrix.

As we know, there are many first order algorithms being used to solve finite dimensional large scale optimization fast, such as accelerated proximal gradient (APG) method [2, 26, 40, 41] and alternating direction method of multipliers (ADMM) [5, 11, 15, 28, 29]. Motivated by the success of these first order algorithms, an APG method in function space (called Fast Inexact Proximal (FIP) method) was proposed to solve the elliptic optimal control problem involving $L^1$-control cost in [38]. It is known that whether the APG method is efficient depends closely on whether the step-length is close enough to the Lipschitz constant, however, the Lipschitz constant is not easy to estimate in usual. So in this paper, we focus on ADMM, which was originally proposed in [9, 16] and has been used broadly in many areas. First, we give a brief overview of ADMM for the following linearly constrained convex optimization problem

$$\begin{cases} \min & \theta_1(x) + \theta_2(y) \\ \text{s.t.} & Ax + By = b, \\ & x \in \mathcal{X}, \\ & y \in \mathcal{Y}, \end{cases} \tag{1.6}$$

where $\theta_1(x) : \mathbb{R}^{n_1} \to \mathbb{R}$ and $\theta_2(y) : \mathbb{R}^{n_2} \to \mathbb{R}$ are convex functions, $A \in \mathbb{R}^{m\times n_1}$, $B \in \mathbb{R}^{m\times n_2}$ and $b \in \mathbb{R}^m$, $\mathcal{X} \subset \mathbb{R}^{n_1}$ and $\mathcal{Y} \subset \mathbb{R}^{n_2}$ are given closed, convex sets. The augmented Lagrangian function of (1.6) is

$$\mathcal{L}_\sigma(x, y, \lambda; \sigma) = \theta_1(x) + \theta_2(y) + (\lambda, Ax + By - b) + \frac{\sigma}{2}\|Ax + By - b\|^2, \tag{1.7}$$

where $\lambda \in \mathbb{R}^m$ is the Lagrange multiplier and $\sigma > 0$ is a penalty parameter. Each iteration of ADMM has three main steps

$$\begin{cases} x^{k+1} = \text{argmin}\left\{\mathcal{L}_\sigma(x, y^k, \lambda^k; \sigma)\middle|\ x \in \mathcal{X}\right\}, \\ y^{k+1} = \text{argmin}\left\{\mathcal{L}_\sigma(x^{k+1}, y, \lambda^k; \sigma)\middle|\ y \in \mathcal{Y}\right\}, \\ \lambda^{k+1} = \lambda^k + \sigma(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{1.8}$$

The advantage of ADMM is that it separates $\theta_1(x)$ and $\theta_2(y)$ into two subproblems, which makes each subproblem in (1.8) could be solved easily. The ADMM algorithm for solving (1.6) has global convergence and sublinear convergence rate at least under some general assumptions.

To apply ADMM type algorithm to ($\mathrm{P}_\lambda$), we introduce an artificial variable $v = \lambda u + y$, which results in

$$\begin{cases} \min \quad J(y, v) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ \text{s.t.} \quad -\Delta y = u \quad \text{in } \Omega, \\ \qquad y = 0 \quad \text{on } \Gamma, \\ \qquad v - \lambda u - y = 0 \quad \text{a.e. on } \Omega, \\ \qquad a \le v \le b \quad \text{a.e. on } \Omega. \end{cases} \tag{$\widehat{\mathrm{P}}_\lambda$}$$

Compared with ($\mathrm{P}_\lambda$) and ($\widetilde{\mathrm{P}}_\lambda$), ($\widehat{\mathrm{P}}_\lambda$) separates the smooth and nonsmooth terms, which makes it more efficiently to take advantage of ADMM.

The ADMM type algorithm has been used in elliptic optimal control problem with control constraints. In [39], the authors proposed a heterogeneous ADMM (hADMM) algorithm. The hADMM algorithm employs two different weighted norms in the augmented term in two subproblems respectively, which is different from the classical ADMM. Also, the authors proved the global convergence and the iteration complexity results $o(\frac{1}{k})$. Inspired by the simpleness, facility for implementation and global convergence rate of the hADMM, we employ it to fully discretized Lavrentiev-regularized problem. Although Lavrentiev-regularized problem can be transformed into a pure control-constrained problem as form ($\widetilde{\mathrm{P}}_\lambda$), it will become ill-conditioned when lambda is very small. Thus we do not apply hADMM to ($\widetilde{\mathrm{P}}_\lambda$), we use its well structure as reference and apply it to ($\widehat{\mathrm{P}}_\lambda$), which possesses well structure as we see in Section 4. For the first subproblem of hADMM, it is equivalent to solve a $2 * 2$ block equation system in each iteration, while using PDAS has to solve a $4 * 4$ block equation system which should be carefully formed based on the active sets in each iteration. For the second subproblem of hADMM, the solution has a closed form, which is very easy to compute.

Moreover, to satisfy the need for more accurate solution, a two-phase strategy is also presented, in which the primal-dual active set (PDAS) method is used as a postprocessor of the hADMM algorithm. It is shown in Section 5 that to get a solution of the same precision, the hADMM algorithm and the two-phase strategy are obviously faster than PDAS method respectively.

The paper is organized as follows. Full discretization is considered in Section 2. Section 3 gives the error estimates of the fully discretized Lavrentiev-regularized problem. In Section 4, we give the frame of the hADMM algorithm and the PDAS method employed to the discretized problems. Two numerical examples are given to verify the error estimates and the efficiency of the proposed algorithm in Section 5. Section 6 contains a brief summary of this paper.

## 2. Full finite element discretization

In order to tackle ($\mathrm{P}_\lambda$) and ($\widehat{\mathrm{P}}_\lambda$) numerically, we consider the full discretization, in which both the state $y$ and the control $u$ are discretized by continuous piecewise linear functions. Let us introduce a family of regular triangulations $\{T_h\}_{h>0}$ of $\Omega$, i.e. $\overline{\Omega} = \bigcup_{T \in T_h} \overline{T}$. With each element $T \in T_h$, we associate two parameters $\rho(T)$

and $R(T)$, where $\rho(T)$ denotes the diameter of the set $T$ and $R(T)$ is the diameter of the largest ball contained in $T$. The mesh size of $T_h$ is defined by $h = \max_{T\in T_h} \rho(T)$. For the upcoming error analysis, we suppose the following standard assumption holds (see [23], [24]).

**Assumption 2.1.** (Regular and quasi-uniform triangulations) The domain $\Omega$ is a open bounded and convex subset of $R^n$, $n = 2, 3$ and its boundary $\Gamma$ is a polygon ($n = 2$) or a polyhedron ($n = 3$). Moreover, there exist two positive constants $\rho$ and $R$ such that

$$\frac{\rho(T)}{R(T)} \le R, \qquad \frac{h}{\rho(T)} \le \rho$$

hold for all $T \in T_h$ and all $h > 0$. Let us define $\overline{\Omega}_h = \bigcup_{T\in T_h} T$, and let $\Omega_h \in \Omega$ and $\Gamma_h$ denote its interior and its boundary, respectively. In the case that $\Omega$ has a $C^{1,1}$-boundary $\Gamma$, we assume that $\overline{\Omega}_h$ is a convex and that all boundary vertices of $\overline{\Omega}_h$ are contained in $\Gamma$, such that

$$|\Omega\backslash\Omega_h| \le ch^2,$$

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

For domains satisfying Assumption 2.1, the following result holds (see [18], Thm. 4.3.1.4, [37], Thm. 2.2.3).

**Lemma 2.2.** *Suppose that $\Omega$ fulfills the condition in Assumption 2.1 and let $f$ be a given function in $L^2(\Omega)$, while $z$ solves*

$$\begin{aligned} -\Delta z &= f \qquad \text{in } \Omega, \\ z &= 0 \qquad \text{on } \Gamma. \end{aligned}$$

*Then $z \in H^2(\Omega)$ and the estimate*

$$\|z\|_{H^2(\Omega)} \le c\|f\|$$

*holds true with a constant $c$ independent of $f$ and $h$.*

The weak formulation of the state equation involved in ($\mathrm{P}_\lambda$) and ($\widehat{\mathrm{P}}_\lambda$)

$$\begin{aligned} -\Delta y &= u \qquad \text{in } \Omega, \\ y &= 0 \qquad \text{on } \Gamma \end{aligned} \tag{2.1}$$

is given by

$$(\nabla y, \nabla z) = (u, z), \qquad \forall z \in H_0^1(\Omega). \tag{2.2}$$

Let a finite dimensional subspace $Z_h$ of $H_0^1(\Omega)$

$$Z_h = \left\{ z_h \in C(\overline{\Omega}) \mid z_h|_T \in P_1 \quad \forall T \in T_h \textit{ and } z_h = 0 \textit{ in } \overline{\Omega}\backslash\Omega_h \right\} \tag{2.3}$$

be the discrete space, where $\mathcal{P}_1$ denotes the space of polynomials whose degree are less than or equal to 1. Let $\{\phi_i(x)\}_{i=1}^{N_h}$ be a basis of $Z_h$ which satisfies the following properties:

$$\phi_i(x) \ge 0, \qquad \|\phi_i(x)\|_\infty = 1, \quad \forall i = 1, 2, ..., N_h, \qquad \sum_{i=1}^{N_h} \phi_i(x) = 1, \quad \forall x \in \Omega_h, \tag{2.4}$$

then (2.2) implies that the weak formulation is satisfied for all basis functions $\{\phi_i(x)\}_{i=1}^{N_h}$, i.e.

$$(\nabla y, \nabla \phi_i) = (u, \phi_i), \qquad \forall i = 1, \cdots, N_h. \tag{2.5}$$

We know from (3.11) that $\bar{v}_\lambda = P_{V_{ad}}(\bar{y}_\lambda - \frac{\lambda}{\alpha} p_\lambda)$, which together with $\bar{u}_\lambda = \frac{\bar{v}_\lambda - \bar{y}_\lambda}{\lambda}$ implies that $\bar{u}_\lambda$ satifies the zero boundary condition. Thus we discretize $y(x)$ and $u(x)$ by the same basis of $Z_h$, i.e.

$$y_h(x) = \sum_{i=1}^{N_h} y_i \phi_i(x) \text{ and } u_h(x) = \sum_{i=1}^{N_h} u_i \phi_i(x), \tag{2.6}$$

where $y_h(x_i) = y_i$ and $u_h(x_i) = u_i$. Then the discrete version of problem ($\mathrm{P}_\lambda$), ($\widetilde{\mathrm{P}}_\lambda$) and ($\widehat{\mathrm{P}}_\lambda$) are denoted by ($\mathrm{P}_{\lambda,h}$), ($\widetilde{\mathrm{P}}_{\lambda,h}$) and ($\widehat{\mathrm{P}}_{\lambda,h}$) respectively,

$$\begin{cases} \min \ \ J_h(y_h, u_h) = \frac{1}{2}\|y_h - y_d\|^2_{L^2(\Omega_h)} + \frac{\alpha}{2}\|u_h\|^2_{L^2(\Omega_h)} \\ \text{s.t.} \ \ (\nabla y_h, \nabla z_h) = (u_h, z_h) \ \ \forall z_h \in Z_h, \\ \qquad a \le \lambda u_h(x) + y_h(x) \le b \quad \text{on } \Omega_\mathrm{h}, \end{cases} \tag{$\mathrm{P}_{\lambda,h}$}$$

$$\begin{cases} \min \ \ J_h(y_h, v_h) = \frac{1}{2}\|y_h - y_d\|^2_{L^2(\Omega_h)} + \frac{\alpha}{2\lambda^2}\|v_h - y_h\|^2_{L^2(\Omega_h)} \\ \text{s.t.} \ \ (\nabla y_h, \nabla z_h) + \frac{1}{\lambda}(y_h, z_h) = \frac{1}{\lambda}(v_h, z_h) \ \ \forall z_h \in Z_h, \\ \qquad a \le v_h(x) \le b \quad \text{on } \Omega_\mathrm{h}, \end{cases} \tag{$\widetilde{\mathrm{P}}_{\lambda,h}$}$$

$$\begin{cases} \min \ \ J_h(y_h, u_h) = \frac{1}{2}\|y_h - y_d\|^2_{L^2(\Omega_h)} + \frac{\alpha}{2}\|u_h\|^2_{L^2(\Omega_h)} \\ \text{s.t.} \ \ (\nabla y_h, \nabla z_h) = (u_h, z_h) \ \ \forall z_h \in Z_h, \\ \qquad v_h - \lambda u_h - y_h = 0 \quad \text{on } \Omega_\mathrm{h}, \\ \qquad a \le v_h(x) \le b \quad \text{on } \Omega_\mathrm{h}. \end{cases} \tag{$\widehat{\mathrm{P}}_{\lambda,h}$}$$

## 3. Error estimates

In this section, we extend the results of [23]. The essential difference between [23] and the present paper is that the discretization method in [23] is variational discretization while this paper considers full discretization, in which both the state and control are discretized by piecewise linear functions. The greatest difficulty that full discretization introduces to the error analysis is that the solution of continuous problem is not feasible for discretized problem. To tackle with this difficulty, we utilize the quasi-interpolation operator and complete the error analysis. It is well known that since projection has to be carried out to get the control in each iteration in variational discretization, which means mesh refinement for the control, the error order of the control of variational discretization is generally higher than that of full discretization. However, the error analysis in this section indicates that the error order of the control of full discretization is not inferior to that of variational discretization because of the employment of the Lavrentiev-regularization. In this section, we give two different error estimates, the first one of which depends on $\lambda$ while the second one of which is uniform in $\lambda$.

### 3.1. Error estimate for fixed $\lambda$

For the error analysis below, we have to use a quasi-interpolation operator $\Pi_h : L^2(\Omega) \to Z_h$, which is defined by

$$\Pi_h v = \sum_{i=1}^{N_h} \pi_i(v)\phi_i(x), \quad \pi_i(v) = \frac{\int_{\Omega_h} v(x)\phi_i(x)dx}{\int_{\Omega_h} \phi_i(x)dx}, \quad \forall v \in L^2(\Omega).$$

Let

$$V_{ad} = \{v \in L^2(\Omega) \mid a \le v \le b \ \text{ a.e. on } \Omega\}$$

and

$$V_{ad,h} = \{v_h = \sum_{i=1}^{N_h} v_i\phi_i(x) | \ a \le v_i \le b \quad \text{on } \Omega_\mathrm{h}\},$$

then there holds

$$v \in V_{ad} \Rightarrow \Pi_h v \in V_{ad,h}, \quad \forall v \in L^2(\Omega).$$

For the interpolation error, the following lemma holds, whose proof can be found in [6, 13].

**Lemma 3.1.** *There exists a constant $C$ independent of $h$ such that*

$$h\|v - \Pi_h v\|_{L^2} + \|v - \Pi_h v\|_{H^{-1}} \le Ch^2\|v\|_{H^1} \qquad \forall v \in H^1(\Omega).$$

First we consider the following variational equation

$$(\nabla w, \nabla z) + \frac{1}{\lambda}(w, z) = (g, z), \qquad \forall z \in H_0^1(\Omega) \tag{3.1}$$

and its discrete version:

$$(\nabla w_h, \nabla z_h) + \frac{1}{\lambda}(w_h, z_h) = (g, z_h), \qquad \forall z_h \in Z_h, \tag{3.2}$$

where $g \in L^2(\Omega)$. We use $w(g)$ and $w_h(g)$ to denote the solution of (3.1) and (3.2) respectively, then the following lemma holds.

