Electronic Research Archive

http://www.aimspress.com/journal/era

ERA, 30(9): 3193–3210.
DOI: 10.3934/era.2022162
Received: 07 March 2020
Revised: 17 May 2022
Accepted: 18 May 2022
Published: 20 June 2022

Research article

An error estimator for spectral method approximation of flow control with state constraint

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

[1] School of Mathematics and Statistics, Xinyang Normal University, No.237 Nanhu Road, Shihe District, Xinyang 464000, China

[2] School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China

* **Correspondence:** Email: hfl_937@sina.com.

Abstract: We consider the spectral approximation of flow optimal control problems with state constraint. The main contribution of this work is to derive an a posteriori error estimator, and show the upper and lower bounds for the approximation error. Numerical experiment shows the efficiency of the error indicator, which will be used to construct reliable adaptive *hp* spectral element approximation for flow control in the future work.

Keywords: spectral method; optimal control; state constraint; stokes equations; upper and lower bounds

1. Introduction

There have been lots of theoretical studies on the optimal control of PDEs with state constraints, which form a foundation for its numerical approximation. Existence and uniqueness of the solutions, Lagrange multipliers, optimality conditions, and important regularity results were derived for control problems with state constraints in the pioneer work [1]. More discussion on these topics can be found in [2, 3]. In respect of numerical methods, the finite element approximation of state constrained control problems has been widely investigated, and we don't try to give a detailed introduction here, one can find more works [4, 5], including pointwise constraints, integral constraint and so on. At the same time, many numerical strategies were developed to provide efficient approximation for these control problems. Bergounioux and Kunisch [6] used the primal-dual active set algorithm to solve state-constrained problems. A semi-smooth Newton method was proposed to compute state-constrained control problems by De los Reyes and Kunisch [7]. Gong and Yan [8] established a mixed variational scheme for control problems with pointwise state constrains, and a direct numerical algorithm was adopted without the optimality conditions.

In recent years, spectral method has been used to approximate control problems. Despite the restriction on the higher regularity of the approximated solutions, spectral method can provide fast convergence rate and high-order accuracy with a smaller number of unknowns, which is significant to successful applications of control problems. The spectral method was considered to solve the control problems with integral state constraint, and a priori error estimates was established in [9]. Chen et al. [10, 11] derived both the a priori and a posteriori error estimates for optimal control with control constraints. However, they only provided the upper bound estimation for the a posteriori error indicators, and numerical tests didn't illustrate the performance of estimators. To investigate the efficiency of adaptive strategy in the spectral method, we have to establish some successful a posteriori error estimators as in the finite element framework. However, there are not much work on these aspects to the best of our knowledge. In this work, we derive a posteriori error estimator, and prove that it can be constructed as the upper and lower bounds of approximation error.

The plan of the article is as follows. Spectral approximation of the control problem is presented, and optimality conditions of the exact problem and discretized problem are provided in the next section. We establish the a posteriori error estimator, and construct it as the upper and lower bounds of the approximation error in section 3. Numerical example confirms the theoretical result, and shows the behaviour of the indicator in section 4.

In this paper, we let $\Omega = (-1, 1)^2$ and denote $\boldsymbol{U} = L^2(\Omega)^2, \boldsymbol{Y} = H_0^1(\Omega)^2, Q = L_0^2(\Omega) = \{q \in L^2(\Omega) \mid \int_\Omega q = 0\}$, where $H_0^1(\Omega)$ and $H^m(\Omega)$ with m being a positive integer are usual Sobolev spaces on Ω. Let C denote a positive constant independent of N, the order of the spectral method.

2. Spectral approximation and optimality conditions

In this section, we state the Galerkin spectral approximation and optimality conditions for the control problem with state constraint. The model under consideration is as follows: find $(\boldsymbol{y}, r, \boldsymbol{u}) \in \boldsymbol{Y}\times Q\times \boldsymbol{U}$ such that

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

where $\boldsymbol{y}_0, \boldsymbol{f} \in L^2(\Omega)^2$, $G_{ad} = \{\boldsymbol{v} \in \boldsymbol{U} \mid \|\, \boldsymbol{v}\, \|_{0,\Omega} \leq d\}$, and d, υ are positive constants. It is necessary to introduce a weak formula of the state equations for spectral approximation of the optimal control. Let

$$\begin{aligned} a(\boldsymbol{w}, \boldsymbol{v}) &= \upsilon \int_\Omega \nabla \boldsymbol{w}\cdot\nabla \boldsymbol{v} \quad \forall\, \boldsymbol{w}, \boldsymbol{v} \in \boldsymbol{Y}, \\ b(\boldsymbol{v}, r) &= \int_\Omega r\, \nabla\cdot \boldsymbol{v} \quad \forall\, (\boldsymbol{v}, r) \in \boldsymbol{Y}\times Q. \end{aligned}$$

Then there exist two positive constants σ and δ such that for any $\boldsymbol{u},\ \boldsymbol{v} \in \boldsymbol{Y},\ q \in Q$,

$$\begin{aligned} |a(\boldsymbol{u}, \boldsymbol{v})| &\leq \sigma \,\|\, \boldsymbol{u}\, \|_{\boldsymbol{Y}} \|\, \boldsymbol{v}\, \|_{\boldsymbol{Y}}, \\ |b(\boldsymbol{v}, q)| &\leq \delta \,\|\, \boldsymbol{v}\, \|_{\boldsymbol{Y}} \|\, q\, \|_Q\, . \end{aligned}$$

By Poincare inequality, we can know that there exists a constant $\gamma > 0$ such that

$$a(\boldsymbol{y}, \boldsymbol{y}) \geq \gamma \|\boldsymbol{y}\|_{\boldsymbol{Y}}^2 \quad \forall\, \boldsymbol{y} \in \boldsymbol{Y}. \tag{2.2}$$

Furthermore, there exists a constant $\beta > 0$ (see, e.g., [12]) such that

$$\sup_{\boldsymbol{v}\in\boldsymbol{Y}} \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{Y}}} \geq \beta \|q\|_Q \quad \forall\, q \in Q. \tag{2.3}$$

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

$$\begin{aligned} a(\boldsymbol{y}(\boldsymbol{u}), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u})) &= (\boldsymbol{u} + \boldsymbol{f}, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \\ b(\boldsymbol{y}(\boldsymbol{u}), \phi) &= 0 \quad \forall\, \phi \in Q. \end{aligned} \tag{2.4}$$

The problem (2.4) is well-posed by Babuška-Brezzi theorem and (2.2)-(2.3), and 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

$$\begin{aligned} \min_{\boldsymbol{y}(\boldsymbol{u})\in G_{ad}} J(\boldsymbol{u}) &= \frac{1}{2}\|\boldsymbol{y}(\boldsymbol{u}) - \boldsymbol{y}_0\|_{0,\Omega}^2 + \frac{\alpha}{2}\|\boldsymbol{u}\|_{0,\Omega}^2, \\ a(\boldsymbol{y}(\boldsymbol{u}), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u})) &= (\boldsymbol{u} + \boldsymbol{f}, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \\ b(\boldsymbol{y}(\boldsymbol{u}), \phi) &= 0 \quad \forall\, \phi \in Q. \end{aligned} \tag{2.5}$$

We can prove existence and uniqueness of the solutions for the control problem (OCP) by reformulating it to a control-constrained problem. In fact, the control problem (OCP) can be equivalent to the control-constrained problem

$$\begin{aligned} \min_{\boldsymbol{u}\in\mathscr{U}_{ad}} J(\boldsymbol{u}) &= \frac{1}{2}\|\boldsymbol{y}(\boldsymbol{u}) - \boldsymbol{y}_0\|_{0,\Omega}^2 + \frac{\alpha}{2}\|\boldsymbol{u}\|_{0,\Omega}^2, \\ a(\boldsymbol{y}(\boldsymbol{u}), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u})) &= (\boldsymbol{u} + \boldsymbol{f}, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \\ b(\boldsymbol{y}(\boldsymbol{u}), \phi) &= 0 \quad \forall\, \phi \in Q, \end{aligned} \tag{2.6}$$

where constraint set $\mathscr{U}_{ad} = \{\boldsymbol{u} \in \boldsymbol{U} : \|G\boldsymbol{u}\|_{0,\Omega} \leq d\}$ with the operator $G : \boldsymbol{u} \to \boldsymbol{y}(\boldsymbol{u})$. It is clear that $\mathscr{U}_{ad}$ is a closed and convex subset in $\boldsymbol{U}$, and the control problem (2.5) has a unique solution by the standard theorem [13].

