Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2017.m1419

Vol. **10**, No. 3, pp. 614-638
August 2017

Spectral Method Approximation of Flow Optimal Control Problems with H^1-Norm State Constraint

Yanping Chen[1,*] and Fenglin Huang[2]

[1] *School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China*
[2] *School of Mathematics and Statistics, Xinyang Normal University, No.237 Nanhu Road, Shihe District, Xinyang 464000, China*

Received 1 July 2014; Accepted (in revised version) 7 June 2016

Abstract. In this paper, we consider an optimal control problem governed by Stokes equations with H^1-norm state constraint. The control problem is approximated by spectral method, which provides very accurate approximation with a relatively small number of unknowns. Choosing appropriate basis functions leads to discrete system with sparse matrices. We first present the optimality conditions of the exact and the discrete optimal control systems, then derive both a priori and a posteriori error estimates. Finally, an illustrative numerical experiment indicates that the proposed method is competitive, and the estimator can indicate the errors very well.

AMS subject classifications: 49J20, 65K15, 65M12, 65M70

Key words: Optimal control, state constraint, stokes equations, legendre polynomials, spectral method.

1. Introduction

Flow control problems can be met frequently in many engineering applications; for instance, leading the velocity to a desired one, minimizing turbulence in a flow, enhancing or deterring mixing and designing the optimal shapes to minimize drag in flows for many industrial devices such as aircraft wings, automobile shapes and boats, etc. Fast and efficient numerical methods play an important role in successful applications of flow control problems. Finite element method has been widely used in solving control problems. One can find a systematic introduction to the theoretical analysis and finite element approximation for control problems including flow control problems in, for example, [16, 20, 23, 24, 27, 28]. However, most of these results focus on control-constrained cases. Recently, control problems with state constraints have been widely discussed. On the theoretical aspects, we refer to [2, 14], and have to mention the works of Casas (see, e.g., [6, 7]). In respect of numerical approximation, many studies have been carried out

*Corresponding author. *Email addresses:* yanpingchen@scnu.edu.cn (Y. P. Chen), hfl_937@sina.com (F. L. Huang)

 ©2017 Global-Science Press

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

to examine finite element analysis for this class of control problems, including point-wise constraint, integral constraint, and L^2-norm constraint (see, e.g., [1,12,19,21,22,25,31]). To compute control problems with state constraints, many numerical strategies are proposed. [3] takes an augmented Lagrangian method to solve state and control constrained problems. Primal-dual active set algorithm is used to solve state constrained problems in [4]. A semi-smooth Newton method is adopted for approximating the regularized point-wise state-constrained optimal control problems in [15]. In [17], the optimal control problem is reformulated to a constrained minimization problem involving only the state, which is characterized by a fourth order variational inequality, then a mixed variational scheme is proposed.

Employing global polynomials as the trial functions, the spectral method has been successfully applied in numerical solutions of PDEs, especially in the field of computational fluid dynamics (see, for example, [8–10, 30], and the references therein). Generally, the solutions to the optimal control problems have limited regularity due to, e.g., the constraints, which leads to the lose of spectral accuracy. However, spectral method enjoys the great superiorities of high-precision and fast convergence rate when the approximated solutions have higher regularity, which is vital to efficient approximation of optimal control problems. Therefore, the spectral method has gained increasing popularity in solving control problems governed by PDEs in the past several years. In [13], spectral method is used to approximate the elliptic control problems with integral control constraint, both a priori and a posteriori error estimates are derived, and numerical tests confirm the efficiency of the proposed method. In [32], a priori error estimates are derived for control problems with integral state constraint. Generally, control problems with state constraints are far more complicated to analyze than control-constrained ones. To our best knowledge, there has been a lack of discussion on the spectral approximation of flow control problems with H^1-norm state constraint, though this class of control problems can be applied in many practical considerations. In fact, for many casting products made by pouring molten metals into mold cavities, the velocity and its derivative of the molten metals should be constrained to ensure the quality of the products. Moreover, it is better to restrict the velocity and its derivative of the flow that passing through the corner. In these cases, we can try to use the control models with H^1-norm constraints for flow velocity.

Let $\Omega \subset \mathbb{R}^2$ be a rectangular domain, C denotes a general positive constant independent of N, the order of the spectral approximation. We shortly describe the structure of our paper as follows: In Section 2, a priori error estimates are derived for Galerkin spectral approximation of the flow control problem. Then, a posteriori error estimates are proposed in Section 3. Finally, an illustrative numerical experiment is presented in Section 4.

2. A priori error estimates

In this section, the optimality conditions are derived, and the spectral approximation of the control problem is presented. Then a priori error estimates are established. Let $\boldsymbol{U} = L^2(\Omega)^2$, $\boldsymbol{Y} = H_0^1(\Omega)^2$, $Q = L_0^2(\Omega) = \left\{q \in L^2(\Omega) \mid \int_\Omega q dx = 0\right\}$. We consider the

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

following control problem: find $(\boldsymbol{y}, r, \boldsymbol{u}) \in \boldsymbol{Y} \times Q \times \boldsymbol{U}$ such that

$$\begin{aligned} &\min_{\boldsymbol{y}(\boldsymbol{u}) \in \boldsymbol{K}} \ J(\boldsymbol{u}) = \frac{1}{2}\|\boldsymbol{y}(\boldsymbol{u}) - \boldsymbol{y}_0\|_{0,\Omega}^2 + \frac{\alpha}{2}\|\boldsymbol{u} - \boldsymbol{u}_0\|_{0,\Omega}^2, \\ &-\Delta \boldsymbol{y}(\boldsymbol{u}) + \nabla r = \boldsymbol{u} \ \text{ in } \Omega, \quad \nabla \cdot \boldsymbol{y}(\boldsymbol{u}) = 0 \ \text{ in } \Omega, \quad \boldsymbol{y}(\boldsymbol{u}) = 0 \ \text{ on } \partial\Omega, \end{aligned} \tag{2.1}$$

where

$$\boldsymbol{K} = \{\boldsymbol{w} \in \boldsymbol{Y} \ \mid \ \|\boldsymbol{w}\|_{1,\Omega} \le d\}, \quad \boldsymbol{y}_0 \in \boldsymbol{U}$$

is the observation, $\boldsymbol{u}_0 \in \boldsymbol{U}$ is a given function, and d, α are positive constants. To consider the spectral approximation of the problem (2.1), we have to give a weak formula for the state equations. Let

$$a(\boldsymbol{y}, \boldsymbol{w}) = \int_\Omega \nabla \boldsymbol{y} \cdot \nabla \boldsymbol{w} dx = \sum_{i,j=1}^{2} \int_\Omega \frac{\partial y_i}{\partial x_j} \cdot \frac{\partial w_i}{\partial x_j} dx, \qquad \forall\, \boldsymbol{y}, \boldsymbol{w} \in \boldsymbol{Y}, \tag{2.2a}$$

$$b(\boldsymbol{v}, r) = \int_\Omega r \nabla \cdot \boldsymbol{v} dx, \qquad \forall\, (\boldsymbol{v}, r) \in \boldsymbol{Y} \times Q, \tag{2.2b}$$

$$(\boldsymbol{u}, \boldsymbol{w}) = \int_\Omega \boldsymbol{u} \cdot \boldsymbol{w} dx, \qquad \forall\, (\boldsymbol{u}, \boldsymbol{w}) \in \boldsymbol{U} \times \boldsymbol{U}. \tag{2.2c}$$

We can see that there exist two constants $\sigma > 0$ and $\delta > 0$ such that

$$| a(\boldsymbol{u}, \boldsymbol{v}) | \le \sigma \, \| \boldsymbol{u} \|_{\boldsymbol{Y}} \| \boldsymbol{v} \|_{\boldsymbol{Y}}, \qquad \forall\, \boldsymbol{u}, \ \boldsymbol{v} \in \boldsymbol{Y}, \tag{2.3a}$$

$$| b(\boldsymbol{v}, q) | \le \delta \, \| \boldsymbol{v} \|_{\boldsymbol{Y}} \| q \|_{Q}, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{Y}, \ q \in Q. \tag{2.3b}$$

Furthermore, it is well known that there exist two positive constants θ, β such that

$$a(\boldsymbol{y}, \boldsymbol{y}) \ge \theta \|\boldsymbol{y}\|_{\boldsymbol{Y}}^2, \qquad \forall\, \boldsymbol{y} \in \boldsymbol{Y}, \tag{2.4a}$$

$$\sup_{\boldsymbol{v} \in \boldsymbol{Y}} \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{Y}}} \ge \beta \|q\|_Q, \qquad \forall\, q \in Q. \tag{2.4b}$$

Now, the standard weak formula for the state equations can be presented as follows: find $(\boldsymbol{y}(\boldsymbol{u}), r(\boldsymbol{u})) \in \boldsymbol{Y} \times Q$ such that

$$a(\boldsymbol{y}(\boldsymbol{u}), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u})) = (\boldsymbol{u}, \boldsymbol{w}), \qquad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \tag{2.5a}$$

$$b(\boldsymbol{y}(\boldsymbol{u}), \phi) = 0, \qquad \forall\, \phi \in Q, \tag{2.5b}$$

which is well-posed by (2.4a), (2.4b), and the Babuška-Brezzi theorem. Then the control problem (2.1) can be restated as follows (OCP): find $(\boldsymbol{y}, r, \boldsymbol{u}) \in \boldsymbol{Y} \times Q \times \boldsymbol{U}$ such that

$$\begin{aligned} &\min_{\boldsymbol{y}(\boldsymbol{u}) \in \boldsymbol{K}} \ J(\boldsymbol{u}) = \frac{1}{2}\|\boldsymbol{y}(\boldsymbol{u}) - \boldsymbol{y}_0\|_{0,\Omega}^2 + \frac{\alpha}{2}\|\boldsymbol{u} - \boldsymbol{u}_0\|_{0,\Omega}^2, \\ &a(\boldsymbol{y}(\boldsymbol{u}), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u})) = (\boldsymbol{u}, \boldsymbol{w}), \qquad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \\ &b(\boldsymbol{y}(\boldsymbol{u}), \phi) = 0, \qquad \forall\, \phi \in Q. \end{aligned}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

Utilizing the theory in [20], the uniqueness of the solution of control problem (OCP) can be derived by reformulating it into a control-constrained problem.

We now derive the optimality conditions for control problem (OCP), and introduce the Lagrange functional $L(\boldsymbol{u},\lambda)$ as follows

$$L(\boldsymbol{u},\lambda)=J(\boldsymbol{u})+\lambda F(\boldsymbol{u}),$$

where

$$F(\boldsymbol{u})=\frac{1}{2}\left(\|\boldsymbol{y}(\boldsymbol{u})\|_{1,\Omega}^2-d^2\right).$$

It can be verified that $J(\boldsymbol{u})$ and $F(\boldsymbol{u})$ are Fréchet differentiable, and the Slater type condition $F(0)<0$ is satisfied. By the results in [5, 11], we can deduce that there exists $\lambda\in\mathbb{R}$ such that

$$(1)\ \ \lambda\geq 0,\quad \lambda F(\boldsymbol{u})=0, \tag{2.6a}$$

$$(2)\ \ \frac{\partial L}{\partial \boldsymbol{u}}(\boldsymbol{u},\lambda)=0, \tag{2.6b}$$

provided $\boldsymbol{u}$ is the solution of control problem (OCP).