**Lemma 3.2.** *Under Assumption 2.1, there exists a constant $C(\Omega)$ independent of $\lambda$ such that*

$$\|w_h(g) - w(g)\|_{L^2(\Omega)} \le C(\Omega)\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right)\|w(g)\|_{H^2(\Omega)}$$

*holds true.*

*Proof.* Let $z = w_h(g) - I_h w(g)$ in (3.1) and $z_h = w_h(g) - I_h w(g)$ in (3.2), then we get

$$(\nabla w(g), \nabla(w_h(g) - I_h w(g))) + \frac{1}{\lambda}(w(g), w_h(g) - I_h w(g)) = (g, w_h(g) - I_h w(g)),$$
$$(\nabla w_h(g), \nabla(w_h(g) - I_h w(g))) + \frac{1}{\lambda}(w_h(g), w_h(g) - I_h w(g)) = (g, w_h(g) - I_h w(g)),$$

where $I_h$ denotes the linear interpolation operator. Subtracting two equalities above, we arrive at

$$(\nabla(w_h(g)-w(g)),\nabla(w_h(g)-I_hw(g)))+\frac{1}{\lambda}(w_h(g)-w(g),w_h(g)-I_hw(g))=0, \tag{3.3}$$

so

$$\begin{aligned}
&\|w_h(g)-w(g)\|^2_{H^1(\Omega)}\\
&\leq (\nabla(w_h(g)-w(g)),\nabla(w_h(g)-w(g)))+\frac{1}{\lambda}\left(w_h(g)-w(g),w_h(g)-w(g)\right)\\
&=(\nabla(w_h(g)-w(g)),\nabla(I_hw(g)-w(g)))+\frac{1}{\lambda}\left(w_h(g)-w(g),I_hw(g)-w(g)\right)\\
&\leq \|\nabla(w_h(g)-w(g))\|\cdot\|\nabla(I_hw(g)-w(g))\|+\frac{1}{\lambda}\|w_h(g)-w(g)\|\cdot\|I_hw(g)-w(g)\|\\
&\leq \|w_h(g)-w(g)\|_{H^1(\Omega)}\cdot\|I_hw(g)-w(g)\|_{H^1(\Omega)}+\frac{1}{\lambda}\|w_h(g)-w(g)\|_{H^1(\Omega)}\cdot\|I_hw(g)-w(g)\|,
\end{aligned} \tag{3.4}$$

where we have used $\frac{1}{\lambda}>1$ .Then we arrive at

$$\begin{aligned}
\|w_h(g)-w(g)\|_{H^1(\Omega)} &\leq \|I_hw(g)-w(g)\|_{H^1(\Omega)}+\frac{1}{\lambda}\|I_hw(g)-w(g)\|\\
&\leq C(\Omega)\left(h+\frac{1}{\lambda}h^2\right)\|w(g)\|_{H^2(\Omega)},
\end{aligned} \tag{3.5}$$

where we used the standard interpolation error estimates, which can be derived from ([33], Thm. 3.1.6). To achieve further estimates on $\|w(g)-w_h(g)\|_{L^2}$, we first let $\phi$ be the solution of

$$(\nabla\phi,\nabla z)=(w-w_h,z),\qquad \forall z\in H_0^1(\Omega), \tag{3.6}$$

then we have

$$\frac{1}{\lambda}(w-w_h,\phi)=\frac{1}{\lambda}(\nabla\phi,\nabla\phi)\geq 0. \tag{3.7}$$

We recall that the following equation holds

$$(\nabla(w-w_h),\nabla z_h)+\frac{1}{\lambda}(w-w_h,z_h)=0,\qquad \forall z_h\in Z_h. \tag{3.8}$$

Let $z=w-w_h$ in (3.6) and $z_h=I_h\phi$ in (3.8), we arrive at

$$\begin{aligned}
\|w-w_h\|^2 &= (\nabla\phi,\nabla(w-w_h))-(\nabla I_h\phi,\nabla(w-w_h))-\frac{1}{\lambda}(I_h\phi,w-w_h)\\
&\leq (\nabla(\phi-I_h\phi),\nabla(w-w_h))-\frac{1}{\lambda}(I_h\phi,w-w_h)+\frac{1}{\lambda}(\phi,w-w_h)\\
&=(\nabla(\phi-I_h\phi),\nabla(w-w_h))+\frac{1}{\lambda}(\phi-I_h\phi,w-w_h)\\
&\leq \|\phi-I_h\phi\|_{H^1}\cdot\|w-w_h\|_{H^1}+\frac{1}{\lambda}\|\phi-I_h\phi\|\cdot\|w-w_h\|\\
&\leq Ch\|\phi\|_{H^2}\cdot\|w-w_h\|_{H^1}+\frac{1}{\lambda}Ch^2\|\phi\|_{H^2}\cdot\|w-w_h\|_{H^1}
\end{aligned}$$

$$\leq Ch\|w-w_h\|\cdot\|w-w_h\|_{H^1}+\frac{1}{\lambda}Ch^2\|w-w_h\|\cdot\|w-w_h\|_{H^1}, \tag{3.9}$$

where we have used (3.7) in the first inequality, standard interpolation error estimates (see [[33], Thm. 3.1.6) in the third inequality and Lemma 2.2 in the last inequality. Then we arrive at

$$\begin{aligned}\|w-w_h\| &\leq C(h+\frac{1}{\lambda}h^2)\|w-w_h\|_{H^1}\\ &\leq C(\Omega)(h+\frac{1}{\lambda}h^2)\left(h+\frac{1}{\lambda}h^2\right)\|w(g)\|_{H^2(\Omega)}\\ &\leq C(\Omega)\left(h^2+\frac{1}{\lambda}h^3+\frac{1}{\lambda^2}h^4\right)\|w(g)\|_{H^2(\Omega)}.\end{aligned} \tag{3.10}$$

□

Let $(\overline{y}_\lambda,\overline{v}_\lambda)$ and $(\overline{y}_{\lambda,h},\overline{v}_{\lambda,h})$ be the solutions of ($\widetilde{\mathrm{P}}_\lambda$) and ($\widetilde{\mathrm{P}}_{\lambda,h}$) respectively, then the optimal system of ($\widetilde{\mathrm{P}}_\lambda$) is:

$$(\nabla\overline{y}_\lambda,\nabla z)+\frac{1}{\lambda}(\overline{y}_\lambda,z)=\frac{1}{\lambda}(\overline{v}_\lambda,z),\qquad \forall z\in H_0^1(\Omega), \tag{3.11a}$$

$$(\nabla p_\lambda,\nabla z)+\frac{1}{\lambda}(p_\lambda,z)=(\overline{y}_\lambda-y_d+\frac{\alpha}{\lambda^2}(\overline{y}_\lambda-\overline{v}_\lambda),z),\quad \forall z\in H_0^1(\Omega), \tag{3.11b}$$

$$\overline{v}_\lambda\in V_{ad},\qquad (\overline{v}_\lambda-\overline{y}_\lambda+\frac{\lambda}{\alpha}p_\lambda,v-\overline{v}_\lambda)\geq 0,\qquad \forall v\in V_{ad}, \tag{3.11c}$$

where $p_\lambda$ denotes the adjoint state. And the optimal system of ($\widetilde{\mathrm{P}}_{\lambda,h}$) is:

$$(\nabla\overline{y}_{\lambda,h},\nabla z_h)+\frac{1}{\lambda}(\overline{y}_{\lambda,h},z_h)=\frac{1}{\lambda}(\overline{v}_{\lambda,h},z_h),\qquad \forall z_h\in Z_h, \tag{3.12a}$$

$$(\nabla p_{\lambda,h},\nabla z_h)+\frac{1}{\lambda}(p_{\lambda,h},z_h)=(\overline{y}_{\lambda,h}-y_d+\frac{\alpha}{\lambda^2}(\overline{y}_{\lambda,h}-\overline{v}_{\lambda,h}),z_h),\ \forall z_h\in Z_h, \tag{3.12b}$$

$$\overline{v}_{\lambda,h}\in V_{ad,h},\qquad (\overline{v}_{\lambda,h}-\overline{y}_{\lambda,h}+\frac{\lambda}{\alpha}p_{\lambda,h},v-\overline{v}_{\lambda,h})\geq 0,\qquad \forall v\in V_{ad,h}, \tag{3.12c}$$

where $p_{\lambda,h}$ denotes the adjoint state. Additionally, Let $y(v)$, $y_h(v)$, $p(v)$, $p^h(v)$ and $p_h(v)$ be the solution of

$$(\nabla y,\nabla z)+\frac{1}{\lambda}(y,z)=\frac{1}{\lambda}(v,z),\quad \forall z\in H_0^1(\Omega), \tag{3.13}$$

$$(\nabla y_h,\nabla z_h)+\frac{1}{\lambda}(y_h,z_h)=\frac{1}{\lambda}(v,z_h),\quad \forall z_h\in Z_h, \tag{3.14}$$

$$(\nabla p,\nabla z)+\frac{1}{\lambda}(p,z)=(y(v)-y_d+\frac{\alpha}{\lambda^2}(y(v)-v),z),\quad \forall z\in H_0^1(\Omega), \tag{3.15}$$

$$(\nabla p^h,\nabla z_h)+\frac{1}{\lambda}(p^h,z_h)=(y(v)-y_d+\frac{\alpha}{\lambda^2}(y(v)-v),z_h),\quad \forall z_h\in Z_h, \tag{3.16}$$

$$(\nabla p_h,\nabla z_h)+\frac{1}{\lambda}(p_h,z_h)=(y_h(v)-y_d+\frac{\alpha}{\lambda^2}(y_h(v)-v),z_h),\quad \forall z_h\in Z_h, \tag{3.17}$$

respectively, then we have $\overline{y}_\lambda=y(\overline{v}_\lambda)$, $p_\lambda=p(\overline{v}_\lambda)$, $\overline{y}_{\lambda,h}=y_h(\overline{v}_{\lambda,h})$, $p_{\lambda,h}=p_h(\overline{v}_{\lambda,h})$. The following corollary can be easily derived from Lemma 3.2.

**Corollary 3.3.** *Suppose that Assumption 2.1 is fullfilled. Then there exists a constant $C(\Omega)$ independent of $\lambda$ such that the following estimate is valid*

$$\|y_h(\overline{v}_\lambda) - \overline{y}_\lambda\| \leq C(\Omega)\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right) \tag{3.18}$$

*In addition*

$$\lambda\|p^h(\overline{v}_\lambda) - p_\lambda\| \leq C(\alpha,\Omega)\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right) \tag{3.19}$$

*holds true with a constant $C(\alpha,\Omega)$ independent of $\lambda$.*

*Proof.* By constructions, $\overline{y}_\lambda = y_\lambda(\overline{v}_\lambda)$ is also the solution of the state equation in $(P_\lambda)$ with $\overline{u}_\lambda = \frac{1}{\lambda}(\overline{v}_\lambda - \overline{y}_\lambda)$ on the right hand side, i.e. it solves (2.1) with $\overline{u}_\lambda$ as inhomogeneity. Therefore, Lemma 2.2 yields

$$\|\overline{y}_\lambda\|_{H^2(\Omega)} \leq c\|\overline{u}_\lambda\| \leq c,$$

where the optimality of $\overline{u}_\lambda$ guarantees its uniform boundedness w.r.t. $\lambda$ in $L^2(\Omega)$. Together with Lemma 3.2, this implies (3.18).

Moreover, again due to $\overline{u}_\lambda = \frac{1}{\lambda}(\overline{v}_\lambda - \overline{y}_\lambda)$, the adjoint state solves

$$\begin{aligned} -\Delta p &= \overline{y}_\lambda - y_d - \frac{1}{\lambda}p + \frac{\alpha}{\lambda}\overline{u}_\lambda && \text{in } \Omega,\\ p &= 0 && \text{on } \Gamma, \end{aligned}$$

and hence, again by Lemma 2.2,

$$\lambda\|p_\lambda\|_{H^2(\Omega)} \leq c(\lambda\|\overline{y}\| + \lambda\|y_d\| + \alpha\|\overline{u}\| + \|p_\lambda\|)$$

follows with a constant $c$ independent of $\lambda$. Thanks to their optimality, $\overline{u}_\lambda$ and $\overline{y}_\lambda$ are uniformly bounded in $L^2(\Omega)$ independent of $\lambda$. Moreover, consider again an arbitrary sequence $\lambda_n$ tending to zero for $n \to \infty$. Then, from [23], we know that the associated sequence of adjoint states converges weakly in $L^2(\Omega)$, giving in turn its uniform boundedness such that $\|p_\lambda\| \leq c$ independent of $\lambda$. Thus, we obtain $\lambda\|p_\lambda\|_{H^2(\Omega)} \leq c$ and consequently, Lemma 3.2 gives the assertion.

□

**Theorem 3.4.** *Suppose that Assumption 2.1 is fulfilled. Let $(\overline{y}_\lambda, \overline{v}_\lambda)$ and $(\overline{y}_{\lambda,h}, \overline{v}_{\lambda,h})$ be the solutions of $(\widetilde{\mathrm{P}}_\lambda)$ and $(\widetilde{\mathrm{P}}_{\lambda,h})$ respectively, then there exists a constant $C(\alpha,\Omega,\lambda_{max})$ independent of $\lambda$ such that*

$$\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\| + \|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|_{H^1(\Omega)} \leq C(\alpha,\Omega,\lambda_{max})\left(\frac{1}{\sqrt{\lambda}}h + \frac{1}{\lambda^2}\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right)\right)$$

*is satisfied.*

*Proof.* Because the solution $\overline{v}_{\lambda,h}$ of $(\widetilde{\mathrm{P}}_{\lambda,h})$ is feasible for $(\widetilde{\mathrm{P}}_\lambda)$, we can insert $\overline{v}_{\lambda,h}$ in (3.11c), which gives

$$(\overline{v}_\lambda - \overline{y}_\lambda + \frac{\lambda}{\alpha}p_\lambda, \overline{v}_{\lambda,h} - \overline{v}_\lambda) \geq 0. \tag{3.20}$$

Let $\widetilde{v} = \Pi_h \overline{v}_\lambda$, where $\Pi_h$ is the quasi-interpolation operator defined above. Then $\widetilde{v}$ is feasible for ($\widetilde{\mathrm{P}}_{\lambda,h}$) and we can insert $\widetilde{v}$ in (3.12c), which gives

$$\left(\overline{v}_{\lambda,h} - \overline{y}_{\lambda,h} + \frac{\lambda}{\alpha} p_{\lambda,h}, \widetilde{v} - \overline{v}_\lambda\right) + \left(\overline{v}_{\lambda,h} - \overline{y}_{\lambda,h} + \frac{\lambda}{\alpha} p_{\lambda,h}, \overline{v}_\lambda - \overline{v}_{\lambda,h}\right) \geq 0. \tag{3.21}$$

Adding (3.20) and (3.21) then yields

$$\left(\overline{v}_{\lambda,h} - \overline{y}_{\lambda,h} + \frac{\lambda}{\alpha} p_{\lambda,h}, \widetilde{v} - \overline{v}_\lambda\right) + \left(\overline{v}_\lambda - \overline{v}_{\lambda,h} - (\overline{y}_\lambda - \overline{y}_{\lambda,h}) + \frac{\lambda}{\alpha}(p_\lambda - p_{\lambda,h}\right), \overline{v}_{\lambda,h} - \overline{v}_\lambda) \geq 0.$$

We can rewrite the inequality above into the following form:

$$\begin{aligned}
0 \leq {} & (\overline{v}_{\lambda,h} - \overline{y}_{\lambda,h} + \frac{\lambda}{\alpha} p_{\lambda,h}, \widetilde{v} - \overline{v}_\lambda) - \| \overline{v}_\lambda - \overline{v}_{\lambda,h} \|^2 + (y_h(\overline{v}_\lambda) - \overline{y}_\lambda, \overline{v}_{\lambda,h} - \overline{v}_\lambda) \\
& + \frac{\lambda}{\alpha}(p_\lambda - p^h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda) + \underbrace{\frac{\lambda}{\alpha}(p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda)}_{I_1} \\
& + \underbrace{(\overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda) + \frac{\lambda}{\alpha}(p_h(\overline{v}_\lambda) - p_{\lambda,h}, \overline{v}_{\lambda,h} - \overline{v}_\lambda)}_{I_2}.
\end{aligned} \tag{3.22}$$

Let $v = \overline{v}_{\lambda,h},\ z_h = p_h(\overline{v}_\lambda) - p_{\lambda,h} \in Z_h$ and $v = \overline{v}_\lambda,\ z_h = p_h(\overline{v}_\lambda) - p_{\lambda,h} \in Z_h$ in (3.14) respectively. Subtracting the two resulted equalities we get

$$\begin{aligned}
& (\nabla \overline{y}_{\lambda,h} - \nabla y_h(\overline{v}_\lambda), \nabla p_h(\overline{v}_\lambda) - \nabla p_{\lambda,h}) + \frac{1}{\lambda}(\overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda), p_h(\overline{v}_\lambda) - p_{\lambda,h}) \\
& = \frac{1}{\lambda}(\overline{v}_{\lambda,h} - \overline{v}_\lambda, p_h(\overline{v}_\lambda) - p_{\lambda,h}).
\end{aligned}$$

Let $v = \overline{v}_\lambda,\ z_h = \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda) \in Z_h$ and $v = \overline{v}_{\lambda,h},\ z_h = \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda) \in Z_h$ in (3.17) respectively. Subtracting the two resulted equalities we arrive at

$$\begin{aligned}
& (\nabla p_h(\overline{v}_\lambda) - \nabla p_{\lambda,h}, \nabla \overline{y}_{\lambda,h} - \nabla \overline{y}_h(\overline{v}_\lambda)) + \frac{1}{\lambda}(p_h(\overline{v}_\lambda) - p_{\lambda,h}, \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)) \\
& = (y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h} + \frac{\alpha}{\lambda^2}(y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h} - \overline{v}_\lambda + \overline{v}_{\lambda,h}), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)).
\end{aligned}$$

So we have

$$\frac{1}{\lambda}(\overline{v}_{\lambda,h} - \overline{v}_\lambda, p_h(\overline{v}_\lambda) - p_{\lambda,h}) = (y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h} + \frac{\alpha}{\lambda^2}(y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h} - \overline{v}_\lambda + \overline{v}_{\lambda,h}), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)).$$

Then we can rewrite $I_2$ in (3.22) as

$$\begin{aligned}
I_2 &= (\overline{y}_\lambda - y_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda) + \frac{\lambda}{\alpha}(p_h(\overline{v}_\lambda) - p_{\lambda,h}, \overline{v}_{\lambda,h} - \overline{v}_\lambda) \\
&= (\overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda) + \frac{\lambda^2}{\alpha}(y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h} + \frac{\alpha}{\lambda^2}(y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h} - \overline{v}_\lambda + \overline{v}_{\lambda,h}), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)) \\
&= -(1 + \frac{\lambda^2}{\alpha})\|y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h}\|^2 + 2(\overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda).
\end{aligned} \tag{3.23}$$