2.1. Spectral method approximation

Considering the spectral approximation of the control problem, we introduce the finite dimensional spaces $\boldsymbol{Y}_N$, M_{N-2}, and $\boldsymbol{U}_N$ to approximate spaces $\boldsymbol{Y}$, Q, and $\boldsymbol{U}$ respectively, where $\boldsymbol{Y}_N = (Q_N^0)^2$ with $Q_N^0 = \{x_N \in Q_N : x_N|_{\partial\Omega} = 0\}$, $M_{N-2} = Q_{N-2} \cap L_0^2(\Omega)$, $\boldsymbol{U}_N = (Q_N)^2$, and Q_N denotes the space of all algebraic polynomials of degree less than or equal to N with respect to each single variable x_i, $(i = 1, 2)$. Letting $L_n(x)$ be the nth-degree Legendre polynomial and $\phi_k(x) = \frac{1}{\sqrt{4k+6}}(L_k(x) - L_{k+2}(x))$ (see [14], for example), then it holds

$$Q_N^0 = \mathrm{span}\{\phi_i(x)\phi_j(y)\}_{i,j=0}^{N-2}, \quad Q_N = \mathrm{span}\{L_i(x)L_j(y)\}_{i,j=0}^{N}, \quad M_{N-2} = Q_{N-2} \cap L_0^2(\Omega). \tag{2.7}$$

We now approximate the state equations as follows

$$\begin{aligned} a(\boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N) &= (\boldsymbol{u}_N + \boldsymbol{f}, \boldsymbol{w}_N) \quad \forall\, \boldsymbol{w}_N \in \boldsymbol{Y}_N \subset \boldsymbol{Y}, \\ b(\boldsymbol{y}_N, \phi_N) &= 0 \quad \forall\, \phi_N \in M_{N-2} \subset Q. \end{aligned} \tag{2.8}$$

It can be known from [12] that there exists a constant $\beta_N = O(N^{-\frac{1}{2}})$ satisfying

$$\sup_{\boldsymbol{v}_N \in \boldsymbol{Y}_N} \frac{b(\boldsymbol{v}_N, q)}{\|\boldsymbol{v}_N\|_{\boldsymbol{Y}}} \geq \beta_N \|q\|_Q \quad \forall\, q \in M_{N-2}. \tag{2.9}$$

It follows from Babuška-Brezzi's theory that problem (2.8) is well-posed, and the Legendre Galerkin spectral approximation of (OCP) can be stated as $(\mathrm{OCP})_\mathrm{N}$:

$$\begin{aligned} \min_{\boldsymbol{y}_N \in G_{ad}^N} \ J_N(\boldsymbol{u}_N) &= \frac{1}{2}\|\boldsymbol{y}_N - \boldsymbol{y}_0\|_{0,\Omega}^2 + \frac{\alpha}{2}\|\boldsymbol{u}_N\|_{0,\Omega}^2, \\ a(\boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N) &= (\boldsymbol{u}_N + \boldsymbol{f}, \boldsymbol{w}_N) \quad \forall\, \boldsymbol{w}_N \in \boldsymbol{Y}_N \subset \boldsymbol{Y}, \\ b(\boldsymbol{y}_N, \phi_N) &= 0 \quad \forall\, \phi_N \in M_{N-2} \subset Q, \end{aligned} \tag{2.10}$$

where $G_{ad}^N = \boldsymbol{Y}_N \cap G_{ad}$.

2.2. *Optimality conditions*

We first derive optimality conditions for the exact problem (OCP) by the techniques and refined results developed in [2], though we can complete the derivation by other strategies. Then the similar conclusion is presented for the discretized problem $(\mathrm{OCP})_\mathrm{N}$.

Lemma 2.1. *The triplet* $(\boldsymbol{y}, r, \boldsymbol{u})$ *is the solution of control problem* (OCP) *if and only if there is a* $(\boldsymbol{y}^*, r^*, \lambda) \in \boldsymbol{Y} \times Q \times \mathbb{R}$ *such that*

$$a(\boldsymbol{y}, \boldsymbol{w}) - b(\boldsymbol{w}, r) = (\boldsymbol{u} + \boldsymbol{f}, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \tag{2.11a}$$

$$b(\boldsymbol{y}, \phi) = 0 \quad \forall\, \phi \in Q, \tag{2.11b}$$

$$a(\boldsymbol{q}, \boldsymbol{y}^*) + b(\boldsymbol{q}, r^*) = ((1+\lambda)\boldsymbol{y} - \boldsymbol{y}_0, \boldsymbol{q}) \quad \forall\, \boldsymbol{q} \in \boldsymbol{Y}, \tag{2.11c}$$

$$b(\boldsymbol{y}^*, \psi) = 0 \quad \forall\, \psi \in Q, \tag{2.11d}$$

$$\boldsymbol{y}^* + \alpha \boldsymbol{u} = 0 \quad \text{in } \Omega, \tag{2.11e}$$

where

$$\lambda = \begin{cases} \text{constant} \geq 0 & \text{if } \|\boldsymbol{y}\|_{0,\Omega} = d, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

Proof. We denote $G : \boldsymbol{u} \to \boldsymbol{y}$ the operator to solve the state equation (2.4), $D(G(\boldsymbol{u}))$ the Gâteaux derivative of G at $\boldsymbol{u}$. Furthermore, let $U = Z = K = L^2(\Omega)^2$, $C = G_{ad}$ respectively and $\boldsymbol{u}_0 = -\boldsymbol{f}$ satisfy the Slater type condition in Theorem 5.2 in [2], then there exists $\boldsymbol{\mu} \in L^2(\Omega)^2$ such that $(\boldsymbol{u}, \boldsymbol{\mu})$ satisfying

$$\begin{aligned} (\boldsymbol{\mu}, \boldsymbol{v} - G(\boldsymbol{u})) &\leq 0 \quad \forall\, \boldsymbol{v} \in G_{ad}, \\ J'(\boldsymbol{u}) + [D(G(\boldsymbol{u}))]^* \boldsymbol{\mu} &= 0, \end{aligned} \tag{2.13}$$

where $(G(\boldsymbol{u}), \boldsymbol{u}) \in \boldsymbol{Y} \times \boldsymbol{U}$ satisfies the control problem (OCP).

In the next analysis, it can be derived that

$$\boldsymbol{\mu} = \lambda \boldsymbol{y}, \tag{2.14}$$

where λ satisfies (2.12). In fact, we can complete the proof by two cases: $\|\boldsymbol{y}\|_{0,\Omega} < d$ and $\|\boldsymbol{y}\|_{0,\Omega} = d$.

In the first case $\|\boldsymbol{y}\|_{0,\Omega} < d$, the following inequality holds by (2.13)

$$(\boldsymbol{\mu}, \boldsymbol{v} - \boldsymbol{y}) \leq 0 \quad \forall\, \boldsymbol{v} \in G_{ad}.$$

Thus we have for all $\boldsymbol{\omega} \in \boldsymbol{U}$ and $\|\omega\|_{0,\Omega} = 1$

$$< \boldsymbol{\mu}, \boldsymbol{\omega} > = \frac{< \boldsymbol{\mu}, (d - \|\boldsymbol{y}\|_{0,\Omega})\boldsymbol{\omega} + \boldsymbol{y} - \boldsymbol{y} >}{d - \|\boldsymbol{y}\|_{0,\Omega}} \leq 0, \tag{2.15}$$

due to $(d - \|\boldsymbol{y}\|_{0,\Omega})\boldsymbol{\omega} + \boldsymbol{y} \in G_{ad}$. Similarly,

$$< \boldsymbol{\mu}, -\boldsymbol{\omega} > \leq 0. \tag{2.16}$$

By (2.15) and (2.16)

$$\boldsymbol{\mu} = 0. \tag{2.17}$$

In the second case $\|\boldsymbol{y}\|_{0,\Omega} = d$, it follows from (2.13) that

$$\begin{aligned} \|\boldsymbol{\mu}\|_{0,\Omega} = \sup_{\boldsymbol{v} \in L^2(\Omega)^2, \boldsymbol{v} \neq 0} \frac{(\boldsymbol{\mu}, \boldsymbol{v})}{\|\boldsymbol{v}\|_{0,\Omega}} = \frac{1}{d} \sup_{\boldsymbol{v} \in G_{ad}} (\boldsymbol{\mu}, \boldsymbol{v}) \leq \frac{1}{d}(\boldsymbol{\mu}, \boldsymbol{y}) \\ \leq \frac{1}{d}\|\boldsymbol{\mu}\|_{0,\Omega}\|\boldsymbol{y}\|_{0,\Omega} = \|\boldsymbol{\mu}\|_{0,\Omega}. \end{aligned}$$

Then, we have

$$(\boldsymbol{\mu}, \boldsymbol{y}) = \|\boldsymbol{\mu}\|_{0,\Omega}\|\boldsymbol{y}\|_{0,\Omega},$$

which implies that

$$\boldsymbol{\mu} = \lambda \boldsymbol{y} \quad \lambda \in \mathbb{R}^1, \tag{2.18}$$

and

$$\lambda \geq 0. \tag{2.19}$$

Then, the identity (2.14) follows from (2.17)–(2.19).

