Journal of Industrial and Management Optimization

https://www.aimspress.com/journal/jimo

JIMO, 22(2): 832–859.
DOI:10.3934/jimo.2026030
Received: 25 September 2025
Revised: 10 December 2025
Accepted: 18 December 2025
Published: 05 January 2026

Research article

Error estimates and superconvergence of mixed covolume approximations for elliptic optimal control problems

Chunjuan Hou[1], Yanping Chen[2,*], Jian Huang[1], Jiawang Liu[1] and Fangfang Qin[2]

[1] National Center for Applied Mathematics in Hunan, Hunan Key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, China

[2] School of Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

* **Correspondence:** Email: ypchen@njupt.edu.cn.

Abstract: In this scholarly article, we analyze the results of error estimates and phenomena of superconvergence associated with the mixed covolume approximation method, which is applied to a particular class of linear elliptic optimal control problems. The control variable is discretized using piecewise constant functions. Additionally, the state and the costate variables are both approximated using the lowest-order Raviart–Thomas (RT_0) mixed finite element method. First, mixed covolume approximation of optimal control problems is constructed. Second, "a priori error estimations" for each variable are computed. Third, a superconvergence result is established, and it is proved that there exists a second-order superconvergence relationship between the centroid interpolation of variable u and its numerical solution. Finally, two carefully designed numerical examples are presented to validate the reliability of the theoretical findings, providing concrete evidence to corroborate the above results and strengthen the coherence of the study's conclusions.

Keywords: mixed covolume method; superconvergence; elliptic equations; optimal control problems; error estimates
Mathematics Subject Classification: 49J20, 65N30

1. Introduction

Partial Differential Equation (PDE)-associated optimal control [1–4] aims to minimize or maximize the objective functional by adjusting control variables in partial differential equations. Extensive research has been conducted on the superconvergence characteristics of finite element approximations and optimal control problems. For the conventional finite element method, relevant findings can be found in references [5–8], while studies on the mixed finite element method are documented in [9–13]. To the best of our knowledge, two types of superconvergence properties related to the control variable exist. One type is characterized by superconvergence between the mean L^2-projection and the approximation process of the control variable, as explored in [5, 7, 11, 12]. The other type involves integration and superconvergence between centroid interpolation techniques and optimal control theory, which is discussed in [6, 9, 10, 13, 14]. In their research work [5, 11], Chen and her co-authors investigated the superconvergence property between the control variable's approximation and the mean L^2-projection. Beyond this, they derived specific superconvergence results by postprocessing techniques. Notably, the superconvergence order of the control variable is limited to $h^{\frac{3}{2}}$ primarily due to the control variable's own low regularity features. In the seminal work [6], Meyer and Rösch explored the finite element approximation of an elliptic optimal control problem with pointwise control constraints. Their investigation developed a novel approach, rigorously demonstrating that the point values of the control within the finite elements are approximated with a high order h^2. In the seminal work [14], Rösch and Vexler extended the pioneering superconvergence approach originally introduced by Meyer and Rösch in [6]. They demonstrated superconvergence results which overcame previous limitations of the technique in [6], which could only accommodate piecewise linear finite elements over triangular meshes. In [9, 10], Chen explored the application of both rectangular and triangular Raviart–Thomas (RT) mixed finite element methods (MFEMs) elliptic optimal control problems. She derived a similar superconvergence result for the control variable, thereby extending the understanding and efficacy of these numerical techniques for optimal control problems. Hou, Leng, and Luan [13] presented a P_0^2-P_1 MFEM specifically tailored to tackle distributed optimal control problems governed by general elliptic equations. They demonstrated that the centroid interpolation technique, paired with the numerical solution of the control variable u, exhibits a superconvergence property, achieving a convergence rate of h^2. Building on this result, they devised an innovative two-grid algorithm to enhance computational efficiency and accuracy for optimal control problems. In recent times, numerous methods have been successfully formulated for solving a broad range of PDEs, such as [15–18].

The mixed covolume method, originally introduced by Russell [19], represents a significant advancement in numerical techniques for solving partial differential equations. The fundamental approach of this method involves establishing a connection between the Petrov–Galerkin scheme and the more conventional finite element Galerkin or mixed methods. This connection is achieved by introducing a crucial transfer operator, denoted as γ_h, which serves the purpose of mapping the trial function space onto the test function space. This innovative technique not only retains the inherent simplicity and ease of implementation associated with finite difference methods but also leverages the high accuracy and precision of finite element methods. Additionally, one of the standout features of the mixed covolume method is its ability to uphold the mass conservation law, which is a critical property in many physical and engineering applications. Despite its promising attributes and potential applications, to the best of our knowledge, the mixed covolume method has not yet been explored or utilized in the context of solving optimal control problems. This gap in application presents an intriguing opportunity for further research and development in the field of numerical methods for optimal control.

The primary objective of this paper is to systematically present and elaborate on the mixed covolume approximation method applied to the elliptic control problem, specifically denoted by equations (1.1)–(1.4). Additionally, we aim to thoroughly discuss and analyze the error estimates of all variables and the phenomenon of superconvergence associated with the optimal control solution. In this context, we focus our attention on a particular class of linear elliptic optimal control problems, which involve the state variables $\boldsymbol{p}$ and y as well as the control variable u. By addressing these aspects, we seek to provide a comprehensive understanding of the numerical techniques, their error estimates, and the convergence properties in the realm of elliptic control problems.

We should consider preserving some important physical properties and physical structures with particular conditions [20, 21]. In the fields of science and engineering, distributed optimal control problems are widely applied in the regulation of system states. The problem (1.1) focused on in this paper is a typical distributed optimal control model with control constraints: its goal is to minimize the weighted sum of the state deviation cost and the control cost under the premise that the control variable u satisfies the non-negativity constraint $u(x) \geq 0$, which corresponds to the physical limitation in practical scenarios that "the control quantity can not be negative", such as the non-negativity of flow rate and energy input.

We consider the following model:

$$\min_{u\in U_{ad}} \left\{ \frac{1}{2}\|\boldsymbol{p}(\cdot) - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y(\cdot) - y_d\|^2 + \frac{\theta}{2}\|u(\cdot)\|^2 \right\} \tag{1.1}$$

under the constraint of the state equation

$$-\mathrm{div}(A(x)\nabla y(x) + c(x)y(x) = f(x) + u(x), \quad x = (x_1, x_2) \in \Omega, \tag{1.2}$$

which admits expression as a first-order system

$$\mathrm{div}\boldsymbol{p}(x) + c(x)y(x) = f(x) + u(x), \quad \boldsymbol{p}(x) = -A(x)\nabla y(x), \quad x \in \Omega \tag{1.3}$$

in conjunction with the boundary condition

$$\frac{\partial y(x)}{\partial \boldsymbol{n}} = 0, \quad x \in \partial\Omega. \tag{1.4}$$

In this context, Ω represents a rectangle within $\mathbb{R}^2$, and $\boldsymbol{n}$ represents the unit outward normal vector of the regional boundary. Let U_{ad} stand for the set of admissible control variables, specified by

$$U_{ad} = \left\{ u(x) : u(x) \in L^2(\Omega) \ and \ u(x) \geq 0 \right\}.$$

We assume that $c(\cdot) \in W^{1,\infty}(\Omega)$, $c(\cdot) \geq k_0 > 0$, $\boldsymbol{p}_d \in (H^2(\Omega))^2$, $y_d \in H^1(\Omega)$, where θ stands for a fixed positive constant, the coefficient $A(x) = (a_{ii}(x))_{i=1,2}$ takes the form of a diagonal matrix function having $a_{ii}(x) \in W^{1,\infty}(\Omega)$, which fulfills the following ellipticity condition:

$$d_*|\rho|^2 \leq \sum_{i,j=1}^{2} a_{ij}(x)\rho_i\rho_j, \forall\ (\rho, x) \in \mathbb{R}^2 \times \bar{\Omega}, d_* > 0.$$

The structure and organization of this paper are outlined as follows. In the second section, we introduce and elaborate on the mixed covolume approximation scheme specifically designed for addressing the optimal control problem delineated by equations (1.1)–(1.4). Additionally, we derive and present the corresponding equivalent optimality conditions that are essential for understanding the underlying theoretical framework. In the third section, we present the comprehensive results derived from the a priori error estimates, which have been calculated and analyzed for every variable involved in the research. In the fourth section, we undertake a detailed superconvergence analysis, focusing particularly on the behavior and properties of the control variable. This analysis aims to provide deeper insights into the convergence characteristics and accuracy of the proposed method. In the fifth section, we illustrate the practical applicability and validate the theoretical findings by presenting two carefully chosen numerical examples. The examples serve to demonstrate the efficacy of the theoretical results discussed in the preceding sections. Finally, in the concluding section, we summarize the key outcomes and contributions of this research. Furthermore, we outline potential avenues for future research and extensions, highlighting areas where further investigation could lead to advancements in the field.

In the present work, the standard notation $W^{m,p}(\Omega)$ represents the Sobolev spaces defined on the domain Ω, where the norm $\|\cdot\|_{m,p}$ is explicitly given by the formula $\|\omega\|_{m,p}^p = \sum_{|\alpha|\le m} \|D^\alpha \omega\|_{L^p(\Omega)}^p$; this norm encapsulates the sum of the L^p norms of all partial derivatives of the function ω up to order m. Additionally, we introduce a semi-norm $|\cdot|_{m,p}$, which is defined by $|\omega|_{m,p}^p = \sum_{|\alpha|=m} \|D^\alpha \omega\|_{L^p(\Omega)}^p$, focusing solely on the derivatives of order exactly m. We also define the subspace $W_0^{m,p}(\Omega)$ as the set of all functions ω in $W^{m,p}(\Omega)$ that vanish on the boundary of Ω, mathematically expressed as $W_0^{m,p}(\Omega) = \{\omega \in W^{m,p}(\Omega) : \omega|_{\partial\Omega} = 0\}$. For the special case when $p = 2$, we denote the Sobolev space $W^{m,2}(\Omega)$ by $H^m(\Omega)$ and its subspace $W_0^{m,2}(\Omega)$ by $H_0^m(\Omega)$. Correspondingly, the norms $\|\cdot\|_{m,2}$ and $\|\cdot\|_{0,2}$ are simplified to $\|\cdot\|_m$ and $\|\cdot\|$, respectively, for ease of notation. Furthermore, we use the symbol C to denote a general positive constant that remains independent of the spatial mesh size h.

2. Mixed covolume approximations

This section will see a mixed covolume approximation scheme for the problem that Eqs (1.1)–(1.4) constructed. Let

$$H(\mathrm{div};\Omega) = \{\boldsymbol{p} \in (L^2(\Omega))^2 : \mathrm{div}\boldsymbol{p} \in L^2(\Omega)\},\ \boldsymbol{V} = \{\boldsymbol{p} \in H(\mathrm{div};\Omega) : \boldsymbol{p}\cdot\boldsymbol{n} = 0 \text{ on } \partial\Omega\},\ W = L^2(\Omega).$$

Assume that $\boldsymbol{p}_d \in \boldsymbol{V}$. Then, we recast the system of equations given by (1.1)–(1.4) into the following weak formulation with the objective to determine a triplet $(\boldsymbol{p}, y, u)$ that resides within the product space $\boldsymbol{V} \times W \times U_{ad}$ satisfying specific conditions that are derived from the original set of equations, thereby transforming them into a more tractable, weak form for further analysis and solution. For $\forall\, \boldsymbol{v} \in \boldsymbol{V},\ \forall\, w \in W$, we get

$$\min_{u\in U_{ad}} \left\{ \frac{1}{2}\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y - y_d\|^2 + \frac{\theta}{2}\|u\|^2 \right\} \tag{2.1}$$

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - y, \mathrm{div}\boldsymbol{v}) = 0, \tag{2.2}$$

$$(\mathrm{div}\boldsymbol{p}, w) + (cy, w) = (f + u, w), \tag{2.3}$$

in which $(\cdot,\cdot)$ represents the inner product defined on $L^2(\Omega)$ or $(L^2(\Omega))^2$.