Differentiate the state Eqs. (2.5) at $\boldsymbol{u}$ in a direction $\boldsymbol{v}$ to get that

$$a(\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v},\boldsymbol{w})-b(\boldsymbol{w},r'(\boldsymbol{u})\boldsymbol{v})=(\boldsymbol{v},\boldsymbol{w}),\qquad \forall\ \boldsymbol{w}\in\boldsymbol{Y}, \tag{2.7a}$$

$$b(\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v},\phi)=0,\qquad \forall\ \phi\in Q. \tag{2.7b}$$

Let $\boldsymbol{w}=\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}$ in (2.5) and use (2.7b), we can conclude that

$$(\nabla\boldsymbol{y},\nabla(\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}))=(\boldsymbol{u},\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}).$$

Furthermore, it can be derived that the equality (2) in (2.6) holds if and only if

$$(\boldsymbol{y}-\boldsymbol{y}_0,\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v})+\lambda(\boldsymbol{y},\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v})+\lambda(\nabla\boldsymbol{y},\nabla(\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}))+\alpha(\boldsymbol{u}-\boldsymbol{u}_0,\boldsymbol{v})=0,\quad \forall\ \boldsymbol{v}\in\boldsymbol{U},$$

or

$$(\boldsymbol{y}-\boldsymbol{y}_0+\lambda(\boldsymbol{u}+\boldsymbol{y}),\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v})+\alpha(\boldsymbol{u}-\boldsymbol{u}_0,\boldsymbol{v})=0,\quad \forall\ \boldsymbol{v}\in\boldsymbol{U}. \tag{2.8}$$

To eliminate the item $\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}$ in (2.8), we introduce the co-state equations as follows

$$a(\boldsymbol{q},\boldsymbol{y}^*)+b(\boldsymbol{q},r^*)=(\boldsymbol{y}-\boldsymbol{y}_0+\lambda(\boldsymbol{u}+\boldsymbol{y}),\boldsymbol{q}),\qquad \forall\ \boldsymbol{q}\in\boldsymbol{Y}, \tag{2.9a}$$

$$b(\boldsymbol{y}^*,\psi)=0,\qquad \forall\ \psi\in Q. \tag{2.9b}$$

Let $\boldsymbol{w}=\boldsymbol{y}^*$ in (2.7a) and $\boldsymbol{q}=\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}$ in (2.9), we can obtain that

$$(\boldsymbol{y}-\boldsymbol{y}_0+\lambda(\boldsymbol{u}+\boldsymbol{y}),\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v})=(\boldsymbol{v},\boldsymbol{y}^*),\qquad \forall\ \boldsymbol{v}\in\boldsymbol{U},$$

associating with (2.8), the equality (2) in (2.6) equals to

$$(\boldsymbol{y}^*+\alpha(\boldsymbol{u}-\boldsymbol{u}_0),\boldsymbol{v})=0,\qquad \forall\ \boldsymbol{v}\in\boldsymbol{U},$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

or

$$\boldsymbol{y}^* + \alpha(\boldsymbol{u} - \boldsymbol{u}_0) = 0.$$

Then it can be derived that the triplet $(\boldsymbol{y}, r, \boldsymbol{u})$ is the solution of control problem (OCP) if and only if there is a

$$(\boldsymbol{y}^*, r^*, \lambda) \in \boldsymbol{Y} \times Q \times \mathbb{R}^1_+ (\mathbb{R}^1_+ = \{c \in \mathbb{R}^1 : c \geq 0\}),$$

such that $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ satisfies the following optimality conditions:

$$a(\boldsymbol{y}, \boldsymbol{w}) - b(\boldsymbol{w}, r) = (\boldsymbol{u}, \boldsymbol{w}), \qquad \forall\ \boldsymbol{w} \in \boldsymbol{Y}, \tag{2.10a}$$
$$b(\boldsymbol{y}, \phi) = 0, \qquad \forall\ \phi \in Q, \tag{2.10b}$$
$$a(\boldsymbol{q}, \boldsymbol{y}^*) + b(\boldsymbol{q}, r^*) = (\boldsymbol{y} - \boldsymbol{y}_0 + \lambda(\boldsymbol{u} + \boldsymbol{y}), \boldsymbol{q}), \qquad \forall\ \boldsymbol{q} \in \boldsymbol{Y}, \tag{2.10c}$$
$$b(\boldsymbol{y}^*, \psi) = 0, \qquad \forall\ \psi \in Q, \tag{2.10d}$$
$$\boldsymbol{y}^* + \alpha(\boldsymbol{u} - \boldsymbol{u}_0) = 0, \qquad \text{in } \Omega, \tag{2.10e}$$

where

$$\lambda = \begin{cases} \text{constant} \geq 0 & \text{if } \|\boldsymbol{y}\|_{1,\Omega} = d, \\ 0 & \text{if } \|\boldsymbol{y}\|_{1,\Omega} < d. \end{cases} \tag{2.11}$$

By the optimality conditions, we can derive the regularity of the optimal control.

Remark 2.1. Let $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ satisfy the optimality conditions (2.10), then the optimal control $\boldsymbol{u} \in H^2(\Omega)^2$, provided the data functions $\boldsymbol{u}_0$, $\boldsymbol{y}_0$ are sufficiently regular (see, e.g., [13]).

Let Q_N denote the space of all algebraic polynomials of degree less than or equal to N with respect to each single variable x_i, $(i = 1, 2)$. The finite dimensional spaces $\boldsymbol{Y}_N$, M_{N-2}, and $\boldsymbol{U}_N$ are used to approximate spaces $\mathbf{Y}$, Q, and $\boldsymbol{U}$ respectively, where $\boldsymbol{Y}_N = (Q_N^0)^2$ with

$$Q_N^0 = \{\boldsymbol{x}_N \in Q_N : \boldsymbol{x}_N \,|_{\partial\Omega} = 0\}, \quad M_{N-2} = Q_{N-2} \cap L_0^2(\Omega), \quad \textit{and} \quad \boldsymbol{U}_N = (Q_N \cap L^2(\Omega))^2.$$

Then we approximate state equations (2.5) as follows:

$$a(\boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N) = (\boldsymbol{u}_N, \boldsymbol{w}_N), \qquad \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N \subset \boldsymbol{Y}, \tag{2.12a}$$
$$b(\boldsymbol{y}_N, \phi_N) = 0, \qquad \forall\ \phi_N \in M_{N-2} \subset Q. \tag{2.12b}$$

It is well known from [8, Chapter 11] and [30, Chapter 7, §3] that there exists a constant $\beta_N = \mathcal{O}(N^{-\frac{1}{2}})$ such that

$$\sup_{\boldsymbol{v}_N \in \boldsymbol{Y}_N} \frac{b(\boldsymbol{v}_N, q_N)}{\|\boldsymbol{v}_N\|_{\boldsymbol{Y}}} \geq \beta_N \|q_N\|_Q, \quad \forall\ q_N \in M_{N-2}. \tag{2.13}$$

Applying the Babuška-Brezzi theorem, problem (2.12) is well-posed.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

Then the $Q_N - Q_{N-2}$ spectral approximation of (OCP) can be stated as $(\mathrm{OCP})_\mathrm{N}$:

$$\min_{\boldsymbol{y}_N \in \boldsymbol{K}_N} J_N(\boldsymbol{u}_N) = \frac{1}{2}\left\|\boldsymbol{y}_N - \boldsymbol{y}_0\right\|_{0,\Omega}^2 + \frac{\alpha}{2}\left\|\boldsymbol{u}_N - \boldsymbol{u}_0\right\|_{0,\Omega}^2, \tag{2.14a}$$

$$a(\boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N) = (\boldsymbol{u}_N, \boldsymbol{w}_N), \qquad \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N \subset \boldsymbol{Y}, \tag{2.14b}$$

$$b(\boldsymbol{y}_N, \phi_N) = 0, \qquad \forall\ \phi_N \in M_{N-2} \subset Q, \tag{2.14c}$$

where $\boldsymbol{K}_N = \boldsymbol{Y}_N \cap \boldsymbol{K}$.

Similarly, the discrete Slater type constraint qualification condition $F_N(0) < 0$ can be satisfied with

$$F_N(u_N) = \frac{1}{2}(\|y_N(u_N)\|_{1,\Omega}^2 - d^2).$$

By the results in [5, 11], the problem $(\mathrm{OCP})_\mathrm{N}$ has a unique solution $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N)$ and that

$$(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N) \in \boldsymbol{Y}_N \times M_{N-2} \times \boldsymbol{U}_N$$

is the solution of $(\mathrm{OCP})_\mathrm{N}$ if and only if there is

$$(\boldsymbol{y}_N^*, r_N^*, \lambda_N) \in \boldsymbol{Y}_N \times M_{N-2} \times \mathbb{R}_+^1,$$

such that $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ satisfies the following optimality conditions:

$$a(\boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N) = (\boldsymbol{u}_N, \boldsymbol{w}_N), \qquad \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N, \tag{2.15a}$$

$$b(\boldsymbol{y}_N, \phi_N) = 0, \qquad \forall\ \phi_N \in M_{N-2}, \tag{2.15b}$$

$$a(\boldsymbol{q}_N, \boldsymbol{y}_N^*) + b(\boldsymbol{q}_N, r_N^*) = (\boldsymbol{y}_N - \boldsymbol{y}_0 + \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \boldsymbol{q}_N), \qquad \forall\ \boldsymbol{q}_N \in \boldsymbol{Y}_N, \tag{2.15c}$$

$$b(\boldsymbol{y}_N^*, \psi_N) = 0, \qquad \forall\ \psi_N \in M_{N-2}, \tag{2.15d}$$

$$\boldsymbol{y}_N^* + \alpha \boldsymbol{u}_N - \alpha P_N \boldsymbol{u}_0 = 0, \qquad \text{in } \Omega, \tag{2.15e}$$

where

$$\lambda_N = \begin{cases} \text{constant} \geq 0 & \text{if } \|\boldsymbol{y}_N\|_{1,\Omega} = d, \\ 0 & \text{if } \|\boldsymbol{y}_N\|_{1,\Omega} < d, \end{cases} \tag{2.16}$$

and P_N is the L^2 orthogonal projection operator, the details will be specified later.

We first present some helpful results for establishing the a priori error estimates. Define the projection operator $P_{1,N}^0 : H_0^1(\Omega)^2 \to (Q_N^0)^2$ (see [8, Chapter 9, §7]), which satisfies for any $\boldsymbol{w} \in H_0^1(\Omega)^2$

$$\int_\Omega \nabla(\boldsymbol{w} - P_{1,N}^0 \boldsymbol{w}) \cdot \nabla \boldsymbol{q}_N = 0, \qquad \forall\ \boldsymbol{q}_N \in (Q_N^0)^2.$$

Then it can be obtained that

$$\|\boldsymbol{w} - P_{1,N}^0 \boldsymbol{w}\|_{k,\Omega} \leq C N^{k-m} \|\boldsymbol{w}\|_{m,\Omega}, \qquad k = 0, 1, \tag{2.17}$$

provided $\boldsymbol{w} \in \left(H^m(\Omega) \cap H_0^1(\Omega)\right)^2$ with $m \geq 1$.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

For any $r \in L^2(\Omega)$, we define $P_N : L^2(\Omega) \to Q_N$ as the L^2 orthogonal projection operator (see [9, Chapter 5, §5.8.2]), which satisfies

$$(r - P_N r, \mu_N) = 0, \quad \forall\ \mu_N \in Q_N.$$

Then we have

$$\|r - P_N r\|_{0,\Omega} \le CN^{-m}\|r\|_{m,\Omega}, \tag{2.18}$$

provided $r \in H^m(\Omega)$ with $m \ge 0$. The result (2.18) can be certainly extended to vector functions.