Similarly let $v = \overline{v}_\lambda, z_h = \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda) \in Z_h$ in (3.16) and (3.17) respectively. Subtracting the two resulted equalities we derive

$$\begin{aligned}&(\nabla p^h(\overline{v}_\lambda) - \nabla p_h(\overline{v}_\lambda), \nabla\overline{y}_{\lambda,h} - \nabla y_h(\overline{v}_\lambda)) + \frac{1}{\lambda}(p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda))\\ &= (y(\overline{v}_\lambda) - y_h(\overline{v}_\lambda) + \frac{\alpha}{\lambda^2}(y(\overline{v}_\lambda) - y_h(\overline{v}_\lambda)), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)).\end{aligned}$$

Let $v = \overline{v}_{\lambda,h}, z_h = p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda) \in Z_h$ and $v = \overline{v}_\lambda, z_h = p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda) \in Z_h$ in (3.14) respectively. Subtracting the two resulted equalities we have

$$\begin{aligned}&(\nabla\overline{y}_{\lambda,h} - \nabla y_h(\overline{v}_\lambda), \nabla p^h(\overline{v}_\lambda) - \nabla p_h(\overline{v}_\lambda)) + \frac{1}{\lambda}(\overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda), p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda))\\ &= \frac{1}{\lambda}(\overline{v}_{\lambda,h} - \overline{v}_\lambda, p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda)).\end{aligned}$$

So we arrive at

$$\frac{1}{\lambda}(\overline{v}_{\lambda,h} - \overline{v}_\lambda, p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda)) = (y(\overline{v}_\lambda) - y_h(\overline{v}_\lambda) + \frac{\alpha}{\lambda^2}(y(\overline{v}_\lambda) - y_h(\overline{v}_\lambda)), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)).$$

Then we can rewrite $I_1$ in (3.22) as

$$\begin{aligned}I_1 &= \frac{\lambda}{\alpha}(p^h(\overline{v}_\lambda) - p_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda)\\ &= \frac{\lambda^2}{\alpha}(\overline{y}_\lambda - y_h(\overline{v}_\lambda) + \frac{\alpha}{\lambda^2}(\overline{y}_\lambda - y_h(\overline{v}_\lambda)), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda))\\ &= (1 + \frac{\lambda^2}{\alpha})(\overline{y}_\lambda - y_h(\overline{v}_\lambda), \overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda)).\end{aligned} \tag{3.24}$$

Inserting (3.23) and (3.24) into (3.22), we get

$$\begin{aligned}0 \le &- \|\overline{v}_\lambda - \overline{v}_{\lambda,h}\|^2 + (y_h(\overline{v}_\lambda) - \overline{y}_\lambda, \overline{v}_{\lambda,h} - \overline{v}_\lambda) + \frac{\alpha}{\lambda}(p_\lambda - p^h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda)\\ &- (1 + \frac{\lambda^2}{\alpha})(y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h}, y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h}) + 2(\overline{y}_{\lambda,h} - y_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda)\\ &- (1 + \frac{\lambda^2}{\alpha})(\overline{y}_\lambda - y_h(\overline{v}_\lambda), y_h(\overline{v}_\lambda) - \overline{y}_{\lambda,h}) + (\overline{v}_{\lambda,h} - \overline{y}_{\lambda,h} + \frac{\lambda}{\alpha}p_{\lambda,h}, \widetilde{v} - \overline{v}_\lambda)\\ = &- [\|\overline{v}_\lambda - \overline{v}_{\lambda,h}\|^2 - 2(\overline{y}_\lambda - \overline{y}_{\lambda,h}, \overline{v}_\lambda - \overline{v}_{\lambda,h}) + \|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2] - \frac{\lambda^2}{\alpha}\|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2\\ &+ (\overline{y}_\lambda - y_h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda) + \frac{\alpha}{\lambda}(p_\lambda - p^h(\overline{v}_\lambda), \overline{v}_{\lambda,h} - \overline{v}_\lambda)\\ &- (1 + \frac{\lambda^2}{\alpha})(\overline{y}_\lambda - \overline{y}_{\lambda,h}, y_h(\overline{v}_\lambda) - \overline{y}_\lambda) + (\overline{v}_{\lambda,h} - \overline{y}_{\lambda,h} + \frac{\lambda}{\alpha}p_{\lambda,h}, \widetilde{v} - \overline{v}_\lambda)\\ = &- \lambda^2\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 - \frac{\lambda^2}{\alpha}\|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2\\ &+ (\overline{y}_\lambda - y_h(\overline{v}_\lambda), \lambda(\overline{u}_{\lambda,h} - \overline{u}_\lambda) + (\overline{y}_{\lambda,h} - \overline{y}_\lambda)) + \frac{\lambda}{\alpha}(p_\lambda - p^h(\overline{v}_\lambda), \lambda(\overline{u}_{\lambda,h} - \overline{u}_\lambda)\\ &+ (\overline{y}_{\lambda,h} - \overline{y}_\lambda)) - (1 + \frac{\lambda^2}{\alpha})(y_h(\overline{v}_\lambda) - \overline{y}_\lambda, \overline{y}_\lambda - \overline{y}_{\lambda,h}) + \frac{\lambda}{\alpha}(\alpha\overline{u}_{\lambda,h} + p_{\lambda,h}, \widetilde{v} - \overline{v}_\lambda).\end{aligned}$$

So we derive

$$\alpha\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2+\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|^2\le\frac{\alpha}{\lambda}(\overline{y}_\lambda-y_h(\overline{v}),\overline{u}_{\lambda,h}-\overline{u}_\lambda)+(p_\lambda-p^h(\overline{v}_\lambda),\overline{u}_{\lambda,h}-\overline{u}_\lambda)$$
$$+\frac{1}{\lambda}(p_\lambda-p^h(\overline{v}_\lambda),\overline{y}_{\lambda,h}-\overline{y}_\lambda)+(y_h(\overline{v}_\lambda)-\overline{y}_\lambda,\overline{y}_{\lambda,h}-\overline{y}_\lambda)$$
$$+\frac{1}{\lambda}(\alpha\overline{u}_{\lambda,h}+p_{\lambda,h},\widetilde{v}-\overline{v}_\lambda).$$

Using Young's inequality we get

$$(\alpha-2k)\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2+(1-2k)\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|^2$$
$$\le\left(\frac{\alpha^2}{k\lambda^2}+\frac{1}{k}\right)\|\overline{y}_\lambda-y_h(\overline{v}_\lambda)\|^2+\left(\frac{1}{k\lambda^2}+\frac{1}{k\lambda^4}\right)\lambda^2\|p_\lambda-p^h(\overline{v}_\lambda)\|^2$$
$$+\frac{1}{\lambda}\|\alpha\overline{u}_{\lambda,h}+p_{\lambda,h}\|_{H^1(\Omega)}\cdot\|\widetilde{v}-\overline{v}_\lambda\|_{H^{-1}(\Omega)},$$

with $k>0$ arbitrary. Then Corollary 3.3 and Lemma 3.1 yield

$$(\alpha-2k)\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2+(1-2k)\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|^2$$
$$\le C(\alpha,\Omega,\lambda_{max})\left[\frac{1}{k\lambda^4}\left(h^2+\frac{1}{\lambda}h^3+\frac{1}{\lambda^2}h^4\right)^2+\frac{1}{\lambda}h^2\right]. \tag{3.25}$$

Let $k=\frac{1}{4}\min(\alpha,1)$ to make $\alpha-2k>0$ and $1-2k>0$, then we arrive at

$$\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|\le C(\alpha,\Omega,\lambda_{max})\left(\frac{1}{\sqrt{\lambda}}h+\frac{1}{\lambda^2}\left(h^2+\frac{1}{\lambda}h^3+\frac{1}{\lambda^2}h^4\right)\right). \tag{3.26}$$

For $\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|_{H^1}$, we have $\forall t>0$,

$$\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|_{H^1}^2\le C\{a(\overline{y}_\lambda-\overline{y}_{\lambda,h},\overline{y}_\lambda-y_h(\overline{u}_\lambda))+a(\overline{y}_\lambda-\overline{y}_{\lambda,h},y_h(\overline{u}_\lambda)-\overline{y}_{\lambda,h})\}$$
$$=C\langle\overline{y}_\lambda-y_h(\overline{u}_\lambda),\overline{u}_\lambda-\overline{u}_{\lambda,h}\rangle_{H_0^1,H^{-1}}$$
$$+C\langle y_h(\overline{u}_\lambda)-\overline{y}_\lambda+\overline{y}_\lambda-\overline{y}_{\lambda,h},\overline{u}_\lambda-\overline{u}_{\lambda,h}\rangle_{H_0^1,H^{-1}}$$
$$\le Ct\|\overline{y}_\lambda-y_h(\overline{u}_\lambda)\|_{H^1}^2+\frac{C}{t}\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2+Ct\|\overline{y}_\lambda-y_h(\overline{u}_\lambda)\|_{H^1}^2$$
$$+\frac{C}{t}\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2+Ct\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|_{H^1}^2+\frac{C}{t}\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2$$
$$=2Ct\|\overline{y}_\lambda-y_h(\overline{u}_\lambda)\|_{H^1}^2+Ct\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|_{H^1}^2+\frac{3C}{t}\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2,$$

which implies

$$(1-Ct)\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|_{H^1}^2\le 2Ct\|\overline{y}_\lambda-y_h(\overline{u}_\lambda)\|_{H^1}^2+\frac{3C}{t}\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|^2. \tag{3.27}$$

We choose $t=\frac{1}{2C}$ to make $1-Ct>0$, then we derive

$$\|\overline{y}_\lambda-\overline{y}_{\lambda,h}\|_{H^1}\le\widehat{C}\{\|\overline{y}_\lambda-y_h(\overline{u}_\lambda)\|_{H^1}+\|\overline{u}_\lambda-\overline{u}_{\lambda,h}\|\}. \tag{3.28}$$

We know from standard error estimates that $\|\overline{y}_\lambda - y_h(\overline{u}_\lambda)\|_{H^1} \le Ch\|\overline{u}_\lambda\|$, which together with (3.26) implies

$$\|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|_{H^1} \le C(\alpha, \Omega, \lambda_{max}) \left( h + \frac{1}{\sqrt{\lambda}} h + \frac{1}{\lambda^2} \left( h^2 + \frac{1}{\lambda} h^3 + \frac{1}{\lambda^2} h^4 \right) \right). \tag{3.29}$$

Since $0 < \lambda < 1$, so $h < \frac{1}{\sqrt{\lambda}} h$. Then the term $h$ can be abandoned from the formula above and we can get the assertion.

□

### 3.2. Error estimate uniform in $\lambda$

We now derive an error estimate which does not depend on $\lambda$. Let $(\overline{y}_\lambda, \overline{u}_\lambda)$ and $(\overline{y}_{\lambda,h}, \overline{u}_{\lambda,h})$ denote the solutions of ($\mathrm{P}_\lambda$) and ($\mathrm{P}_{\lambda,h}$) respectively, then the optimal system of ($\mathrm{P}_\lambda$) is:

$$(\nabla \overline{y}_\lambda, \nabla z) = (\overline{u}_\lambda, z) \quad \forall z \in H_0^1(\Omega), \tag{3.30a}$$
$$(\nabla p_\lambda, \nabla z) = (\overline{y}_\lambda - y_d - \mu_a + \mu_b, z), \ \forall z \in H_0^1(\Omega), \tag{3.30b}$$
$$\alpha \overline{u}_\lambda + p_\lambda + \lambda \mu_b - \lambda \mu_a = 0 \quad \text{a.e. in } \Omega, \tag{3.30c}$$
$$(\mu_a, a - \lambda \overline{u}_\lambda - \overline{y}_\lambda) = (\mu_b, \lambda \overline{u}_\lambda + \overline{y}_\lambda - b) = 0, \tag{3.30d}$$
$$\mu_a(x) \ge 0, \quad \mu_b(x) \ge 0 \quad \text{a.e. in } \Omega, \tag{3.30e}$$
$$a \le \lambda \overline{u}_\lambda + \overline{y}_\lambda \le b \quad \text{a.e. in } \Omega, \tag{3.30f}$$

where $p_\lambda$ is the adjoint state and $\mu_a, \ \mu_b$ are Lagrange multipliers associated to the regularized pointwise state constraints in ($\mathrm{P}_\lambda$). Similarly, the optimal system of ($\mathrm{P}_{\lambda,h}$) is:

$$(\nabla \overline{y}_{\lambda,h}, \nabla z_h) = (\overline{u}_{\lambda,h}, z_h) \quad \forall z_h \in Z_h, \tag{3.31a}$$
$$(\nabla p_{\lambda,h}, \nabla z_h) = (\overline{y}_{\lambda,h} - y_d - \mu_{a,h} + \mu_{b,h}, z_h), \ \forall z_h \in Z_h, \tag{3.31b}$$
$$\alpha \overline{u}_{\lambda,h} + p_{\lambda,h} + \lambda \mu_{b,h} - \lambda \mu_{a,h} = 0 \quad \text{a.e. in } \Omega, \tag{3.31c}$$
$$(\mu_{a,h}, a - \lambda \overline{u}_{\lambda,h} - \overline{y}_{\lambda,h}) = (\mu_{b,h}, \lambda \overline{u}_{\lambda,h} + \overline{y}_{\lambda,h} - b) = 0, \tag{3.31d}$$
$$\mu_{a,h}(x) \ge 0, \quad \mu_{b,h}(x) \ge 0 \quad \text{a.e. in } \Omega, \tag{3.31e}$$
$$a \le \lambda \overline{u}_{\lambda,h} + \overline{y}_{\lambda,h} \le b \quad \text{a.e. in } \Omega, \tag{3.31f}$$

where $p_{\lambda,h}$ is the adjoint state and $\mu_{a,h}, \ \mu_{b,h}$ are Lagrange multipliers. We consider a sequence of positive real numbers $\lambda_k$ tending to zero for $k \to \infty$. We use ($\mathrm{P_k}$) to denote the regularized problems associated to $\lambda_k$ and their solutions are denoted by $(\overline{y}_k, \overline{u}_k)$ with an adjoint state $p_k$ and Lagrange multipliers $\mu_{ak}, \ \mu_{bk}$. To begin with, we give the following lemma which focuses on the boundedness of the Lagrangian multipliers. Since upper bound and lower bound exist simultaneously in the problem we consider, the proof of the following lemma encounter some difficulties compared with the situation with only one bound. However, we utilize the fact that at least one of the two multipliers is equal to zero and complete the proof.

**Lemma 3.5.** *Under Assumption 1.1, the sequence of Lagrange multipliers $\{\mu_{bk}\}$ and $\{\mu_{ak}\}$ are uniformly bounded in $L^1(\Omega)$.*

*Proof.* Let $u_1 = \min(\widehat{u}, 0), \ u_2 = \max(\widehat{u}, 0) \in L^2(\Omega)$, then we have $u_1(x) \le 0, \ u_2(x) \ge 0$ a.e. in $\Omega$. Then from the maximum principle for the state equation, we have $(Su_1)(x) < b, \ a < (Su_2)(x) \ \forall x \in \overline{\Omega}$. So $\forall \lambda \ge 0$, there exists $\tau_1, \ \tau_2 > 0$ such that

$$\begin{aligned} &\lambda u_1(x) + (Su_1)(x) \le b - \tau_1 \quad \text{a.e. in } \Omega, \\ &a + \tau_2 \le \lambda u_2(x) + (Su_2)(x) \quad \text{a.e. in } \Omega. \end{aligned} \tag{3.32}$$

Let $\widehat{u}_{1,k} = u_1 - \overline{u}_k,\ \widehat{u}_{2,k} = \overline{u}_k - u_2$, then using (3.32) we arrive at

$$\begin{aligned} &\tau_1 + \lambda_k \overline{u}_k(x) + (S\overline{u}_k)(x) - b \le -(\lambda_k \widehat{u}_{1,k}(x) + (S\widehat{u}_{1,k})(x)) \quad \text{a.e. in } \Omega, \\ &\tau_2 + a - \lambda_k \overline{u}_k(x) - (S\overline{u}_k)(x) \le -(\lambda_k \widehat{u}_{2,k}(x) + (S\widehat{u}_{2,k})(x)) \quad \text{a.e. in } \Omega. \end{aligned} \tag{3.33}$$

We multiply the two formulas in (3.33) by $\mu_{bk}$ and $\mu_{ak}$ respectively, which implies

$$\begin{aligned} &\int_\Omega \tau_1 \mu_{bk} dx \le \int_\Omega -(\lambda_k \widehat{u}_{1,k} + S\widehat{u}_{1,k})\mu_{bk} dx \quad \text{a.e. in } \Omega, \\ &\int_\Omega \tau_2 \mu_{ak} dx \le \int_\Omega -(\lambda_k \widehat{u}_{2,k} + S\widehat{u}_{2,k})\mu_{ak} dx \quad \text{a.e. in } \Omega. \end{aligned} \tag{3.34}$$

Since (3.30c) is equivalent to

$$\int_\Omega (\alpha \overline{u}_k + G^*(G\overline{u}_k - y_d + \mu_{bk} - \mu_{ak}) + \lambda_k \mu_{bk} - \lambda_k \mu_{ak}) z dx = 0, \quad \forall z \in L^2(\Omega). \tag{3.35}$$

We know that at least one of $\mu_{bk}$ and $\mu_{ak}$ is 0. When $\mu_{ak}$ is 0, let $z = \widehat{u}_{1,k}$ in (3.35), then we arrive at

$$\int_\Omega -(\lambda_k \widehat{u}_{1,k} + G\widehat{u}_{1,k})\mu_{bk} dx = \int_\Omega (\alpha \overline{u}_k + G^*(G\overline{u}_k - y_d))\widehat{u}_{1,k} dx. \tag{3.36}$$

When $\mu_{bk}$ is 0, let $z = \widehat{u}_{2,k}$ in (3.35), then we get

$$\int_\Omega -(\lambda_k \widehat{u}_{2,k} + G\widehat{u}_{2,k})\mu_{ak} dx = \int_\Omega -(\alpha \overline{u}_k + G^*(G\overline{u}_k - y_d))\widehat{u}_{2,k} dx. \tag{3.37}$$

Together with (3.34), we arrive at

$$\begin{aligned} &\int_\Omega \tau_1 \mu_{bk} dx \le ((\alpha + \|G\|^2)\|\overline{u}_k\| + \|G\|\|y_d\|)(\|u_1\| + \|\overline{u}_k\|), \\ &\int_\Omega \tau_2 \mu_{ak} dx \le ((\alpha + \|G\|^2)\|\overline{u}_k\| + \|G\|\|y_d\|)(\|u_2\| + \|\overline{u}_k\|). \end{aligned} \tag{3.38}$$

From the optimality of $\overline{u}_k$, we know the uniform boundedness of $\overline{u}_k$ in $L^2(\Omega)$. So we know that $\{\mu_{bk}\}$ and $\{\mu_{ak}\}$ are uniformly bounded in $L^1(\Omega)$. □

Similarly to Lemma 3.5, we can prove the uniform boundedness of $\|\mu_{ak,h}\|_{L^1(\Omega)}$ and $\|\mu_{bk,h}\|_{L^1(\Omega)}$ w.r.t $h, \lambda$ by replacing $S$ by $S_h$ and $G$ by $G_h$.