It can be derived from (2.4) that for any $\boldsymbol{v} \in L^2(\Omega)^2$

$$a(\boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}. \tag{2.20}$$

By $\boldsymbol{y} - \boldsymbol{y}_0 + \lambda \boldsymbol{y} \in L^2(\Omega)^2$, we can introduce the co-state equation as follows

$$\begin{aligned} a(\boldsymbol{q}, \boldsymbol{y}^*) + b(\boldsymbol{q}, r^*) = (\boldsymbol{y} - \boldsymbol{y}_0 + \lambda \boldsymbol{y}, \boldsymbol{q}) \quad \forall\, \boldsymbol{q} \in \boldsymbol{Y}. \\ b(\boldsymbol{y}^*, \psi) = 0 \quad \forall\, \psi \in Q. \end{aligned} \tag{2.21}$$

Letting $\boldsymbol{w} = \boldsymbol{y}^*$ in (2.20) and $\boldsymbol{q} = \boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}$ in (2.21), we have that for all $\boldsymbol{v} \in L^2(\Omega)^2$

$$\begin{aligned} & < J'(\boldsymbol{u}) + [DG(\boldsymbol{u})]^* \boldsymbol{\mu}, \boldsymbol{v} > \\ = & (\boldsymbol{y} - \boldsymbol{y}_0, \boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}) + (\alpha \boldsymbol{u}, \boldsymbol{v}) + \lambda(\boldsymbol{y}, \boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}) \\ = & (\boldsymbol{y} - \boldsymbol{y}_0 + \lambda \boldsymbol{y}, \boldsymbol{y}'(\boldsymbol{u})\boldsymbol{v}) + (\alpha \boldsymbol{u}, \boldsymbol{v}) = (\alpha \boldsymbol{u} + \boldsymbol{y}^*, \boldsymbol{v}). \end{aligned} \tag{2.22}$$

Then (2.11)-(2.12) follows from (2.4), (2.13), (2.19), (2.21)-(2.22). Furthermore, lemma 2.1 can be proved as soon as the uniqueness of the solution is derived for (2.11). In fact, letting both $(\boldsymbol{y}_1, r_1, \boldsymbol{u}_1, \boldsymbol{y}_1^*, r_1^*, \lambda)$ and $(\boldsymbol{y}_2, r_2, \boldsymbol{u}_2, \boldsymbol{y}_2^*, r_2^*, \lambda)$ satisfy (2.11), then we have

$$\begin{aligned} a(\boldsymbol{y}_1 - \boldsymbol{y}_2, \boldsymbol{w}) - b(\boldsymbol{w}, r_1 - r_2) &= (\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \\ b(\boldsymbol{y}_1 - \boldsymbol{y}_2, \phi) &= 0 \quad \forall\, \phi \in Q, \\ a(\boldsymbol{q}, \boldsymbol{y}_1^* - \boldsymbol{y}_2^*) + b(\boldsymbol{q}, r_1^* - r_2^*) &= (\boldsymbol{y}_1 - \boldsymbol{y}_2 + \lambda_1 \boldsymbol{y}_1 - \lambda_2 \boldsymbol{y}_2, \boldsymbol{q}) \quad \forall\, \boldsymbol{q} \in \boldsymbol{Y}, \\ b(\boldsymbol{y}_1^* - \boldsymbol{y}_2^*, \psi) &= 0 \quad \forall\, \psi \in Q. \end{aligned}$$

It follows that

$$-\alpha(\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{u}_1 - \boldsymbol{u}_2) - (\boldsymbol{y}_1 - \boldsymbol{y}_2, \boldsymbol{y}_1 - \boldsymbol{y}_2) = (\lambda_1 \boldsymbol{y}_1 - \lambda_2 \boldsymbol{y}_2, \boldsymbol{y}_1 - \boldsymbol{y}_2). \tag{2.23}$$

By (2.13)-(2.14), it holds that for all $\boldsymbol{v} \in G_{ad}$

$$\lambda_1(\boldsymbol{y}_1, \boldsymbol{v} - \boldsymbol{y}_1) \le 0, \quad \lambda_2(\boldsymbol{y}_2, \boldsymbol{v} - \boldsymbol{y}_2) \le 0. \tag{2.24}$$

It follows from (2.23) and (2.24) that

$$\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{0,\Omega} + \|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_{0,\Omega} = 0,$$

which implies $\boldsymbol{u}_1 = \boldsymbol{u}_2, \boldsymbol{y}_1 = \boldsymbol{y}_2$. Furthermore, it can be deduced that $r_1 = r_2, \boldsymbol{y}_1^* = \boldsymbol{y}_2^*, r_1^* = r_2^*, \lambda_1 = \lambda_2$. Then we can complete the proof of lemma 2.1 as we argue above. □

Similarly, we can derive optimality conditions for the discretized control problem (2.10), and obtain the following result.

Lemma 2.2. *The control problem* $(\mathrm{OCP})_{\mathrm{N}}$ *has a unique solution, and the triple* $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N) \in \boldsymbol{Y}_N \times M_{N-2} \times \boldsymbol{U}_N$ *is the solution of* $(\mathrm{OCP})_{\mathrm{N}}$ *if and only if there is a* $(\boldsymbol{y}_N^*, r_N^*, \lambda_N) \in \boldsymbol{Y}_N \times M_{N-2} \times \mathbb{R}$ *such that*

$$a(\boldsymbol{y}_N, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N) = (\boldsymbol{u}_N + \boldsymbol{f}, \boldsymbol{w}_N) \quad \forall\, \boldsymbol{w}_N \in \boldsymbol{Y}_N, \tag{2.25a}$$

$$b(\boldsymbol{y}_N, \phi_N) = 0 \quad \forall\, \phi_N \in M_{N-2}, \tag{2.25b}$$

$$a(\boldsymbol{q}_N, \boldsymbol{y}_N^*) + b(\boldsymbol{q}_N, r_N^*) = ((1 + \lambda_N)\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{q}_N) \quad \forall\, \boldsymbol{q}_N \in \boldsymbol{Y}_N, \tag{2.25c}$$

$$b(\boldsymbol{y}_N^*, \psi_N) = 0 \quad \forall\, \psi_N \in M_{N-2}. \tag{2.25d}$$

$$\boldsymbol{y}_N^* + \alpha \boldsymbol{u}_N = 0 \quad \text{in } \Omega, \tag{2.25e}$$

where

$$\lambda_N = \begin{cases} \text{constant} \ge 0 & \text{if } \| \boldsymbol{y}_N(\boldsymbol{u}_N) \|_{0,\Omega} = d, \\ 0 & \text{otherwise.} \end{cases} \tag{2.26}$$

Remark 2.1. *(see, e.g., [10]). The optimal control* $\boldsymbol{u} \in H^2(\Omega)^2$ *if the initial data functions* $\boldsymbol{y}_0, \boldsymbol{f} \in L^2(\Omega)^2$ *by (2.11).*

3. A posteriori error estimator

In this section, the a posteriori error estimates are derived, and an error indicator is established for the spectral approximation of the control problem. We first recall two important spectral projection operators and more details can be found in [15].

Lemma 3.1. *Let $P^0_{1,N} : H^1_0(\Omega)^2 \to (Q^0_N)^2$ be the projection operator, satisfying for any $\boldsymbol{w} \in H^1_0(\Omega)^2$*

$$\int_\Omega \nabla(\boldsymbol{w} - P^0_{1,N}\boldsymbol{w}) \cdot \nabla \boldsymbol{v}_N = 0 \quad \forall\, \boldsymbol{v}_N \in \boldsymbol{Y}_N.$$

If $\boldsymbol{w} \in (H^m(\Omega) \cap H^1_0(\Omega))^2$ with $m \geq 1$, then

$$\|\boldsymbol{w} - P^0_{1,N}\boldsymbol{w}\|_{k,\Omega} \leq CN^{k-m}\|\boldsymbol{w}\|_{m,\Omega} \qquad k = 0, 1. \tag{3.1}$$

Lemma 3.2. *Define $P_N : L^2(\Omega) \to Q_N$ being the L^2 orthogonal projection operator, which satisfies for any $r \in L^2(\Omega)$*

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

If $r \in H^m(\Omega)$ with $m \geq 0$, then

$$\|r - P_N r\|_{0,\Omega} \leq CN^{-m}\|r\|_{m,\Omega}.$$

The following lemma is helpful for further analysis.