According to the findings presented in [1], it can be deduced that the optimal control problem defined by Eqs (2.1)–(2.3) possesses a unique solution, denoted as the triplet $(\boldsymbol{p}, y, u)$. Furthermore, it is established that a given triplet $(\boldsymbol{p}, y, u)$ qualifies as the solution to the system of Eqs (2.1)–(2.3) under the stringent condition that there exists a corresponding co-state pair $(\boldsymbol{q}, z)$, where $(\boldsymbol{q}, z)$ belongs to the product space $\boldsymbol{V} \times W$. Only when this co-state pair $(\boldsymbol{q}, z)$ is present does the $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ fulfill the necessary optimality conditions:

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \mathrm{div}\boldsymbol{v}) = 0,\ \forall\, \boldsymbol{v} \in \boldsymbol{V}, \tag{2.4}$$

$$(\mathrm{div}\boldsymbol{p}, w) + (cy, w) = (f + u, w),\ \forall\, w \in W, \tag{2.5}$$

$$(A^{-1}\boldsymbol{q}, \boldsymbol{v}) - (z, \mathrm{div}\boldsymbol{v}) = -(\boldsymbol{p} - \boldsymbol{p}_d, \boldsymbol{v}),\ \forall\, \boldsymbol{v} \in \boldsymbol{V}, \tag{2.6}$$

$$(\mathrm{div}\boldsymbol{q}, w) + (cz, w) = (y - y_d, w),\ \forall\, w \in W, \tag{2.7}$$

$$(\theta u + z, \tilde{u} - u) \geq 0,\ \forall\, \tilde{u} \in U_{ad}. \tag{2.8}$$

The inequality (2.8) is capable of being written as

$$u(x) = \max\{0, -z(x)\}/\theta. \tag{2.9}$$

Let $\mathcal{T}_h = \bigcup K_B$ represent a uniform triangulation of the domain Ω, where each K_B denotes a specific triangle whose barycenter (geometric center) is located at point B (refer to Figure 1 for a visual illustration). In this context, the parameter h is defined as the maximum value among the diameters h_{K_B} of all the triangles K_B within the triangulation, effectively representing the largest triangle size in the mesh. We further introduce a notation for the nodes: $P_1, P_2, \cdots, P_{N_s}$ are designated as those nodes that are situated within the interior of the domain Ω, while $P_{N_s+1}, P_{N_s+2}, \cdots, P_N$ are identified as the nodes that lie on the boundary of the domain Ω. This distinction between interior and boundary nodes is crucial for the subsequent analysis and computational procedures involving the triangulated domain.

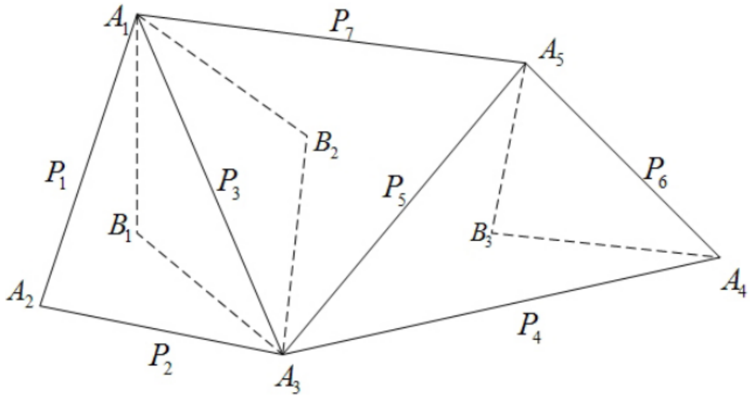

Figure 1. Primal and dual domains.

Next, we construct the dual partition $\mathcal{T}_h^*$. The dual grid fundamentally represents the integration of interior quadrilaterals, which are four-sided polygons located within the core area of the grid, and border triangles, which are three-sided polygons situated along the outer edges or boundaries of the grid. Referring to Figure 1 for a more clearer understanding, the interior node P_3 is precisely located on the shared side of the two triangles, specifically K_{B_1}, which is represented by the triangle $\Delta A_1A_2A_3$, and K_{B_2}, which is represented by the triangle $\Delta A_1A_3A_5$. This positioning indicates that the quadrilateral $A_1B_2A_3B_1$ serves as the dual element associated with the node P_3. In a similar vein, for a boundary node such as P_6, the corresponding dual element takes the form of a triangle, specifically $\Delta A_5B_3A_4$ in this particular instance. More generally, let us consider K_P^*, represented by the dashed quadrilateral in Figure 1, which is a dual element. This dual element is part of the union of two primal elements, namely K_L (which corresponds to the triangle $\Delta A_1A_2A_3$ as depicted in Figure 1) and K_R (which corresponds to the triangle $\Delta A_1A_3A_5$). This framework helps in understanding the intricate relationships between the primal and dual elements within the grid structure.

Specify the function space $\overline{\boldsymbol{Y}}_h = \{(p_{1h}, p_{2h}) : p_{1h}, p_{2h} \textit{ is constant on } K_P^* \cap K_i,\ i = R, L\}$ and the transfer operator, denoted as γ_h, functions as a mapping mechanism that transitions elements from the finite-dimensional space $\boldsymbol{V}_h$ to the dual space $\overline{\boldsymbol{Y}}_h$: $\gamma_h : \boldsymbol{V}_h \to \overline{\boldsymbol{Y}}_h$

$$\gamma_h \boldsymbol{v}_h = \sum_{m=1}^{N_s} \left(\boldsymbol{v}_h|_{K_L}(P_m)\chi^*_{K_m^* \cap K_L} + \boldsymbol{v}_h|_{K_R}(P_m)\chi^*_{K_m^* \cap K_R} \right),$$

where χ_Q^* specifically denotes the characteristic function associated with the set Q. This function serves as a mathematical indicator that uniquely identifies and characterizes the elements within the set Q, distinguishing them from those outside of it.

Based on the partition $\mathcal{T}_h$, which serves as the foundational discretization framework, we employ the lowest-order Raviart-Thomas (RT_0) mixed finite element space, denoted as $\boldsymbol{V}_h \times W_h$, as our trial function space. This choice is supported by the work presented in [23], which provides a comprehensive theoretical and practical basis for the utilization of such mixed spaces in numerical simulations.

$$\begin{aligned} \boldsymbol{V}_h &:= \{\boldsymbol{v}_h \in \boldsymbol{V} : \forall\, K \in \mathcal{T}_h,\ \boldsymbol{v}_h|_K = (b_1 + b_0x_1, b_2 + b_0x_2)\}, \\ W_h &:= \{w_h \in W : \forall\, K \in \mathcal{T}_h,\ w_h|_K \textit{ is constant}\}. \end{aligned}$$

Moreover, the approximated space associated with the control is defined by

$$U_h := W_h \cap U_{ad}.$$

Prior to presenting the mixed finite element scheme, it is essential to introduce several key operators that will be utilized in the subsequent discussions. To begin with, we define the standard $L^2(\Omega)$-projection , as referenced in [23], denoted by P_h: $W \to W_h$, for any given function ϱ that resides within the space W, the projection P_h ensures a particular relationship holds true:

$$(P_h\varrho - \varrho, w_h) = 0, \quad \forall\, w_h \in W_h, \tag{2.10}$$

$$\|\varrho - P_h\varrho\|_{0,\rho} \le Ch\|\varrho\|_{1,\rho}, \quad 2 \le \rho \le \infty, \ \forall\, \varrho \in W^{1,\rho}(\Omega). \tag{2.11}$$

Secondly, recall the Fortin projection, as detailed in the references [23] and [24], denoted by Λ_h : $\boldsymbol{V} \to \boldsymbol{V}_h$. For any vector $\boldsymbol{\zeta}$ that belongs to the vector space $\boldsymbol{V}$, the following condition is satisfied:

$$(\mathrm{div}(\Lambda_h\boldsymbol{\zeta} - \boldsymbol{\zeta}), w_h) = 0, \quad \forall\, w_h \in W_h, \tag{2.12}$$

$$\|\boldsymbol{\zeta} - \Lambda_h\boldsymbol{\zeta}\| \le Ch\|\boldsymbol{\zeta}\|_1, \quad \forall\, \boldsymbol{\zeta} \in (H^1(\Omega))^2, \tag{2.13}$$

$$\|\mathrm{div}(\boldsymbol{\zeta} - \Lambda_h\boldsymbol{\zeta})\| \le Ch\|\mathrm{div}\boldsymbol{\zeta}\|_1, \quad \forall\, \mathrm{div}\boldsymbol{\zeta} \in H^1(\Omega). \tag{2.14}$$

Regarding any continuous function $g(x)$ that belongs to the space of continuous functions $C(\Omega)$, if $x \in K_B$, we proceed to define a corresponding interpolation function denoted as g_I:

$$g_I(x) = g(B).$$

In [6], the authors have proved the estimation below associated with the numerical integration.

Lemma 2.1. *Suppose g is a function that belongs to the space $H^2(K_{B_i})$. For every index i in the given set, we can assert the following:*

$$\left|\int_{K_{B_i}} (g(x) - g(B_i))\mathrm{d}x\right| \le Ch^2 \sqrt{|K_{B_i}|} \cdot |g|_{H^2(K_{B_i})}$$

and

$$\sum_i \left|\int_{K_{B_i}} (g(x) - g(B_i))\mathrm{d}x\right| \le Ch^2 \left(\sum_i |g|^2_{H^2(K_{B_i})}\right)^{1/2},$$

where $|K_{B_i}|$ =meas(K_{B_i}).

We make the assumption that the inverse matrix A^{-1} possesses the property of Lipschitz continuity, meaning that it adheres to the Lipschitz condition, namely

$$\left|A^{-1} - \overline{A^{-1}}\right| \le Ch, \tag{2.15}$$

where $A^{-1} = diag\{1/a_{11}, 1/a_{22}\}$ and $\overline{A^{-1}} = diag\{(1/a_{11})_I, (1/a_{22})_I\}$.