**Theorem 3.6.** *Let $(\overline{y}_\lambda, \overline{u}_\lambda)$ and $(\overline{y}_{\lambda,h}, \overline{u}_{\lambda,h})$ be the solutions of ($\mathrm{P}_\lambda$) and ($\mathrm{P}_{\lambda,h}$) respectively, then there exists some $0 < h_0 \le 1$ such that*

$$\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\| + \|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|_{H^1(\Omega)} \le Ch^{1-\frac{n}{4}}, \quad \forall\ 0 < h \le h_0$$

*holds, where $n$ denotes the dimension of $\Omega$ and $C > 0$ is a positive constant which is independent of $\lambda$.*

*Proof.* Subtracting (3.30c) and (3.31c), we get:

$$\alpha(\overline{u}_\lambda - \overline{u}_{\lambda,h}) + (p_\lambda - p_{\lambda,h}) + \lambda(\mu_b - \mu_{b,h}) - \lambda(\mu_a - \mu_{a,h}) = 0. \tag{3.39}$$

Multiplying the formula above by $\overline{u}_\lambda - \overline{u}_{\lambda,h}$, we derive

$$\begin{aligned}\alpha\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 &= (p_{\lambda,h} - p_\lambda, \overline{u}_\lambda - \overline{u}_{\lambda,h}) - (\lambda(\mu_b - \mu_{b,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) \\ &\quad +(\lambda(\mu_a - \mu_{a,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) \\ &= -(\lambda(\mu_b - \mu_{b,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) + (\lambda(\mu_a - \mu_{a,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) \\ &\quad +(p^h - p_\lambda, \overline{u}_\lambda - \overline{u}_{\lambda,h}) + (p_{\lambda,h} - p^h, \overline{u}_\lambda - \overline{u}_{\lambda,h}),\end{aligned} \tag{3.40}$$

where $p^h$ is the solution of

$$(\nabla p^h, \nabla z_h) = (\overline{y}_\lambda - y_d - \mu_a + \mu_b, z_h) \quad \forall z_h \in Z_h, \tag{3.41}$$

$y^h$ is the solution of

$$(\nabla y^h, \nabla z_h) = (\overline{u}_\lambda, z_h) \quad \forall z_h \in Z_h. \tag{3.42}$$

Let $z_h = y^h - \overline{y}_{\lambda,h} \in Z_h$ in the formula which we get by subtracting (3.31b) and (3.41), then we arrive at

$$(\nabla(p_{\lambda,h} - p^h), \nabla(y^h - \overline{y}_{\lambda,h})) = (\overline{y}_{\lambda,h} - \overline{y}_\lambda + \mu_{b,h} - \mu_b - \mu_{a,h} + \mu_a, y^h - \overline{y}_{\lambda,h}). \tag{3.43}$$

Similarly, let $z_h = p_{\lambda,h} - p^h \in Z_h$ in the formula which we get by subtracting (3.42) and (3.31a), then we derive

$$(\nabla(y^h - \overline{y}_{\lambda,h}), \nabla(p_{\lambda,h} - p^h)) = (\overline{u}_\lambda - \overline{u}_{\lambda,h}, p_{\lambda,h} - p^h). \tag{3.44}$$

So we can get

$$\begin{aligned}(p_{\lambda,h} - p^h, \overline{u}_\lambda - \overline{u}_{\lambda,h}) &= (\overline{y}_{\lambda,h} - \overline{y}_\lambda, y^h - \overline{y}_{\lambda,h}) + \underbrace{(\mu_b, \overline{y}_{\lambda,h} - y^h)}_{I} + \underbrace{(\mu_{b,h}, y^h - \overline{y}_{\lambda,h})}_{II} \\ &\quad + \underbrace{(\mu_a, y^h - \overline{y}_{\lambda,h})}_{III} + \underbrace{(\mu_{a,h}, \overline{y}_{\lambda,h} - y^h)}_{IV}.\end{aligned}$$

For the term $I$, since $\overline{y}_{\lambda,h} \le b - \lambda\overline{u}_{\lambda,h}$ and $\mu_b \ge 0$, we derive

$$(\mu_b, \overline{y}_{\lambda,h} - y^h) \le (\mu_b, b - \lambda\overline{u}_{\lambda,h} - y^h - b + \lambda\overline{u}_\lambda + \overline{y}_\lambda) \tag{3.45}$$

$$= (\mu_b, \lambda(\overline{u}_\lambda - \overline{u}_{\lambda,h})) + (\mu_b, \overline{y}_\lambda - y^h). \tag{3.46}$$

For the term $II$, because of $\overline{y}_\lambda \le b - \lambda\overline{u}_\lambda$ and $\mu_{b,h} \ge 0$, we have

$$\begin{aligned}(\mu_{b,h}, y^h - \overline{y}_{\lambda,h}) &= (\mu_{b,h}, \overline{y}_\lambda - \overline{y}_{\lambda,h}) + (\mu_{b,h}, y^h - \overline{y}_\lambda) \\ &\le (\mu_{b,h}, b - \lambda\overline{u}_\lambda - \overline{y}_{\lambda,h} - b + \lambda\overline{u}_{\lambda,h} + \overline{y}_{\lambda,h}) + (\mu_{b,h}, y^h - \overline{y}_\lambda) \\ &\le (\mu_{b,h}, \lambda(\overline{u}_{\lambda,h} - \overline{u}_\lambda)) + (\mu_{b,h}, y^h - \overline{y}_\lambda).\end{aligned} \tag{3.47}$$

For the term $III$, based on $-\overline{y}_{\lambda,h} \le -a + \lambda\overline{u}_{\lambda,h}$ and $\mu_a \ge 0$, we arrive at

$$\begin{aligned}(\mu_a, y^h - \overline{y}_{\lambda,h}) &\le (\mu_a, y^h - a + \lambda\overline{u}_{\lambda,h} + a - \lambda\overline{u}_\lambda - \overline{y}_\lambda) \\ &= (\mu_a, \lambda(\overline{u}_{\lambda,h} - \overline{u}_\lambda)) + (\mu_a, y^h - \overline{y}_\lambda)).\end{aligned} \tag{3.48}$$

For the term $IV$, following from $-\overline{y}_\lambda \leq -a + \lambda \overline{u}_\lambda$ and $\mu_{a,h} \geq 0$, we have

$$\begin{aligned}(\mu_{a,h}, \overline{y}_{\lambda,h} - y^h) &= (\mu_{a,h}, \overline{y}_{\lambda,h} - \overline{y}_\lambda) + (\mu_{a,h}, \overline{y}_\lambda - y^h) \\ &\leq (\mu_{a,h}, \overline{y}_{\lambda,h} - a + \lambda \overline{u}_\lambda + a - \lambda \overline{u}_{\lambda,h} - \overline{y}_{\lambda,h}) + (\mu_{a,h}, \overline{y}_\lambda - y^h) \\ &\leq (\mu_{a,h}, \lambda(\overline{u}_\lambda - \overline{u}_{\lambda,h})) + (\mu_{a,h}, \overline{y}_\lambda - y^h). \end{aligned} \tag{3.49}$$

Inserting (3.45), (3.47), (3.48) and (3.49) into (3.40), we get

$$\begin{aligned}\alpha\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 &\leq -(\lambda(\mu_b - \mu_{b,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) + (\lambda(\mu_a - \mu_{a,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) \\ &\quad + (p^h - p_\lambda, \overline{u}_\lambda - \overline{u}_{\lambda,h}) + (\overline{y}_{\lambda,h} - \overline{y}_\lambda, y^h - \overline{y}_{\lambda,h}) \\ &\quad + (\lambda(\mu_b - \mu_{b,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) - (\lambda(\mu_a - \mu_{a,h}), \overline{u}_\lambda - \overline{u}_{\lambda,h}) \\ &\quad + (\mu_b, \overline{y}_\lambda - y^h) + (\mu_{b,h}, y^h - \overline{y}_\lambda) + (\mu_a, y^h - \overline{y}_\lambda) + (\mu_{a,h}, \overline{y}_\lambda - y^h) \\ &= -\|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2 + (\overline{y}_{\lambda,h} - \overline{y}_\lambda, y^h - \overline{y}_\lambda) + (p^h - p_\lambda, \overline{u}_\lambda - \overline{u}_{\lambda,h}) \\ &\quad + (\mu_b - \mu_{b,h} - \mu_a + \mu_{a,h}, \overline{y}_\lambda - y^h),\end{aligned}$$

which gives

$$\begin{aligned}&\alpha\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 + \|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2 \\ &\leq (\overline{y}_{\lambda,h} - \overline{y}_\lambda, y^h - \overline{y}_\lambda) + (p^h - p_\lambda, \overline{u}_\lambda - \overline{u}_{\lambda,h}) + (\mu_b - \mu_{b,h} - \mu_a + \mu_{a,h}, \overline{y}_\lambda - y^h) \\ &\leq \frac{1}{2}\|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2 + \frac{1}{2}\|y^h - \overline{y}_\lambda\|^2 + \frac{2}{\alpha}\|p^h - p_\lambda\|^2 + \frac{\alpha}{2}\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 \\ &\quad + \|\mu_b - \mu_{b,h} - \mu_a + \mu_{a,h}\|_{L^1(\Omega)} \cdot \|\overline{y}_\lambda - y^h\|_{L^\infty(\Omega)}.\end{aligned}$$

Then we arrive at

$$\begin{aligned}&\alpha\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 + \|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|^2 \\ &\leq \|\overline{y}_\lambda - y^h\|^2 + \frac{4}{\alpha}\|p_\lambda - p^h\|^2 + 2\|\mu_b - \mu_{b,h} - \mu_a + \mu_{a,h}\|_{L^1(\Omega)} \cdot \|\overline{y}_\lambda - y^h\|_{L^\infty(\Omega)}.\end{aligned}$$

It is shown in [7] that the following formula holds

$$\|p_\lambda - p^h\|^2 \leq h^{4-n}\left(\|\overline{y}_\lambda - y_d\|^2 + \|\mu_a\|_{L^1}^2 + \|\mu_b\|_{L^1}^2\right). \tag{3.50}$$

Through standard finite element error estimates and the fact that $\|\overline{u}_\lambda\|$ is bounded independent of $\lambda$ resulting from the optimality of $\overline{u}_\lambda$, we know that $\|\overline{y}_\lambda - y^h\|^2 \leq Ch^4$ and $\|\overline{y}_\lambda - y^h\|_{L^\infty(\Omega)} \leq Ch^{2-\frac{n}{2}}$. Together with Lemma 3.5, we have the following estimation for $\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|$,

$$\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 \leq C(h^4 + h^{4-n} + h^{2-\frac{n}{2}}), \tag{3.51}$$

which implies $\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\| \leq Ch^{1-\frac{n}{4}}$. Then as the proof of Theorem 3.4, we can get

$$\|\overline{y}_\lambda - \overline{y}_{\lambda,h}\|_{H^1} \leq C(h + h^{1-\frac{n}{4}}), \tag{3.52}$$

which gives the assertion.

□

In addition, if we assume that $\overline{u}_\lambda$ is uniformly bounded in $L^\infty(\Omega)$, then from [12] we know that

$$\|\overline{y}_\lambda - y^h\|_{L^\infty(\Omega)} \leq Ch^2|\log(h)|^2\|\overline{u}_\lambda\|_{L^\infty(\Omega)}. \tag{3.53}$$

Then from the proof of Theorem 3.6 we have

$$\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\|^2 \leq C(h^4 + h^{4-n} + h^2|\log(h)|^2), \tag{3.54}$$

which implies the following corollary.

**Corollary 3.7.** *Assume that the sequence of optimal solutions to* ($\mathrm{P}_\lambda$) *for* $\lambda \downarrow 0$*, denoted by* $\{\overline{u}_\lambda\}$*, is uniformly bounded in* $L^\infty(\Omega)$*, and assume further that the solution of* (2.1) *satisfies* $y \in W^{2,q}(\Omega)$ *for all* $1 \leq q < \infty$ *if* $u \in L^\infty(\Omega)$*. Then the sequence of solutions of* ($\mathrm{P}_{\lambda,h}$)*, denoted by* $\{\overline{u}_{\lambda,h}\}$ *satisfies*

$$\|\overline{u}_\lambda - \overline{u}_{\lambda,h}\| \leq C\max\{h|\log(h)|, h^{2-\frac{n}{2}}\}, \qquad \forall\ 0 < h \leq h_0$$

*where* $n$ *denotes the dimension of* $\Omega$ *and* $C$ *is a constant independent of* $\lambda$ *and* $h$*.*

### 3.3. Analysis for error estimates

The main novelty with respect to the error estimates of our paper is that we prove the error order of full discretization is not inferior to that of variational discretization, which has been stated in detail in introduction. The overall error consists of two parts: one arising from the regularization and another caused by the discretization. We know from [36] that for the error resulted from Lavrentiev-regularization, the following theorem holds

**Theorem 3.8.** *Let* $(y^*, u^*)$ *and* $(\overline{y}_\lambda, \overline{u}_\lambda)$ *be the solutions of* (P) *and* ($\mathrm{P}_\lambda$)*, then the following error estimate holds*

$$\|u^* - \overline{u}_\lambda\| \leq c\sqrt{\lambda},$$

*where* $c$ *is a constant independent of* $\lambda$*.*

Combining Theorem 3.8 with Theorem 3.4 and Corollary 3.7, we arrive at the following results for the overall error.

$$\|u^* - \overline{u}_{\lambda,h}\| \leq C_1\left(\sqrt{\lambda} + \frac{1}{\sqrt{\lambda}}h + \frac{1}{\lambda^2}\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right)\right) \tag{3.55}$$

and

$$\|u^* - \overline{u}_{\lambda,h}\| \leq C_2\left(\sqrt{\lambda} + \max\{h|\log(h)|, h^{2-\frac{n}{2}}\}\right), \tag{3.56}$$

where $n = 2,\ 3$ denotes the dimension of $\Omega$ and $C_1$, $C_2$ are positive constants independent of $\lambda$ and $h$. As we said in Introduction, the error order of full discretization is not inferior to that of variational discretization. It is clear from (3.55) and (3.56) that when $\lambda$ is fixed, both two error estimates decrease as $h$ declines until reaching a lower bound resulting from term $\sqrt{\lambda}$, i.e. Lavrentiev regularization. However, for fixed $h$, the first error estimate may decrease also may increase as $\lambda$ declines because term $\sqrt{\lambda}$ and term $\frac{1}{\sqrt{\lambda}}h + \frac{1}{\lambda^2}\left(h^2 + \frac{1}{\lambda}h^3 + \frac{1}{\lambda^2}h^4\right)$ exist simultaneously. While the second error estimate may decline until reaching a lower bound also may remain unchanged as $\lambda$ decreases. These statements declare that for fixed $h$, it is not the smaller $\lambda$ the better. Additionally, both (3.55) and (3.56) give an upper bound for the error, which one is a better estimate also depends on the values of $C_1$ and $C_2$. So different problems may have various error variation trend. We could verify the statement above through the numerical experiments in Section 5.