Lemma 2.1. (Chapter 3, 10). *For each divergence-free function $\boldsymbol{y} \in \boldsymbol{Y} \cap H^m(\Omega)^2$ with $m \ge 1$, there exists a divergence-free function $\Phi_N \in \boldsymbol{Y}_N$ such that*

$$\|\boldsymbol{y} - \Phi_N\|_{1,\Omega} \le CN^{1-m}\|\boldsymbol{y}\|_{m,\Omega}. \tag{2.19}$$

Lemma 2.2. *Let $(\boldsymbol{y}_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, \lambda_N)$ satisfy optimality conditions* (2.15)*, then we have*

$$\max\left\{\|\boldsymbol{u}_N\|_{0,\Omega}, \|\boldsymbol{y}_N\|_{1,\Omega}, \|\boldsymbol{y}_N^*\|_{1,\Omega}, |\lambda_N|\right\} \le C. \tag{2.20}$$

Proof. We can find that

$$\frac{1}{2}\int_\Omega |\boldsymbol{y}_N - \boldsymbol{y}_0|^2 dx + \frac{\alpha}{2}\int_\Omega |\boldsymbol{u}_N - \boldsymbol{u}_0|^2 dx \le J_N(0) \le C\int_\Omega (|\boldsymbol{y}_0|^2 + |\boldsymbol{u}_0|^2) dx. \tag{2.21}$$

It follows from (2.15e) and (2.21) that

$$\|\boldsymbol{y}_N\|_{1,\Omega} + \|\boldsymbol{y}_N^*\|_{0,\Omega} + \|\boldsymbol{u}_N\|_{0,\Omega} \le C. \tag{2.22}$$

To estimate $\|\boldsymbol{y}_N^*\|_{1,\Omega}$, we divide the discussions into two cases

$$\|\boldsymbol{y}_N\|_{1,\Omega} < d \qquad \text{and} \qquad \|\boldsymbol{y}_N\|_{1,\Omega} = d.$$

If $\|\boldsymbol{y}_N\|_{1,\Omega} < d$, we have $\lambda_N = 0$ by (2.16). Let $\boldsymbol{q}_N = \boldsymbol{y}_N^*$ in (2.15), then

$$\|\boldsymbol{y}_N^*\|_{1,\Omega} \le C. \tag{2.23}$$

If $\|\boldsymbol{y}_N\|_{1,\Omega} = d$, then let

$$\boldsymbol{q}_N = \boldsymbol{y}_N^* - \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)\boldsymbol{y}_N$$

in (2.15), it holds

$$\begin{aligned} &a(\boldsymbol{y}_N^* - \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)\boldsymbol{y}_N, \boldsymbol{y}_N^*) \\ =&(\boldsymbol{y}_N - \boldsymbol{y}_0 + \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \boldsymbol{y}_N^* - \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)\boldsymbol{y}_N) \\ =&(\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N^* - \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)\boldsymbol{y}_N), \end{aligned}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

where we use (2.15a) and (2.15b) to derive that $(\nabla \boldsymbol{y}_N, \nabla \boldsymbol{y}_N) = (\boldsymbol{u}_N, \boldsymbol{y}_N)$, which leads to $(\boldsymbol{u}_N + \boldsymbol{y}_N, \boldsymbol{y}_N) = d^2$. Thus, with the help of inequality (2.22), we can deduce that

$$\begin{aligned} a(\boldsymbol{y}_N^*, \boldsymbol{y}_N^*) &= \left(\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N^* - \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)\boldsymbol{y}_N\right) + a\left(\frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)\boldsymbol{y}_N, \boldsymbol{y}_N^*\right) \\ &= (\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N^*) - \frac{1}{d^2}(\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N)(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N) + \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{u}_N + \boldsymbol{y}_N)a(\boldsymbol{y}_N, \boldsymbol{y}_N^*) \\ &\le C\|\boldsymbol{y}_N^*\|_{1,\Omega}. \end{aligned} \tag{2.24}$$

By (2.23) and (2.24), we obtain

$$\|\boldsymbol{y}_N^*\|_{1,\Omega} \le C. \tag{2.25}$$

Similarly, the discussions are divided into two cases to estimate $|\lambda_N|$.

If $\|\boldsymbol{y}_N\|_{1,\Omega} < d$, then $|\lambda_N| \le C$ by (2.16).

If $\|\boldsymbol{y}_N\|_{1,\Omega} = d$, let $\boldsymbol{q}_N = \boldsymbol{y}_N$ in (2.15), we derive

$$\lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N, \boldsymbol{y}_N) = a(\boldsymbol{y}_N, \boldsymbol{y}_N^*) - (\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N).$$

It follows from above two cases that

$$|\lambda_N| \le C. \tag{2.26}$$

Then the lemma follows from (2.22), (2.25), and (2.26). □

Lemma 2.3. ([8,30]). *Let $\Omega \subset \mathbb{R}^2$ be a rectangular domain, and $(\boldsymbol{\epsilon}, \rho) \in \boldsymbol{Y} \times Q$ be the solution of the following equations*

$$\begin{aligned} &a(\boldsymbol{\epsilon}, \boldsymbol{w}) \pm b(\boldsymbol{w}, \rho) = (\Phi, \boldsymbol{w}), && \forall\ \boldsymbol{w} \in \boldsymbol{Y}, \\ &b(\boldsymbol{\epsilon}, \phi) = 0, && \forall\ \phi \in Q. \end{aligned}$$

Then it holds

$$\|\boldsymbol{\epsilon}\|_{2,\Omega} + \|\rho\|_{1,\Omega} \le C\|\Phi\|_{0,\Omega}.$$

Moreover, if

$$(\boldsymbol{\epsilon}, \rho) \in (H_0^1(\Omega) \cap H^m(\Omega))^2 \times (L_0^2(\Omega) \cap H^{m-1}(\Omega)),$$

we can derive that

$$\begin{aligned} &\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_N\|_{1,\Omega} \le CN^{1-m}(\|\boldsymbol{\epsilon}\|_{m,\Omega} + \|\rho\|_{m-1,\Omega}), \\ &\|\rho - \rho_N\|_{0,\Omega} \le CN^{\frac{3}{2}-m}(\|\boldsymbol{\epsilon}\|_{m,\Omega} + \|\rho\|_{m-1,\Omega}), \\ &\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_N\|_{0,\Omega} \le CN^{-m}(\|\boldsymbol{\epsilon}\|_{m,\Omega} + \|\rho\|_{m-1,\Omega}), \end{aligned}$$

where $\boldsymbol{\epsilon}_N$ and ρ_N are the $Q_N - Q_{N-2}$ approximations of $\boldsymbol{\epsilon}$ and ρ respectively.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

We now introduce the following auxiliary system: find $(\boldsymbol{y}_N(\boldsymbol{u}), r_N(\boldsymbol{u}), \boldsymbol{y}_N^*(\boldsymbol{u}), r_N^*(\boldsymbol{u}))$ such that

$$a(\boldsymbol{y}_N(\boldsymbol{u}), \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N(\boldsymbol{u})) = (\boldsymbol{u}, \boldsymbol{w}_N), \qquad \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N, \tag{2.27a}$$

$$b(\boldsymbol{y}_N(\boldsymbol{u}), \phi_N) = 0, \qquad \forall\ \phi_N \in M_{N-2}, \tag{2.27b}$$

$$a(\boldsymbol{q}_N, \boldsymbol{y}_N^*(\boldsymbol{u})) + b(\boldsymbol{q}_N, r_N^*(\boldsymbol{u})) = (\boldsymbol{y} - \boldsymbol{y}_0 + \lambda(\boldsymbol{u} + \boldsymbol{y}), \boldsymbol{q}_N), \qquad \forall\ \boldsymbol{q}_N \in \boldsymbol{Y}_N, \tag{2.27c}$$

$$b(\boldsymbol{y}_N^*(\boldsymbol{u}), \psi_N) = 0, \qquad \forall\ \psi_N \in M_{N-2}. \tag{2.27d}$$

Note that $(\boldsymbol{y}_N(\boldsymbol{u}), r_N(\boldsymbol{u}), \boldsymbol{y}_N^*(\boldsymbol{u}), r_N^*(\boldsymbol{u}))$ is the spectral approximation of $(\boldsymbol{y}, r, \boldsymbol{y}^*, r^*)$. By lemma 2.3, we have

$$\|\boldsymbol{y} - \boldsymbol{y}_N(\boldsymbol{u})\|_{1,\Omega} \le CN^{1-m}\Big(\|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega}\Big), \tag{2.28a}$$

$$\|r - r_N(\boldsymbol{u})\|_{0,\Omega} \le CN^{\frac{3}{2}-m}\Big(\|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega}\Big), \tag{2.28b}$$

$$\|\boldsymbol{y} - \boldsymbol{y}_N(\boldsymbol{u})\|_{0,\Omega} \le CN^{-m}\Big(\|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega}\Big), \tag{2.28c}$$

and

$$\|\boldsymbol{y}^* - \boldsymbol{y}_N^*(\boldsymbol{u})\|_{1,\Omega} \le CN^{1-m}\Big(\|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega}\Big), \tag{2.29a}$$

$$\|r^* - r_N^*(\boldsymbol{u})\|_{0,\Omega} \le CN^{\frac{3}{2}-m}\Big(\|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega}\Big), \tag{2.29b}$$

$$\|\boldsymbol{y}^* - \boldsymbol{y}_N^*(\boldsymbol{u})\|_{0,\Omega} \le CN^{-m}\Big(\|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega}\Big), \tag{2.29c}$$

provided

$$(\boldsymbol{y}, \boldsymbol{y}^*, r, r^*) \in [(H_0^1(\Omega) \cap H^m(\Omega))^2]^2 \times [L_0^2(\Omega) \cap H^{m-1}(\Omega)]^2.$$

With the help of the auxiliary system, some important results can be derived. It follows from (2.15) and (2.27) that

$$a(\boldsymbol{y}_N(\boldsymbol{u}) - \boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N(\boldsymbol{u}) - r_N) = (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{w}_N), \qquad \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N, \tag{2.30a}$$

$$b(\boldsymbol{y}_N(u) - \boldsymbol{y}_N, \phi_N) = 0, \qquad \forall\ \phi_N \in \boldsymbol{M}_{N-2}. \tag{2.30b}$$

Then it holds

$$b(\boldsymbol{y}_N(u) - \boldsymbol{y}_N, r_N(\boldsymbol{u}) - r_N) = 0. \tag{2.31}$$

Let $\boldsymbol{w}_N = \boldsymbol{y}_N(\boldsymbol{u}) - \boldsymbol{y}_N$ in (2.30), we have

$$\|\boldsymbol{y}_N(\boldsymbol{u}) - \boldsymbol{y}_N\|_{1,\Omega} \le C\|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}. \tag{2.32}$$

According to (2.13), (2.30), and (2.32), we obtain

$$\begin{aligned}\beta_N\|r_N(\boldsymbol{u}) - r_N\|_{0,\Omega} &\le \sup_{\boldsymbol{w}_N \in \boldsymbol{Y}_N} \frac{b(\boldsymbol{w}_N, r_N(\boldsymbol{u}) - r_N)}{\|\boldsymbol{w}_N\|_{1,\Omega}} \\ &= \sup_{\boldsymbol{w}_N \in \boldsymbol{Y}_N} \frac{1}{\|\boldsymbol{w}_N\|_{1,\Omega}}[a(\boldsymbol{y}_N(\boldsymbol{u}) - \boldsymbol{y}_N, \boldsymbol{w}_N) - (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{w}_N)] \\ &\le C\|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}.\end{aligned} \tag{2.33}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

By (2.15) and (2.27), we have that for all $\boldsymbol{q}_N \in \boldsymbol{Y}_N$, $\psi_N \in \boldsymbol{M}_{N-2}$

$$\begin{aligned} &a(\boldsymbol{q}_N, \boldsymbol{y}_N^*(u) - \boldsymbol{y}_N^*) + b(\boldsymbol{q}_N, r_N^*(u) - r_N^*) \\ =&(\boldsymbol{y} - \boldsymbol{y}_N + \lambda(\boldsymbol{u} + \boldsymbol{y}) - \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \boldsymbol{q}_N), \qquad (2.34a) \\ &b(\boldsymbol{y}_N^*(u) - \boldsymbol{y}_N^*, \psi_N) = 0. \qquad (2.34b) \end{aligned}$$

Let $\boldsymbol{q}_N = \boldsymbol{y}_N^*(\boldsymbol{u}) - \boldsymbol{y}_N^*$, we derive that

$$\begin{aligned} \|\boldsymbol{y}_N^*(\boldsymbol{u}) - \boldsymbol{y}_N^*\|_{1,\Omega} &\le C\|\boldsymbol{y} - \boldsymbol{y}_N + \lambda(\boldsymbol{u} + \boldsymbol{y}) - \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N)\|_{0,\Omega} \\ &\le C\left(\|\boldsymbol{y} - \boldsymbol{y}_N\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + |\lambda - \lambda_N|\right). \qquad (2.35) \end{aligned}$$

By (2.13) and (2.34), we have

$$\begin{aligned} \beta_N\|r_N^*(\boldsymbol{u}) - r_N^*\|_{0,\Omega} &\le \sup_{\boldsymbol{q}_N \in \boldsymbol{Y}_N} \frac{b\left(\boldsymbol{q}_N, r_N^*(\boldsymbol{u}) - r_N^*\right)}{\|\boldsymbol{q}_N\|_{1,\Omega}} \\ = \sup_{\boldsymbol{q}_N \in \boldsymbol{Y}_N} &\frac{-a(\boldsymbol{q}_N, \boldsymbol{y}_N^*(\boldsymbol{u}) - \boldsymbol{y}_N^*) + (\boldsymbol{y} - \boldsymbol{y}_N, \boldsymbol{q}_N) + (\lambda(\boldsymbol{u} + \boldsymbol{y}) - \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \boldsymbol{q}_N)}{\|\boldsymbol{q}_N\|_{1,\Omega}} \\ \le &C\left(\|\boldsymbol{y}_N^*(\boldsymbol{u}) - \boldsymbol{y}_N^*\|_{1,\Omega} + \|\boldsymbol{y} - \boldsymbol{y}_N\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + |\lambda - \lambda_N|\right) \\ \le &C\left(\|\boldsymbol{y} - \boldsymbol{y}_N\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + |\lambda - \lambda_N|\right). \qquad (2.36) \end{aligned}$$

A priori error estimates can be presented as soon as the following two lemmas are proved.