Let $|||\boldsymbol{v}_h|||_h^2 := (\boldsymbol{v}_h, \gamma_h \boldsymbol{v}_h)$, $\forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h$. Then the mixed covolume approximation for the system of equations represented by (2.1)–(2.3) can be formulated as follows: the objective is to determine a triplet $(\boldsymbol{p}_h, y_h, u_h)$, where each component belongs to the respective discrete function spaces $\boldsymbol{V}_h$, W_h, and U_h, so that

$$\min_{u_h \in U_h} \left\{\frac{1}{2}|||\boldsymbol{p}_h - \Lambda_h \boldsymbol{p}_d|||_h^2 + \frac{1}{2}\|y_h - y_d\|^2 + \frac{\theta}{2}\|u_h\|^2\right\} \tag{2.16}$$

$$(\overline{A^{-1}}\boldsymbol{p}_h, \gamma_h \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0,\ \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.17}$$

$$(\mathrm{div}\boldsymbol{p}_h, w_h) + (cy_h, w_h) = (f + u_h, w_h),\ \forall\ w_h \in W_h. \tag{2.18}$$

Once more, there is a unique solution $(\boldsymbol{p}_h, y_h, u_h)$ for the optimal control problem as delineated by the equations (2.16)–(2.18), and a triplet $(\boldsymbol{p}_h, y_h, u_h)$ constitutes the solution of (2.16)–(2.18) if and only if a co-state $(\boldsymbol{q}_h, z_h) \in \boldsymbol{V}_h \times W_h$ exists where $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ fulfills the following optimality conditions $(\forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ \forall\ w_h \in W_h)$:

$$(\overline{A^{-1}}\boldsymbol{p}_h, \gamma_h \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \tag{2.19}$$

$$(\mathrm{div}\boldsymbol{p}_h, w_h) + (cy_h, w_h) = (f + u_h, w_h), \tag{2.20}$$

$$(\overline{A^{-1}}\boldsymbol{q}_h, \gamma_h \boldsymbol{v}_h) - (z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h - \Lambda_h \boldsymbol{p}_d, \gamma_h \boldsymbol{v}_h), \tag{2.21}$$

$$(\mathrm{div}\boldsymbol{q}_h, w_h) + (cz_h, w_h) = (y_h - y_d, w_h), \tag{2.22}$$

$$(\theta u_h + z_h, \tilde{u}_h - u_h) \ge 0,\ \forall\ \tilde{u}_h \in U_h. \tag{2.23}$$

In a manner analogous to the formulation presented in Eq (2.14), the control inequality delineated by Eq (2.23) can similarly be represented in an equivalent form.

$$u_h = \max\{0, -z_h\}/\theta.$$

We explicitly define the discrete solution, which is represented as $(\boldsymbol{p}_h(\tilde{u}), y_h(\tilde{u}), \boldsymbol{q}_h(\tilde{u}), z_h(\tilde{u}))$, and this solution resides within the product space $(\boldsymbol{V}_h \times W_h)^2$. This particular solution is intricately associated with the parameter $\tilde{u}$, for $\forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h$, $\forall\ w_h \in W_h$; it is designed to satisfy a set of predefined equations:

$$(\overline{A^{-1}}\boldsymbol{p}_h(\tilde{u}), \gamma_h\boldsymbol{v}_h) - (y_h(\tilde{u}), \mathrm{div}\boldsymbol{v}_h) = 0, \tag{2.24}$$

$$(\mathrm{div}\boldsymbol{p}_h(\tilde{u}), w_h) + (cy_h(\tilde{u}), w_h) = (f + \tilde{u}, w_h), \tag{2.25}$$

$$(\overline{A^{-1}}\boldsymbol{q}_h(\tilde{u}), \gamma_h\boldsymbol{v}_h) - (z_h(\tilde{u}), \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(\tilde{u}) - \Lambda_h\boldsymbol{p}_d, \gamma_h\boldsymbol{v}_h), \tag{2.26}$$

$$(\mathrm{div}\boldsymbol{q}_h(\tilde{u}), w_h) + (cz_h(\tilde{u}), w_h) = (y_h(\tilde{u}) - y_d, w_h). \tag{2.27}$$

Thus, as we have clearly defined in our previous discussions, the exact solution, along with its closely aligned approximation, can be expressed in the following manner:

$$(\boldsymbol{p}, y, \boldsymbol{q}, z) = (\boldsymbol{p}(u), y(u), \boldsymbol{q}(u), z(u)),$$
$$(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h) = (\boldsymbol{p}_h(u_h), y_h(u_h), \boldsymbol{q}_h(u_h), z_h(u_h)).$$

From the reference cited as [25], we obtain the following results of Lemma 2.2–Lemma 2.6, which provide a solid foundation for our subsequent discussions and conclusions.

Lemma 2.2. *Three positive constants K_0, K_1, and K_2 (which are independent of h) exist such that*

$$(\boldsymbol{v}_h, \gamma_h\boldsymbol{v}_h) \geq K_0\|\boldsymbol{v}_h\|^2,\ \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.28}$$

$$(A^{-1}\boldsymbol{v}_h, \gamma_h\boldsymbol{v}_h) \geq K_1\|\boldsymbol{v}_h\|^2,\ \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.29}$$

$$(\overline{A^{-1}}\boldsymbol{v}_h, \gamma_h\boldsymbol{v}_h) \geq K_2\|\boldsymbol{v}_h\|^2,\ \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{2.30}$$

Lemma 2.3. *For any $\boldsymbol{p} \in (H^1(\Omega))^2$ and $\boldsymbol{v}_h \in \boldsymbol{V}_h$, we have*

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}_h - \gamma_h\boldsymbol{v}_h) \leq Ch\|\boldsymbol{p}\|_1\|\boldsymbol{v}_h\|. \tag{2.31}$$

Lemma 2.4. *For any vectors $\boldsymbol{v}_h$ and $\boldsymbol{p}_h$ that belong to the finite-dimensional subspace $\boldsymbol{V}_h$, we can get the following relationship:*

$$\|\gamma_h\boldsymbol{v}_h\| \leq C\|\boldsymbol{v}_h\|, \tag{2.32}$$

$$(\overline{A^{-1}}\boldsymbol{p}_h, \gamma_h\boldsymbol{v}_h) = (\overline{A^{-1}}\boldsymbol{v}_h, \gamma_h\boldsymbol{p}_h). \tag{2.33}$$

Lemma 2.5. *A positive constant C with no dependence on h exists such that*

$$\|\boldsymbol{p} - \gamma_h\Lambda_h\boldsymbol{p}\| \leq Ch\|\boldsymbol{p}\|_1,\ \forall\ \boldsymbol{p} \in (H^1(\Omega))^2. \tag{2.34}$$

Lemma 2.6. *Suppose that $\boldsymbol{p}$ belongs to the Sobolev space $(H^1(\Omega))^2$, then, for an arbitrary $\boldsymbol{v}$ that also resides in the same space $(H^1(\Omega))^2$, we can assert the following property:*

$$|(A^{-1}\boldsymbol{p}, (I - \gamma_h)\Lambda_h\boldsymbol{v})| \leq Ch^2\|\boldsymbol{p}\|_1\|\boldsymbol{v}\|_1. \tag{2.35}$$

3. Error estimates

In this particular section of our discussion, we will systematically undertake the task of deriving comprehensive results pertaining to error estimates. These error estimates will be meticulously analyzed for every variable involved in our study.

Lemma 3.1. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z) \in (\boldsymbol{V} \times W)^2$ denote the solution corresponding to (2.4)–(2.8), while $(\boldsymbol{p}_h(u), y_h(u),$
$\boldsymbol{q}_h(u), z_h(u)) \in (\boldsymbol{V}_h \times W_h)^2$ serve as the solution of (2.24)–(2.27) when $\tilde{u} = u$. Suppose $\boldsymbol{p}$, $\boldsymbol{q}$ belong to the space $(H^1(\Omega))^2$, and y and z are in space $H^1(\Omega)$, then we can establish the following result:*

$$\|y - y_h(u)\| + \|\boldsymbol{p} - \boldsymbol{p}_h(u)\| \le Ch, \tag{3.1}$$
$$\|z - z_h(u)\| + \|\boldsymbol{q} - \boldsymbol{q}_h(u)\| \le Ch. \tag{3.2}$$

Proof. Let

$$\begin{aligned}
\boldsymbol{p} - \boldsymbol{p}_h &= \boldsymbol{p} - \Lambda_h \boldsymbol{p} + \Lambda_h \boldsymbol{p} - \boldsymbol{p}_h(u) =: \lambda_{\boldsymbol{p}} + e_{\boldsymbol{p}}, \\
y - y_h &= y - P_h y + P_h y - y_h(u) =: \lambda_y + e_y, \\
\boldsymbol{q} - \boldsymbol{q}_h &= \boldsymbol{q} - \Lambda_h \boldsymbol{q} + \Lambda_h \boldsymbol{q} - \boldsymbol{q}_h(u) =: \lambda_{\boldsymbol{q}} + e_{\boldsymbol{q}}, \\
z - z_h &= z - P_h z + P_h z - z_h(u) =: \lambda_z + e_z.
\end{aligned}$$

For $\forall\ w_h \in W_h, \boldsymbol{v}_h \in \boldsymbol{V}_h$, subtracting (2.24)–(2.27) form (2.4)–(2.7), then, by employing the equations denoted as (2.10) and (2.12), we are able to derive the error equations as follows:

$$(\overline{A^{-1}} e_{\boldsymbol{p}}, \gamma_h \boldsymbol{v}_h) - (e_y, \mathrm{div}\boldsymbol{v}_h) = -(A^{-1}\boldsymbol{p}, \boldsymbol{v}_h) + (\overline{A^{-1}} \Lambda_h \boldsymbol{p}, \gamma_h \boldsymbol{v}_h), \tag{3.3}$$
$$(\mathrm{div} e_{\boldsymbol{p}}, w_h) + (c(\lambda_y + e_y), w_h) = 0, \tag{3.4}$$
$$(\overline{A^{-1}} e_{\boldsymbol{q}}, \gamma_h \boldsymbol{v}_h) - (e_z, \mathrm{div}\boldsymbol{v}_h) = -(A^{-1}\boldsymbol{q}, \boldsymbol{v}_h) + (\overline{A^{-1}} \Lambda_h \boldsymbol{q}, \gamma_h \boldsymbol{v}_h) - (\boldsymbol{p} - \boldsymbol{p}_d, \boldsymbol{v}_h) + (\boldsymbol{p}_h - \Lambda_h \boldsymbol{p}_d, \gamma_h \boldsymbol{v}_h), \tag{3.5}$$
$$(\mathrm{div} e_{\boldsymbol{q}}, w_h) + (c(\lambda_z + e_z), w_h) = (e_y, w_h). \tag{3.6}$$

Setting $v_h = e_{\boldsymbol{p}}$ and $w_h = e_y$ in equations (3.3) and (3.4), respectively, and then combining the two equations yields

$$(\overline{A^{-1}} e_{\boldsymbol{p}}, \gamma_h e_{\boldsymbol{p}}) + (c e_y, e_y) = -(A^{-1}\boldsymbol{p}, e_{\boldsymbol{p}}) + (\overline{A^{-1}} \Lambda_h \boldsymbol{p}, \gamma_h e_{\boldsymbol{p}}) - (c\lambda_y, e_y). \tag{3.7}$$

Close attention should be paid to the term on the right side of equation (3.7). Using the results of (2.11), (2.13), (2.15), (2.31), and the Cauchy–Schwarz inequality, it follows that

$$\begin{aligned}
&-(A^{-1}\boldsymbol{p}, e_{\boldsymbol{p}}) + (\overline{A^{-1}} \Lambda_h \boldsymbol{p}, \gamma_h e_{\boldsymbol{p}}) - (c\lambda_y, e_y) \\
=&-(A^{-1}\boldsymbol{p}, e_{\boldsymbol{p}} - \gamma_h e_{\boldsymbol{p}}) - ((A^{-1} - \overline{A^{-1}})\boldsymbol{p}, \gamma_h e_{\boldsymbol{p}}) - (\overline{A^{-1}}(\boldsymbol{p} - \Lambda_h \boldsymbol{p}), \gamma_h e_{\boldsymbol{p}}) - (c\lambda_y, e_y) \\
\le& Ch\|\boldsymbol{p}\|_1 \|e_{\boldsymbol{p}}\| + Ch\|\boldsymbol{p}\| \cdot \|\gamma_h e_{\boldsymbol{p}}\| + Ch\|\alpha\|_{0,\infty}\|\boldsymbol{p}\|_1 \|\gamma_h e_{\boldsymbol{p}}\| + Ch\|c\|_{0,\infty}\|y\|_1\|e_y\| \\
\le& Ch\|\boldsymbol{p}\|_1 \|e_{\boldsymbol{p}}\| + Ch\|c\|_{0,\infty}\|y\|_1\|e_y\| \\
\le& Ch^2(\|\boldsymbol{p}\|_1^2 + \|y\|_1^2) + \frac{K_2}{2}\|e_{\boldsymbol{p}}\|^2 + \frac{k_0}{2}\|e_y\|^2.
\end{aligned} \tag{3.8}$$