## 4. Heterogeneous ADMM and two-phase strategy

The total error of utilizing numerical methods to solve PDE constrained problem consists of two parts: discretization error and the error of algorithm for discretized problem. The error order of piecewise linear finite element method in terms of mesh size is $O(h)$, which makes the discretization error account for the main part. Thus using heterogeneous ADMM (hADMM), which is a fast and efficient first order algorithm, to get a solution of moderate precision is sufficient. Heterogeneous ADMM is different from the classical ADMM, where two different norms are applied in the first two subproblems. However, in order to satisfy the need for more accurate solution, a two-phase strategy is also presented, in which the PDAS method is used as a postprocessor of the hADMM algorithm. However, we should emphasize that here the 'accurate' refers to the KKT precision of the numerical algorithm but not the error between exact solution and numerical solution.

To rewrite the discretized problem into a matrix-vector form, we define the following matrices

$$K_h = \left( \int_{\Omega_h} \nabla\phi_i \cdot \nabla\phi_j \ dx \right)_{i,j=1}^{N_h} \quad \text{and} \quad M_h = \left( \int_{\Omega_h} \phi_i \cdot \phi_j \ dx \right)_{i,j=1}^{N_h}, \tag{4.1}$$

where $K_h$ and $M_h$ denote the finite element stiffness matrix and mass matrix respectively. Let

$$y_{d,h}(x) = \sum_{i=1}^{N_h} y_d^i \phi_i(x) \tag{4.2}$$

be the nodal projection of $y_d$ onto $Z_h$, where $y_d^i = y_d(x^i)$. The lump mass matrix $W_h$ is defined by

$$W_h = \text{diag}\left( \int_{\Omega_h} \phi_i(x) \ dx \right)_{i=1}^{N_h}, \tag{4.3}$$

which is a diagonal matrix. Actually, each principal diagonal element of $W_h$ is twice as the counterpart of $M_h$. For the mass matrix $M_h$ and the lump mass matrix $W_h$, the following proposition hold.

**Proposition 4.1** ([44], Tab. 1)**.** *$\forall\ z \in \mathbb{R}^{N_h}$, the following inequalities hold:*

$$\|z\|_{M_h}^2 \le \|z\|_{W_h}^2 \le c\|z\|_{M_h}^2, \quad \textit{where} \quad c = \begin{cases} 4 & if \quad n = 2, \\ 5 & if \quad n = 3. \end{cases}$$

For simplicity, we use the symbol before discretization to denote the column vectors of the coefficients of the functions with respect to the basis $\{\phi_i(x)\}_{i=1}^{N_h}$ which are discretized above, for example, $y = (y_1, y_2, \cdots, y_{N_h})^T \in \mathbb{R}^{N_h}$. Then we can rewrite the problem ($\mathrm{P}_{\lambda,h}$) and ($\widehat{\mathrm{P}}_{\lambda,h}$) into a matrix-vector form, which are the actual versions we apply the hADMM algorithm and PDAS method to respectively

$$\begin{cases} \min\limits_{y,u,v \in \mathbb{R}^{N_h}} & J_h(y,u) = \dfrac{1}{2}\|y - y_d\|_{M_h}^2 + \dfrac{\alpha}{2}\|u\|_{M_h}^2 \\ \quad \text{s.t.} & K_h y = M_h u, \\ & v - \lambda u - y = 0, \\ & v \in [a,b]^{N_h}. \end{cases} \tag{$\widehat{\mathrm{P}}'_{\lambda,h}$}$$

$$\left\{\begin{aligned} &\min_{y,u\in\mathbb{R}^{N_h}} && J_h(y,u)=\frac{1}{2}\|y-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2 \\ &\quad\text{s.t.} && K_h y = M_h u, \\ & && \lambda u + y\in[a,b]^{N_h}. \end{aligned}\right. \tag{${\rm P'}_{\lambda,h}$}$$

In the process of implementation, if a solution with moderate accuracy is sufficient, hADMM algorithm is applied. In addition, if more accurate solution ('accurate' refers to the KKT precision of the numerical algorithm but not the error between exact solution and numerical solution) is required, a two-phase strategy is employed, in which the PDAS method is used as a postprocessor of the hADMM algorithm. The following two subsections focus on the hADMM algorithm and the PDAS method respectively.

## 4.1. Two ADMM-type algorithms for ($\widehat{\rm P}'_{\lambda,h}$)

Since the stiffness matrix $K_h$ and the mass matrix $M_h$ are symmetric positive definite matrices, we can rewrite ($\widehat{\rm P}'_{\lambda,h}$) into the reduced form

$$\left\{\begin{aligned} &\min_{u,v\in\mathbb{R}^{N_h}} && J_h(y,u)=\frac{1}{2}\|K_h^{-1}M_h u-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2 \\ &\quad\text{s.t.} && v-\lambda u - K_h^{-1}M_h u = 0, \\ & && v\in[a,b]^{N_h}. \end{aligned}\right. \tag{${\rm R}\widehat{\rm P}'_{\lambda,h}$}$$

In order to show the differences between our hADMM and classical ADMM, we give the details of these two algorithms respectively. First, let us focus on classical ADMM.

### *4.1.1. Classical ADMM*

We can see from the content below that the first subproblem of classical ADMM has to solve a $3*3$ block equation system. It can be reduced into a $2*2$ block equation system, however, it will introduce additional computation of $M_h^{-1}$. More importantly, classical ADMM algorithm is not mesh independent.

The augmented Lagrangian function of (${\rm R}\widehat{\rm P}'_{\lambda,h}$) is:

$$\begin{aligned} L_\sigma(v,u;\mu) = &\frac{1}{2}\|K_h^{-1}M_h u-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2+(\mu, v-\lambda u-K_h^{-1}M_h u) \\ &+\frac{\sigma}{2}\|v-\lambda u-K_h^{-1}M_h u\|^2+\delta_{[a,b]^{N_h}}(v), \end{aligned} \tag{4.4}$$

where $\mu\in\mathbb{R}^{N_h}$ is the Lagrange multiplier and $\sigma>0$ is a penalty parameter. We give the three main steps at $k$-th iteration.

$$\left\{\begin{aligned} &\text{step 1}: u^{k+1}=\arg\min_u L_\sigma(v^k,u;\mu^k) \\ &\text{step 2}: v^{k+1}=\arg\min_v L_\sigma(v,u^{k+1};\mu^k) \\ &\text{step 3}: \mu^{k+1}=\mu^k+\sigma(v^{k+1}-\lambda u^{k+1}-y^{k+1}) \end{aligned}\right.$$

Now let us give the details about two subproblems with respect to $u$ and $v$ respectively. The first subproblem is equivalent to the following problem

$$\begin{aligned} \min_{y,u\in\mathbb{R}^{N_h}} \quad & \frac{1}{2}\|y-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2+(\mu^k,v^k-\lambda u-y)+\frac{\sigma}{2}\|v^k-\lambda u-y\|^2 \\ \text{s.t.} \quad & K_h y-M_h u=0, \end{aligned} \tag{4.5}$$

whose Lagrangian function is

$$L_1(y,u;p)=\frac{1}{2}\|y-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2+(\mu^k,v^k-\lambda u-y)+\frac{\sigma}{2}\|v^k-\lambda u-y\|^2+(p,K_h y-M_h u),$$

where $p$ is the Lagrangian multiplier corresponding to the equality constraint $K_h y-M_h u=0$. Then the KKT conditions of (4.5) are

$$\begin{cases} M_h(y-y_d)-\mu^k-\sigma(v^k-\lambda u-y)+{K_h}^T p=0 \\ \alpha M_h u-\lambda\mu^k-\lambda\sigma(v^k-\lambda u-y)-{M_h}^T p=0 \\ K_h y-M_h u=0 \end{cases}$$

$$\Leftrightarrow \begin{bmatrix} M_h+\sigma I & \lambda\sigma I & {K_h}^T \\ \lambda\sigma I & \lambda^2\sigma I+\alpha M_h & -{M_h}^T \\ K_h & -M_h & 0 \end{bmatrix} \begin{bmatrix} y^{k+1} \\ u^{k+1} \\ p^{k+1} \end{bmatrix} = \begin{bmatrix} M_h y_d+\mu^k+\sigma v^k \\ \lambda(\mu^k+\sigma v^k) \\ 0 \end{bmatrix}. \tag{4.6}$$

The second subproblem is equivalent to the following problem

$$\begin{aligned} \min_{v\in\mathbb{R}^{N_h}} \quad & (\mu,v-\lambda u^{k+1}-y^{k+1})+\frac{\sigma}{2}\|v-\lambda u^{k+1}-y^{k+1}\|^2 \\ \text{s.t.} \quad & v\in[a,b]^{N_h}, \end{aligned} \tag{4.7}$$

whose object function is a quadratic function, so it has a closed form solution

$$v^{k+1}=\Pi_{[a,b]^{N_h}}\left(\lambda u^{k+1}+y^{k+1}-\frac{\mu^k}{\sigma}\right). \tag{4.8}$$

### *4.1.2. Heterogeneous ADMM (hADMM)*

The essential difference between hADMM and classical ADMM is that the former adopts two different weighted norms in two subproblems in each iteration. It is clear from the content below that the first subproblem of hADMM only has to solve a $2*2$ block system without any additional computations, which can be solved by generalized minimal residual (GMRES) with preconditioning matrix, and the second subproblem has a closed form solution. More importantly, The numerical results in Section 5 indicate that our hADMM algorithm is mesh independent, while classical ADMM is not. Additionally, to construct the relation between the continuous problem and discretized problem, proposing hADMM algorithm is a natural idea. Following the hADMM proposed in [39], we define the weighted augmented Lagrangian function of ($\mathrm{R}\widehat{\mathrm{P}}'_{\lambda,h}$) as

$$\begin{aligned} \widetilde{L}_\sigma(v,u;\mu)=&\frac{1}{2}\|K_h^{-1}M_h u-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2+(\mu,v-\lambda u-K_h^{-1}M_h u)_{M_h} \\ &+\frac{\sigma}{2}\|v-\lambda u-K_h^{-1}M_h u\|_{M_h}^2+\delta_{[a,b]^{N_h}}(v), \end{aligned} \tag{4.9}$$

where $\mu\in\mathbb{R}^{N_h}$ is the Lagrange multiplier and $\sigma>0$ is a penalty parameter. Based on (4.9), we give the following ADMM-type algorithm

$$\begin{cases} \text{step 1}: u^{k+1} = \arg\min\limits_{u} \widetilde{L}_\sigma(v^k, u; \mu^k) \\ \text{step 2}: v^{k+1} = \arg\min\limits_{v} \widetilde{L}_\sigma(v, u^{k+1}; \mu^k) \\ \text{step 3}: \mu^{k+1} = \mu^k + \sigma(v^{k+1} - \lambda u^{k+1} - y^{k+1}), \end{cases}$$

which is a modification of classical ADMM. Now let us give the details about two subproblems with respect to $u$ and $v$ respectively. The first subproblem is equivalent to the following problem

$$\begin{aligned} \min_{y,u\in\mathbb{R}^{N_h}} \quad & \frac{1}{2}\|y-y_d\|^2_{M_h} + \frac{\alpha}{2}\|u\|^2_{M_h} + (\mu^k, v^k - \lambda u - y)_{M_h} + \frac{\sigma}{2}\|v^k - \lambda u - y\|^2_{M_h} \\ \text{s.t.} \quad & K_h y - M_h u = 0, \end{aligned} \tag{4.10}$$

whose Lagrangian function is

$$L_2(y,u;p) = \frac{1}{2}\|y-y_d\|^2_{M_h} + \frac{\alpha}{2}\|u\|^2_{M_h} + (\mu^k, v^k - \lambda u - y)_{M_h} + \frac{\sigma}{2}\|v^k - \lambda u - y\|^2_{M_h} + (p, K_h y - M_h u),$$

where $p$ is the Lagrangian multiplier corresponding to the equality constraint $K_h y - M_h u = 0$. Since the smoothness of (4.10), solving it is equivalent to solving the following linear system

$$\begin{bmatrix} (1+\sigma)M_h & \lambda\sigma M_h & K_h^T \\ \lambda\sigma M_h & (\lambda^2\sigma+\alpha)M_h & -{M_h}^T \\ K_h & -M_h & 0 \end{bmatrix} \begin{bmatrix} y^{k+1} \\ u^{k+1} \\ p^{k+1} \end{bmatrix} = \begin{bmatrix} M_h(y_d + \mu^k + \sigma v^k) \\ \lambda M_h(\mu^k + \sigma v^k) \\ 0 \end{bmatrix}, \tag{4.11}$$

from which we derive that

$$u^{k+1} = \frac{1}{\lambda^2\sigma+\alpha}(p^{k+1} - \lambda\sigma y^{k+1} + \lambda(\mu^k + \sigma v^k)). \tag{4.12}$$

Then (4.11) could be reduced into the following equation system without any additional calculation.

$$\begin{bmatrix} (1+\frac{\sigma\alpha}{\lambda^2\sigma+\alpha})M_h & \frac{\lambda\sigma}{\lambda^2\sigma+\alpha}M_h + K_h^T \\ -\frac{\lambda\sigma}{\lambda^2\sigma+\alpha}M_h - K_h & \frac{1}{\lambda^2\sigma+\alpha}M_h \end{bmatrix} \begin{bmatrix} y^{k+1} \\ p^{k+1} \end{bmatrix} = \begin{bmatrix} M_h y_d + \frac{\alpha}{\lambda^2\sigma+\alpha}M_h(\mu^k + \sigma v^k) \\ -\frac{\lambda}{\lambda^2\sigma+\alpha}M_h(\mu^k + \sigma v^k) \end{bmatrix}. \tag{4.13}$$

It is seen that the hADMM only has to solve a $2*2$ block equation system in the first subproblem in each iteration. We should emphasize here that writing the optimality conditions in the form of an antisymmetric matrix can make it more convenient for the design of the preconditioning matrix and the equation system can be solved by GMRES with preconditioner. (4.13) can also be written into a symmetric matrix, however, some of the principle elements of the coefficient matrix will be negative.

Even if applying the modified ADMM, the second subproblem, which is equivalent to the following problem

$$\begin{aligned} \min_{v\in\mathbb{R}^{N_h}} \quad & (\mu, v - \lambda u^{k+1} - y^{k+1})_{M_h} + \frac{\sigma}{2}\|v - \lambda u^{k+1} - y^{k+1}\|^2_{M_h} \\ \text{s.t.} \quad & v \in [a,b]^{N_h}, \end{aligned} \tag{4.14}$$

does not have a closed form solution. Utilizing the relationship between $M_h$ and $W_h$, we replace the term $\frac{\sigma}{2}\|v - \lambda u^{k+1} - y^{k+1}\|^2_{M_h}$ by $\frac{\sigma}{2}\|v - \lambda u^{k+1} - y^{k+1}\|^2_{W_h}$, where $W_h$ is the lump mass matrix defined in (4.3). Then the second subproblem is transformed to the following optimization problem

$$\begin{aligned} \min_{v\in\mathbb{R}^{N_h}} \quad & (\mu^k, v-\lambda u^{k+1}-y^{k+1})_{M_h} + \frac{\sigma}{2}\|v-\lambda u^{k+1}-y^{k+1}\|^2_{W_h} \\ \text{s.t.} \quad & v\in[a,b]^{N_h}, \end{aligned} \tag{4.15}$$

whose solution has the following closed form

$$v^{k+1} = \Pi_{[a,b]^{N_h}}\left(\lambda u^{k+1} + y^{k+1} - \frac{{W_h}^{-1}M_h\mu^k}{\sigma}\right). \tag{4.16}$$

Although this will introduce the computation of $W_h^{-1}$, $W_h$ is a diagonal matrix, whose inverse will not cost much computation.

Based on the content above, we give the frame of the hADMM algorithm:

**Algorithm 1** Heterogeneous ADMM (hADMM) algorithm for ($\mathrm{R}\widehat{\mathrm{P}}'_{\lambda,h}$)

Initialization: Give initial point $(v^0,\mu^0)\in\mathbb{R}^{N_h}\times\mathbb{R}^{N_h}$ and a tolerant parameter $\tau>0$. Set $k=0$.

**Step 1** Compute $(y^{k+1},u^{k+1})$ through solving the following equation system

$$\begin{bmatrix} (1+\frac{\sigma\alpha}{\lambda^2\sigma+\alpha})M_h & \frac{\lambda\sigma}{\lambda^2\sigma+\alpha}M_h + K_h^T \\ -\frac{\lambda\sigma}{\lambda^2\sigma+\alpha}M_h - K_h & \frac{1}{\lambda^2\sigma+\alpha}M_h \end{bmatrix}\begin{bmatrix} y^{k+1} \\ p^{k+1}\end{bmatrix} = \begin{bmatrix} M_h y_d + \frac{\alpha}{\lambda^2\sigma+\alpha}M_h(\mu^k+\sigma v^k) \\ -\frac{\lambda}{\lambda^2\sigma+\alpha}M_h(\mu^k+\sigma v^k)\end{bmatrix}.$$

Compute $u^{k+1}$ as follows

$$u^{k+1} = \frac{1}{\lambda^2\sigma+\alpha}(p^{k+1} - \lambda\sigma y^{k+1} + \lambda(\mu^k+\sigma v^k)).$$

**Step 2** Compute $v^{k+1}$ as follows

$$v^{k+1} = \Pi_{[a,b]^{N_h}}\left(\lambda u^{k+1} + y^{k+1} - \frac{{W_h}^{-1}M_h\mu^k}{\sigma}\right).$$

**Step 3** Compute $\mu^{k+1}$ as follows

$$\mu^{k+1} = \mu^k + \sigma(v^{k+1} - \lambda u^{k+1} - y^{k+1}).$$

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

For the convergence result of the heterogeneous ADMM algorithm, we have the following theorem.