Lemma 2.4. *Let* $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ *and* $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ *satisfy the optimality conditions* (2.10) *and* (2.15) *respectively. Assume that*

$$\begin{aligned} &(\boldsymbol{y}, \boldsymbol{y}^*, r, r^*, \boldsymbol{u}, \boldsymbol{u}_0) \\ \in&[(H_0^1(\Omega) \cap H^m(\Omega))^2]^2 \times [L_0^2(\Omega) \cap H^{m-1}(\Omega)]^2 \times [H^{m-1}(\Omega)^2]^2, \quad m > 1. \end{aligned}$$

Then we have for N *large enough*

$$\begin{aligned} |\lambda - \lambda_N| \le &C(N^{1-m}\|\boldsymbol{y}\|_{m,\Omega} + \|\boldsymbol{y} - \boldsymbol{y}_N(\boldsymbol{u})\|_{0,\Omega} + \|\boldsymbol{y}^* - \boldsymbol{y}_N^*(\boldsymbol{u})\|_{0,\Omega} \\ &+ \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + \|\boldsymbol{u}_0 - P_N\boldsymbol{u}_0\|_{0,\Omega}). \end{aligned}$$

Proof. According to (2.10) and (2.15), we conclude that for all $\boldsymbol{q}_N \in \boldsymbol{Y}_N$, $\psi_N \in \boldsymbol{M}_{N-2}$

$$\begin{aligned} &a(\boldsymbol{q}_N, \boldsymbol{y}^* - \boldsymbol{y}_N^*) + b(\boldsymbol{q}_N, r^* - r_N^*) \\ =&(\boldsymbol{y} - \boldsymbol{y}_N + \lambda(\boldsymbol{u} + \boldsymbol{y}) - \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \boldsymbol{q}_N), \\ &b(\boldsymbol{y}^* - \boldsymbol{y}_N^*, \psi_N) = 0. \end{aligned}$$

Let $\boldsymbol{q}_N = \Phi_N$ satisfy lemma 2.1, we have

$$\begin{aligned} &(\lambda - \lambda_N)(\boldsymbol{y}, \boldsymbol{y}) \\ =&a(\Phi_N, \boldsymbol{y}^* - \boldsymbol{y}_N^*) + b(\Phi_N, \boldsymbol{r}^* - \boldsymbol{r}_N^*) - (\boldsymbol{y} - \boldsymbol{y}_N, \Phi_N) \\ &- (\lambda(\boldsymbol{u} + \boldsymbol{y}) - \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \Phi_N) + (\lambda - \lambda_N)(\boldsymbol{y}, \boldsymbol{y}) \\ =&a(\Phi_N, \boldsymbol{y}^* - \boldsymbol{y}_N^*) - (\boldsymbol{y} - \boldsymbol{y}_N, \Phi_N) - \lambda(\boldsymbol{u} - \boldsymbol{u}_N, \Phi_N) - (\lambda - \lambda_N)(\boldsymbol{u}_N, \Phi_N - \boldsymbol{y}) \\ &- (\lambda - \lambda_N)(\boldsymbol{u}_N - \boldsymbol{u}, \boldsymbol{y}) - (\lambda - \lambda_N)(\boldsymbol{u}, \boldsymbol{y}) - \lambda(\boldsymbol{y} - \boldsymbol{y}_N, \Phi_N) \\ &- (\lambda - \lambda_N)(\boldsymbol{y}_N - \boldsymbol{y}, \Phi_N) - (\lambda - \lambda_N)(\boldsymbol{y}, \Phi_N - \boldsymbol{y}). \end{aligned}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

It follows that

$$\begin{aligned}&(\lambda-\lambda_N)(\boldsymbol{u}+\boldsymbol{y},\boldsymbol{y})\\ =&a(\Phi_N,\boldsymbol{y}^*-\boldsymbol{y}_N^*)-(\boldsymbol{y}-\boldsymbol{y}_N,\Phi_N)-\lambda(\boldsymbol{u}-\boldsymbol{u}_N,\Phi_N)-(\lambda-\lambda_N)(\boldsymbol{u}_N,\Phi_N-\boldsymbol{y})\\ &-(\lambda-\lambda_N)(\boldsymbol{u}_N-\boldsymbol{u},\boldsymbol{y})-\lambda(\boldsymbol{y}-\boldsymbol{y}_N,\Phi_N)\\ &-(\lambda-\lambda_N)(\boldsymbol{y}_N-\boldsymbol{y},\Phi_N)-(\lambda-\lambda_N)(\boldsymbol{y},\Phi_N-\boldsymbol{y}).\end{aligned}\tag{2.37}$$

Furthermore, it holds that by (2.10a)

$$\begin{aligned}a(\Phi_N,\boldsymbol{y}^*-\boldsymbol{y}_N^*)&=a(\Phi_N-\boldsymbol{y},\boldsymbol{y}^*-\boldsymbol{y}_N^*)+a(\boldsymbol{y},\boldsymbol{y}^*-\boldsymbol{y}_N^*)\\ &=a(\Phi_N-\boldsymbol{y},\boldsymbol{y}^*-\boldsymbol{y}_N^*)+b(\boldsymbol{y}^*-\boldsymbol{y}_N^*,r)+(\boldsymbol{u},\boldsymbol{y}^*-\boldsymbol{y}_N^*)\\ &=a(\Phi_N-\boldsymbol{y},\boldsymbol{y}^*-\boldsymbol{y}_N^*)+(\boldsymbol{u}-\nabla r,\boldsymbol{y}^*-\boldsymbol{y}_N^*).\end{aligned}\tag{2.38}$$

If $\|\boldsymbol{y}\|_{1,\Omega}=d$, we can deduce that $(\boldsymbol{u}+\boldsymbol{y},\boldsymbol{y})=d^2$ by (2.10a) and (2.10b). Furthermore, it can be derived from (2.10e), (2.15e), (2.19)-(2.20), (2.32) and (2.37)-(2.38) that

$$\begin{aligned}&|\lambda-\lambda_N|\\ \le&C\left(\|\boldsymbol{y}-\Phi_N\|_{1,\Omega}+\|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|_{0,\Omega}+\|\boldsymbol{y}-\boldsymbol{y}_N\|_{0,\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\right)\\ \le&C\left(N^{1-m}\|\boldsymbol{y}\|_{m,\Omega}+\|\boldsymbol{y}-\boldsymbol{y}_N(\boldsymbol{u})\|_{0,\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}+\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}\right).\end{aligned}\tag{2.39}$$

If $\|\boldsymbol{y}\|_{1,\Omega}<d$ and $\|\boldsymbol{y}_N\|_{1,\Omega}<d$, then

$$\lambda=\lambda_N=0.\tag{2.40}$$

If $\|\boldsymbol{y}\|_{1,\Omega}<d$ and $\|\boldsymbol{y}_N\|_{1,\Omega}=d$, let $\boldsymbol{q}_N=\boldsymbol{y}_N$ in (2.34) and $\boldsymbol{w}_N=\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}_N^*$ in (2.15a), we have

$$\begin{aligned}\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N,\boldsymbol{y}_N)=&-a(\boldsymbol{y}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}_N^*)+(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N)\\ =&-(\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}_N^*)+(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N),\end{aligned}$$

associating with (2.10e), (2.15e), (2.20), and (2.32), we deduce that

$$\begin{aligned}&|\lambda-\lambda_N|=|\lambda_N|\le C\|\boldsymbol{y}-\boldsymbol{y}_N\|_{0,\Omega}+C\|\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}_N^*\|_{0,\Omega}\\ \le&C\left(\|\boldsymbol{y}-\boldsymbol{y}_N(\boldsymbol{u})\|_{0,\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}+\|\boldsymbol{y}^*-\boldsymbol{y}_N^*(\boldsymbol{u})\|_{0,\Omega}+\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}\right).\end{aligned}\tag{2.41}$$

Then the lemma can be completed from (2.39)-(2.41). □

Lemma 2.5. *Let* $(\boldsymbol{y},r,\boldsymbol{u},\boldsymbol{y}^*,r^*,\lambda)$ *and* $(\boldsymbol{y}_N,r_N,\boldsymbol{u}_N,\boldsymbol{y}_N^*,r_N^*,\lambda_N)$ *satisfy the optimality conditions* (2.10) *and* (2.15) *respectively. Assume that*

$$\begin{aligned}&(\boldsymbol{y},\boldsymbol{y}^*,r,r^*,\boldsymbol{u},\boldsymbol{u}_0)\\ \in&[(H_0^1(\Omega)\cap H^m(\Omega))^2]^2\times[L_0^2(\Omega)\cap H^{m-1}(\Omega)]^2\times[H^{m-1}(\Omega)^2]^2,\quad m>1.\end{aligned}$$

Then we have for N *large enough*

$$\begin{aligned}\|P_N\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\le&C(\|\boldsymbol{y}^*-\boldsymbol{y}_N^*(\boldsymbol{u})\|_{0,\Omega}+\|\boldsymbol{y}-\boldsymbol{y}_N(\boldsymbol{u})\|_{1,\Omega}+\|\boldsymbol{u}-P_N\boldsymbol{u}\|_{0,\Omega}\\ &+\|r-P_{N-2}r\|_{0,\Omega})+CN^{1-m}(\|\boldsymbol{y}\|_{m,\Omega}+\|\boldsymbol{u}_0\|_{m-1,\Omega}).\end{aligned}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

Proof. By (2.11) and (2.16), we have that

$$-\lambda(\boldsymbol{y},\boldsymbol{y}-\boldsymbol{y}_N)_{H^1}\le 0,\qquad \lambda_N(\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}_N)_{H^1}\le 0. \tag{2.42}$$

Furthermore, it follows from (2.10e), (2.15e), and (2.30) that

$$\begin{aligned}
&\alpha\|P_N\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2=\alpha(P_N\boldsymbol{u}-\boldsymbol{u}_N,P_N\boldsymbol{u}-\boldsymbol{u}_N)=\alpha(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{u}-\boldsymbol{u}_N)\\
=&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*)+\big(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0)\big)\\
=&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)+\big(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}_N^*(\boldsymbol{u})\big)+\big(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0)\big)\\
=&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)+\big(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}_N^*(\boldsymbol{u})\big)+\big(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0)\big)\\
=&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)+a\big(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N,\boldsymbol{y}_N^*-\boldsymbol{y}_N^*(\boldsymbol{u})\big)\\
&+\big(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0)\big),
\end{aligned}$$

associating with (2.10a), (2.15a), (2.34), and (2.42), we have

$$\begin{aligned}
&\alpha\|P_N\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2\\
=&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)-(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)\\
&-(\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)+(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0))\\
=&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)-(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)\\
&+(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y},\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)-\lambda(\boldsymbol{y},\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)_{H^1}+\lambda b(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N,r)\\
&+\lambda_N(\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y})_{H^1}+\lambda_N(\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}_N)_{H^1}+(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0))\\
\le&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)-(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)\\
&+(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y},\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)-\lambda(\boldsymbol{y},\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y})_{H^1}+\lambda b(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N,r)\\
&+\lambda_N(\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y})_{H^1}+(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0)).
\end{aligned}$$