Note the left-side term of (3.7), and combine with the conclusion (2.30) and the assumption of $c(x) \geq k_0 > 0$; it can be known that

$$(\overline{A^{-1}}e_{\boldsymbol{p}}, \gamma_h e_{\boldsymbol{p}}) + (ce_y, e_y) \geq K_2\|e_{\boldsymbol{p}}\|^2 + K_0\|e_y\|^2.$$

Subsequently, combining the result of the above inequality with (3.8), we deduce that

$$\|e_{\boldsymbol{p}}\| + \|e_y\| \leq Ch(\|\boldsymbol{p}\|_1 + \|y\|_1). \tag{3.9}$$

Taking $\boldsymbol{v}_h = e_{\boldsymbol{q}}$ and $w_h = e_z$ in (3.5) and (3.6), respectively, and then summing the two equations gives

$$\begin{aligned}(\overline{A^{-1}}e_{\boldsymbol{q}}, \gamma_h e_{\boldsymbol{q}}) + (ce_z, e_z) =& -(A^{-1}\boldsymbol{q}, e_{\boldsymbol{q}}) + (\overline{A^{-1}}\Lambda_h\boldsymbol{q}, \gamma_h e_{\boldsymbol{q}}) + (\boldsymbol{p}_h(u) - \Lambda_h\boldsymbol{p}_d, \gamma_h e_{\boldsymbol{q}})\\ &- (\boldsymbol{p} - \boldsymbol{p}_d, e_{\boldsymbol{q}}) - (c\lambda_z, e_z) + (\lambda_y + e_y, e_z)\\ =& -(A^{-1}\boldsymbol{q}, e_{\boldsymbol{q}}) + (\overline{A^{-1}}\Lambda_h\boldsymbol{q}, \gamma_h e_{\boldsymbol{q}}) + (\boldsymbol{p}_h(u) - \boldsymbol{p} + \boldsymbol{p}_d - \Lambda_h\boldsymbol{p}_d, \gamma_h e_{\boldsymbol{q}})\\ &+ (\boldsymbol{p} - \boldsymbol{p}_d, \gamma_h e_{\boldsymbol{q}} - e_{\boldsymbol{q}}) - (c\lambda_z, e_z) + (\lambda_y + e_y, e_z).\end{aligned}$$

Analogous to (3.9), we can get

$$\|e_{\boldsymbol{q}}\| + \|e_z\| \leq Ch(\|\boldsymbol{p}\|_1 + \|y\|_1 + \|\boldsymbol{q}\|_1 + \|z\|_1 + \|\boldsymbol{p}_d\|_1) + C\|e_{\boldsymbol{p}}\| + C\|e_y\|. \tag{3.10}$$

At this point, by utilizing (3.9), (3.10), (2.11), and (2.13), we have successfully completed the proof. □

Lemma 3.2. *Suppose $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h)$ be the solution of equations denoted by (2.19)–(2.23) and let the solution corresponding to (2.24)–(2.27) be $(\boldsymbol{p}_h(u), y_h(u), \boldsymbol{q}_h(u), z_h(u))$. We then proceed to establish the following relationship when $\tilde{u}$ is specifically set to u:*

$$\|\boldsymbol{p}_h(u) - \boldsymbol{p}_h\| + \|y_h(u) - y_h\| \leq C\|u - u_h\|, \tag{3.11}$$

$$\|\boldsymbol{q}_h(u) - \boldsymbol{q}_h\| + \|z_h(u) - z_h\| \leq C\|u - u_h\|. \tag{3.12}$$

Proof. From the equations denoted (2.19)–(2.22) as well as those denoted by (2.24)–(2.27), for any arbitrary w_h that belongs to W_h and $\boldsymbol{v}_h$ belongs to $\boldsymbol{V}_h$, we can obtain

$$(\overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}_h), \gamma_h\boldsymbol{v}_h) - (y_h(u) - y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \tag{3.13}$$

$$(\mathrm{div}(\boldsymbol{p}_h(u) - \boldsymbol{p}_h), w_h) + (c(y_h(u) - y_h), w_h) = (u - u_h, w_h), \tag{3.14}$$

$$(\overline{A^{-1}}(\boldsymbol{q}_h(u) - \boldsymbol{q}_h), \gamma_h\boldsymbol{v}_h) - (z_h(u) - z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(u) - \boldsymbol{p}_h, \gamma_h\boldsymbol{v}_h), \tag{3.15}$$

$$(\mathrm{div}(\boldsymbol{q}_h(u) - \boldsymbol{q}_h), w_h) + (c(z_h(u) - z_h), w_h) = (y_h(u) - y_h, w_h). \tag{3.16}$$

By setting $\boldsymbol{v}_h = \boldsymbol{p}_h(u) - \boldsymbol{p}_h$ in equation (3.13) and $w_h = y_h(u) - y_h$ in equation (3.14), respectively, we carefully substitute these definitions into their respective equations. Following this step, we then proceed to add the two resulting equations together, thereby deriving

$$(\overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}_h), \gamma_h(\boldsymbol{p}_h(u) - \boldsymbol{p}_h)) + (c(y_h(u) - y_h), y_h(u) - y_h) = (u - u_h, y_h(u) - y_h). \tag{3.17}$$

Employing (3.17), (2.30), the hypothesis concerning c, and the fundamental principles underlying the Cauchy–Schwarz inequality, we readily derive (3.11).

In Eqs (3.15)–(3.16), we select $\boldsymbol{v}_h = \boldsymbol{q}_h(u) - \boldsymbol{q}_h$, $w_h = z_h(u) - z_h$, respectively. Then we aggregate the outcomes of these two equations by summing them together. By applying the same approach as in (3.17), we deduce that

$$\|\boldsymbol{q}_h(u) - \boldsymbol{q}_h\| + \|z_h(u) - z_h\| \leq C\|\boldsymbol{p}_h(u) - \boldsymbol{p}_h\| + C\|y_h(u) - y_h\|. \tag{3.18}$$

Consequently, Lemma 3.2 can be established on the basis of inequalities (3.11) and (3.18). □

Theorem 3.1. *Suppose u is the solution to the system of equations denoted by (2.4)–(2.8) and let u_h be the corresponding solution to the system of equations represented by (2.19)–(2.23), respectively. If the vectors $\boldsymbol{p}$, $\boldsymbol{q} \in (H^1(\Omega))^2$, the scalar functions y, $z \in H^1(\Omega)$, then we can establish the following relationship:*

$$\|u - u_h\| \leq Ch.$$

Proof. In the context of our analysis, we specifically choose $\tilde{u}$ to be equal to u_h and assign $\tilde{u}_h$ to be $P_h u$. These selections are made for the equations labeled as (2.8) and (2.23), respectively. Thereby we can obtain two inequalities below:

$$(\theta u + z, u_h - u) \geq 0 \tag{3.19}$$

and

$$(\theta u_h + z_h, P_h u - u_h) \geq 0. \tag{3.20}$$

By (3.19), (3.20), and (2.10), it can be found that

$$\begin{aligned}
\theta\|u - u_h\|^2 =&\theta(u - u_h, u - u_h) \\
=&(\theta u + z, u - u_h) + (z_h(u) - z, u - u_h) \\
&+ (z_h - z_h(u), u - u_h) + (\theta u_h + z_h, u_h - u) \\
\leq&(z_h(u) - z, u - u_h) + (z_h - z_h(u), u - u_h) \\
&+ (\theta u_h + z_h, u_h - P_h u) + (\theta u_h + z_h, P_h u - u) \\
\leq&(z_h(u) - z, u - u_h) + (z_h - z_h(u), u - u_h).
\end{aligned} \tag{3.21}$$

In (3.14) and (3.15), set $w_h = z_h - z_h(u)$ and $\boldsymbol{v}_h = \boldsymbol{p}_h(u) - \boldsymbol{p}_h$, respectively, and then subtracting the two equations yields

$$\begin{aligned}
(z_h - z_h(u), u - u_h) =&(z_h - z_h(u), c(y_h(u) - y_h)) \\
&- (\gamma_h(\boldsymbol{p}_h(u) - \boldsymbol{p}_h), \overline{A^{-1}}(\boldsymbol{q}_h(u) - \boldsymbol{q}_h)) - (\gamma_h(\boldsymbol{p}_h(u) - \boldsymbol{p}_h), \boldsymbol{p}_h(u) - \boldsymbol{p}_h).
\end{aligned} \tag{3.22}$$

Select $\boldsymbol{v}_h = \boldsymbol{q}_h(u) - \boldsymbol{q}_h$ in (3.13), and let $w_h = y_h(u) - y_h$ in (3.16), respectively. Then we find the result by summing the two equations:

$$(\gamma_h(\boldsymbol{q}_h(u) - \boldsymbol{q}_h), \overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}_h)) + (y_h(u) - y_h, c(z_h(u) - z_h)) = (y_h(u) - y_h, y_h(u) - y_h). \tag{3.23}$$

By (2.28), (2.33), (3.22), and (3.23), we can get

$$(z_h - z_h(u), u - u_h) = -(\gamma_h(\boldsymbol{p}_h(u) - \boldsymbol{p}_h), \boldsymbol{p}_h(u) - \boldsymbol{p}_h) - (y_h(u) - y_h, y_h(u) - y_h) \\ \leq -\|y_h(u) - y_h\|^2 - K_0\|\boldsymbol{p}_h(u) - \boldsymbol{p}_h\|^2 \leq 0. \tag{3.24}$$

From (3.21), (3.24), and the Cauchy inequality, we can deduce that

$$\|u - u_h\| \leq C\|z - z_h(u)\|. \tag{3.25}$$

By combining (3.25) with (3.2), we bring the proof of the theorem to a close. □

Applying Theorem 3.1 as well as Lemmas 3.1–3.2, one readily arrives at the following theorem:

Theorem 3.2. *Take* $(\boldsymbol{p}, y, \boldsymbol{q}, z) \in (\boldsymbol{V} \times W)^2$ *to denote the solution of (2.4)–(2.8) and* $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h) \in (\boldsymbol{V}_h \times W_h)^2$ *to be the solution corresponding to (2.19)–(2.23), respectively. If* $y,\ z \in H^1(\Omega)$ *and* $\boldsymbol{p},\ \boldsymbol{q} \in (H^1(\Omega))^2$*, we then obtain*

$$\|y - y_h\| + \|\boldsymbol{p} - \boldsymbol{p}_h\| + \|z - z_h\| + \|\boldsymbol{q} - \boldsymbol{q}_h\| \leq Ch.$$

4. Superconvergence

In this part of the text, we will proceed to derive a superconvergent result specifically for the control variable u. Following the reference [6], It is possible for us to classify the triangles $K_B \in \mathcal{T}_h$ into two sets, $\mathcal{T}_1$ and $\mathcal{T}_2$:

$$\mathcal{T}_1 = \{K_{B_i} : u \text{ only belongs to } W^{1,\infty}(K_{B_i})\},\ \ \mathcal{T}_2 = \{K_{B_i} : u \in H^2(K_{B_i})\}.$$

To achieve a superconvergent result, it is imperative that we introduce an additional assumption pertaining to the exact control solution of u:

$$\text{meas}(\mathcal{T}_1) \leq Ch. \tag{4.1}$$

The assumption denoted as (4.1) presents a significant challenge in terms of verification, primarily due to the intricate nature of the conditions it encompasses. However, despite these inherent difficulties, it has been observed that this assumption holds true in a multitude of practical scenarios.