Lemma 3.3. *Let $(\boldsymbol{y}_N, \boldsymbol{u}_N, \boldsymbol{y}^*_N, \lambda_N)$ be the solution of (2.25), then*

$$\max\{\|\boldsymbol{u}_N\|_{0,\Omega}, \|\boldsymbol{y}_N\|_{1,\Omega}, \|\boldsymbol{y}^*_N\|_{1,\Omega}, |\lambda_N|\} \leq C.$$

Proof. By $(\boldsymbol{y}_N, \boldsymbol{u}_N, \boldsymbol{y}^*_N, \lambda_N)$ satisfying the optimality conditions (2.25), we have

$$\begin{aligned} J_N(\boldsymbol{u}_N) &= \frac{1}{2}\|\boldsymbol{y}_N - \boldsymbol{y}_0\|^2_{0,\Omega} + \frac{\alpha}{2}\|\boldsymbol{u}_N\|^2_{0,\Omega} \leq J_N(-P_N\boldsymbol{f}) \\ &= \frac{1}{2}\|\boldsymbol{y}_N(-P_N f) - \boldsymbol{y}_0\|^2_{0,\Omega} + \frac{\alpha}{2}\| - P_N\boldsymbol{f}\|^2_{0,\Omega} = \frac{1}{2}\|\boldsymbol{y}_0\|^2_{0,\Omega} + \frac{\alpha}{2}\|P_N\boldsymbol{f}\|^2_{0,\Omega} \leq C, \end{aligned}$$

associating with (2.25a) and (2.25e), it holds that

$$\|\boldsymbol{y}_N\|_{1,\Omega} + \|\boldsymbol{y}^*_N\|_{0,\Omega} + \|\boldsymbol{u}_N\|_{0,\Omega} \leq C. \tag{3.2}$$

The discussion on estimating $\|\boldsymbol{y}^*_N\|_{1,\Omega}$ can be divided into two cases: $\|\boldsymbol{y}_N\|_{0,\Omega} < d$ and $\|\boldsymbol{y}_N\|_{0,\Omega} = d$. The first case $\|\boldsymbol{y}_N\|_{0,\Omega} < d$, we have $\lambda_N = 0$. Letting $\boldsymbol{q}_N = \boldsymbol{y}^*_N$ in (2.25c), we have

$$\|\boldsymbol{y}^*_N\|_{1,\Omega} \leq C. \tag{3.3}$$

The second case $\|\boldsymbol{y}_N\|_{0,\Omega} = d$, let $\boldsymbol{q}_N = \boldsymbol{y}^*_N - \frac{1}{d^2}(\boldsymbol{y}^*_N, \boldsymbol{y}_N)\boldsymbol{y}_N$ in (2.25c). Then it holds

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

It follows that

$$
\begin{aligned}
a(\boldsymbol{y}_N^*, \boldsymbol{y}_N^*) &= (\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N^* - \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{y}_N)\boldsymbol{y}_N) \\
&\quad + a(\frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{y}_N)\boldsymbol{y}_N, \boldsymbol{y}_N^*) \\
&= (\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N^*) - \frac{1}{d^2}(\boldsymbol{y}_N - \boldsymbol{y}_0, \boldsymbol{y}_N)(\boldsymbol{y}_N^*, \boldsymbol{y}_N) \\
&\quad + \frac{1}{d^2}(\boldsymbol{y}_N^*, \boldsymbol{y}_N)a(\boldsymbol{y}_N, \boldsymbol{y}_N^*) \\
&\leq \varepsilon\|\boldsymbol{y}_N^*\|_{1,\Omega}^2 + \frac{C^2}{4\varepsilon}.
\end{aligned} \tag{3.4}
$$

It follows from (3.3) and (3.4) that

$$\|\boldsymbol{y}_N^*\|_{1,\Omega} \leq C. \tag{3.5}$$

Similarly, we can estimate $|\lambda_N|$ by two cases: $\|\boldsymbol{y}_N\|_{0,\Omega} < d$ and $\|\boldsymbol{y}_N\|_{0,\Omega} = d$.
If $\|\boldsymbol{y}_N\|_{0,\Omega} < d$, then $|\lambda_N| \leq C$.
If $\|\boldsymbol{y}_N\|_{0,\Omega} = d$, letting $\boldsymbol{q}_N = \boldsymbol{y}_N$ in (2.25c), we derive

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

Therefore, it can be derived that

$$|\lambda_N| \leq C. \tag{3.6}$$

Then we can complete the proof by (3.2), (3.5), and (3.6). □

3.1. Upper error bound

In this subsection, we establish the a posteriori error estimator and prove that it provides an upper bound for the discretization errors. It is convenient to introduce an 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

$$a(\boldsymbol{y}(\boldsymbol{u}_N), \boldsymbol{w}) - b(\boldsymbol{w}, r(\boldsymbol{u}_N)) = (\boldsymbol{u}_N + P_N\boldsymbol{f}, \boldsymbol{w}) \quad \forall\, \boldsymbol{w} \in \boldsymbol{Y}, \tag{3.7a}$$

$$b(\boldsymbol{y}(\boldsymbol{u}_N), \phi) = 0 \quad \forall\, \phi \in Q, \tag{3.7b}$$

$$a(\boldsymbol{q}, \boldsymbol{y}^*(\boldsymbol{u}_N)) + b(\boldsymbol{q}, r^*(\boldsymbol{u}_N)) = ((1 + \lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0, \boldsymbol{q}) \quad \forall\, \boldsymbol{q} \in \boldsymbol{Y}, \tag{3.7c}$$

$$b(\boldsymbol{y}^*(\boldsymbol{u}_N), \psi) = 0 \quad \forall\, \psi \in Q. \tag{3.7d}$$

By utilizing lemma 3.1–3.3, we can get the following lemmas.

Lemma 3.4. *Let $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ and $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ be the solutions of (2.11) and (2.25) respectively. Then we have that*

$$
\begin{aligned}
|\lambda - \lambda_N| \leq C(&\|\boldsymbol{y}^*(\boldsymbol{u}_N) - \boldsymbol{y}_N^*\|_{0,\Omega} + \|\boldsymbol{y}(\boldsymbol{u}_N) - \boldsymbol{y}_N\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} \\
&+ N^{-1}\|\boldsymbol{y}_0 - P_N\boldsymbol{y}_0\|_{0,\Omega} + N^{-1}\|\boldsymbol{f} - P_N\boldsymbol{f}\|_{0,\Omega}).
\end{aligned} \tag{3.8}
$$

Proof. By (2.11) and (3.7), we have

$$a(\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{w})-b(\boldsymbol{w},r-r(\boldsymbol{u}_N))=(\boldsymbol{u}-\boldsymbol{u}_N+\boldsymbol{f}-P_N\boldsymbol{f},\boldsymbol{w})\quad\forall\,\boldsymbol{w}\in Y,\tag{3.9a}$$

$$b(\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N),\phi)=0\quad\forall\,\phi\in Q,\tag{3.9b}$$

$$a(\boldsymbol{q},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+b(\boldsymbol{q},r^*-r^*(\boldsymbol{u}_N))=(\boldsymbol{y}-\boldsymbol{y}_N+\lambda\boldsymbol{y}-\lambda_N\boldsymbol{y}_N+P_N\boldsymbol{y}_0-\boldsymbol{y}_0,\boldsymbol{q})\quad\forall\,\boldsymbol{q}\in Y,\tag{3.9c}$$

$$b(\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N),\psi)=0\quad\forall\,\psi\in Q.\tag{3.9d}$$

The following analysis is divided into three cases.