Noting that

$$(\nabla\cdot(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N),\mu)=0\quad\forall\ \mu\in Q_{N-2}.$$

Thus, with the help of (2.20), (2.32), and lemma 2.4, we can deduce that

$$\begin{aligned}
&\alpha\parallel P_N\boldsymbol{u}-\boldsymbol{u}_N\parallel_{0,\Omega}^2+\|\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N\|_{0,\Omega}^2\\
\le&(P_N\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}^*)+(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y},\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N)\\
&+((\lambda_N-\lambda)\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y})_{H^1}-\lambda(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y})_{H^1}\\
&+(P_N\boldsymbol{u}-\boldsymbol{u}_N,\alpha(\boldsymbol{u}_0-P_N\boldsymbol{u}_0))+\inf_{\mu\in Q_{N-2}}b(\lambda(\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N),r-\mu).\\
\le&\varepsilon\left(\|P_N\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2+\|\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N\|_{0,\Omega}^2\right)+C_1(\varepsilon)\left(\|\boldsymbol{y}^*-\boldsymbol{y}_N^*(\boldsymbol{u})\|_{0,\Omega}^2\right.\\
&+\|\boldsymbol{y}-\boldsymbol{y}_N(\boldsymbol{u})\|_{1,\Omega}^2+\|P_N\boldsymbol{u}-\boldsymbol{u}\|_{0,\Omega}^2+\|r-P_{N-2}r\|_{0,\Omega}^2\\
&\left.+\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}^2+N^{2-2m}\|\boldsymbol{y}\|_{m,\Omega}^2\right),
\end{aligned}\tag{2.43}$$

where

$$C_1(\varepsilon)=C^2/4\varepsilon,\quad\varepsilon>0.$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

is a sufficiently small constant. Then the lemma can be proved from (2.43). □

The main result of this section can now be stated.

Theorem 2.1. *Let* $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ *and* $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ *satisfy the optimality conditions* (2.10) and (2.15) respectively. Assume that

$$\begin{aligned}&(\boldsymbol{y}, \boldsymbol{y}^*, r, r^*, \boldsymbol{u}, \boldsymbol{u}_0)\\ \in&[(H_0^1(\Omega)\cap H^m(\Omega))^2]^2 \times [L_0^2(\Omega)\cap H^{m-1}(\Omega)]^2 \times [H^{m-1}(\Omega)^2]^2, \quad m>1.\end{aligned}$$

Then we have for N large enough

$$\begin{aligned}&\|\boldsymbol{y}-\boldsymbol{y}_N\|_{1,\Omega} + \|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|_{1,\Omega} + N^{-\frac{1}{2}}\|r-r_N\|_{0,\Omega}\\ &\quad + N^{-\frac{1}{2}}\|r^*-r_N^*\|_{0,\Omega} + |\lambda-\lambda_N| + \|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega} \le CN^{1-m}\xi,\end{aligned}$$

with

$$\xi = \|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega} + \|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega} + \|\boldsymbol{u}\|_{m-1,\Omega} + \|\boldsymbol{u}_0\|_{m-1,\Omega}.$$

Proof. It follows from (2.28)-(2.29), (2.32)-(2.33), (2.35)-(2.36), lemma 2.4, and lemma 2.5 that

$$\begin{aligned}&\|\boldsymbol{y}-\boldsymbol{y}_N\|_{1,\Omega} + \|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|_{1,\Omega} + N^{-\frac{1}{2}}\|r-r_N\|_{0,\Omega}\\ &\quad + N^{-\frac{1}{2}}\|r^*-r_N^*\|_{0,\Omega} + |\lambda-\lambda_N| + \|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\\ \le&\|\boldsymbol{y}-\boldsymbol{y}_N(\boldsymbol{u})\|_{1,\Omega} + \|\boldsymbol{y}_N(\boldsymbol{u})-\boldsymbol{y}_N\|_{1,\Omega} + \|\boldsymbol{y}^*-\boldsymbol{y}_N^*(\boldsymbol{u})\|_{1,\Omega} + \|\boldsymbol{y}_N^*(\boldsymbol{u})-\boldsymbol{y}_N^*\|_{1,\Omega}\\ &\quad + N^{-\frac{1}{2}}\|r-r_N(\boldsymbol{u})\|_{0,\Omega} + N^{-\frac{1}{2}}\|r_N(\boldsymbol{u})-r_N\|_{0,\Omega} + N^{-\frac{1}{2}}\|r^*-r_N^*(\boldsymbol{u})\|_{0,\Omega}\\ &\quad + N^{-\frac{1}{2}}\|r_N^*(\boldsymbol{u})-r_N^*\|_{0,\Omega} + |\lambda-\lambda_N| + \|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\\ \le&CN^{1-m}(\|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega}) + C\|\boldsymbol{u}-P_N\boldsymbol{u}\|_{0,\Omega} + C\|\boldsymbol{u}_N-P_N\boldsymbol{u}\|_{0,\Omega}\\ &\quad + CN^{1-m}\Big(\|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega}\Big) + C|\lambda-\lambda_N|\\ \le&CN^{1-m}\Big(\|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega} + \|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega} + \|\boldsymbol{u}\|_{m-1,\Omega}\Big)\\ &\quad + C\|\boldsymbol{u}_N-P_N\boldsymbol{u}\|_{0,\Omega} + C|\lambda-\lambda_N|\\ \le&CN^{1-m}\Big(\|\boldsymbol{y}\|_{m,\Omega} + \|r\|_{m-1,\Omega} + \|\boldsymbol{y}^*\|_{m,\Omega} + \|r^*\|_{m-1,\Omega} + \|\boldsymbol{u}\|_{m-1,\Omega} + \|\boldsymbol{u}_0\|_{m-1,\Omega}\Big),\end{aligned}$$

which concludes the proof. □

3. A posteriori error estimates

In this section, we derive a posteriori error estimates for the spectral approximation of control problem (OCP), which is helpful for constructing equivalent a posteriori error estimates, and developing *hp* spectral element approximation and adaptive algorithm for control problems.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

It follows from (2.17) that

$$\|\boldsymbol{w} - P_{1,N}^0 \boldsymbol{w}\|_{0,\Omega} \leq CN^{-m}\|\boldsymbol{w}\|_{m,\Omega}, \tag{3.1}$$

provided

$$\boldsymbol{w} \in (H^m(\Omega) \cap H_0^1(\Omega))^2, \quad m \geq 1.$$

To derive a posteriori error estimates, we introduce the following auxiliary system: find

$$(\boldsymbol{y}(\boldsymbol{u}_N), \boldsymbol{r}(\boldsymbol{u}_N), \boldsymbol{y}^*(\boldsymbol{u}_N), \boldsymbol{r}^*(\boldsymbol{u}_N))$$

such that

$$\begin{aligned}
&a(\boldsymbol{y}(\boldsymbol{u}_N), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u}_N)) = (\boldsymbol{u}_N, \boldsymbol{w}), && \forall\ \boldsymbol{w} \in \boldsymbol{Y}, && \text{(3.2a)}\\
&b(\boldsymbol{y}(\boldsymbol{u}_N), \phi) = 0, && \forall\ \phi \in Q, && \text{(3.2b)}\\
&a(\boldsymbol{q}, \boldsymbol{y}^*(\boldsymbol{u}_N)) + b(\boldsymbol{q}, r^*(\boldsymbol{u}_N)) = (\boldsymbol{y}_N - \boldsymbol{y}_0 + \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N), \boldsymbol{q}), && \forall\ \boldsymbol{q} \in \boldsymbol{Y}, && \text{(3.2c)}\\
&b(\boldsymbol{y}^*(\boldsymbol{u}_N), \psi) = 0, && \forall\ \psi \in Q. && \text{(3.2d)}
\end{aligned}$$

Then, it can be derived by (2.4b), (2.10a)–(2.10b), and (3.2a)–(3.2b) that

$$\|\boldsymbol{y} - \boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega} + \|r - r(\boldsymbol{u}_N)\|_{0,\Omega} \leq C\|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}. \tag{3.3}$$

Similarly, it holds that

$$\begin{aligned}
&\|\boldsymbol{y}^* - \boldsymbol{y}^*(\boldsymbol{u}_N)\|_{1,\Omega} + \|r^* - r^*(\boldsymbol{u}_N)\|_{0,\Omega}\\
\leq& C\|\boldsymbol{y} - \boldsymbol{y}_N\|_{0,\Omega} + C\|\lambda(\boldsymbol{u} + \boldsymbol{y}) - \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N)\|_{0,\Omega}\\
\leq& C(\|\boldsymbol{y}(\boldsymbol{u}_N) - \boldsymbol{y}_N\|_{0,\Omega} + |\lambda - \lambda_N| + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}).
\end{aligned} \tag{3.4}$$

The following two lemmas are helpful for deriving the a posteriori error estimates.

Lemma 3.1. *Let* $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ *and* $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ *satisfy the optimality conditions* (2.10) *and* (2.15) *respectively. Then we have that*

$$\begin{aligned}
&\|\boldsymbol{y}(\boldsymbol{u}_N) - \boldsymbol{y}_N\|_{1,\Omega} + \|r(\boldsymbol{u}_N) - r_N\|_{0,\Omega}\\
\leq& CN^{-1}\|\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N\|_{0,\Omega} + C\|\nabla \cdot \boldsymbol{y}_N\|_{0,\Omega}, && \text{(3.5a)}\\
&\|\boldsymbol{y}^*(\boldsymbol{u}_N) - \boldsymbol{y}_N^*\|_{1,\Omega} + \|r^*(\boldsymbol{u}_N) - r_N^*\|_{0,\Omega}\\
\leq& CN^{-1}\|\Delta \boldsymbol{y}_N^* + \nabla r_N^* + \boldsymbol{y}_N - \boldsymbol{y}_0 + \lambda_N(\boldsymbol{u}_N + \boldsymbol{y}_N)\|_{0,\Omega} + C\|\nabla \cdot \boldsymbol{y}_N^*\|_{0,\Omega}. && \text{(3.5b)}
\end{aligned}$$

Proof. It follows from (2.15a) and (3.2a) that

$$a(\boldsymbol{y}(\boldsymbol{u}_N) - \boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r(\boldsymbol{u}_N) - r_N) = 0, \quad \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N,$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

associating with (3.2), we have

$$\begin{aligned}
&\theta\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}^2 \le a(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)\\
=&a(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)-b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)\\
&+b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)-a(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N))\\
&+b(P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),r(\boldsymbol{u}_N)-r_N)\\
=&a\Big(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)\Big)\\
&-b\Big(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),r(\boldsymbol{u}_N)-r_N\Big)\\
&+b\Big(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N\Big)\\
=&\Big(\nabla(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),\nabla(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N))\Big)\\
&-\Big(\nabla\cdot(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)),r(\boldsymbol{u}_N)-r_N\Big)-b(\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)\\
=&\Big(\triangle\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)\Big)-b(\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N).
\end{aligned}$$

Thus, it holds

$$\begin{aligned}
&\theta\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}^2\\
\le&CN^{-1}\|\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N\|_{0,\Omega}\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}+\|\nabla\cdot\boldsymbol{y}_N\|_{0,\Omega}\|r(\boldsymbol{u}_N)-r_N\|_{0,\Omega}\\
\le&C_1(\varepsilon)\Big(N^{-2}\|\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N\|_{0,\Omega}^2+\|\nabla\cdot\boldsymbol{y}_N\|_{0,\Omega}^2\Big)\\
&+\varepsilon\Big(\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}^2+\|r(\boldsymbol{u}_N)-r_N\|_{0,\Omega}^2\Big).
\end{aligned}\tag{3.6}$$