Next, we bring to mind a result by Grisvard [26].

Lemma 4.1. *[26] Assume the domain* Ω *is convex. For every function* $\psi \in \Omega$*, consider the equation*

$$-\text{div}(A(x)\nabla\phi) + c(x)\phi = \psi \ \ in\ \Omega \tag{4.2}$$

with $\phi|_{\partial\Omega} = 0$*. Then there exists a unique solution* $\phi \in H_0^1(\Omega) \cap H^2(\Omega)$*. Moreover, there exists a constant* $C > 0$ *such that*

$$\|\phi\|_2 \leq C\|\psi\|. \tag{4.3}$$

Lemma 4.2. *Assume $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h)$ represent the solution delineated by (2.19)–(2.23). Suppose $(\boldsymbol{p}_h(u_I), y_h(u_I), \boldsymbol{q}_h(u_I), z_h(u_I))$ denote the solution outlined in (2.24)–(2.27), wherein $\tilde{u} = u_I$. We can get*

$$||y_h(u_I) - y_h|| + ||\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h|| \leq C||u_I - u_h||, \tag{4.4}$$

$$||z_h(u_I) - z_h|| + ||\boldsymbol{q}_h(u_I) - \boldsymbol{q}_h|| \leq C||u_I - u_h||. \tag{4.5}$$

Proof. We derive the error equations below from (2.19)–(2.22) and (2.24)–(2.27):

$$(\overline{A^{-1}}(\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h), \gamma_h \boldsymbol{v}_h) - (y_h(u_I) - y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \ \forall \ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.6}$$

$$(\mathrm{div}(\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h), w_h) + (c(y_h(u_I) - y_h), w_h) = (u_I - u_h, w_h), \ \forall \ w_h \in W_h, \tag{4.7}$$

$$(\overline{A^{-1}}(\boldsymbol{q}_h(u_I) - \boldsymbol{q}_h), \gamma_h \boldsymbol{v}_h) - (z_h(u_I) - z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h, \gamma_h \boldsymbol{v}_h), \ \forall \ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.8}$$

$$(\mathrm{div}(\boldsymbol{q}_h(u_I) - \boldsymbol{q}_h), w_h) + (c(z_h(u) - z_h), w_h) = (y_h(u_I) - y_h, w_h), \ \forall \ w_h \in W_h. \tag{4.9}$$

Set $\boldsymbol{v}_h = \boldsymbol{p}_h(u_I) - \boldsymbol{p}_h$ in equation (4.6), and $w_h = y_h(u_I) - y_h$ in expression (4.7), respectively. Subsequently, add these two resulting equations together:

$$\begin{aligned}(\overline{A^{-1}}(\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h), \gamma_h(\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h)) + (c(y_h(u_I) - y_h), y_h(u_I) - y_h)&\\ = (u_I - u_h, y_h(u_I) - y_h).&\end{aligned} \tag{4.10}$$

By use of (4.10), (2.30), the Cauchy–Schwarz inequality, and the assumption regarding c, we easily get (4.4).

This is accomplished by selecting $\boldsymbol{v}_h = \boldsymbol{q}_h(u_I) - \boldsymbol{q}_h$ for the first equation and $w_h = z_h(u_I) - z_h$ for the second equation in the respective formulas (4.8) and (4.9). Following this substitution, we then add the two resulting equations together and use the same technique as (4.10); it can be shown that

$$||z_h(u_I) - z_h|| + ||\boldsymbol{q}_h(u_I) - \boldsymbol{q}_h|| \leq C||y_h(u_I) - y_h|| + C||\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h||, \tag{4.11}$$

which brings the proof to a close. □

Lemma 4.3. *Assume $(\boldsymbol{p}_h(u), y_h(u), \boldsymbol{q}_h(u), z_h(u))$ and $(\boldsymbol{p}_h(u_I), y_h(u_I), \boldsymbol{q}_h(u_I), z_h(u_I))$ act as the respective solutions of the equations denoted by (2.24)–(2.27), where the variable $\tilde{u}$ is set to u in the first case and to u_I in the second case. Consequently, we can establish the following relationship:*

$$||y_h(u) - y_h(u_I)|| + ||z_h(u) - z_h(u_I)|| \leq Ch^2. \tag{4.12}$$

Proof. From (2.24)–(2.27), for $\forall \ w_h \in W_h, \forall \ \boldsymbol{v}_h \in \boldsymbol{V}_h$, we derive error expressions below:

$$(\overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}_h(u_I)), \gamma_h \boldsymbol{v}_h) - (y_h(u) - y_h(u_I), \mathrm{div}\boldsymbol{v}_h) = 0, \tag{4.13}$$

$$(\mathrm{div}(\boldsymbol{p}_h(u) - \boldsymbol{p}_h(u_I)), w_h) + (c(y_h(u) - y_h(u_I)), w_h) = (u - u_I, w_h), \tag{4.14}$$

$$(\overline{A^{-1}}(\boldsymbol{q}_h(u) - \boldsymbol{q}_h(u_I)), \gamma_h \boldsymbol{v}_h) - (z_h(u) - z_h(u_I), \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(u) - \boldsymbol{p}_h(u_I), \gamma_h \boldsymbol{v}_h), \tag{4.15}$$

$$(\mathrm{div}(\boldsymbol{q}_h(u) - \boldsymbol{q}_h(u_I)), w_h) + (c(z_h(u) - z_h(u_I)), w_h) = (y_h(u) - y_h(u_I), w_h). \tag{4.16}$$

Notice that

$$(u - u_I, w_h) = \int_{\mathcal{T}_1} (u - u_I) w_h \mathrm{d}x + \int_{\mathcal{T}_2} (u - u_I) w_h \mathrm{d}x.$$

Through the Schwarz inequality, assumption of (4.1), and Lemma 2.1, we can get

$$\begin{aligned}\left|\int_{\mathcal{T}_1}(u-u_I)w_h\mathrm{d}x\right| \le &\sum_{K_{B_i}\in\mathcal{T}_1}\int_{K_{B_i}}|u(x)-u(S_i)|\cdot|w_h|\mathrm{d}x\\ \le& Ch\sum_{K_{B_i}\in\mathcal{T}_1}\|u\|_{W^{1,\infty}(K_{B_i})}\int_{K_{B_i}}|w_h|\mathrm{d}x\\ \le& Ch\|u\|_{W^{1,\infty}(\Omega)}\sum_{K_{B_i}\in\mathcal{T}_1}\sqrt{|K_{B_i}|}\left(\int_{K_{B_i}}|w_h|^2\mathrm{d}x\right)^{\frac{1}{2}}\\ \le& Ch^{1.5}\|u\|_{W^{1,\infty}(\Omega)}\|w_h\| \end{aligned}\tag{4.17}$$

and

$$\begin{aligned}\left|\int_{\mathcal{T}_2}(u-u_I)w_h\mathrm{d}x\right| =& \left|\sum_{K_{B_i}\in\mathcal{T}_2}\int_{K_{B_i}}(u(x)-u(B_i))w_h(x)\mathrm{d}x\right|\\ =&\left|\sum_{K_{B_i}\in\mathcal{T}_2}w_h(B_i)\int_{K_{B_i}}(u(x)-u(B_i))\mathrm{d}x\right|\\ \le&\sum_{K_{B_i}\in\mathcal{T}_2}Ch^2|w_h(B_i)|\sqrt{|K_{B_i}|}\cdot|u|_{H^2(K_{B_i})}\\ \le& Ch^2|u|_{H^2(\mathcal{T}_2)}\|w_h\|. \end{aligned}\tag{4.18}$$

Thus, from (4.17) and (4.18), we easily find that

$$|(u-u_I,w_h)|\le Ch\|w_h\|.\tag{4.19}$$

Using (4.19) and the same estimates as Lemma 4.2, we conclude that

$$\|\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)\|+\|y_h(u)-y_h(u_I)\|\le Ch,\tag{4.20}$$

$$\|\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)\|+\|z_h(u)-z_h(u_I)\|\le C\|\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)\|+C\|y_h(u)-y_h(u_I)\|.\tag{4.21}$$

Suppose ϕ is the solution to the equation denoted by (4.2), where ψ is defined as $\psi=y_h(u)-y_h(u_I)$. It follows directly from the given Eqs (4.2), (2.12), (2.10), Green's formula, (4.13), and (4.14) that

$$\begin{aligned}\|y_h(u)-y_h(u_I)\|^2 =&(y_h(u)-y_h(u_I),-\mathrm{div}(A\nabla\phi))+(y_h(u)-y_h(u_I),c\phi)\\ =&-(y_h(u)-y_h(u_I),\mathrm{div}(\Lambda_h(A\nabla\phi)))+(c(y_h(u)-y_h(u_I)),\phi)\\ =&-(\overline{A^{-1}}(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),\gamma_h\Lambda_h(A\nabla\phi))+(c(y_h(u)-y_h(u_I)),\phi)\\ =&(\overline{A^{-1}}(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))\\ &+((A^{-1}-\overline{A^{-1}})(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),A\nabla\phi)\\ &+(\mathrm{div}(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),\phi)+(c(y_h(u)-y_h(u_I)),\phi)\\ =&(\overline{A^{-1}}(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))\end{aligned}$$

$$+ ((A^{-1} - \overline{A^{-1}})(\boldsymbol{p}_h(u) - \boldsymbol{p}_h(u_I)), A\nabla\phi)$$
$$+ (u - u_I, P_h\phi) + (c(y_h(u) - y_h(u_I)), \phi - P_h\phi). \tag{4.22}$$

We decompose the integral expression $(u - u_I, P_h\phi)$ into two distinct and manageable parts:

$$(u - u_I, P_h\phi) = \int_{\mathcal{T}_1} (u - u_I)P_h\phi \mathrm{d}x + \int_{\mathcal{T}_2} (u - u_I)P_h\phi \mathrm{d}x. \tag{4.23}$$

It follows directly from the established assumption denoted as (4.1) by invoking the classical embedding theorem represented as $\|\phi\|_{0,\infty} \leq C\|\phi\|_2$ (C is a constant). Additionally, using the result $\|P_h\phi\|_{0,\infty} \leq \|\phi\|_{0,\infty}$, we can deduce

$$\begin{aligned}\left|\int_{\mathcal{T}_1} (u - u_I)P_h\phi \mathrm{d}x\right| \leq& \sum_{K_{B_i}\in\mathcal{T}_1} \int_{K_{B_i}} |u(x) - u(B_i)| \cdot |P_h\phi| \mathrm{d}x \\ \leq& Ch \sum_{K_{B_i}\in\mathcal{T}_1} \|u\|_{W^{1,\infty}(K_{B_i})} \|P_h\phi\|_{L^\infty(K_{B_i})} \int_{K_{B_i}} \mathrm{d}x \\ \leq& Ch\|u\|_{W^{1,\infty}(\Omega)} \|P_h\phi\|_{L^\infty(\Omega)} \sum_{K_{B_i}\in\mathcal{T}_1} |K_{B_i}| \\ \leq& Ch^2\|u\|_{W^{1,\infty}(\Omega)} \|\phi\|_{0,\infty} \\ \leq& Ch^2\|u\|_{W^{1,\infty}(\Omega)} \|\phi\|_2. \end{aligned} \tag{4.24}$$