If $\|\boldsymbol{y}\|_{0,\Omega}=d$, let $\boldsymbol{q}=\boldsymbol{y}$ in (3.9c). Then we have

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

Let $\boldsymbol{w}=\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)$ in (2.11a), we have

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

Thus it can be deduced that

$$(\lambda-\lambda_N)d^2=(\boldsymbol{u}+\boldsymbol{f},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))-(1+\lambda_N)(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y})+(\boldsymbol{y}_0-P_N\boldsymbol{y}_0,\boldsymbol{y}).\tag{3.10}$$

If $\|\boldsymbol{y}_N\|_{0,\Omega}=d$, let $\boldsymbol{q}=\boldsymbol{y}(\boldsymbol{u}_N)$ in (3.9c) to obtain that

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

Let $\boldsymbol{w}=\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)$ in (3.7a) to derive that

$$\begin{aligned}&(\boldsymbol{y}-\boldsymbol{y}_N+\lambda\boldsymbol{y}-\lambda\boldsymbol{y}_N+\lambda\boldsymbol{y}_N-\lambda_N\boldsymbol{y}_N,\boldsymbol{y}_N+\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)\\&=(\boldsymbol{u}_N+P_N\boldsymbol{f},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+(\boldsymbol{y}_0-P_N\boldsymbol{y}_0,\boldsymbol{y}(\boldsymbol{u}_N)).\end{aligned}$$

It follows that

$$\begin{aligned}(\lambda-\lambda_N)d^2=&(\boldsymbol{u}_N+P_N\boldsymbol{f},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))-(1+\lambda)(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N))\\&-(\lambda-\lambda_N)(\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)+(\boldsymbol{y}_0-P_N\boldsymbol{y}_0,\boldsymbol{y}(\boldsymbol{u}_N)-P^0_{1,N}\boldsymbol{y}(\boldsymbol{u}_N)).\end{aligned}\tag{3.11}$$

If $\|\boldsymbol{y}\|_{0,\Omega}<d$ and $\|\boldsymbol{y}_N\|_{0,\Omega}<d$, it follows from (2.12) and (2.26) that $\lambda=\lambda_N=0$, which implies the identity

$$|\lambda-\lambda_N|=0.\tag{3.12}$$

Letting $\boldsymbol{w}=\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)$ in (3.9a) and $\phi=r-r(\boldsymbol{u}_N)$ in (3.9b), we have that

$$\begin{aligned}a(\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))&=(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))+(\boldsymbol{f}-P_N\boldsymbol{f},\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)-P^0_{1,N}(\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)))\\&\le\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}+CN^{-1}\|\boldsymbol{f}-P_N\boldsymbol{f}\|_{0,\Omega}\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}\\&\le(\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}+CN^{-1}\|\boldsymbol{f}-P_N\boldsymbol{f}\|_{0,\Omega})\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega},\end{aligned}$$

which implies that

$$\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}+\|r-r(\boldsymbol{u}_N)\|_{0,\Omega}\le C\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}+CN^{-1}\|\boldsymbol{f}-P_N\boldsymbol{f}\|_{0,\Omega}.\tag{3.13}$$

Then, we can derive from (2.11e), (2.25e), (3.10)–(3.13), and lemma 3.3 that

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

which implies the lemma 3.4. □

Lemma 3.5. *Let $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ be the solution of (2.11), and $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ be the solution of (2.25). Then we have that*

$$\begin{aligned}&\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}+\|r(\boldsymbol{u}_N)-r_N\|_{0,\Omega}\\ &\le CN^{-1}\|\nu\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N+P_N\boldsymbol{f}\|_{0,\Omega}+C\|\nabla\cdot\boldsymbol{y}_N\|_{0,\Omega},\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}&\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}+\|r^*(\boldsymbol{u}_N)-r_N^*\|_{0,\Omega}\\ &\le CN^{-1}\|\nu\Delta\boldsymbol{y}_N^*+\nabla r_N^*+(1+\lambda_N)\boldsymbol{y}_N-P_N\boldsymbol{y}_0\|_{0,\Omega}+C\|\nabla\cdot\boldsymbol{y}_N^*\|_{0,\Omega}.\end{aligned} \tag{3.15}$$

Proof. By (2.25) and (3.7), we have

$$\begin{aligned}&\gamma\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}^2\le a(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N)\\ &=a(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N))\\ &\quad-b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N),r(\boldsymbol{u}_N)-r_N)+b(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)\\ &=(\nu\Delta\boldsymbol{y}_N-\nabla r_N-\nu\Delta\boldsymbol{y}(\boldsymbol{u}_N)+\nabla r(\boldsymbol{u}_N),\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N))\\ &\quad-b(\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)\\ &=(\nu\triangle\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N+P_N\boldsymbol{f},\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N-P_{1,N}^0(\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N))-b(\boldsymbol{y}_N,r(\boldsymbol{u}_N)-r_N)\\ &\le CN^{-1}\|\nu\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N+P_N\boldsymbol{f}\|_{0,\Omega}\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}+C\|\nabla\cdot\boldsymbol{y}_N\|_{0,\Omega}\|r(\boldsymbol{u}_N)-r_N\|_{0,\Omega}.\end{aligned} \tag{3.16}$$

Noting that

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

It follows that

$$\begin{aligned}\beta\|r(\boldsymbol{u}_N)-r_N\|_{0,\Omega}&\le\sup_{\forall\,\boldsymbol{w}\in\boldsymbol{Y}}\frac{|\,b(\boldsymbol{w},r(\boldsymbol{u}_N)-r_N)\,|}{\|\boldsymbol{w}\|_{1,\Omega}}\\ &\le CN^{-1}\|\nu\Delta\boldsymbol{y}_N-\nabla r_N+\boldsymbol{u}_N+P_N\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{1,\Omega}.\end{aligned} \tag{3.17}$$

Similarly, we can derive from (2.25) and (3.7) that

$$\begin{aligned}\gamma\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}^2&\le CN^{-1}\|\nu\Delta\boldsymbol{y}_N^*+\nabla r_N^*+(1+\lambda_N)\boldsymbol{y}_N-P_N\boldsymbol{y}_0\|_{0,\Omega}\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{1,\Omega}\\ &\quad+C\|\nabla\cdot\boldsymbol{y}_N^*\|_{0,\Omega}\|r^*(\boldsymbol{u}_N)-r_N^*\|_{0,\Omega},\end{aligned} \tag{3.18}$$

and

$$
\begin{aligned}
\beta\|r^*(\boldsymbol{u}_N) - r_N^*\|_{0,\Omega} &\le \sup_{\forall\, \boldsymbol{q}\in \boldsymbol{Y}} \frac{|\, b(\boldsymbol{q}, r^*(\boldsymbol{u}_N) - r_N^*)\,|}{\|\boldsymbol{q}\|_{1,\Omega}} \\
&\le CN^{-1}\|\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0\|_{0,\Omega} + C\|\boldsymbol{y}^*(\boldsymbol{u}_N) - \boldsymbol{y}_N^*\|_{1,\Omega}.
\end{aligned} \tag{3.19}
$$

Then, we can complete the proof by (3.16)–(3.19) and Cauchy's inequality with ϵ. □

The main result of this subsection can now be stated as follows.

Theorem 3.1. *Let* $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ *and* $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ *be the solutions of (2.11) and (2.25) respectively. Then we have that*

$$\|e\| \le C(\eta + \theta), \tag{3.20}$$

where the total approximation error $\|e\|$ *is defined by*

$$
\begin{aligned}
\|e\| = \|\boldsymbol{y} - \boldsymbol{y}_N\|_{1,\Omega} + \|r - r_N\|_{0,\Omega} + \|\boldsymbol{y}^* - \boldsymbol{y}_N^*\|_{1,\Omega} \\
+ \|r^* - r_N^*\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + |\lambda - \lambda_N|,
\end{aligned} \tag{3.21}
$$

the estimator η *is given below*

$$
\begin{aligned}
&\eta = \eta_1 + \eta_2 + \eta_3 + \eta_4, \\
&\eta_1 = N^{-1}\|\nu\Delta\boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N\boldsymbol{f}\|_{0,\Omega}, \quad \eta_2 = \|\nabla\cdot\boldsymbol{y}_N\|_{0,\Omega}, \\
&\eta_3 = N^{-1}\|\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0\|_{0,\Omega}, \quad \eta_4 = \|\nabla\cdot\boldsymbol{y}_N^*\|_{0,\Omega},
\end{aligned} \tag{3.22}
$$

and θ *is presented as follows*

$$\theta = N^{-1}\|\boldsymbol{y}_0 - P_N\boldsymbol{y}_0\|_{0,\Omega} + N^{-1}\|\boldsymbol{f} - P_N\boldsymbol{f}\|_{0,\Omega}. \tag{3.23}$$

Proof. By (2.11e), (2.25e), and (3.9), we have

$$
\begin{aligned}
&\alpha\|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}^2 = \alpha(\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{u} - \boldsymbol{u}_N) \\
&= (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* - \boldsymbol{y}^*(\boldsymbol{u}_N)) - (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* - \boldsymbol{y}^*(\boldsymbol{u}_N) - \alpha\boldsymbol{u} + \alpha\boldsymbol{u}_N) \\
&= (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* - \boldsymbol{y}^*(\boldsymbol{u}_N)) + (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}^*(\boldsymbol{u}_N) + \alpha\boldsymbol{u}) - (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* + \alpha\boldsymbol{u}_N) \\
&= (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* - \boldsymbol{y}^*(\boldsymbol{u}_N)) + (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}^*(\boldsymbol{u}_N) - \boldsymbol{y}^*) \\
&= (\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* - \boldsymbol{y}^*(\boldsymbol{u}_N)) - (\boldsymbol{y} - \boldsymbol{y}_N + \lambda\boldsymbol{y} - \lambda_N\boldsymbol{y}_N, \boldsymbol{y} - \boldsymbol{y}(\boldsymbol{u}_N)) \\
&\quad + (\boldsymbol{y}_0 - P_N y_0, \boldsymbol{y} - \boldsymbol{y}(\boldsymbol{u}_N)) + (\boldsymbol{f} - P_N\boldsymbol{f}, \boldsymbol{y}^* - \boldsymbol{y}^*(\boldsymbol{u}_N)).
\end{aligned} \tag{3.24}
$$