Noting that

$$\begin{aligned}
&b\left(\boldsymbol{w},r(\boldsymbol{u}_N)-r_N\right)=b\left(P_{1,N}^0\boldsymbol{w},r(\boldsymbol{u}_N)-r_N\right)+b\left(\boldsymbol{w}-P_{1,N}^0\boldsymbol{w},r(\boldsymbol{u}_N)-r_N\right)\\
=&a\left(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,P_{1,N}^0\boldsymbol{w}\right)+b\left(\boldsymbol{w}-P_{1,N}^0\boldsymbol{w},r(\boldsymbol{u}_N)-r_N\right)\\
=&-a\left(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{w}-P_{1,N}^0\boldsymbol{w}\right)+a\left(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{w}\right)+b\left(\boldsymbol{w}-P_{1,N}^0\boldsymbol{w},r(\boldsymbol{u}_N)-r_N\right)\\
=&(-\Delta\boldsymbol{y}_N+\nabla r_N-\boldsymbol{u}_N,\boldsymbol{w}-P_{1,N}^0\boldsymbol{w})+a\left(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{w}\right).
\end{aligned}$$

Therefore, we can derive by (2.4b)

$$\begin{aligned}
\beta\|r(\boldsymbol{u}_N)-r_N\|_{0,\Omega}&\le\sup_{\boldsymbol{w}\in Y}\frac{|\,b(\boldsymbol{w},r(\boldsymbol{u}_N)-r_N)\,|}{\|\boldsymbol{w}\|_{1,\Omega}}\\
&\le CN^{-1}\|\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N\|_{0,\Omega}+C\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}.
\end{aligned}\tag{3.7}$$

Similarly, apply (2.15c) and (3.2c) to obtain

$$a(\boldsymbol{q}_N,\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*)+b(\boldsymbol{q}_N,r^*(\boldsymbol{u}_N)-r_N^*)=0,\quad\forall\,\boldsymbol{q}_N\in\boldsymbol{Y}_N.$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

It follows that

$$\begin{aligned}
&\theta\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}^2 \le a\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*,\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\big)\\
=&a\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*,\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\big)+b\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*,r^*(\boldsymbol{u}_N)-r_N^*\big)\\
&-b\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*,r^*(\boldsymbol{u}_N)-r_N^*\big)-a\big(P_{1,N}^0(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*),\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\big)\\
&-b\big(P_{1,N}^0(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*),r^*(\boldsymbol{u}_N)-r_N^*\big)\\
=&a\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*-P_{1,N}^0(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*),\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\big)\\
&+b\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*-p_{1,N}^0(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*),r^*(\boldsymbol{u}_N)-r_N^*\big)\\
&-b\big(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*,r^*(\boldsymbol{u}_N)-r_N^*\big)\\
=&\big(\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*-p_{1,N}^0(\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*)\big)\\
&+b\big(\boldsymbol{y}_N^*,r^*(\boldsymbol{u}_N)-r_N^*\big)\\
\le&CN^{-1}\|\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N)\|_{0,\Omega}\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}\\
&+C\|\nabla\cdot\boldsymbol{y}_N^*\|_{0,\Omega}\|r^*(\boldsymbol{u}_N)-r_N^*\|_{0,\Omega}\\
\le&\varepsilon(\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}^2+\|r^*(\boldsymbol{u}_N)-r_N^*\|_{0,\Omega}^2)+C_1(\varepsilon)\|\nabla\cdot\boldsymbol{y}_N^*\|_{0,\Omega}^2\\
&+C_1(\varepsilon)N^{-2}\|\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N)\|_{0,\Omega}^2. \qquad (3.8)
\end{aligned}$$

Noting that

$$\begin{aligned}
&b(\boldsymbol{q},r^*(\boldsymbol{u}_N)-r_N^*)=b\Big(P_{1,N}^0\boldsymbol{q},r^*(\boldsymbol{u}_N)-r_N^*\Big)+b\Big(\boldsymbol{q}-P_{1,N}^0\boldsymbol{q},r^*(\boldsymbol{u}_N)-r_N^*\Big)\\
=&-a\Big(P_{1,N}^0\boldsymbol{q},\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\Big)+b\Big(\boldsymbol{q}-p_{1,N}^0\boldsymbol{q},r^*(\boldsymbol{u}_N)-r_N^*\Big)\\
=&a\Big(\boldsymbol{q}-P_{1,N}^0\boldsymbol{q},\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\Big)-a\Big(\boldsymbol{q},\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\Big)+b\Big(\boldsymbol{q}-P_{1,N}^0\boldsymbol{q},r^*(\boldsymbol{u}_N)-r_N^*\Big)\\
=&\Big(\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{q}-P_{1,N}^0\boldsymbol{q}\Big)-a(\boldsymbol{q},\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*)\\
\le&CN^{-1}\|\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N)\|_{0,\Omega}\|\boldsymbol{q}\|_{1,\Omega}\\
&+C\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}\|\boldsymbol{q}\|_{1,\Omega}.
\end{aligned}$$

Therefore, we can derive by (2.4b)

$$\begin{aligned}
&\beta\|r^*(\boldsymbol{u}_N)-r_N^*\|_{0,\Omega}\le\sup_{\boldsymbol{q}\in\boldsymbol{Y}}\frac{|\,b(\boldsymbol{q},r^*(\boldsymbol{u}_N)-r_N^*)\,|}{\|\boldsymbol{q}\|_{1,\Omega}}\\
\le&CN^{-1}\|\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N)\|_{0,\Omega}+C\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}. \qquad (3.9)
\end{aligned}$$

Then the lemma can be proved from (3.6)-(3.9). □

Lemma 3.2. *Let* $(\boldsymbol{y},r,\boldsymbol{u},\boldsymbol{y}^*,r^*,\lambda)$ *and* $(\boldsymbol{y}_N,r_N,\boldsymbol{u}_N,\boldsymbol{y}_N^*,r_N^*,\lambda_N)$ *satisfy the optimality conditions* (2.10) *and* (2.15) *respectively. Then we have that*

$$\begin{aligned}
|\,\lambda-\lambda_N\,|\le&C\Big(\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}+\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}\\
&+\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\Big). \qquad (3.10)
\end{aligned}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

Proof. Applying (2.10) and (3.2) to deduce that

$$a(\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{w})-b(\boldsymbol{w},r-r(\boldsymbol{u}_N))=(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{w}), \qquad \forall\, \boldsymbol{w}\in Y, \tag{3.11a}$$

$$b(\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N),\phi)=0, \qquad \forall\, \phi\in Q \tag{3.11b}$$

$$\begin{aligned} &a(\boldsymbol{q},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+b(\boldsymbol{q},r^*-r^*(\boldsymbol{u}_N))\\ &\quad=(\boldsymbol{y}-\boldsymbol{y}_N+\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{q}), \qquad \forall\, \boldsymbol{q}\in Y, \end{aligned} \tag{3.11c}$$

$$b(\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N),\psi)=0, \qquad \forall\, \psi\in Q. \tag{3.11d}$$

Then we can prove (3.10) by three cases. If $\|\boldsymbol{y}\|_{1,\Omega}=d$, set $\boldsymbol{q}=\boldsymbol{y}$ in (3.11) to obtain

$$a(\boldsymbol{y},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))=(\boldsymbol{y}-\boldsymbol{y}_N+\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}).$$

Set $\boldsymbol{w}=\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)$ in (2.10) to derive

$$a(\boldsymbol{y},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))=(\boldsymbol{u},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)).$$

Then

$$(\boldsymbol{y}-\boldsymbol{y}_N+\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y})=(\boldsymbol{u},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)).$$

It follows that

$$\begin{aligned} &(\lambda-\lambda_N)(\boldsymbol{u}+\boldsymbol{y},\boldsymbol{y})\\ &=-(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y})+\lambda_N(\boldsymbol{u}_N-\boldsymbol{u}+\boldsymbol{y}_N-\boldsymbol{y},\boldsymbol{y})+(\boldsymbol{u},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)). \end{aligned} \tag{3.12}$$

If $\|\boldsymbol{y}\|_{1,\Omega}<d$ and $\|\boldsymbol{y}_N\|_{1,\Omega}=d$, set $\boldsymbol{q}=\boldsymbol{y}(\boldsymbol{u}_N)$ in (3.11) to derive that

$$a(\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))=(\boldsymbol{y}-\boldsymbol{y}_N-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}(\boldsymbol{u}_N)),$$

due to $\lambda=0$ by (2.11). Furthermore, set $\boldsymbol{w}=\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)$ in (3.2a) to deduce that

$$a(\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))=(\boldsymbol{u}_N,\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)).$$

Then

$$(\boldsymbol{y}-\boldsymbol{y}_N-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}(\boldsymbol{u}_N))=(\boldsymbol{u}_N,\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)).$$

It follows that

$$\begin{aligned} \lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N,\boldsymbol{y}_N)=&(\boldsymbol{y}-\boldsymbol{y}_N-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)\\ &+(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N)-(\boldsymbol{u}_N,\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)). \end{aligned} \tag{3.13}$$

If $\|\boldsymbol{y}\|_{1,\Omega}<d$ and $\|\boldsymbol{y}_N\|_{1,\Omega}<d$, then $\lambda=\lambda_N=0$. Therefore, we obtain

$$|\,\lambda-\lambda_N\,|=0. \tag{3.14}$$

By (2.10e), (2.15e), (2.20), (3.3), and (3.12)-(3.14), we deduce that

$$\begin{aligned} &|\,\lambda-\lambda_N\,|\\ \le& C\|\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{y}-\boldsymbol{y}_N\|_{0,\Omega}+C\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\\ \le& C\|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|_{0,\Omega}+C\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\\ \le& C\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}+C\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}, \end{aligned}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

which concludes the proof. □

We can now present the desired result.

Theorem 3.1. *Let* $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ *and* $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ *satisfy the optimality conditions* (2.10) *and* (2.15) *respectively. Then we have that*

$$\begin{aligned}&\|\boldsymbol{y}-\boldsymbol{y}_N\|_{1,\Omega}+\|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|_{1,\Omega}+\|r-r_N\|_{0,\Omega}\\&\quad+\|r^*-r_N^*\|_{0,\Omega}+|\lambda-\lambda_N|+\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\le C(\eta+\theta),\end{aligned}\tag{3.15}$$

where

$$\theta=\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega},$$

and the estimator

$$\begin{aligned}\eta=&N^{-1}\|\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N\|_{0,\Omega}+\|\nabla\cdot\boldsymbol{y}_N\|_{0,\Omega}\\&+N^{-1}\|\Delta\boldsymbol{y}_N^*+\nabla r_N^*+\boldsymbol{y}_N-\boldsymbol{y}_0+\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N)\|_{0,\Omega}+\|\nabla\cdot\boldsymbol{y}_N^*\|_{0,\Omega}.\end{aligned}$$

Proof. By (2.10e), (2.15e), and (3.11), we have

$$\begin{aligned}&\alpha\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2=\alpha(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{u}-\boldsymbol{u}_N)\\
=&(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))-(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)-\alpha(\boldsymbol{u}-\boldsymbol{u}_0)+\alpha(\boldsymbol{u}_N-\boldsymbol{u}_0))\\
=&(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}^*(\boldsymbol{u}_N)+\alpha(\boldsymbol{u}-\boldsymbol{u}_0))-(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*+\alpha(\boldsymbol{u}_N-\boldsymbol{u}_0))\\
=&(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}^*)-(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*+\alpha(\boldsymbol{u}_N-\boldsymbol{u}_0))\\
=&(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))-(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))\\
&-(\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))-(\boldsymbol{u}-\boldsymbol{u}_N,\alpha P_N\boldsymbol{u}_0-\alpha\boldsymbol{u}_0)\\
=&(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))-(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))\\
&-(\boldsymbol{u}-\boldsymbol{u}_N,\alpha P_N\boldsymbol{u}_0-\alpha\boldsymbol{u}_0)-(\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}-\boldsymbol{y}_N)\\
&-(\lambda(\boldsymbol{u}+\boldsymbol{y})-\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N),\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)).\end{aligned}\tag{3.16}$$