In a manner analogous to the derivation presented in Eq (4.18), by leveraging the inequality $\|P_h\phi\| \leq C\|\phi\|$, we can derive that

$$\left|\int_{\mathcal{T}_2} (u - u_I)P_h\phi \mathrm{d}x\right| \leq Ch^2\|P_h\phi\| \cdot |u|_{H^2(\mathcal{T}_2)} \leq Ch^2\|\phi\|_2|u|_{H^2(\mathcal{T}_2)}. \tag{4.25}$$

By use of (4.22)–(4.25), the Cauchy inequality, (2.11), (2.13), and (2.15), we get the following result:

$$\|y_h(u) - y_h(u_I)\|^2 \leq C(h(\|y_h(u) - y_h(u_I)\| + \|\boldsymbol{p}_h(u) - \boldsymbol{p}_h(u_I)\|) + h^2)\|\phi\|_2. \tag{4.26}$$

Substituting (4.20) into (4.26) and using (4.3), we derive

$$\|y_h(u) - y_h(u_I)\| \leq Ch^2. \tag{4.27}$$

Next, we estimate $\|z_h(u) - z_h(u_I)\|$. Let ϕ denote the solution associated with (4.2), where $\psi = z_h(u) - z_h(u_I)$. It can be observed from (4.2), (2.12), (2.10), Green's formula, (4.13), (4.15), and (4.16) that

$$\begin{aligned}\|z_h(u) - z_h(u_I)\|^2 =& (z_h(u) - z_h(u_I), -\mathrm{div}(A\nabla\phi)) + (z_h(u) - z_h(u_I), c\phi) \\ =& -(z_h(u) - z_h(u_I), \mathrm{div}(\Lambda_h(A\nabla\phi))) + (c(z_h(u) - z_h(u_I)), \phi) \\ =& -(\overline{A^{-1}}(\boldsymbol{q}_h(u) - \boldsymbol{q}_h(u_I)), \gamma_h\Lambda_h(A\nabla\phi)) - (\boldsymbol{p}_h(u) - \boldsymbol{p}_h(u_I), \gamma_h\Lambda_h(A\nabla\phi)) \\ &+ (c(z_h(u) - z_h(u_I)), \phi)\end{aligned}$$

$$
\begin{aligned}
=&(\overline{A^{-1}}(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)), A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))+((A^{-1}-\overline{A^{-1}})(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)), A\nabla\phi)\\
&+(\mathrm{div}(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)),\phi)+(c(z_h(u)-z_h(u_I)),\phi)-(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I),\gamma_h\Lambda_h(A\nabla\phi))\\
=&(\overline{A^{-1}}(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)), A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))+((A^{-1}-\overline{A^{-1}})(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)), A\nabla\phi)\\
&+(y_h(u)-y_h(u_I),P_h\phi)+(c(z_h(u)-z_h(u_I)),\phi-P_h\phi)\\
&+(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I),A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))-(A^{-1}(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),A^2\nabla\phi)\\
=&(\overline{A^{-1}}(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)), A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))+((A^{-1}-\overline{A^{-1}})(\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)), A\nabla\phi)\\
&+(y_h(u)-y_h(u_I),\phi)+(c(z_h(u)-z_h(u_I)),\phi-P_h\phi)\\
&+(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I),A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))+((\overline{A^{-1}}-A^{-1})(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),A^2\nabla\phi)\\
&+(\overline{A^{-1}}(\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)),\gamma_h\Lambda_h(A^2\nabla\phi)-A^2\nabla\phi)-(y_h(u)-y_h(u_I),\mathrm{div}(A^2\nabla\phi)). \quad (4.28)
\end{aligned}
$$

Similar to (4.26), we prove that

$$
\begin{aligned}
\|z_h(u)-z_h(u_I)\|^2 \le & C\|y_h(u)-y_h(u_I)\|\\
&+Ch(\|z_h(u)-z_h(u_I)\|+\|\boldsymbol{p}_h(u)-\boldsymbol{p}_h(u_I)\|+\|\boldsymbol{q}_h(u)-\boldsymbol{q}_h(u_I)\|)\|\phi\|_2. \quad (4.29)
\end{aligned}
$$

Thus, the proof can be completed by (4.20)–(4.21), (4.27), (4.29), and (4.3). □

Lemma 4.4. *Let the solution of (2.4)–(2.8) be* $(\boldsymbol{p}, y, \boldsymbol{q}, z) \in (\boldsymbol{V}\times W)^2$ *and assume the solution of (2.24)–(2.27) is* $(\boldsymbol{p}_h(u), y_h(u), \boldsymbol{q}_h(u), z_h(u)) \in (\boldsymbol{V}_h \times W_h)^2$*, where* $\tilde{u}$ *is set to be equal to u. If the functions* $\boldsymbol{p},\ \boldsymbol{q} \in (H^2(\Omega))^2 \cap (W^{1,\infty}(\Omega))^2$*, and the functions* $y,\ z \in H^1(\Omega)$*, then we can establish the following result:*

$$
\|P_h y - y_h(u)\| + \|P_h z - z_h(u)\| \le Ch^2. \quad (4.30)
$$

Proof. For $\forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ \forall\ w_h \in W_h$, we rephrase the equation presented in (3.3)–(3.6) by reformulating them in a new format:

$$
\begin{aligned}
&(\overline{A^{-1}}(\boldsymbol{p}-\boldsymbol{p}_h(u)),\gamma_h\boldsymbol{v}_h)-(e_y,\mathrm{div}\boldsymbol{v}_h)=-(A^{-1}\boldsymbol{p},\boldsymbol{v}_h)+(\overline{A^{-1}}\boldsymbol{p},\gamma_h\boldsymbol{v}_h), && (4.31)\\
&(\mathrm{div}e_{\boldsymbol{p}},w_h)+(ce_y,w_h)=-(c\lambda_y,w_h), && (4.32)\\
&(\overline{A^{-1}}(\boldsymbol{q}-\boldsymbol{q}_h(u)),\gamma_h\boldsymbol{v}_h)-(e_z,\mathrm{div}\boldsymbol{v}_h)=-(A^{-1}\boldsymbol{q},\boldsymbol{v}_h)+(\overline{A^{-1}}\boldsymbol{q},\gamma_h\boldsymbol{v}_h)\\
&\qquad-(\boldsymbol{p}-\boldsymbol{p}_d,\boldsymbol{v}_h)+(\boldsymbol{p}_h(u)-\Pi_h\boldsymbol{p}_d,\Lambda_h\boldsymbol{v}_h), && (4.33)\\
&(\mathrm{div}e_{\boldsymbol{q}},w_h)+(ce_z,w_h)=-(c\lambda_z,w_h)+(e_y,w_h). && (4.34)
\end{aligned}
$$

Now, we shall estimate (4.30) in both Part I and Part II.

Part I. Let ϕ be the solution associated with (4.2), where $\psi = e_y$. We easily conclude that

$$
\begin{aligned}
(e_y,\psi)=&(e_y,-\mathrm{div}(A\nabla\phi))+(e_y,c\phi)=-(e_y,\mathrm{div}(\Lambda_h(A\nabla\phi)))+(ce_y,\phi)\\
=&-(A^{-1}\boldsymbol{p},\Lambda_h(A\nabla\phi))+(\overline{A^{-1}}\boldsymbol{p},\gamma_h\Lambda_h(A\nabla\phi))-(\overline{A^{-1}}(\boldsymbol{p}-\boldsymbol{p}_h(u)),\gamma_h\Lambda_h(A\nabla\phi))+(ce_y,\phi)\\
=&-(A^{-1}\boldsymbol{p},\Lambda_h(A\nabla\phi))+(\overline{A^{-1}}\boldsymbol{p},\gamma_h\Lambda_h(A\nabla\phi)-\Lambda_h(A\nabla\phi))+(\overline{A^{-1}}\boldsymbol{p},\Lambda_h(A\nabla\phi))\\
&+(\overline{A^{-1}}(\boldsymbol{p}-\boldsymbol{p}_h(u)),A\nabla\phi-\gamma_h\Lambda_h(A\nabla\phi))-(\overline{A^{-1}}(\boldsymbol{p}-\boldsymbol{p}_h(u)),A\nabla\phi)+(ce_y,\phi)\\
=&((A^{-1}-\overline{A^{-1}})\boldsymbol{p},A\nabla\phi-\Lambda_h(A\nabla\phi))+(\overline{A^{-1}}\boldsymbol{p},\gamma_h\Lambda_h(A\nabla\phi)-\Lambda_h(A\nabla\phi))-((A^{-1}-\overline{A^{-1}})\boldsymbol{p},A\nabla\phi)
\end{aligned}
$$

$$+ (\overline{A^{-1}}(\boldsymbol{p} - \boldsymbol{p}_h(u)), A\nabla\phi - \gamma_h \Lambda_h(A\nabla\phi)) + ((A^{-1} - \overline{A^{-1}})(\boldsymbol{p} - \boldsymbol{p}_h(u)), A\nabla\phi)$$
$$- (\boldsymbol{p} - \boldsymbol{p}_h(u), \nabla\phi) + (ce_y, \phi) =: \sum_{k=1}^{7} J_k, \tag{4.35}$$

where we used (4.2), (2.12), (2.10), (4.31), and (4.32).

From (2.13), (2.15), (2.34), and the Cauchy inequality, we see that

$$\begin{aligned} J_1 + J_4 + J_5 \leq & Ch^2\|\boldsymbol{p}\| \cdot \|A\nabla\phi\|_1 + Ch\|\boldsymbol{p} - \boldsymbol{p}_h(u)\| \cdot \|A\nabla\phi\| + Ch\|\boldsymbol{p} - \boldsymbol{p}_h(u)\| \cdot \|A\nabla\phi\|_1 \\ \leq & C(h^2 + h\|\boldsymbol{p} - \boldsymbol{p}_h(u)\|)\|\phi\|_2. \end{aligned} \tag{4.36}$$

For J_2, it can be estimated by (2.13), (2.15), Lemmas 2.5–2.6, and the Cauchy inequality that

$$\begin{aligned} J_2 = & ((\overline{A^{-1}} - A^{-1})\boldsymbol{p}, \gamma_h \Lambda_h(A\nabla\phi) - A\nabla\phi) + ((\overline{A^{-1}} - A^{-1})\boldsymbol{p}, A\nabla\phi - \Lambda_h(A\nabla\phi)) \\ & + (A^{-1}\boldsymbol{p}, \gamma_h \Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) \\ \leq & Ch^2\|\boldsymbol{p}\|_1 \|A\nabla\phi\|_1. \end{aligned} \tag{4.37}$$