It is clear that by (2.12) and (2.26)

$$\lambda(\boldsymbol{y}, \boldsymbol{y}_N - \boldsymbol{y}) \le 0, \quad \lambda_N(\boldsymbol{y}_N, \boldsymbol{y} - \boldsymbol{y}_N) \le 0, \tag{3.25}$$

which implies that

$$-(\lambda\boldsymbol{y} - \lambda_N\boldsymbol{y}_N, \boldsymbol{y} - \boldsymbol{y}_N) \le 0. \tag{3.26}$$

It follows from (3.24) and (3.26) that

$$
\begin{aligned}
\alpha\|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega}^2 \le\ &(\boldsymbol{u} - \boldsymbol{u}_N, \boldsymbol{y}_N^* - \boldsymbol{y}^*(\boldsymbol{u}_N)) - (\boldsymbol{y} - \boldsymbol{y}_N, \boldsymbol{y} - \boldsymbol{y}(\boldsymbol{u}_N)) \\
&- (\lambda\boldsymbol{y} - \lambda_N\boldsymbol{y}_N, \boldsymbol{y}_N - \boldsymbol{y}(\boldsymbol{u}_N)) + (\boldsymbol{y}_0 - P_N y_0, \boldsymbol{y} - \boldsymbol{y}(\boldsymbol{u}_N)) \\
&+ (\boldsymbol{f} - P_N\boldsymbol{f}, \boldsymbol{y}^* - \boldsymbol{y}^*(\boldsymbol{u}_N)),
\end{aligned}
$$

associating with lemma 3.1 and lemma 3.3, we have

$$\begin{aligned}
&\alpha\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2+\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}^2\\
\le\ &(\boldsymbol{u}-\boldsymbol{u}_N,\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N))+(\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N),\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))-\lambda(\boldsymbol{y}-\boldsymbol{y}_N,\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N))\\
&-(\lambda-\lambda_N)(\boldsymbol{y}_N,\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N))+(\boldsymbol{y}_0-P_N y_0,\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N))\\
&+(\boldsymbol{f}-P_N\boldsymbol{f},\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N))\\
\le\ &\epsilon(\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}^2+\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{1,\Omega}^2+\|\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{1,\Omega}^2+|\lambda-\lambda_N|)\\
&+C_1(\epsilon)(\|\boldsymbol{y}^*(\boldsymbol{u}_N)-\boldsymbol{y}_N^*\|_{0,\Omega}^2+\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{0,\Omega}^2)\\
&+C_1(\epsilon)N^{-2}(\|\boldsymbol{y}_0-P_N\boldsymbol{y}_0\|_{0,\Omega}^2+\|\boldsymbol{f}-P_N\boldsymbol{f}\|_{0,\Omega}^2).
\end{aligned}\tag{3.27}$$

It can be derived from (3.9), (3.13), and lemma 3.3 that

$$\begin{aligned}
&\|\boldsymbol{y}^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{1,\Omega}+\|r^*-r^*(\boldsymbol{u}_N)\|_{0,\Omega}\\
\le\ &C\|\boldsymbol{y}-\boldsymbol{y}_N\|_{0,\Omega}+C\|\lambda\boldsymbol{y}-\lambda_N\boldsymbol{y}_N\|_{0,\Omega}+CN^{-1}\|\boldsymbol{y}_0-P_N\boldsymbol{y}_0\|_{0,\Omega}\\
\le\ &C\|\boldsymbol{y}(\boldsymbol{u}_N)-\boldsymbol{y}_N\|_{0,\Omega}+C|\lambda-\lambda_N|+C\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}\\
&+CN^{-1}\|\boldsymbol{y}_0-P_N\boldsymbol{y}_0\|_{0,\Omega}+CN^{-1}\|\boldsymbol{f}-P_N\boldsymbol{f}\|_{0,\Omega},
\end{aligned}\tag{3.28}$$

associating with (3.13), (3.27), and lemma 3.4, it can be derived that

$$\begin{aligned}
\|\boldsymbol{u}-\boldsymbol{u}_N\|_{0,\Omega}+\|\boldsymbol{y}-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega}\le\ &C(\|\boldsymbol{y}_N^*-\boldsymbol{y}^*(\boldsymbol{u}_N)\|_{0,\Omega}+\|\boldsymbol{y}_N-\boldsymbol{y}(\boldsymbol{u}_N)\|_{0,\Omega})\\
&+CN^{-1}(\|\boldsymbol{y}_0-P_N\boldsymbol{y}_0\|_{0,\Omega}+\|\boldsymbol{f}-P_N\boldsymbol{f}\|_{0,\Omega}).
\end{aligned}\tag{3.29}$$

Then the theorem follows from lemma 3.4, lemma 3.5, (3.13), and (3.28)-(3.29). □

3.2. Lower error bound

A lower bound for the error $\|e\|$ is established to investigate the sharpness of the indicator in this subsection. We first need some polynomial inverse estimates presented in the following lemma, which can be found in [16].

Lemma 3.6. *Let $\alpha,\beta\in\mathbb{R}$ satisfy $-1<\alpha<\beta$ and $\delta\in[0,1]$. Then there exist constant $C_1, C_2=C_2(\alpha,\beta)$, and $C_3=C_3(\delta)$ such that for all $z_N\in Q_N$*

$$\int_\Omega|\nabla z_N(x)|^2\Phi_\Omega(x)dx\le C_1N^2\int_\Omega z_N^2(x)dx,\tag{3.30a}$$

$$\int_\Omega z_N^2(x)\Phi_\Omega^\alpha(x)dx\le C_2N^{2(\beta-\alpha)}\int_\Omega z_N^2(x)\Phi_\Omega^\beta(x)dx,\tag{3.30b}$$

$$\int_\Omega|\nabla z_N(x)|^2\Phi_\Omega^{2\delta}(x)dx\le C_3N^{2(2-\delta)}\int_\Omega z_N^2(x)\Phi_\Omega^\delta(x)dx,\tag{3.30c}$$

where $\Phi_\Omega(x)$ is the distance function defined by

$$\Phi_\Omega(x):=dist(x,\partial\Omega).$$

In the subsequent analysis, we establish the lower bound for discretization error $\|e\|$ by using lemma 3.6. In fact, letting $\alpha = 0, \beta = \gamma \in (\frac{1}{2}, 1]$ in (3.30b), we can derive that

$$\begin{aligned} &N^{-2} \int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 dx \\ &\leq CN^{2\gamma-2} \int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^\gamma dx. \end{aligned} \tag{3.31}$$

Denote $\boldsymbol{\vartheta} = (\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f})\Phi_\Omega^\gamma$, then we have

$$\begin{aligned} &\int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^\gamma dx \\ =& \int_\Omega (\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}) \cdot \boldsymbol{\vartheta} dx \\ =& \int_\Omega (\nu\Delta(\boldsymbol{y}_N - \boldsymbol{y}) + \nabla(r - r_N) + (\boldsymbol{u}_N - \boldsymbol{u}) + (P_N \boldsymbol{f} - \boldsymbol{f})) \cdot \boldsymbol{\vartheta} dx \\ \leq& C|\boldsymbol{\vartheta}|_{1,\Omega}(\|\boldsymbol{y} - \boldsymbol{y}_N\|_{1,\Omega} + \|r - r_N\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + N^{-1}\|\boldsymbol{f} - P_N \boldsymbol{f}\|_{0,\Omega}). \end{aligned} \tag{3.32}$$

Furthermore, it follows from (3.30b) and (3.30c) that

$$\begin{aligned} |\boldsymbol{\vartheta}|_{1,\Omega}^2 =& \int_\Omega |\nabla((\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f})\Phi_\Omega^\gamma)|^2 dx \\ \leq& C \int_\Omega |\nabla(\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f})|^2 \Phi_\Omega^{2\gamma} dx \\ &+ C \int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 |\nabla \Phi_\Omega^\gamma|^2 dx \\ \leq& CN^{2(2-\gamma)} \int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^\gamma dx \\ &+ C \int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^{2\gamma-2} dx \\ \leq& CN^{2(2-\gamma)} \int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^\gamma dx. \end{aligned}$$

Thus, we have

$$|\boldsymbol{\vartheta}|_{1,\Omega} \leq CN^{(2-\gamma)} \Big(\int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^\gamma dx \Big)^{\frac{1}{2}}. \tag{3.33}$$