It follows from (2.10a) and (2.42) that

$$\begin{aligned}-\lambda(\boldsymbol{u}+\boldsymbol{y},\boldsymbol{y}-\boldsymbol{y}_N)=&-\lambda(\boldsymbol{y},\boldsymbol{y}-\boldsymbol{y}_N)_{H^1}+\lambda b(\boldsymbol{y}-\boldsymbol{y}_N,r)\\
\le&\lambda b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r-r_N).\end{aligned}\tag{3.17}$$

Furthermore, it can be derived from (2.42) and (3.2a) that

$$\begin{aligned}&\lambda_N(\boldsymbol{u}_N+\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}_N)\\
=&\lambda_N(\boldsymbol{u}_N,\boldsymbol{y}-\boldsymbol{y}_N)+\lambda_N(\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}_N)+\lambda_N\left(\nabla\boldsymbol{y}_N,\nabla(\boldsymbol{y}-\boldsymbol{y}_N)\right)-\lambda_N\left(\nabla\boldsymbol{y}_N,\nabla(\boldsymbol{y}-\boldsymbol{y}_N)\right)\\
=&\lambda_N\left(\nabla\boldsymbol{y}(\boldsymbol{u}_N),\nabla(\boldsymbol{y}-\boldsymbol{y}_N)\right)-\lambda_N b(\boldsymbol{y}-\boldsymbol{y}_N,r(\boldsymbol{u}_N))\\
&+\lambda_N(\boldsymbol{y}_N,\boldsymbol{y}-\boldsymbol{y}_N)_{H^1}-\lambda_N(\nabla\boldsymbol{y}_N,\nabla(\boldsymbol{y}-\boldsymbol{y}_N))\\
\le&\lambda_N(\nabla(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),\nabla(\boldsymbol{y}-\boldsymbol{y}_N))-\lambda_N b(\boldsymbol{y}-\boldsymbol{y}_N,r(\boldsymbol{u}_N))\\
=&\lambda_N(\nabla(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),\nabla(\boldsymbol{y}-\boldsymbol{y}_N))-\lambda_N b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N).\end{aligned}\tag{3.18}$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

It follows from (3.16)-(3.18) that

$$\begin{aligned}
&\alpha\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2+\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}^2\\
\le&(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+(\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))\\
&-(\boldsymbol{u}-\boldsymbol{u}_N,\alpha P_N\boldsymbol{u}_0-\alpha\boldsymbol{u}_0)+\lambda b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r-r_N)\\
&+\lambda_N(\nabla(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),\nabla(\boldsymbol{y}-\boldsymbol{y}_N))-\lambda_N b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)\\
&-(\lambda-\lambda_N)(\boldsymbol{u},\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N))-\lambda_N(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N))\\
&-((\lambda-\lambda_N)\boldsymbol{y},\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N))-\lambda_N(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N))\\
\le&\varepsilon(\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2+\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}^2+|\lambda-\lambda_N|^2+\|r-r(u_N)\|_{0,\Omega}^2)+C_1(\varepsilon)\\
&\cdot\left(\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}^2+\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}^2+\|r(u_N)-r_N\|_{0,\Omega}^2+\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}^2\right),
\end{aligned}$$

associating with (3.3) and (3.10), we have

$$\begin{aligned}
&\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2+\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}^2\\
\le&C\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}^2+C\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}^2\\
&+C\|r(u_N)-r_N\|_{0,\Omega}^2+C\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}^2.
\end{aligned}\tag{3.19}$$

It follows from (3.3)-(3.4), (3.10), and (3.19) that

$$\begin{aligned}
&\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}+\|r-r(\boldsymbol{u}_N)\|_{0,\Omega}+\|\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{1,\Omega}\\
&\quad+\|r^*-r^*(\boldsymbol{u}_N)\|_{0,\Omega}+|\lambda-\lambda_N|+\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\\
\le&C\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}+C\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}\\
&\quad+C\|r(u_N)-r_N\|_{0,\Omega}+C\|\boldsymbol{u}_0-P_N\boldsymbol{u}_0\|_{0,\Omega}.
\end{aligned}\tag{3.20}$$

Then the theorem can be completed from (3.5a), (3.5b), and (3.20). □

4. Numerical Example

In this section, an illustrative numerical example is performed to confirm the efficiency of the proposed method, and to show that the a posteriori error estimator can indicate the errors well. We consider the following control problem on $\Omega=(-1,1)^2$: find

$$(\boldsymbol{y},r,\boldsymbol{u})\in\boldsymbol{Y}\times Q\times\boldsymbol{U},$$

such that

$$\begin{aligned}
&\min_{\boldsymbol{y}(\boldsymbol{u})\in\boldsymbol{K}}\left\{\frac{1}{2}\|\boldsymbol{y}(\boldsymbol{u})-\boldsymbol{y}_0\|_{0,\Omega}^2+\frac{1}{2}\|\boldsymbol{u}-\boldsymbol{u}_0\|_{0,\Omega}^2\right\},\\
&-\Delta\boldsymbol{y}(\boldsymbol{u})+\nabla r=\boldsymbol{u}\ \text{ in }\Omega,\quad \nabla\cdot\boldsymbol{y}(\boldsymbol{u})=0\ \text{ in }\Omega,\quad \boldsymbol{y}(\boldsymbol{u})=0\ \text{ on }\partial\Omega,
\end{aligned}\tag{4.1}$$

with

$$\boldsymbol{K}=\{\boldsymbol{w}\in\boldsymbol{Y}\ \mid\ \|\boldsymbol{w}\|_{1,\Omega}\le 1\}.$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

The data of this example are as follows

$$\boldsymbol{y}_0 = \frac{-\pi^4 + 0.8\pi^2 + 1.8}{\sqrt{6+8\pi^2}}\boldsymbol{\omega} + \frac{1}{\sqrt{6+8\pi^2}}\boldsymbol{h}_1, \quad \boldsymbol{u}_0 = \frac{2\pi^2}{\sqrt{6+8\pi^2}}(\boldsymbol{\omega} + \boldsymbol{h}_2),$$

where

$$\boldsymbol{\omega} = ((1+\cos\pi x_1)\sin\pi x_2, -\sin\pi x_1(1+\cos\pi x_2)),$$
$$\boldsymbol{h}_1 = (1.6\pi^2\cos\pi x_1\sin\pi x_2, 2\pi^4\sin\pi x_1\cos\pi x_2), \quad \boldsymbol{h}_2 = (\cos\pi x_1\sin\pi x_2,\ 0).$$

The exact solutions are given by follows

$$\boldsymbol{y} = \frac{1}{\sqrt{6+8\pi^2}}\boldsymbol{\omega}, \quad \boldsymbol{y}^* = \frac{\pi^2}{\sqrt{6+8\pi^2}}\boldsymbol{\omega}, \quad r = \frac{\pi}{\sqrt{6+8\pi^2}}\sin\pi x_1\sin\pi x_2,$$
$$r^* = \frac{\pi^3}{\sqrt{6+8\pi^2}}\sin\pi x_1\sin\pi x_2, \quad \boldsymbol{u} = -\boldsymbol{y}^* + \boldsymbol{u}_0, \quad \lambda = 0.8.$$

To solve the discrete system, Arrow-Hurwicz algorithm is proposed (see, for example, [18] and [26]). Note that the Lagrange multipler in Step 5 is calculated by a variant of

$$\lambda_N^{k+1} = P_{[0,+\infty)}(\lambda_N^k + \rho F_N(u_N))$$

with $P_{[0,+\infty)}$ being the projection operator from $\mathbb{R} \to [0,+\infty)$, more details can be found from the saddle-point search algorithm in [18].

We denote

$$\boldsymbol{y}_N = (y_{N1}, y_{N2}), \qquad \boldsymbol{y}_N^* = (y_{N1}^*, y_{N2}^*), \qquad \boldsymbol{u}_N = (u_{N1}, u_{N2}).$$

Let $L_n(x)$ be the nth-degree Legendre polynomial (see [29] and [30], for example) and

$$\phi_k(x) = \frac{1}{\sqrt{4k+6}}(L_k(x) - L_{k+2}(x)), \quad \boldsymbol{Y}_N = (H_N)^2, \quad H_N = \mathrm{span}\{\phi_i(x)\phi_j(y)\}_{i,j=0}^{N-2},$$
$$Q_{N-2} = \mathrm{span}\{L_i(x)L_j(y); i,j = 0,\cdots,N-2\}, \qquad M_{N-2} = Q_{N-2} \cap L_0^2(\Omega).$$

Then we can present the approximations as follows: for m=1,2,

$$y_{Nm} = \sum_{i,j=0}^{N-2} y_{ijm}\phi_i(x_1)\phi_j(x_2), \qquad y_{Nm}^* = \sum_{i,j=0}^{N-2} y_{ijm}^*\phi_i(x_1)\phi_j(x_2),$$
$$u_{Nm} = \sum_{i,j=0}^{N} u_{ijm}L_i(x_1)L_j(x_2), \qquad r_N = \sum_{i,j=0}^{N-2} r_{ij}L_i(x_1)L_j(x_2),$$
$$r_N^* = \sum_{i,j=0}^{N-2} r_{ij}^*L_i(x_1)L_j(x_2), \qquad r_{00} = 0, \quad r_{00}^* = 0.$$

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

Algorithm 4.1 Arrow-Hurwicz Algorithm: The main steps of the algorithm are presented as follows.

1. Let $k = 0$, give a step size $\rho > 0$, the initial values λ_N^0 and $\boldsymbol{u}_N^0$.
2. Let $l = 0$ and $\boldsymbol{u}_N^{k,0} = \boldsymbol{u}_N^k$.
3. Compute the state equations
$$\begin{aligned} &a(\boldsymbol{y}_N^{k,l}, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N^{k,l}) = (\boldsymbol{u}_N^{k,l}, \boldsymbol{w}_N), && \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N, \\ &b(\boldsymbol{y}_h^{k,l}, \phi_N) = 0, && \forall\ \phi_N \in M_{N-2}. \end{aligned}$$
4. Compute the co-state equations
$$\begin{aligned} &a(\boldsymbol{y}_N^{*k,l}, \boldsymbol{w}_h) + b(r_N^{*k,l}, \boldsymbol{w}_N) = \left((1+\lambda_N^k)\boldsymbol{y}_N^{k,l} - \boldsymbol{y}_0, \boldsymbol{w}_N\right), && \forall\ \boldsymbol{w}_N \in \boldsymbol{Y}_N, \\ &b(\boldsymbol{y}_N^{*k,l}, \psi_N) = 0, && \forall\ \psi_N \in M_{N-2}. \end{aligned}$$
5. Set
$$\boldsymbol{u}_N^{k,l+1} = \boldsymbol{u}_N^{k,l} - P_N\left(\boldsymbol{y}_N^{*k,l} + \alpha(\boldsymbol{u}_N^{k,l} - \boldsymbol{u}_0)\right).$$
if
$$\|\boldsymbol{u}_N^{k,l+1} - \boldsymbol{u}_N^{k,l}\|_{0,\Omega} > \mathrm{Tol}_{\boldsymbol{u}},$$
let $l = l + 1$ and then we turn to *Step 3*.
6. Set $\lambda_N^{k+1} = max\left\{0, \lambda_N^k + \rho\left(\|\boldsymbol{y}_N^{k,l}\|_{1,\Omega} - d\right)\right\}$.
7. Stop if $|\lambda_N^{k+1} - \lambda_N^k| < \mathrm{Tol}_\lambda$ and output
$$\boldsymbol{u}_N = \boldsymbol{u}_N^{k,l}, \quad \boldsymbol{y}_N = \boldsymbol{y}_N^{k,l}, \quad r_N = r_N^{k,l}, \quad \lambda_N^k,$$
else, let $\boldsymbol{u}_N^{k+1} = \boldsymbol{u}_N^{k,l+1}$, $k = k + 1$, and turn to *Step 2*.