**Theorem 4.2.** *([39], Thm. 4.5) Let $(y^*,u^*,v^*,p^*,\mu^*)$ be the KKT point of ($\widehat{\mathrm{P}}'_{\lambda,h}$). $\{(u^k,v^k,\mu^k)\}$ is generated by Algorithm 1 with the associated state $\{y^k\}$ and adjoint state $\{p^k\}$, then we have*

$$\lim_{k\to\infty}\{\|u^k-u^*\| + \|v^k-v^*\| + \|\mu^k-\mu^*\|\} = 0,$$
$$\lim_{k\to\infty}\{\|y^k-y^*\| + \|p^k-p^*\|\} = 0.$$

## 4.2. Primal-Dual Active Set method as postprocessor

As we have said above, the error of utilizing numerical methods to solve PDE constrained problem consists of two parts: discretization error and the error of algorithm for discretized problem, in which the discretization error account for the main part. Algorithms of high precision do not make much sense but waste computations in practice. In general, using hADMM algorithm to get a solution of moderate precision is sufficient. Although algorithms of high precision are not necessary, we also provide a two-phase strategy to satisfy the requirement for numerical solution of high precision, in which the PDAS method is used as a postprocessor of the hADMM algorithm. The PDAS method was used to solve control constrained elliptic optimal control problem in [4]. In [21], the authors show its relation to semismooth Newton method, which can be used to prove its local superlinear convergence. We employ the PDAS method to ($\mathrm{P'}_{\lambda,h}$), whose Lagrangian function is:

$$\widehat{L}(v,u;\mu)=\frac{1}{2}\|y-y_d\|_{M_h}^2+\frac{\alpha}{2}\|u\|_{M_h}^2+(p,K_hy-M_hu)+(\mu_a,a-\lambda u-y)+(\mu_b,\lambda u+y-b),$$

where $\mu_a,\mu_b\in\mathbb{R}^{N_h}$ are the Lagrange multipliers. Then the KKT conditions of ($\mathrm{P'}_{\lambda,h}$) are

$$\begin{cases} M_h(y-y_d)+K_h^Tp-\mu_a+\mu_b=0,\\ \alpha M_hu-{M_h}^Tp-\lambda\mu_a+\lambda\mu_b=0,\\ K_hy-M_hu=0,\\ \mu_a\geq 0,\quad a-\lambda u-y\leq 0,\quad (\mu_a,a-\lambda u-y)=0,\\ \mu_b\geq 0,\quad \lambda u+y-b\leq 0,\quad (\mu_b,\lambda u+y-b)=0, \end{cases} \tag{4.17}$$

which can be equivalently rewritten as

$$\begin{cases} M_h(y-y_d)+K_h^Tp-\mu_a+\mu_b=0,\\ \alpha M_hu-{M_h}^Tp-\lambda\mu_a+\lambda\mu_b=0,\\ K_hy-M_hu=0,\\ \min(\mu_a,\lambda u+y-a)=\mu_a+\min(0,\lambda u+y-a-\mu_a)=0\\ \min(\mu_b,b-\lambda u-y)=\mu_b+\min(0,b-\lambda u-y-\mu_b)=0. \end{cases} \tag{4.18}$$

Let $\mu=\mu_b-\mu_a$, then (4.19) can be reduced into the following $4*4$ block system

$$\begin{cases} M_h(y-y_d)+K_h^Tp+\mu=0,\\ \alpha M_hu-{M_h}^Tp+\lambda\mu=0,\\ K_hy-M_hu=0,\\ \mu-\max(0,\mu+\lambda u+y-b)-\min(0,\lambda u+y-a-\mu)=0. \end{cases} \tag{4.19}$$

We define the active and inactive sets as

$$\mathcal{A}_{a,h}=\{i\in\{1,2,...,N_h\}:\lambda u_i+y_i+\mu_i-a<0\}, \tag{4.20}$$
$$\mathcal{A}_{b,h}=\{i\in\{1,2,...,N_h\}:\lambda u_i+y_i+\mu_i-b>0\}, \tag{4.21}$$
$$\mathcal{I}=\{1,2,\cdots,N_h\}\backslash(\mathcal{A}_{a,h}\cup\mathcal{A}_{b,h}) \tag{4.22}$$

and note that the following properties hold

$$\begin{aligned} &\lambda u_i+y_i=a \quad \text{on } \mathcal{A}_{a,h}, \qquad \lambda u_i+y_i=b \quad \text{on } \mathcal{A}_{b,h},\\ &\mu_i<0 \ \text{ on } \mathcal{A}_{a,h},\quad \mu_i>0 \quad \text{on } \mathcal{A}_{b,h},\quad \mu_i=0 \quad \text{on } \mathcal{I}. \end{aligned} \tag{4.23}$$

Let

$$(E_a)_{ij}=\begin{cases}1 & i=j \text{ and } i\in\mathcal{A}_{a,h},\\ 0 & \text{else,}\end{cases}\qquad (E_b)_{ij}=\begin{cases}1 & i=j \text{ and } i\in\mathcal{A}_{b,h},\\ 0 & \text{else,}\end{cases} \tag{4.24}$$

then we can rewrite the optimal system (4.19) into a linear system

$$\begin{bmatrix} M_h & 0 & E_a+E_b & K_h^T \\ 0 & \alpha M_h & \lambda(E_a+E_b) & -M_h{}^T \\ E_a+E_b & \lambda(E_a+E_b) & I-E_a-E_b & 0 \\ K_h & -M_h & 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ u \\ \mu \\ p \end{bmatrix} = \begin{bmatrix} M_h y_d \\ 0 \\ E_a a + E_b b \\ 0 \end{bmatrix}. \tag{4.25}$$

It is shown in [4] that whether the two consecutive active sets equal is a termination criterion for the primal-dual active set method. Following the content above, we give the frame of the PDAS method:

**Algorithm 2** Primal-Dual Active Set (PDAS) algorithm for ($\mathrm{P'}_{\lambda,h}$)

Initialization: Choose initial point $y^0$, $u^0$, $p^0$ and $\mu^0 \in \mathbb{R}^{N_h}$; Set $k=0$.

**Step 1** Determine the following subsets of $\{1,2,...,N_h\}$ (Active and Inactive sets)

$$\begin{aligned} \mathcal{A}_{a,h}^k &= \{i \in \{1,2,...,N_h\} : \lambda u_i^k + y_i^k + \mu_i^k - a < 0\}, \\ \mathcal{A}_{b,h}^k &= \{i \in \{1,2,...,N_h\} : \lambda u_i^k + y_i^k + \mu_i^k - b > 0\}, \\ \mathcal{I}^k &= \{1,2,...,N_h\} \backslash (\mathcal{A}_{a,h}^k \cup \mathcal{A}_{b,h}^k). \end{aligned}$$

**Step 2** Determine $E_a^k$ and $E_b^k$ through (4.24) and solve the following system

$$\begin{bmatrix} M_h & 0 & E_a^k+E_b^k & K_h^T \\ 0 & \alpha M_h & \lambda(E_a^k+E_b^k) & -M_h{}^T \\ E_a^k+E_b^k & \lambda(E_a^k+E_b^k) & I-E_a^k-E_b^k & 0 \\ K_h & -M_h & 0 & 0 \end{bmatrix} \begin{bmatrix} y^{k+1} \\ u^{k+1} \\ \mu^{k+1} \\ p^{k+1} \end{bmatrix} = \begin{bmatrix} M_h y_d \\ 0 \\ E_a^k a + E_b^k b \\ 0 \end{bmatrix}.$$

**Step 3** If $k>1$, $\mathcal{A}_{a,h}^{k+1} = \mathcal{A}_{a,h}^k$ and $\mathcal{A}_{b,h}^{k+1} = \mathcal{A}_{b,h}^k$ or a termination criterion is met, Stop; else, set $k := k+1$ and go to Step 1.

For the convergence result of the PDAS method, we have the following theorem. For more details, we refer to [14, 42, 43].

**Theorem 4.3.** *Let $(u^k, y^k)$ be generated by Algorithm 2, if the initialization $(u^0, y^0)$ is sufficiently close to the solution $(u^*, y^*)$ of ($\mathrm{P'}_{\lambda,h}$), then $(u^k, y^k)$ converge superlinearly to $(u^*, y^*)$.*

## 5. Numerical result

In this section, two numerical experiments are considered. All calculations were performed using MATLAB (R2013a) on a PC with Intel (R) Core (TM) i7-4790K CPU (4.00GHz), whose operation system is 64-bit Windows 7.0 and RAM is 16.0 GB.

In the hADMM algorithm, the accuracy of a numerical solution is measured by the following residual

$$\eta_{\mathrm{A}} = \max\{r_1,\ r_2,\ r_3,\ r_4,\ r_5\}, \tag{5.1}$$

where

$$\begin{aligned} r_1 &= \|M_h(y-y_d) + K_h p - M_h \mu\|, \\ r_2 &= \|\alpha M_h u - M_h p - \lambda M_h \mu\|, \\ r_3 &= \|v - \Pi_{[a,b]}(v - M_h \mu)\|, \\ r_4 &= \|K_h y - M_h u\|, \\ r_5 &= \|v - \lambda u - y\|_{M_h}. \end{aligned}$$

TABLE 1. The $L^2$ error $\|u_r^* - \overline{u}_{\lambda,h}\|$ for Example 5.1.

| $h$ | $\lambda$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $10^{-2}$ | $10^{-2.5}$ | $10^{-3}$ | $10^{-3.5}$ | $10^{-4}$ | $10^{-4.5}$ | $10^{-5}$ | $10^{-5.5}$ | $10^{-6}$ |
| $\frac{2.5\sqrt{2}}{2^4}$ | 9.8613 | 5.2420 | 4.7054 | 4.7211 | 4.7837 | 4.8108 | 4.8215 | 4.8250 | 4.8261 |
| $\frac{2.5\sqrt{2}}{2^5}$ | 9.6385 | 4.2513 | 1.9925 | 1.5824 | 1.6055 | 1.6280 | 1.6360 | 1.6385 | 1.6391 |
| $\frac{2.5\sqrt{2}}{2^6}$ | 9.6298 | 4.2403 | 1.8113 | 0.7893 | 0.5401 | 0.5360 | 0.5438 | 0.5471 | 0.5484 |
| $\frac{2.5\sqrt{2}}{2^7}$ | 9.5821 | 4.2302 | 1.7991 | 0.7626 | 0.3373 | 0.2073 | 0.1956 | 0.1997 | 0.2019 |
| $\frac{2.5\sqrt{2}}{2^8}$ | 9.5361 | 4.1992 | 1.7624 | 0.7588 | 0.3188 | 0.0767 | 0.0660 | 0.0659 | – |

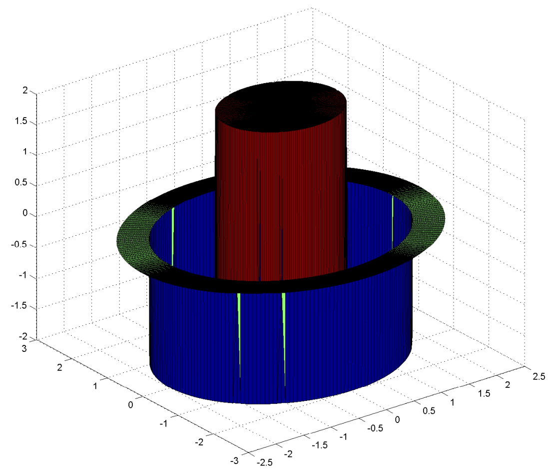

FIGURE 1. Figure of the desired state $y_d$ on the grid of size $h = \frac{14\sqrt{2}}{2^6}$.

Similarly, in the PDAS method, the accuracy of a numerical solution is measured by

$$\eta_{\mathrm{P}} = \max\{\gamma_1,\ \gamma_2,\ \gamma_3,\ \gamma_4\}, \tag{5.2}$$

where

$$\begin{aligned}
\gamma_1 &= \|M_h(y - y_d) + K_h p + \mu_{ab}\|, \\
\gamma_2 &= \|\alpha M_h u - M_h p + \lambda \mu_{ab}\|, \\
\gamma_3 &= \|K_h y - M_h u\|, \\
\gamma_4 &= \|\mu_{ab} - \max\{0, \mu_{ab} + \lambda u + y - b\} - \min\{0, \lambda u + y - a + \mu_{ab}\}\|.
\end{aligned}$$

Let $\epsilon$ be a given accuracy tolerance, then the terminal condition is $\eta_{\mathrm{A}}(\eta_{\mathrm{P}}) < \epsilon$.

To present the finite element error estimates results, it is convenient to introduce the experimental order of convergence (EOC), which for some positive error functional $E(h)$ with $h > 0$ is defined as follows: Given two

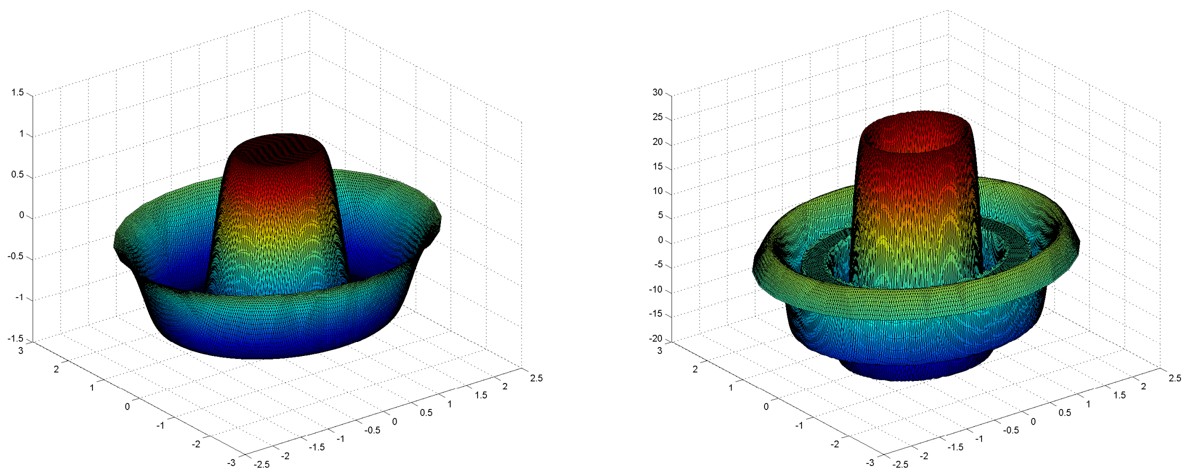


FIGURE 2. Figures of numerical state and control on the grid of size $h = \frac{14\sqrt{2}}{2^6}$ with $\lambda = 10^{-4.5}$. (a) Numerical state $y_{\lambda,h}$ and (b) numerical control $u_{\lambda,h}$.

TABLE 2. The EOC for Example 5.1.

| $h$ | $\lambda$ | | | | |
|---|---|---|---|---|---|
| | $10^{-4}$ | $10^{-4.5}$ | $10^{-5}$ | $10^{-5.5}$ | $10^{-6}$ |
| $\frac{2.5\sqrt{2}}{2^4}$ | 1.5751 | 1.5632 | 1.5593 | 1.5582 | 1.5580 |
| $\frac{2.5\sqrt{2}}{2^5}$ | 1.5717 | 1.6028 | 1.5890 | 1.5825 | 1.5796 |
| $\frac{2.5\sqrt{2}}{2^6}$ | 1.3677 | 1.3705 | 1.4752 | 1.4540 | 1.4416 |
| $\frac{2.5\sqrt{2}}{2^7}$ | 1.3947 | 1.4344 | 1.5674 | 1.5995 | – |

grid sizes $h_1 \neq h_2$, let

$$\mathrm{EOC} := \frac{\log E(h_1) - \log E(h_2)}{\log h_1 - \log h_2}.$$

It follows from this definition that if $E(h) = O(h^\gamma)$ then EOC $\approx \gamma$. The error functional $E(\cdot)$ investigated in the present section is given by

$$E(h) := \|u^* - \overline{u}_{\lambda,h}\|_{L^2(\Omega)}.$$

In this section, we use two adjacent values of $h$ to compute the EOC.