By use of (2.11), (2.15), Lemma 2.1, and using the Cauchy inequality, we can perform the estimation of J_3 as

$$\begin{aligned} J_3 = & -((1/a_{11} - (1/a_{11})_I)p_1, a_{11}\phi_{x_1}) - ((1/a_{22} - (1/a_{22})_I)p_2, a_{22}\phi_{x_2}) \\ = & (1/a_{11} - (1/a_{11})_I, P_h(p_1 a_{11}\phi_{x_1}) - p_1 a_{11}\phi_{x_1}) + (1/a_{22} - (1/a_{22})_I, P_h(p_2 a_{22}\phi_{x_2}) - p_2 a_{22}\phi_{x_2}) \\ & - (1/a_{11} - (1/a_{11})_I, P_h(p_1 a_{11}\phi_{x_1})) - (1/a_{22} - (1/a_{22})_I, P_h(p_2 a_{22}\phi_{x_2})) \\ \leq & Ch^2\|p_1 a_{11}\phi_{x_1}\|_1 + Ch^2\|p_2 a_{22}\phi_{x_2}\|_1 + Ch^2\|P_h(p_1 a_{11}\phi_{x_1})\| + Ch^2\|P_h(p_2 a_{22}\phi_{x_2})\| \\ \leq & Ch^2\|\boldsymbol{p}\|_{1,\infty}\|\phi\|_2. \end{aligned} \tag{4.38}$$

For J_6 and J_7, it follows from Green's formula, (4.32), (2.10), (2.11), (2.14) and Cauchy inequality that

$$\begin{aligned} J_6 + J_7 = & (\mathrm{div}(\boldsymbol{p} - \boldsymbol{p}_h(u)), \phi) + (ce_y, \phi) \\ = & (\mathrm{div} r_{\boldsymbol{p}}, \phi - P_h\phi) + (\mathrm{div} e_{\boldsymbol{p}} + ce_y, \phi - P_h\phi) - ((c - P_h c)\lambda_y, P_h\phi) \\ = & (\mathrm{div} r_{\boldsymbol{p}}, \phi - P_h\phi) + (ce_y, \phi - P_h\phi) - ((c - P_h c)\lambda_y, P_h\phi) \\ \leq & Ch^2\|\mathrm{div}\boldsymbol{p}\|_1\|\phi\|_1 + Ch\|e_y\| \cdot \|\phi\|_1 + Ch^2\|c\|_{1,\infty}\|y\|_1\|\phi\| \\ \leq & Ch^2\|\phi\|_2 + Ch\|e_y\| \cdot \|\phi\|_2. \end{aligned} \tag{4.39}$$

Substituting (4.36)–(4.39) into (4.35), using (4.3) and (3.1), we have

$$\|e_y\| \leq Ch^2 + Ch\|\boldsymbol{p} - \boldsymbol{p}_h(u)\| \leq Ch^2. \tag{4.40}$$

Part II. Let ϕ serve as the solution of the equation represented by (4.2), where $\psi = e_z$. Following a similar approach to the one employed in deriving equation (4.35), we proceed to analyze that

$$\begin{aligned} (e_z, \psi) = & (e_z, -\mathrm{div}(A\nabla\phi)) + (e_z, c\phi) = -(e_z, \mathrm{div}(\Lambda_h(A\nabla\phi))) + (ce_z, \phi) \\ = & -(A^{-1}\boldsymbol{q}, \Pi_h(A\nabla\phi)) + (\overline{A^{-1}}\boldsymbol{q}, \gamma_h \Lambda_h(A\nabla\phi)) - (\overline{A^{-1}}(\boldsymbol{q} - \boldsymbol{q}_h(u)), \gamma_h \Lambda_h(a\nabla\phi)) + (ce_z, \phi) \end{aligned}$$

$$
\begin{aligned}
&= -(A^{-1}\boldsymbol{q}, \Lambda_h(A\nabla\phi)) + (\overline{A^{-1}}\boldsymbol{q}, \gamma_h\Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) + (\overline{A^{-1}}\boldsymbol{q}, \Lambda_h(A\nabla\phi)) \\
&\quad - (\overline{A^{-1}}(\boldsymbol{q} - \boldsymbol{q}_h(u)), A\nabla\phi) + (\overline{A^{-1}}(\boldsymbol{q} - \boldsymbol{q}_h(u)), A\nabla\phi - \gamma_h\Lambda_h(A\nabla\phi)) + (ce_z, \phi) \\
&= ((A^{-1} - \overline{A^{-1}})\boldsymbol{q}, A\nabla\phi - \Lambda_h(A\nabla\phi)) - ((A^{-1} - \overline{A^{-1}})\boldsymbol{q}, A\nabla\phi) + (\overline{A^{-1}}\boldsymbol{q}, \gamma_h\Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) \\
&\quad + ((A^{-1} - \overline{A^{-1}})(\boldsymbol{q} - \boldsymbol{q}_h(u)), A\nabla\phi) + (\overline{A^{-1}}(\boldsymbol{q} - \boldsymbol{q}_h(u)), A\nabla\phi - \gamma_h\Lambda_h(A\nabla\phi)) \\
&\quad - (\boldsymbol{q} - \boldsymbol{q}_h(u), \nabla\phi) + (ce_z, \phi) - (\boldsymbol{p} - \boldsymbol{p}_d, \Lambda_h(A\nabla\phi)) + (\boldsymbol{p}_h(u) - \Lambda_h\boldsymbol{p}_d, \gamma_h\Lambda_h(A\nabla\phi)) =: \sum_{i=1}^{9} D_i.
\end{aligned} \tag{4.41}
$$

Similar to (4.36)–(4.39), we have

$$
\sum_{i=1}^{5} D_i \leq C(h^2 + h\|\boldsymbol{q} - \boldsymbol{q}_h(u)\|)\|\phi\|_2 \tag{4.42}
$$

and

$$
\begin{aligned}
D_6 + D_7 &= (\mathrm{div}(\boldsymbol{q} - \boldsymbol{q}_h(u)), \phi) + (ce_z, \phi) \\
&= (\mathrm{div} r_{\boldsymbol{q}}, \phi - P_h\phi) + (ce_z, \phi - P_h\phi) - ((c - P_hc)\lambda_z, P_h\phi) + (e_y, P_h\phi) \\
&\leq Ch^2(\|\mathrm{div}\boldsymbol{q}\|_1 + \|z\|_1)\|\phi\|_2 + Ch\|e_z\| \cdot \|\phi\|_2 + C\|e_y\| \cdot \|\phi\|.
\end{aligned} \tag{4.43}
$$

Notice that

$$
\begin{aligned}
D_8 + D_9 &= (\boldsymbol{p}_h(u) - \Lambda_h\boldsymbol{p}_d - \boldsymbol{p} + \boldsymbol{p}_d, \gamma_h\Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) + (\boldsymbol{p} - \boldsymbol{p}_d, \gamma_h\Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) \\
&\quad + (\boldsymbol{p}_h(u) - \Lambda_h\boldsymbol{p}_d, \Lambda_h(A\nabla\phi)) - (\boldsymbol{p} - \boldsymbol{p}_d, \Lambda_h(A\nabla\phi)) \\
&= (\boldsymbol{p}_h(u) - \boldsymbol{p} + \boldsymbol{p}_d - \Lambda_h\boldsymbol{p}_d, \gamma_h\Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) + (\boldsymbol{p} - \boldsymbol{p}_d, \gamma_h\Lambda_h(A\nabla\phi) - \Lambda_h(A\nabla\phi)) \\
&\quad + (\boldsymbol{p}_d - \Lambda_h\boldsymbol{p}_d, \Lambda_h(A\nabla\phi)) - (\boldsymbol{p}_h(u) - \boldsymbol{p}, \Lambda_h(A\nabla\phi) - A\nabla\phi) + (\boldsymbol{p}_h(u) - \boldsymbol{p}, A\nabla\phi).
\end{aligned} \tag{4.44}
$$

Using (4.31) and (2.12), we conclude that

$$
\begin{aligned}
(\boldsymbol{p}_h(u) - \boldsymbol{p}, A\nabla\phi) &= (A^{-1}(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi) \\
&= ((A^{-1} - \overline{A^{-1}})(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi) - (\overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi - \Lambda_h(A^2\nabla\phi)) \\
&\quad + (\overline{A^{-1}}(\boldsymbol{p} - \boldsymbol{p}_h(u)), \Lambda_h(A^2\nabla\phi) - \gamma_h\Lambda_h(A^2\nabla\phi)) + (\overline{A^{-1}}(\boldsymbol{p} - \boldsymbol{p}_h(u)), \gamma_h\Lambda_h(A^2\nabla\phi)) \\
&= ((A^{-1} - \overline{A^{-1}})(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi) - (\overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi - \Lambda_h(A^2\nabla\phi)) \\
&\quad + (\overline{A^{-1}}(\boldsymbol{p} - \boldsymbol{p}_h(u)), \Lambda_h(A^2\nabla\phi) - \gamma_h\Lambda_h(A^2\nabla\phi)) + (e_y, \mathrm{div}\Lambda_h(A^2\nabla\phi)) \\
&\quad - (A^{-1}\boldsymbol{p}, \Lambda_h(A^2\nabla\phi)) + (\overline{A^{-1}}\boldsymbol{p}, \gamma_h\Lambda_h(A^2\nabla\phi))
\end{aligned}
$$

$$
\begin{aligned}
&= ((A^{-1} - \overline{A^{-1}})(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi) - (\overline{A^{-1}}(\boldsymbol{p}_h(u) - \boldsymbol{p}), A^2\nabla\phi - \Lambda_h(A^2\nabla\phi)) \\
&\quad + (\overline{A^{-1}}(\boldsymbol{p} - \boldsymbol{p}_h(u)), \Lambda_h(A^2\nabla\phi) - \gamma_h\Lambda_h(A^2\nabla\phi)) + (e_y, \mathrm{div}(A^2\nabla\phi)) \\
&\quad + ((A^{-1} - \overline{A^{-1}})\boldsymbol{p}, A^2\nabla\phi - \gamma_h\Lambda_h(A^2\nabla\phi)) - ((A^{-1} - \overline{A^{-1}})\boldsymbol{p}, A^2\nabla\phi) \\
&\quad + (A^{-1}\boldsymbol{p}, \gamma_h\Lambda_h(A^2\nabla\phi) - \Lambda_h(A^2\nabla\phi)).
\end{aligned} \tag{4.45}
$$

Moreover, from [27], we know that

$$(\boldsymbol{p}_d - \Lambda_h \boldsymbol{p}_d, \Lambda_h(A\nabla\phi)) \leq Ch^2\|\boldsymbol{p}_d\|_2\|\Lambda_h(A\nabla\phi)\| \leq Ch^2\|\boldsymbol{p}_d\|_2\|\phi\|_2. \tag{4.46}$$

Similar to (4.36)–(4.37), we prove that

$$D_8 + D_9 \leq Ch\|\boldsymbol{p} - \boldsymbol{p}_h(u)\| \cdot \|\phi\|_2 + Ch^2(\|\boldsymbol{p}_d\|_2 + \|\boldsymbol{p}\|_{1,\infty})\|\phi\|_2. \tag{4.47}$$

Substituting (4.42), (4.43), and (4.47) into (4.41), we derive

$$\|e_z\| \leq Ch^2 + Ch\|z - z_h(u)\| + Ch\|\boldsymbol{p} - \boldsymbol{p}_h(u)\| + Ch\|\boldsymbol{q} - \boldsymbol{q}_h(u)\| + C\|e_y\|. \tag{4.48}$$

Thus, the proof can be completed by (4.48), (4.40), (3.3), and (3.6). □

Theorem 4.1. *Let the solution corresponding to (2.4)–(2.8) is (u, y, z), and the solution served as (2.19)–(2.23) be (u_h, y_h, z_h); u_I is the centroid interpolation of u, P_h denotes L^2-projection, $z \in H^2(\Omega)$. Suppose all assumptions from the preceding lemmas hold. Then, we obtain*

$$\|u_I - u_h\| \leq Ch^2, \tag{4.49}$$

$$\|P_h z - z_h\| \leq Ch^2, \tag{4.50}$$

$$\|P_h y - y_h\| \leq Ch^2. \tag{4.51}$$

Proof. From [6, 10], we know the following inequality

$$\theta\|u_I - u_h\|^2 = \theta(u_I - u_h, u_I - u_h) \leq (u_I - u_h, z_h - z_I). \tag{4.52}$$