The numerical results are presented in Table 1, which shows us that the errors decrease quickly and the spectral accuracy is achieved. It also confirms that the spectral method provides very efficient approximation to the state-constrained control problems with a relatively small number of unknowns when the approximated solutions have higher regularity. In Table 1, the total error is defined by

$$\begin{aligned} \|e\| =& \|\boldsymbol{y} - \boldsymbol{y}_N\|_{1,\Omega} + \|\boldsymbol{y}^* - \boldsymbol{y}_N^*\|_{1,\Omega} + \|r - r_N\|_{0,\Omega} \\ &+ \|r^* - r_N^*\|_{0,\Omega} + |\lambda - \lambda_N| + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}. \end{aligned}$$

The table suggests that the error estimator can indicate the total error $\|e\|$ very well, which

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

Table 1: The values of discretization errors and indicator η.

N	4	8	12	16	20
$\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}$	5.87747e-1	2.72787e-3	1.51816e-6	1.26441e-7	1.15424e-7
$\|\boldsymbol{y}-\boldsymbol{y}_N\|_{1,\Omega}$	3.99351e-1	3.26411e-3	2.62428e-6	2.38535e-8	3.09446e-8
$\|r-r_N\|_{0,\Omega}$	2.80428e-1	2.97252e-3	2.49986e-6	5.26836e-9	5.26836e-9
$\|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|_{1,\Omega}$	3.31942	2.99138e-2	2.49747e-5	1.99920e-7	2.30848e-7
$\|r^*-r_N^*\|_{0,\Omega}$	2.67124	2.93414e-2	2.46714e-5	1.19209e-7	1.03238e-7
$\|\lambda-\lambda_N\|$	7.17934e-1	6.25525e-5	1.54084e-9	1.57468e-9	1.57469e-9
$\|e\|$	7.97612	6.82822e-2	5.62900e-5	4.76266e-7	4.87298e-7
η	8.86459	6.97704e-2	5.30573e-5	1.73363e-7	1.83494e-7
θ	6.14604e-1	1.43256e-3	5.26302e-7	1.27647e-7	1.27647e-7

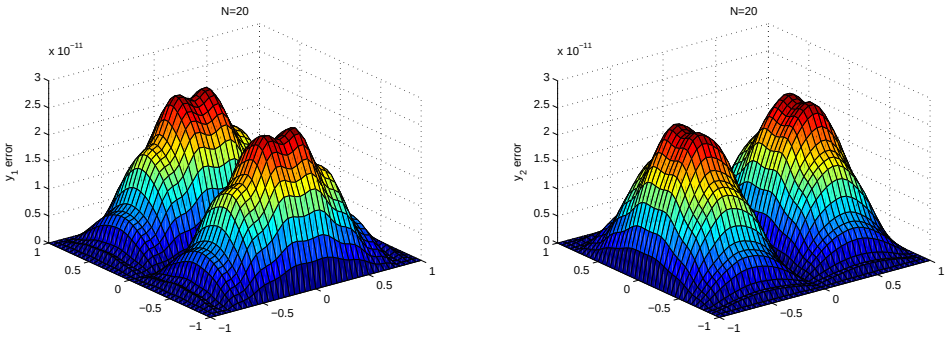


Figure 1: Point-wise error values of the velocity $\boldsymbol{y}$.

is important for developing *hp*-adaptive spectral element method for optimal control problems.

Moreover, in Figs. 1 and 2, we plot the point-wise error values of $\boldsymbol{y}$, $\boldsymbol{u}$, and r with $N = 20$. The two figures show that the final point-wise errors of $\boldsymbol{y}$, $\boldsymbol{u}$, r are of magnitude 10^{-11}, 10^{-9}, and 10^{-11} respectively.

5. Conclusions and future works

In this paper, a Legendre-Galerkin spectral method is investigated to approximate flow control problems with H^1-norm state constraint, and both a priori and a posteriori error estimates are derived. Numerical experiment shows that spectral method for this class of control problems can be efficient and the error estimator can work well. In addition, spectral method can be used to approximate other control problems with state constraints and control-state constraints. Furthermore, it is not clear whether *hp*-adaptive spectral element method can be used as widely as the adaptive finite element method for solving optimal control problems.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

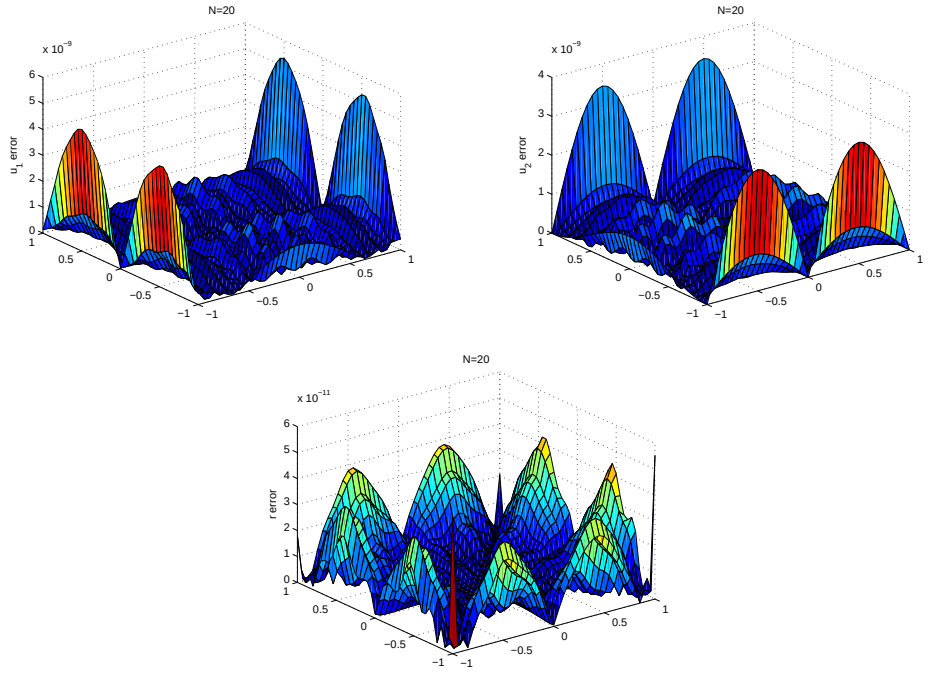


Figure 2: Point-wise error values of the control $\boldsymbol{u}$ and the pressure r.

Acknowledgments The first author was supported by the National Natural Science Foundation of China under Grant No. 11671157 and No. 91430104. The second author was supported by the National Natural Science Foundation of China under Grant No. 11601466 and the Nanhu Scholars Program for Young Scholars of XYNU, and the Doctoral Scientific Research Startup Fund of Xinyang Normal University(2016). Moreover, this work was also supported by the National Natural Science Foundation of China under Grant No. 11401347. The authors are grateful to the referees for their helpful and profound comments and advices.

References

[1] O. Benedix and B. Vexler, *A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints*, Comput. Optim. Appl., 44(1) (2009), pp. 3–25.

[2] M. Bergounioux and K. Kunisch, *On the structure of Lagrange multipliers for state-constrained optimal control problems*, Syst. Control Lett., 48 (2003), pp. 169–176.

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

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

[5] V. Barbu and T. Precupanu, Convexity and Optimization in Banach Spaces, Springer, 2012.

[6] E. Casas, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control Optim., 24(6) (1986), pp. 1309–1318.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

[7] E. Casas, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control Optim., 31(4) (1993), pp. 993–1006.
[8] C. Canuto, M.Y. Hussaini, A. Quarteroni, and T.A. Zang, Spectral Methods in Fluid Dynamics, Springer-Verlag, New York, 1988.
[9] C. Canuto, M.Y. Hussaini, A. Quarteroni, and T.A. Zang, Spectral Methods: Fundamentals in Single Domains, Springer-Verlag, Berlin, 2006.
[10] C. Canuto, M.Y. Hussaini, A. Quarteroni, and T.A. Zang, Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics, Springer-Verlag, Berlin, 2007.
[11] F.H. Clarke, Optimization and Nonsmooth Analysis, John Wiley Sons, Amsterdam, 1983.
[12] S. Cherednichenko and A. Rösch, *Error estimates for the discretization of elliptic control problems with pointwise control and state constraints*, Comput. Optim. Appl., 44(1) (2009), pp. 27–55.
[13] Y. Chen, N. Yi, and W.B. Liu, *A Legendre Galerkin spectral method for optimal control problems governed by elliptic equations*, SIAM J. Numer. Anal., 46(5) (2008), pp. 2254–2275.
[14] J.C. De Los Reyes and R. Griesse, *State-constrained optimal control of the three-dimensionl stationary Navier-Stokes equations* , J. Math. Anal. Appl., 343 (2008), pp. 257–272.
[15] J.C. De Los Reyes and K. Kunisch, *A semi-smooth Newton method for regularized state-constrained optimal control of the Navier-Stokes equations* , Comput., 78 (2006), pp. 287–309.
[16] M.D. Gunzburger and L.S. Hou, *Finite-dimensional approximation of a class of constrained nonlinear optimal control problems*, SIAM J. Control Optim., 34(3) (1996), pp. 1001–1043.
[17] W. Gong and N.N. Yan, *A mixed finite element scheme for optimal control problems with pointwise state constraints*, J. Sci. Comput., 46 (2011), pp. 182–203.
[18] R. Glowinski, J.L. Lions, and R. Trémolières, *Numerical analysis of variational inequalities*, Amsterdam, North-Holland, 1981.
[19] R.H.W. Hoppe and M. Kieweg, *A posteriori error estimation of finite element approximations of pointwise state constrained distributed control problems*, J. Numer. Math., 17(3) (2009), pp. 219–244.
[20] J. L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.
[21] W. B. Liu, W. Gong and N. Yan, *A new finite element approximation of a state-constrained optimal control problem*, J. Comp. Math., 27(1) (2009), pp. 97–114.
[22] W. B. Liu, D. P. Yang, L. Yuan, and C. Q. Ma, *Finite element approximations of an optimal control problem with integral state constraint*, SIAM J. Numer. Anal., 48(3) (2010), pp. 1163–1185.
[23] W. B. Liu and N. Yan, *A posteriori error estimates for control problems governed by Stokes equations*, SIAM J. Numer. Anal., 40(5) (2002), pp. 1850–1869.
[24] W.B. Liu and N. Yan, *Adaptive finite element methods for optimal control governed by PDEs*, Science Press, Beijing, 2008.
[25] D. Meidner, R. Rannacher, and B. Vexler, *A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time*, SIAM J. Control Optim., 49(5) (2011), pp. 1961–1997.
[26] H. F. Niu and D. P. Yang, *Finite element analysis of optimal control problem governed by Stokes equations with L^2-norm state-constriants*, J. Comput. Math., 29(5) (2011), pp. 589–604.
[27] P. Neittaanmaki and D. Tiba, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications*, Marcel Dekker, New York, 1994
[28] O.Pironneau, Optimal Shape Design for Elliptic Systems, Springer-Verlag, Berlin, 1984.
[29] J. Shen, *On fast direct poisson solver, inf-sup constant and iterative Stokes solver by Legendre-Galerkin method*, J. Comput. Phys., 116(1) (1995), pp. 184–188.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press

[30] J. Shen and T. Tang, Spectral and High-Order Methods with Applications, Science Press, Beijing, 2006.
[31] L. Yuan and D. P. Yang, *A posteriori error estimate of optimal control problem of PDE with integral constraint for state*, J. Comput. Math., 27(4) (2009), pp. 525–542.
[32] J. W. Zhou and D. P. Yang, *Spectral mixed Galerkin method for state constrained optimal control problem governed by the first bi-harmonic equation*, Int. J. Comput. Math., 88(14) (2011), pp. 2988–3011.

https://doi.org/10.4208/nmtma.2017.m1419 Published online by Cambridge University Press