In both two examples, hADMM algorithm and two-phase strategy are employed to get numerical solutions of different precision respectively, *i.e.* the iteration is terminated with different $\epsilon$. Their convergence behavior are both compared with PDAS method, which is a special semi-smooth Newton method (see [21]). There are

TABLE 3. The convergence behavior of our hADMM algorithm and PDAS (a special semi-smooth Newton method) for Example 5.1.

| $h$ | #dofs | $\lambda$ | | hADMM | PDAS |
|---|---|---|---|---|---|
| $\frac{2.5\sqrt{2}}{2^6}$ | 18 977 | $10^{-4}$ | Iter | 22 | 16 |
| | | | Residual $\eta$ | 9.08e−03 | 8.97e−03 |
| | | | Time/s | 1.68 | 11.48 |
| | | $10^{-4.5}$ | Iter | 22 | 17 |
| | | | Residual $\eta$ | 9.33e−03 | 9.02e−03 |
| | | | Time/s | 1.65 | 12.33 |
| | | $10^{-5}$ | Iter | 22 | 18 |
| | | | Residual $\eta$ | 9.49e−03 | 8.54e−03 |
| | | | Time/s | 1.68 | 13.38 |
| $\frac{2.5\sqrt{2}}{2^7}$ | 76 353 | $10^{-4.5}$ | Iter | 25 | 33 |
| | | | Residual $\eta$ | 6.42e−03 | 7.93e−03 |
| | | | Time/s | 32.50 | 155.19 |
| | | $10^{-5}$ | Iter | 25 | 34 |
| | | | Residual $\eta$ | 4.53e−03 | 7.58e−03 |
| | | | Time/s | 32.80 | 162.01 |
| | | $10^{-5.5}$ | Iter | 25 | 34 |
| | | | Residual $\eta$ | 3.12e−03 | 9.35e−03 |
| | | | Time/s | 32.14 | 167.03 |
| $\frac{2.5\sqrt{2}}{2^8}$ | 306 305 | $10^{-5}$ | Iter | 28 | 68 |
| | | | Residual $\eta$ | 4.12e-03 | 4.67e−03 |
| | | | Time/s | 214.49 | 2938.8 |
| | | $10^{-5.5}$ | Iter | 28 | 70 |
| | | | Residual $\eta$ | 2.61e−03 | 8.93e−03 |
| | | | Time/s | 214.17 | 3049.3 |
| | | $10^{-6}$ | Iter | 28 | 72 |
| | | | Residual $\eta$ | 2.35e−03 | 6.72e−03 |
| | | | Time/s | 214.52 | 3148.9 |

four tables in both two examples. The first one in each example gives the $L^2$ error of the control and the second one shows the convergence order for fixed $\lambda$ through EOC defined above. The last two tables focus on the convergence behavior, including the times of iteration, residual $\eta$ and time, of the hADMM algorithm and the two-phase strategy compared with the PDAS method respectively. In the last two tables, '#dofs' denotes the dimension of the control variable on each grid level, 'iter' represents the times of iteration and 'residual' represents the precision $\eta$ of the numerical algorithm, which is defined above. In Table 4 and Table 8, two sub columns in the column of 'two-phase strategy' record the convergence behavior of two phases, i.e. hADMM and PDAS, respectively.

Table 4. The convergence behavior of our two-phase strategy and PDAS (a special semi-smooth Newton method) for Example 5.1.

| $h$ | #dofs | $\lambda$ | | Two-phase strategy (hADMM \| PDAS) | PDAS |
|---|---|---|---|---|---|
| $\frac{2.5\sqrt{2}}{2^6}$ | 18 977 | $10^{-4}$ | Iter | 22 \| 4 | 19 |
| | | | Residual $\eta$ | 9.07e−03 \| 3.27e−14 | 3.28e−14 |
| | | | Time/s | 4.62 (1.80 \| 2.82) | 13.64 |
| | | $10^{-4.5}$ | Iter | 22 \| 5 | 20 |
| | | | Residual $\eta$ | 9.32e−03 \| 4.64e−14 | 4.65e−14 |
| | | | Time/s | 5.11 (1.64 \| 3.47) | 14.46 |
| | | $10^{-5}$ | Iter | 22 \| 5 | 22 |
| | | | Residual $\eta$ | 9.48e−03 \| 3.12e−14 | 3.12e−14 |
| | | | Time/s | 5.19 (1.65 \| 3.54) | 16.35 |
| $\frac{2.5\sqrt{2}}{2^7}$ | 76 353 | $10^{-4.5}$ | Iter | 25 \| 6 | 37 |
| | | | Residual $\eta$ | 6.44e−03 \| 8.11e−14 | 8.11e−14 |
| | | | Time/s | 60.92 (33.17 \| 27.75) | 173.74 |
| | | $10^{-5}$ | Iter | 25 \| 6 | 38 |
| | | | Residual $\eta$ | 4.55e−03 \| 8.22e−14 | 8.23e−14 |
| | | | Time/s | 59.78 (33.16 \| 26.62) | 180.97 |
| | | $10^{-5.5}$ | Iter | 25 \| 7 | 40 |
| | | | Residual $\eta$ | 3.12e−03 \| 8.19e−14 | 8.19e−14 |
| | | | Time/s | 64.64 (33.05 \| 31.59) | 195.93 |
| $\frac{2.5\sqrt{2}}{2^8}$ | 306 305 | $10^{-5}$ | Iter | 28 \| 8 | 73 |
| | | | Residual $\eta$ | 4.12e−03 \| 1.43e−13 | 1.43e−13 |
| | | | Time/s | 542.99 (214.55 \| 328.44) | 3152.4 |
| | | $10^{-5.5}$ | Iter | 28 \| 9 | 75 |
| | | | Residual $\eta$ | 2.62e−03 \| 1.36e−13 | 1.36e−13 |
| | | | Time/s | 580.03 (214.80 \| 365.23) | 3241.4 |
| | | $10^{-6}$ | Iter | 28 \| 10 | 77 |
| | | | Residual $\eta$ | 2.34e−03 \| 1.58e−13 | 1.59e−13 |
| | | | Time/s | 615.18 (214.75 \| 400.43) | 3334.2 |

**Example 5.1.** We consider $\Omega = B_{\frac{5}{2}}(0)$ as the test domain and set $a = -1,\ b = 1,\ \alpha = 10^{-3}$ and $\sigma = 11$ in the first example. The desired state is defined by

$$y_d(r) = \begin{cases} 2 & 0 \le r \le 1, \\ -2 & 1 < r \le 2, \\ 0 & 2 < r \le 2.5. \end{cases}$$

When the exact solution is not known, using numerical solution as relative exact solution is a common method. For more details, one can see [24]. In our practice implementation, we choose $h = \frac{2.5\sqrt{2}}{2^8}$ and $\lambda = 10^{-6}$. When $h = \frac{2.5\sqrt{2}}{2^8}$, the scale of data is 306 305, which results in a large scale discretized problem. When lambda is too small, the problem will be ill-conditioned and the error will increase on the contrary from the error analysis in Section 3. Through testing with different lambda, e.g. $\lambda = 10^{-5.5}$, $10^{-6.5}$ and $10^{-7}$, we find that $\lambda = 10^{-6}$

TABLE 5. The $L^2$ error $\|u^* - \overline{u}_{\lambda,h}\|$ for Example 5.2.

| $h$ | $\lambda$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $10^{-2}$ | $10^{-2.5}$ | $10^{-3}$ | $10^{-3.5}$ | $10^{-4}$ | $10^{-4.5}$ | $10^{-5}$ | $10^{-5.5}$ | $10^{-6}$ |
| $\frac{14\sqrt{2}}{2^5}$ | 2.0358 | 2.0296 | 2.0218 | 2.0256 | 2.0260 | 2.0263 | 2.0264 | 2.0264 | 2.0265 |
| $\frac{14\sqrt{2}}{2^6}$ | 7.0136e−1 | 6.9802e−1 | 6.9544e−1 | 6.9136e−1 | 6.9012e−1 | 6.8968e−1 | 6.8966e−1 | 6.8963e−1 | 6.8958e−1 |
| $\frac{14\sqrt{2}}{2^7}$ | 2.8032e−1 | 2.5438e−1 | 2.4976e−1 | 2.4906e−1 | 2.4885e−1 | 2.4880e−1 | 2.4881e−1 | 2.4881e−1 | 2.4878e−1 |
| $\frac{14\sqrt{2}}{2^8}$ | 2.1943e−1 | 9.5802e−2 | 8.9552e−2 | 8.8678e−2 | 8.8670e−2 | 8.8656e−2 | 8.8652e−2 | 8.8651e−2 | 8.8648e−2 |
| $\frac{14\sqrt{2}}{2^9}$ | 1.9627e−1 | 4.2603e−2 | 3.8927e−2 | 3.2657e−2 | 3.1490e−2 | 3.1436e−2 | 3.1418e−2 | 3.1404e−2 | 3.1401e−2 |

TABLE 6. The EOC for Example 5.2.

| $h$ | $\lambda$ | | | | |
|---|---|---|---|---|---|
| | $10^{-4}$ | $10^{-4.5}$ | $10^{-5}$ | $10^{-5.5}$ | $10^{-6}$ |
| $\frac{14\sqrt{2}}{2^5}$ | 1.5537 | 1.5548 | 1.5550 | 1.5550 | 1.5552 |
| $\frac{14\sqrt{2}}{2^6}$ | 1.4716 | 1.4709 | 1.4708 | 1.4708 | 1.4708 |
| $\frac{14\sqrt{2}}{2^7}$ | 1.4888 | 1.4887 | 1.4888 | 1.4888 | 1.4887 |
| $\frac{14\sqrt{2}}{2^8}$ | 1.4936 | 1.4958 | 1.4966 | 1.4972 | 1.4973 |

is an appropriate choice. We give the $L^2$ errors $\|u_r^* - \overline{u}_{\lambda,h}\|$ on grids of different sizes with nine different values of $\lambda$ from $10^{-2}$ to $10^{-6}$ in Table 1. As an example, the figures of the desired state $y_d$, the numerical state $y_{\lambda,h}$ and numerical control $u_{\lambda,h}$ on the grid of size $h = \frac{2.5\sqrt{2}}{2^6}$ with $\lambda = 10^{-4.5}$ are displayed in Figures 1 and 2. If a solution with moderate accuracy is enough, hADMM algorithm is employed and compared with PDAS method. Both two algorithms are terminated when $\eta_{\mathrm{A}}(\eta_{\mathrm{P}}) < 10^{-2}$ in this case and the corresponding numerical results are displayed in Table 3. In addition, if more accurate solution is required, we employ the two-phase strategy and compare it with PDAS method. In this case, both two algorithms are terminated when $\eta_{\mathrm{A}}(\eta_{\mathrm{P}}) < 10^{-13}$ and the numerical results are shown in Table 4.

From Table 1, we can see that for fixed $\lambda$, the error decreases as $h$ declines at first until it reaches a bound resulted from regularization. When $h$ is fixed, the error declines as $\lambda$ decreases generally, while the error shows a rising trend with the last few values of $\lambda$. The numerical results in Table 1 declares that for fixed $h$ error may increase as $\lambda$ decreases, which verify the error estimates in Section 3. Table 2 shows the convergence order for fixed $\lambda$. Tables 3 and 4 show that both the hADMM algorithm and the two-phase strategy are much faster than PDAS method especially when the finite element grid size $h$ is very small. The numerical results in the last two tables verify the efficiency of the hADMM algorithm and the two-phase strategy. Thus there are reasons to believe that the efficiency of hADMM and two-phase strategy will be more obvious when the finite element grid size $h$ gets smaller.

We would like to point out that the number of iteration of both hADMM and PDAS are independent of the discretization level. First, from Table 3, it is easy to see that when $\lambda = 10^{-5}$ the number of iteration of hADMM for three discretization levels are 22, 25 and 28 respectively. More importantly, as we know, the Semi-smooth Newton method is locally superlinearly convergent, so we would like to point out that the mesh independence of PDAS is also expressed locally. That is to say the number of iteration of PDAS from moderate accuracy to

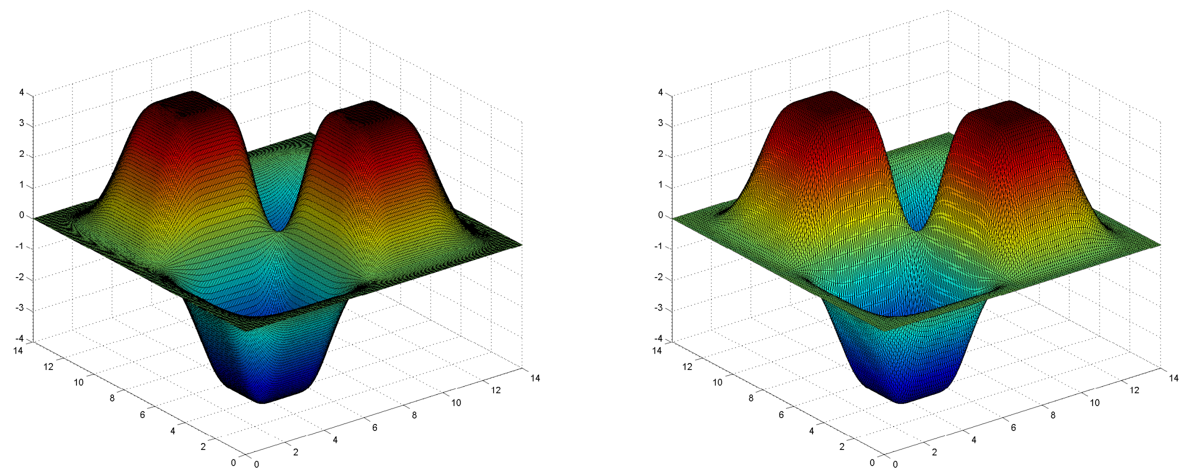

FIGURE 3. Figures of exact and numerical state on the grid of size $h = \frac{14\sqrt{2}}{2^7}$ with $\lambda = 10^{-4.5}$. (a) Exact state $y^*$ and (b) numerical state $y_{\lambda,h}$.

high accuracy (*e.g.* $10^{-3}$ to $10^{-13}$) is independent of the discretization level, which can be embodied by our numerical results. For Example 5.1, we could subtract the last column of Tables 3 and 4. When $\lambda = 10^{-5}$, the number of iteration of PDAS for three discretization levels when accuracy improving from $10^{-3}$ to $10^{-13}$ are 4, 4 and 5 respectively.

**Example 5.2.** We consider $\Omega = [0, 14]^2$ as the test domain and set $\alpha = 10^{-3},\ a = -4,\ b = 4,\ \sigma = 0.5$ and define $g(x)$ as

$$g(x) = \begin{cases} \frac{1}{6}x^3 + \frac{1}{8\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - \frac{1}{4\pi^2}x & x \in [0,1)\,, \\ -\left(\frac{1}{6}x^3 + \frac{1}{8\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - x^2 + \left(1 - \frac{1}{4\pi^2}\right)x - \frac{1}{3} + \frac{1}{2\pi^2}\right) & x \in [1,3)\,, \\ \frac{1}{6}x^3 + \frac{1}{8\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 2x^2 + \left(8 - \frac{1}{4\pi^2}\right)x - \frac{26}{3} + \frac{1}{\pi^2} & x \in [3,4)\,, \\ 2 & x \in [4,5)\,, \\ -\left(\frac{1}{3}x^3 + \frac{1}{4\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 5x^2 + \left(25 - \frac{1}{2\pi^2}\right)x - \frac{131}{3} + \frac{5}{2\pi^2}\right) & x \in [5,6)\,, \\ \frac{1}{3}x^3 + \frac{1}{4\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 7x^2 + \left(47 - \frac{1}{2\pi^2}\right)x - \frac{301}{3} + \frac{7}{2\pi^2} & x \in [6,8)\,, \\ -\left(\frac{1}{3}x^3 + \frac{1}{4\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 9x^2 + \left(81 - \frac{1}{2\pi^2}\right)x - 241 + \frac{9}{2\pi^2}\right) & x \in [8,9)\,, \\ -2 & x \in [9,10)\,, \\ \frac{1}{6}x^3 + \frac{1}{8\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 5x^2 + \left(50 - \frac{1}{4\pi^2}\right)x - \frac{506}{3} + \frac{5}{2\pi^2} & x \in [10,11)\,, \\ -\left(\frac{1}{6}x^3 + \frac{1}{8\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 6x^2 + \left(71 - \frac{1}{4\pi^2}\right)x - 275 + \frac{3}{\pi^2}\right) & x \in [11,13)\,, \\ \frac{1}{6}x^3 + \frac{1}{8\pi^3}\cos\left(2\pi x - \frac{\pi}{2}\right) - 7x^2 + \left(98 - \frac{1}{4\pi^2}\right)x - \frac{1372}{3} + \frac{7}{2\pi^2} & x \in [13,14]\,. \end{cases}$$

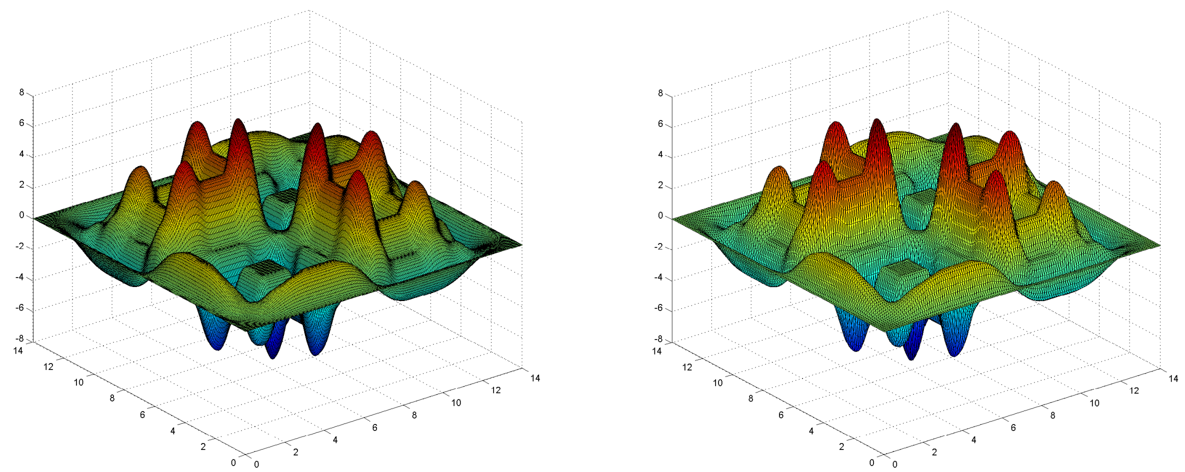

FIGURE 4. Figures of exact and numerical control on the grid of size $h = \frac{14\sqrt{2}}{2^7}$ with $\lambda = 10^{-4.5}$. (a) Exact control $u^*$ and (b) numerical control $u_{\lambda,h}$.