It follows from (3.31)–(3.33) that

$$\begin{aligned} \eta_1 &= N^{-1}\|\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}\|_{0,\Omega} \\ &\leq CN^{\gamma-1} \Big(\int_\Omega |\nu\Delta \boldsymbol{y}_N - \nabla r_N + \boldsymbol{u}_N + P_N \boldsymbol{f}|^2 \Phi_\Omega^\gamma dx \Big)^{\frac{1}{2}} \\ &\leq CN(\|\boldsymbol{y} - \boldsymbol{y}_N\|_{1,\Omega} + \|r - r_N\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_N\|_{0,\Omega} + N^{-1}\|\boldsymbol{f} - P_N \boldsymbol{f}\|_{0,\Omega}). \end{aligned} \tag{3.34}$$

It is clear that

$$\eta_2^2 = \int_\Omega (\nabla \cdot \boldsymbol{y}_N)^2 = \int_\Omega (\nabla \cdot \boldsymbol{y}_N - \nabla \cdot \boldsymbol{y})^2 \leq C\|\boldsymbol{y} - \boldsymbol{y}_N\|_{1,\Omega}^2.$$

Then, we have

$$\eta_2 \le C\|\boldsymbol{y} - \boldsymbol{y}_N\|_{1,\Omega}. \tag{3.35}$$

Similarly, letting $\alpha = 0, \beta = \gamma \in (\frac{1}{2}, 1]$ in (3.30b), it holds that

$$\begin{aligned} &N^{-2}\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 dx \\ &\le CN^{2\gamma-2}\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 \Phi_\Omega^\gamma dx. \end{aligned} \tag{3.36}$$

Let $\boldsymbol{\vartheta}^* = (\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0)\Phi_\Omega^\gamma$ to obtain that

$$\begin{aligned} &\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 \Phi_\Omega^\gamma dx \\ &= \int_\Omega (\nu\Delta(\boldsymbol{y}_N^* - \boldsymbol{y}^*) + \nabla(r_N^* - r^*) + ((1+\lambda_N)\boldsymbol{y}_N - (1+\lambda)\boldsymbol{y}) + (\boldsymbol{y}_0 - P_N\boldsymbol{y}_0))\cdot\boldsymbol{\vartheta}^* dx \\ &\le C|\boldsymbol{\vartheta}^*|_{1,\Omega}(\|\boldsymbol{y}^* - \boldsymbol{y}_N^*\|_{1,\Omega} + \|r^* - r_N^*\|_{0,\Omega} + |\lambda - \lambda_N| \\ &\quad + \|\boldsymbol{y} - \boldsymbol{y}_N\|_{0,\Omega} + N^{-1}\|\boldsymbol{y}_0 - P_N\boldsymbol{y}_0\|_{0,\Omega}). \end{aligned} \tag{3.37}$$

Furthermore, it can be derived from (3.30b) and (3.30c) that

$$\begin{aligned} |\boldsymbol{\vartheta}^*|_{1,\Omega}^2 &\le C\int_\Omega |\nabla(\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0)|^2 \Phi_\Omega^{2\gamma} dx \\ &\quad + C\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 |\nabla\Phi_\Omega^\gamma|^2 dx \\ &\le CN^{2(2-\gamma)}\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 \Phi_\Omega^\gamma dx. \end{aligned}$$

which implies the inequality

$$|\boldsymbol{\vartheta}^*|_{1,\Omega} \le CN^{(2-\gamma)}\Big(\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 \Phi_\Omega^\gamma dx\Big)^{\frac{1}{2}}. \tag{3.38}$$

It follows from (3.36)–(3.38) that

$$\begin{aligned} \eta_3 &= N^{-1}\|\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0\|_{0,\Omega} \\ &\le CN^{\gamma-1}\Big(\int_\Omega |\nu\Delta\boldsymbol{y}_N^* + \nabla r_N^* + (1+\lambda_N)\boldsymbol{y}_N - P_N\boldsymbol{y}_0|^2 \Phi_\Omega^\gamma dx\Big)^{\frac{1}{2}} \\ &\le CN(\|\boldsymbol{y}^* - \boldsymbol{y}_N^*\|_{1,\Omega} + \|r^* - r_N^*\|_{0,\Omega} + |\lambda - \lambda_N| \\ &\quad + \|\boldsymbol{y} - \boldsymbol{y}_N\|_{0,\Omega} + N^{-1}\|\boldsymbol{y}_0 - P_N\boldsymbol{y}_0\|_{0,\Omega}). \end{aligned} \tag{3.39}$$

Similar to η_2, we have

$$\eta_4 \le C\|\boldsymbol{y}^* - \boldsymbol{y}_N^*\|_{1,\Omega}. \tag{3.40}$$

Then the estimation of lower error bound can be stated as following theorem.

Theorem 3.2. *Let $(\boldsymbol{y}, r, \boldsymbol{u}, \boldsymbol{y}^*, r^*, \lambda)$ and $(\boldsymbol{y}_N, r_N, \boldsymbol{u}_N, \boldsymbol{y}_N^*, r_N^*, \lambda_N)$ be the solutions of (2.11) and (2.25) respectively. Then we have that*

$$\frac{1}{N}\eta \le C(\|e\| + \theta), \tag{3.41}$$

where e and θ are defined in theorem 3.1.

Proof. The theorem follows from (3.34)-(3.35) and (3.39)-(3.40). □

4. Numerical experiment

In this section, we carry out a numerical example for the control problem (OCP) to investigate whether the indicator η tends to zero at the same rate as the error $\|e\|$. We are interested in the following model: find $(\boldsymbol{y}, r, \boldsymbol{u}) \in \boldsymbol{Y} \times Q \times \boldsymbol{U}$ such that

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

where $G_{ad} = \{\boldsymbol{v} \in \boldsymbol{U} \mid \| \boldsymbol{v} \|_{0,\Omega} \le d\}$. The data of this example are as follows

$$\boldsymbol{y}_0 = \Delta \boldsymbol{y}^* + \nabla r^* + (1+\lambda)\boldsymbol{y}, \quad \boldsymbol{f} = -\Delta \boldsymbol{y} + \nabla r - \boldsymbol{u},$$

and $d = \sqrt{6}\pi$. The exact solutions are given by

$$\begin{aligned} \boldsymbol{y} = \boldsymbol{y}^* &= (-\pi(1+\cos\pi x_1)\sin\pi x_2, \pi\sin\pi x_1(1+\cos\pi x_2)), \\ r = \pi^2\cos\pi x_1 \sin\pi x_2, &\quad r^* = \pi^2 \sin\pi x_1\cos\pi x_2, \quad \lambda = 0.2. \end{aligned}$$

We solve the discrete system (2.25) via Arrow-Hurwicz algorithm (see, for example, [17] and [18]).

Arrow-Hurwicz Algorithm: we describe the main steps of the algorithm as follows.

- *Step 1:* Let $k = 0$, choose a step size $\rho > 0$, give the initial values λ_N^0 and $\boldsymbol{u}_N^0$.
- *Step 2:* Let $l = 0$ and $\boldsymbol{u}_N^{k,0} = \boldsymbol{u}_N^k$.
- *Step 3:* Solve the state equations

$$\begin{aligned} a(\boldsymbol{y}_N^{k,l}, \boldsymbol{w}_N) - b(\boldsymbol{w}_N, r_N^{k,l}) = (\boldsymbol{u}_N^{k,l} + \boldsymbol{f}, \boldsymbol{w}_N) & \quad \forall\, \boldsymbol{w}_N \in \boldsymbol{Y}_N, \\ b(\boldsymbol{y}_h^{k,l}, \phi_N) = 0 & \quad \forall\, \phi_N \in M_{N-2}, \end{aligned}$$

and the co-state equations

$$\begin{aligned} a(\boldsymbol{y}_N^{*k,l}, \boldsymbol{w}_h) + b(r_N^{*k,l}, \boldsymbol{w}_N) = ((1+\lambda_N^k)\boldsymbol{y}_N^{k,l} - \boldsymbol{y}_0, \boldsymbol{w}_N) & \quad \forall\, \boldsymbol{w}_N \in \boldsymbol{Y}_N, \\ b(\boldsymbol{y}_N^{*k,l}, \psi_N) = 0, & \quad \forall\, \psi_N \in M_{N-2}. \end{aligned}$$

- *Step 4:* Let $\boldsymbol{u}_N^{k,l+1} = \boldsymbol{u}_N^{k,l} - P_N(\boldsymbol{y}_N^{*k,l} + \alpha \boldsymbol{u}_N^{k,l})$. 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*.

- *Step 5:* $\lambda_N^{k+1} = max\{0, \lambda_N^k + \rho(\|\boldsymbol{y}_N^{k,l}\|_{0,\Omega} - d)\}$.

Table 1. The values of discretization errors, indicator η, and θ.

N	4	8	12	16
$\|\|\boldsymbol{u}-\boldsymbol{u}_N\|\|_{0,\Omega}$	1.34228	5.24377e-3	2.92731e-6	1.46001e-7
$\|\|\boldsymbol{y}-\boldsymbol{y}_N\|\|_{1,\Omega}$	8.47190	6.22992e-2	5.06155e-5	9.72122e-7
$\|\|r-r_N\|\|_{0,I}$	7.29581	6.14368e-2	5.16530e-5	2.38419e-7
$\|\|\boldsymbol{y}^*-\boldsymbol{y}_N^*\|\|_{1,\Omega}$	8.49494	6.23009e-2	5.06195e-5	1.04945e-6
$\|\|r^*-r_N^*\|\|_{0,\Omega}$	7.29581	6.14371e-2	5.16530e-5	2.92002e-7
$\|\lambda-\lambda_N\|$	0.20000	1.15717e-4	1.56612e-10	2.92860e-12
$\|\|e\|\|$	33.1007	2.52833e-1	2.07468e-4	2.69799e-6
η	42.9463	3.18312e-1	2.40411e-4	1.49619e-6
θ	5.55155	6.44712e-3	1.58181e-6	2.79281e-7

- *Step 6:* 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*.

Denote $\boldsymbol{y}_N=(y_{N1},y_{N2})$, $\boldsymbol{y}_N^*=(y_{N1}^*,y_{N2}^*)$, and $\boldsymbol{u}_N=(u_{N1},u_{N2})$, we have the following expressions by the property of (2.7)

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

$$r_N=\sum_{i,j=0}^{N-2}r_{ij}L_i(x_1)L_j(x_2),\quad r_N^*=\sum_{i,j=0}^{N-2}r_{ij}^*L_i(x_1)L_j(x_2),\quad m=1,2,\quad r_{00}=0,\quad r_{00}^*=0.$$

The numerical results are presented in Table 1. The table shows that the errors decrease rapidly with a relatively small number of unknowns, which is important in a number of applications. We further plot the error $||e||$ and the error indicator η versus the ploynomial degree N in Figure 1, where the longitudinal axis is in logarithmic scale. It can be observed from the figure that the two curves are very close to each other, which implies that the indicator is nearly equivalent to the error $||e||$. Moreover, it seems that the error $||e||$ and the indicator η decay with the rate $\frac{10^4}{10^{0.625N}}$, which shows that the proposed method for the control problem is very efficient and the spectral accuracy is achieved. Compared with the case of finite element method, it can be seen that the indicator developed in this work can be implemented more simply and can provide successful estimation for the errors with less computational load, which is helpful for developing the *hp* adaptive spectral element method for the optimal control problems.

5. Conclusions

In this paper, the upper and lower bounds of approximation error are provided with the help of the a posteriori error indicator. The illustrative numerical experiment shows the performance of the error estimator. In our future work, we hope to extend these results to adaptive method in the *hp* spectral element framework, which will be compared with the adaptive *hp* finite element method for optimal control problems.

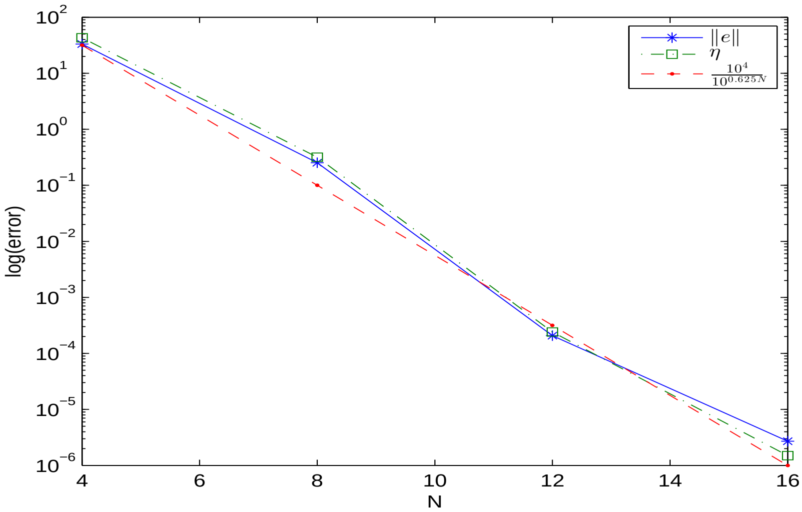

Figure 1. The total error $||e||$ and error estimator η.

Acknowledgments

The authors are grateful to the referees for their useful comments and suggestions, which lead to improvements of the presentation. The authors were supported by the State Key Program of National Natural Science Foundation of China (11931003), the National Natural Science Foundation of China (Grant No. 11601466 and 41974133), the Nanhu Scholars Program for Young Scholars of XYNU, the Postgraduate Education Reform and Quality Improvement Project of Henan Province (Grant No. YJS2022ZX34), the National Natural Science Foundation of Henan (Grant No. 202300410343), the Training Plan of Young Key Teachers in Universities of Henan (Grant No. 2018GGJS096), and the Scientific Research Fund Project for Young Scholars of XYNU (Grant No. 2020-QN-050).

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. E. Casas, Control of an elliptic problem with pointwise state constraints, *SIAM J. Control Optim.*, **24** (1986), 1309–1318. https://doi.org/10.1137/0324078
2. E. Casas, Boundary control of semilinear elliptic equations with pointwise state constraints, *SIAM J. Control Optim.*, **31** (1993), 993–1006 . https://doi.org/10.1137/0331044
3. F. Tröltzsch, *Optimal Control of Partial Differential Equations. Theory, Methods, and Applications*, Grad. Stud. Math. 112, American Mathematical Society, Providence, RI, 2010.
4. O. Benedix, B. Vexler, A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints, *Comput. Optim. Appl.*, **44** (2009), 3–25. https://doi.org/10.1007/s10589-008-9200-y

5. W. Liu, D. Yang, L. Yuan, C. Ma, Finite elemtnt approximations of an optimal control problem with integral state constraint, *SIAM J. Numer. Anal.*, **48** (2010), 1163–1185. https://doi.org/10.1137/080737095
6. M. Bergounioux, K. Kunisch, Primal-dual strategy for state-constrained optimal control problems, *Comput. Optim. Appl.*, **22** (2002), 193–224. https://doi.org/10.1023/A:1015489608037
7. J. De los Reyes, K. Kunisch, A semi-smooth newton method for regularized state-constrained optimal control of the Navier-Stokes equations, *Computing*, **78** (2006), 287–309. https://doi.org/10.1007/s00607-006-0183-1
8. W. Gong, N. Yan, A mixed finite element scheme for optimal control problems with pointwise state constraints, *J. Sci. Comput.*, **46** (2011), 182–203. https://doi.org/10.1007/s10915-010-9392-z
9. J. Zhou, D. Yang, Spectral mixed Galerkin method for state constrained optimal control problem governed by the first bi-harmonic equation, *Int. J. Comput. Math.*, **88** (2011), 2988–3011. https://doi.org/10.1080/00207160.2011.563845
10. Y. Chen, N. Yi, W.B. Liu, A Legendre Galerkin spectral method for optimal control problems governed by elliptic equations, *SIAM J. Numer. Anal.*, **46** (2008), 2254–2275. https://doi.org/10.1137/070679703
11. Y. Chen, F. Huang, N. YI , W. B. Liu, A Legendre-Galerkin spectral method for optimal control problems governed by Stokes equations, *SIAM J. Numer. Anal.*, **49** (2011), 1625–1648. https://doi.org/10.1137/080726057
12. C. Bernardi, Y. Maday, Uniform inf-sup conditions for the spectral discretization of the Stokes problem, *Math. Models Methods Appl. Sci.*, **9** (1999), 395–414. https://doi.org/10.1142/S0218202599000208
13. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.
14. J. Shen, Efficient Spectral-Galerkin Method I: Direct Solvers for the Second and Fourth Order Equations Using Legendre Polynomials, *SIAM J. Sci. Comput.*, **15** (1994), 1489–1505. https://doi.org/10.1137/0915089
15. C. Canuto, M. Hussaini, A. Quarteroni, T. Zang, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, Berlin, 1988.
16. J. Melenk, B. Wohlmuth, On residual-based a posteriori error estimation in hp-FEM, *Adv. Comput. Math.*, **15** (2001), 311–331. https://doi.org/10.1023/A:1014268310921
17. R. Glowinski, J. Lions, R. Trémolières, *Numerical analysis of variational inequalities*, Amsterdam, North-Holland, 1981.
18. H. Niu and D. Yang, Finite element analysis of optimal control problem governed by Stokes equations with L^2-norm state-constriants, *J. Comput. Math.*, **29** (2011), 589–604. https://www.jstor.org/stable/43693666

AIMS Press