Let $y^*(x) = -g(x_1)g(x_2)$,

$$\mu_a = \begin{cases} 0.1\sin(\pi x_1)\sin(\pi x_2) & x \in (4,5)\times(4,5) \text{ or } x \in (9,10)\times(9,10), \\ 0 & \text{else}, \end{cases}$$

$$\mu_b = \begin{cases} -0.1\sin(\pi x_1)\sin(\pi x_2) & x \in (4,5)\times(9,10) \text{ or } x \in (9,10)\times(4,5), \\ 0 & \text{else}, \end{cases}$$

then from the optimal condition we arrive at

$$\begin{aligned} u^*(x) &= -\Delta y^* = g^{(2)}(x_1)g(x_2) + g(x_1)g^{(2)}(x_2), \\ p &= -\alpha u^*, \\ y_d &= y^* + \mu_b - \mu_a + \Delta p. \end{aligned}$$

The exact solution is known in this example and the $L^2$ errors $\|u^* - \overline{u}_{\lambda,h}\|$ on grids of different sizes with nine different values of $\lambda$ from $10^{-2}$ to $10^{-6}$ are given in Table 5. As an example, the figures of exact state $y^*$ and numerical state $y_{\lambda,h}$, exact control $u^*$ and numerical control $u_{\lambda,h}$ on the grid of size $h = \frac{14\sqrt{2}}{2^7}$ with $\lambda = 10^{-4.5}$ are displayed in Figures 3 and 4. As stated in Example 5.1, if a solution with moderate accuracy is sufficient, both hADMM and PDAS are terminated when $\eta_{\rm A}(\eta_{\rm P}) < 10^{-3}$ and the corresponding numerical results are displayed in Table 7. Moreover, if more accurate solution is required, we employ the two-phase strategy and compare it with PDAS. Both two algorithms are terminated when $\eta_{\rm A}(\eta_{\rm P}) < 10^{-13}$ in this case and the numerical results are given in Table 8.

Table 5 shows that when $\lambda$ is fixed, the error declines as $h$ decreases until the error is up to a lower bound caused by the regularization. While for a fixed $h$, the error declines as $\lambda$ decreases. The data in Table 5 verify the error estimates in Section 3. The last two tables in this example are similar to their counterparts in Example 5.1. It is easy to see from the numerical results that the hADMM algorithm and the two-phase strategy are faster than PDAS method especially when the finite element grid size $h$ is very small, which verifies the efficiency of the hADMM algorithm and the two-phase strategy.

TABLE 7. The convergence behavior of hADMM and PDAS (a special semi-smooth Newton method) for Example 5.2.

| $h$ | #dofs | $\lambda$ | | hADMM | PDAS |
|---|---|---|---|---|---|
| $\frac{14\sqrt{2}}{2^7}$ | 16 129 | $10^{-4}$ | Iter | 31 | 3 |
| | | | Residual $\eta$ | 6.84e−04 | 3.66e−04 |
| | | | Time/s | 0.39 | 1.55 |
| | | $10^{-4.5}$ | Iter | 31 | 3 |
| | | | Residual $\eta$ | 7.06e−04 | 4.01e−04 |
| | | | Time/s | 0.36 | 1.46 |
| | | $10^{-5}$ | Iter | 30 | 3 |
| | | | Residual $\eta$ | 9.12e−04 | 4.12e−04 |
| | | | Time/s | 0.34 | 1.44 |
| $\frac{14\sqrt{2}}{2^8}$ | 65 025 | $10^{-4.5}$ | Iter | 30 | 6 |
| | | | Residual $\eta$ | 9.39e−04 | 3.18e−04 |
| | | | Time/s | 4.01 | 17.25 |
| | | $10^{-5}$ | Iter | 31 | 6 |
| | | | Residual $\eta$ | 6.48e−04 | 5.71e−04 |
| | | | Time/s | 4.19 | 16.88 |
| | | $10^{-5.5}$ | Iter | 31 | 6 |
| | | | Residual $\eta$ | 4.69e−04 | 3.85e−04 |
| | | | Time/s | 4.25 | 17.21 |
| $\frac{14\sqrt{2}}{2^9}$ | 261 121 | $10^{-5}$ | Iter | 30 | 14 |
| | | | Residual $\eta$ | 5.76e−04 | 2.56e−04 |
| | | | Time/s | 80.38 | 267.73 |
| | | $10^{-5.5}$ | Iter | 30 | 14 |
| | | | Residual $\eta$ | 9.10e−04 | 4.24e−04 |
| | | | Time/s | 80.14 | 264.62 |
| | | $10^{-6}$ | Iter | 30 | 14 |
| | | | Residual $\eta$ | 9.15e−04 | 5.37e−04 |
| | | | Time/s | 81.23 | 267.02 |

Through the numerical results of Example 5.2, we could also see that the number of iteration of both hADMM and PDAS are independent of the discretization level. First, from Table 7, it is easy to see that when $\lambda = 10^{-5}$ the number of iteration of hADMM for three discretization levels are 30, 31 and 30 respectively. Furthermore, it is well known that the Semi-smooth Newton method is locally superlinearly convergent, so we would like to point out that the mesh independence of PDAS is also expressed locally. For Example 5.2, we could subtract the last column of Tables 7 and 8. It is not difficult to see that the difference is independent of the discretization level. For example, when $\lambda = 10^{-5}$, the number of iteration of PDAS for three discretization levels when accuracy improving from $10^{-4}$ to $10^{-13}$ are 3, 4 and 6 respectively.

TABLE 8. The convergence behavior of the two-phase strategy and PDAS (a special semi-smooth Newton method) for Example 5.2.

| $h$ | #dofs | $\lambda$ | | two-phase strategy (hADMM \| PDAS) | PDAS |
|---|---|---|---|---|---|
| $\frac{14\sqrt{2}}{2^7}$ | 16 129 | $10^{-4}$ | Iter | 31 \| 4 | 6 |
| | | | Residual $\eta$ | 6.84e−04 \| 8.64e−14 | 8.67e−14 |
| | | | Time/s | 2.30 (0.36 \| 1.94) | 2.94 |
| | | $10^{-4.5}$ | Iter | 31 \| 4 | 7 |
| | | | Residual $\eta$ | 7.06e−04 \| 8.69e−14 | 8.68e−14 |
| | | | Time/s | 2.32 (0.37 \| 1.95) | 3.49 |
| | | $10^{-5}$ | Iter | 30 \| 4 | 6 |
| | | | Residual $\eta$ | 9.12e−04 \| 7.84e−14 | 7.84e−14 |
| | | | Time/s | 2.32 (0.36 \| 1.96) | 2.97 |
| $\frac{14\sqrt{2}}{2^8}$ | 65 025 | $10^{-4.5}$ | Iter | 30 \| 6 | 11 |
| | | | Residual $\eta$ | 9.39e−04 \| 3.58e−13 | 3.57e−13 |
| | | | Time/s | 22.40 (4.81 \| 17.59) | 32.87 |
| | | $10^{-5}$ | Iter | 31 \| 6 | 10 |
| | | | Residual $\eta$ | 6.48e−04 \| 2.87e−13 | 2.87e−13 |
| | | | Time/s | 22.18 (4.53 \| 17.65) | 29.51 |
| | | $10^{-5.5}$ | Iter | 31 \| 6 | 10 |
| | | | Residual $\eta$ | 4.69e−04 \| 4.38e−13 | 4.37e−13 |
| | | | Time/s | 22.50 (4.78 \| 17.72) | 29.73 |
| $\frac{14\sqrt{2}}{2^9}$ | 261 121 | $10^{-5}$ | Iter | 30 \| 8 | 20 |
| | | | Residual $\eta$ | 5.76e−04 \| 7.73e−13 | 7.73e−13 |
| | | | Time/s | 231.48 (80.94 \| 150.54) | 385.78 |
| | | $10^{-5.5}$ | Iter | 30 \| 8 | 21 |
| | | | Residual $\eta$ | 9.10e−04 \| 8.33e−13 | 8.34e−13 |
| | | | Time/s | 231.88 (80.94 \| 150.94) | 402.96 |
| | | $10^{-6}$ | Iter | 30 \| 8 | 22 |
| | | | Residual $\eta$ | 9.15e−04 \| 9.53e−13 | 9.53e−13 |
| | | | Time/s | 232.51 (80.28 \| 152.23) | 420.21 |

## 6. Conclusion

In this paper, state-constrained elliptic control problems are considered, where the Lagrange multipliers associated to the state constraints are only measure functions. To tackle this difficulty, we utilize Lavrentiev regularization. After that, the regularized problem is discretized by full finite element discretization, in which both the state and control are discretized by piecewise linear functions. We derive error analysis of the overall error resulted from regularization and discretization. To solve the discretized problem efficiently, a heterogeneous alternating direction method of multipliers (hADMM) is proposed. If more accurate solution is required, a two-phase strategy is proposed, in which the primal-dual active set (PDAS) method is used as a postprocessor of the hADMM. Numerical results not only verify the analysis results of error estimate but also show the efficiency of the proposed algorithm.

*Acknowledgements.* We would like to thank Dr. Long Chen very much for the contribution of the FEM package iFEM [10] in Matlab. Also we are grateful for valuable suggestions of our colleagues.

## References


[1] J.J. Alibert and J.P. Raymond, Boundary control of semilinear elliptic equations with discontinuous leading coefficients and unbounded controls. *Numer. Func. Anal. Opt.* **18** (1997) 235–250.

[2] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2** (2009) 183–202.

[3] M. Bergounioux and K. Kunisch, Augemented Lagrangian techniques for elliptic state constrained optimal control problems. *SIAM J. Control Optim.* **35** (1997) 1524–1543.

[4] M. Bergounioux and K. Kunisch, Primal-dual strategy for state-constrained optimal control problems. *Comput. Optim. Appl.* **22** (2002) 193–224.

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

[6] C. Carstensen, Quasi-interpolation and *a posteriori* error analysis in finite element methods. *ESAIM: M2AN* **33** (1999) 1187–1202.

[7] E. Casas, $L^2$ estimates for the finite element method for the Dirichlet problem with singular data. *Numer. Math.* **47** (1985) 627–632.

[8] E. Casas, Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control Optim.* **31** (1993) 993–1006.

[9] T.F. Chan and R. Glowinski, Finite Element Approximation and Iterative Solution of a Class of Mildly Non-Linear Elliptic Equations. Computer Science Department, Stanford University, Stanford (1978).

[10] L. Chen, iFEM: An Integrated Finite Element Methods Package in MATLAB. Technical Report. University of California at Irvine, Irvine (2009).

[11] L. Chen, D.F. Sun and K.C. Toh, An efficient inexact symmetric Gauss–Seidel based majorized ADMM for high-dimensional convex composite conic programming. *Math. Program.* **161** (2017) 237–270.

[12] K. Deckelnick and M. Hinze, Numerical analysis of a control and state constrained elliptic control problem with piecewise constant control approximations. *Numer. Math. Adv. Appl.* Springer (2008) 597–604.

[13] J.C. De Los Reyes, C. Meyer and B. Vexler, Finite element error analysis for state-constrained optimal control of the Stokes equations. *WIAS* (2008).

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

[15] M. Fazel, T.K. Pong, D.F. Sun and P. Tseng, Hankel matrix rank minimization with applications to system identification and realization. *SIAM J. Matrix Anal. Appl.* **34** (2013) 946–977.

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

[17] W. Gong and N. Yan, A mixed finite element scheme for optimal control problems with pointwise state constraints. *J. Sci. Comput.* **46** (2011) 182–203.

[18] P. Grisvard, Elliptic problems in nonsmooth domains. Pitman Advanced Pub. Program (1985).

[19] M. Hintermüller and K. Kunisch, Feasible and noninterior path-following in constrained minimization with low multiplier regularity. *SIAM J. Control Optim.* **45** (2006) 1198–1221.

[20] M. Hintermüller and K. Kunisch, Path-following methods for a class of constrained minimization problems in function space. *SIAM J. Optim.* **17** (2006) 159–187.

[21] M. Hintermüller, K. Ito and K. Kunisch, The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13** (2002) 865–888.

[22] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.* **30** (2005) 45–61.

[23] M. Hinze and M. Christian, Variational discretization of Lavrentiev-regularized state constrained elliptic optimal control problems. *Comput. Optim. Appl.* **46** (2010) 487–510.

[24] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich. Optimization with PDE Constraints. Springer, Netherlands (2009).

[25] K. Itoa and K. Kunisch, Semi-smooth Newton methods for state-constrained optimal control problems. *Syst. Control Lett.* **50** (2003) 221–228.

[26] K.F. Jiang, D.F. Sun and K.C. Toh, An inexact accelerated proximal gradient method for large scale linearly constrained convex SDP. *SIAM J. Optim.* **22** (2012) 1042–1064.

[27] K. Kunisch, K. Liang and X. Lu, Optimal control for an elliptic system with polygonal state constraints. *SIAM J. Control Optim.* **48** (2010) 5053–5072.

[28] X.D. Li, D.F. Sun and K.C. Toh, QSDPNAL: A Two-Phase Newton-CG Proximal Augmented Lagrangian Method for Convex Quadratic Semidefinite Programming Problems. Preprint arXiv:1512.08872 (2015).

[29] X.D. Li, D.F. Sun and K.C. Toh, A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions. *Math. Program.* **155** (2016) 333–373.

[30] W.B. Liu and N.N. Yan, Adaptive Finite Element Methods for Optimal Control Governed by PDEs. Science Press, Beijing (2008).

[31] C. Meyer, U. Prüfert and F. Tröltzsch, On two numerical methods for state-constrained elliptic control problems. *Optim. Method Softw.* **22** (2007) 871–899.

[32] C. Meyer, A. Rösch and F. Tröltzsch, Optimal Control of PDEs with regularized pointwise state constraints. *Comput. Optim. Appl.* **33** (2006) 209–228.

[33] G.C. Philippe, The Finite Element Method for Elliptic Problems. North-Holland Publ. Company (1978).

[34] J.W. Pearson, S. Martin and A.J. Wathen, Preconditioners for state-constrained optimal control problems with Moreau-Yosida penalty function. *Numer. Linear Algebra Appl.* **21** (2014) 81–97.

[35] M. Porcelli, V. Simoncini and M. Stoll, Preconditioning PDE-Constrained Optimization with $L^1$-Sparsity and Control Constraints. Preprint arXiv:1611.07201 (2016).

[36] U. Prüfert, F. Tröltzsch and M. Weiser, The convergence of an interior point method for an elliptic control problem with mixed control-state constraints. *Comput. Optim. Appl.* **39** (2008) 183–218.

[37] Z. Wu, J. Yin and C. Wang, Elliptic and parabolic equations. In vol. 118 of *Springer Proceedings in Mathematics & Statistics* (1994).

[38] A. Schindele and A. Borzì, Proximal methods for elliptic optimal control problems with sparsity cost functional. *Appl. Math.* **7** (2016) 967.

[39] X.L. Song and B. Yu, A two-phase strategy for control constrained elliptic optimal control problems. *Numer. Linear Algebra Appl.* (to appear).

[40] K.C. Toh and S. Yun, An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. *Pac. J. Optim.* **6** (2010) 615-640.

[41] P. Tseng, On accelerated proximal gradient methods for convex–concave optimization. *SIAM J. Optim.* (2008).

[42] M. Ulbrich, *Nonsmooth Newton-like Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. Ph.D. thesis, Habilitation thesis, Fakultät für Mathematik, Technische Universität München (2002).

[43] M. Ulbrich, Semismooth newton methods for operator equations in function spaces. *SIAM J. Optim.* **13** (2002) 805–841.

[44] A.J. Wathen, Realistic eigenvalue bounds for the Galerkin mass matrix. *IMA J. Numer. Anal.* **7** (1987) 449–457.