Notice that

$$\begin{aligned}(u_I - u_h, z_h - z_I) =&(u_I - u_h, z_h - z_h(u_I)) + (u_I - u_h, z_h(u_I) - z_h(u)) \\ &+ (u_I - u_h, z_h(u) - z) + (u_I - u_h, z - z_I) \\ =&(u_I - u_h, z_h - z_h(u_I)) + (u_I - u_h, z_h(u_I) - z_h(u)) \\ &+ (u_I - u_h, z_h(u) - P_h z) + (u_I - u_h, z - z_I).\end{aligned} \tag{4.53}$$

By setting $\boldsymbol{v}_h = \boldsymbol{q}_h(u_I) - \boldsymbol{q}_h$ in equation (4.6) and subsequently defining $w_h = z_h(u_I) - z_h$ in equation (4.7), we proceed to assign $\boldsymbol{v}_h = \boldsymbol{p}_h(u_I) - \boldsymbol{p}_h$ in equation (4.8) and finally specify $w_h = y_h(u_I) - y_h$ in equation (4.9). Using (2.28) and (2.33), it becomes evident that

$$\begin{aligned}(u_I - u_h, z_h - z_h(u_I)) &= -\|y_h(u_I) - y_h\|^2 - (\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h, \gamma_h(\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h)) \\ &\leq -C_0\|\boldsymbol{p}_h(u_I) - \boldsymbol{p}_h\|^2 - \|y_h(u_I) - y_h\|^2 \\ &\leq 0.\end{aligned} \tag{4.54}$$

Similar to (4.18), it it easy to get

$$(u_I - u_h, z - z_I) \leq Ch^2\|u_I - u_h\| \cdot |z|_{H^2(\Omega)}. \tag{4.55}$$

By referring to the equations (4.53)–(4.55) and leveraging the well-established Cauchy inequality along with the insights provided by Lemmas 4.3 and 4.4, we are able to complete the proof for the proposition labeled as 4.49.

From (4.4), (4.5), (4.12), (4.30), (4.49), and the fundamental principle of triangle inequality, it becomes an uncomplicated task to get the conclusions (4.50) and (4.51). The theorem's proof is finalized.

□

5. Numerical experiments

In this section of our discussion, we present, as detailed below, two distinct examples that are specifically designed to effectively elucidate the theoretical results we have previously outlined. Specifically, the control function, denoted as u, was meticulously discretized using piecewise constant functions, a method chosen for its simplicity and effectiveness in this context. On the other hand, the state variables, represented by $(y, \boldsymbol{p})$, and the costate variables, denoted as $(z, \boldsymbol{q})$, were approximated using the lowest-order Raviart–Thomas mixed finite element functions. To provide a clear and consistent framework for these examples, we have defined the domain of our analysis as $\Omega = [0, 1] \times [0, 1]$.

Table 1. The error outcomes for state variables.

h	$\|\|y - y_h\|\|$	Rate	$\|\|P_h y - y_h\|\|$	Rate	$\|\|\boldsymbol{p} - \boldsymbol{p}_h\|\|$	Rate
1 / 10	1.92624e-04	–	6.30811e-05	–	2.02128e-03	–
1 / 20	9.28889e-05	1.0522	1.60668e-05	1.9731	1.02668e-03	0.9773
1 / 40	4.59799e-05	1.0145	4.03602e-06	1.9931	5.15490e-04	0.9940
1 / 80	2.29306e-05	1.0037	1.01023e-06	1.9983	2.58018e-04	0.9985
1 / 160	1.14579e-05	1.0009	2.52633e-07	1.9996	1.29043e-04	0.9996
1 / 320	5.72799e-06	1.0002	6.31632e-08	1.9999	6.45260e-05	0.9999
1 / 640	2.86388e-06	1.0001	1.57911e-08	2.0000	3.22636e-05	1.0000

Table 2. The error findings corresponding to costate variables.

h	$\|\|z - z_h\|\|$	Rate	$\|\|P_h z - z_h\|\|$	Rate	$\|\|\boldsymbol{q} - \boldsymbol{q}_h\|\|$	Rate
1 / 10	1.83366e-04	–	3.28416e-04	–	2.03310e-03	–
1 / 20	9.16506e-05	1.0005	8.50535e-05	1.9491	1.02826e-03	0.9835
1 / 40	4.58223e-05	1.0001	2.14691e-05	1.9861	5.15691e-04	0.9956
1 / 80	2.29108e-05	1.0000	5.38087e-06	1.9963	2.58044e-04	0.9989
1 / 160	1.14554e-05	1.0000	1.34609e-06	1.9991	1.29047e-04	0.9997
1 / 320	5.72768e-06	1.0000	3.36577e-07	1.9998	6.45264e-05	0.9999
1 / 640	2.86384e-06	1.0000	8.41476e-08	1.9999	3.22636e-05	1.0000

Example 1. We tackle the optimal control problem of the two-dimensional elliptic type given below with $\theta = 1$ and $c = 1$.

$$\min_{u \in U_{ad}} \frac{1}{2} \left\{ \|\boldsymbol{p}(x) - \boldsymbol{p}_d\|^2 + \|y(x) - y_d\|^2 + \|u(x)\|^2 \right\} \tag{5.1}$$

under the condition of the state equation

$$\mathrm{div}\boldsymbol{p}(x) + y(x) = f(x) + u(x), \quad \boldsymbol{p}(x) = -A(x)\nabla y(x), \tag{5.2}$$

where

$$
\begin{aligned}
A(x) &= \begin{pmatrix} 1 + x_1^2 & 0 \\ 0 & 1 + x_2^2 \end{pmatrix}, \\
y(x) &= x_1^2 \cdot x_2^2 (1 - x_1)^2 (1 - x_2)^2 , \\
u(x) &= y(x) = x_1^2 \cdot x_2^2 (1 - x_1)^2 (1 - x_2)^2 , \\
\boldsymbol{p}_d &= \boldsymbol{p}(x) = \begin{pmatrix} -2(1 + x_1^2) \cdot x_1(1 - x_1)(1 - 2x_1)(1 - x_2)^2 \cdot x_2^2 \\ -2x_1^2(1 - x_1)^2 \cdot x_2(1 - x_2)(1 + x_2^2) \cdot (1 - 2x_2) \end{pmatrix}, \\
\boldsymbol{q} &= -A(x)\nabla z = \begin{pmatrix} (1 + x_1^2) \cdot 2x_1(1 - x_1)(1 - 2x_1)x_2^2(1 - x_2)^2 \\ (1 + x_2^2) \cdot x_1^2(1 - x_1)^2 \cdot 2x_2(1 - x_2)(1 - 2x_2) \end{pmatrix}, \\
z(x) &= -y(x) = -x_1^2 \cdot x_2^2 (1 - x_1)^2 (1 - x_2)^2 , \\
y_d &= y(x) - z(x) - \operatorname{div} \boldsymbol{q}(x), \\
f(x) &= \operatorname{div} \boldsymbol{p}(x) + y(x) - u(x).
\end{aligned}
$$

Table 3. The error results for the control u.

h	$\|\|u - u_h\|\|$	Rate	$\|\|u_I - u_h\|\|$	Rate
1 / 10	1.83366e-04	–	9.62000e-06	–
1 / 20	9.16506e-05	1.0005	2.51811e-06	1.9337
1 / 40	4.58223e-05	1.0001	6.37258e-07	1.9824
1 / 80	2.29108e-05	1.0000	1.59810e-07	1.9955
1 / 160	1.14554e-05	1.0000	3.99837e-08	1.9989
1 / 320	5.72768e-06	1.0000	9.99788e-09	1.9997
1 / 640	2.86384e-06	1.0000	2.49959e-09	1.9999

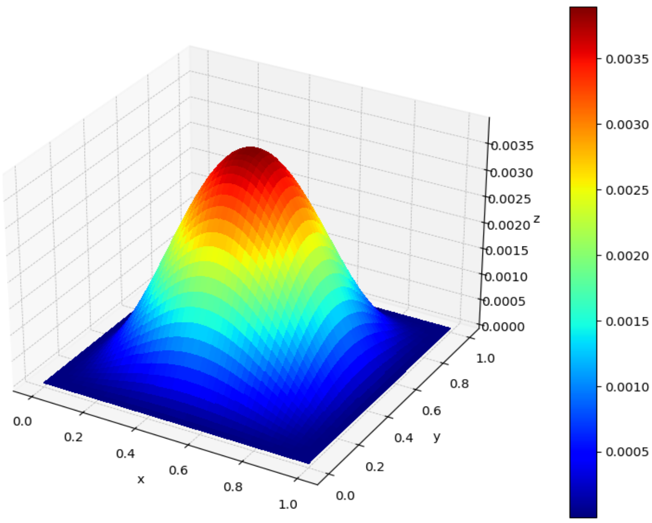

Figure 2. The contour of the numerical solution for u over a 160 × 160 triangle mesh.

In Table 1–Table 3, we put forward the detailed error metrics associated with the state variables, costate variables, and control variables. These error metrics have been derived from a systematic sequence of meshes that are refined uniformly. Specifically, the error norms $||y-y_h||$, $||\boldsymbol{p}-\boldsymbol{p}_h||$, $||z-z_h||$, $||\boldsymbol{q}-\boldsymbol{q}_h||$, and $||u-u_h||$ are all shown to exhibit first-order error estimation results, but discretization error of $||P_hy-y_h||$,$||P_hz-z_h||$ and $||u_I-u_h||$ exhibit a convergence result of order $O(h^2)$. In Figure 2–Figure 4, the detailed profiles associated with the numerical solutions for the variables u, z as well as $\boldsymbol{p}$ are presented. As is evident from Table 1–Table 3, the theoretical outcomes put forward in this study are fully validated. The data presented therein offers clear confirmation of the analytical conclusions, with all key findings aligning closely with the theoretical projections outlined earlier.

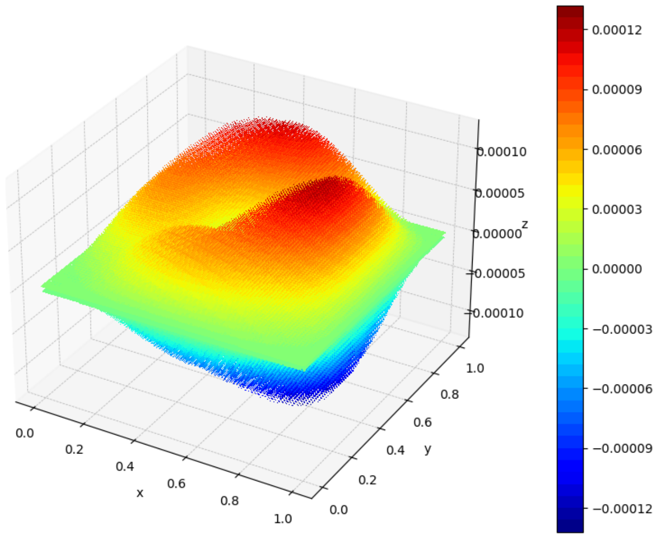

Figure 3. The contour of the numerical solution for $\boldsymbol{p}$ over a 160×160 triangle mesh.

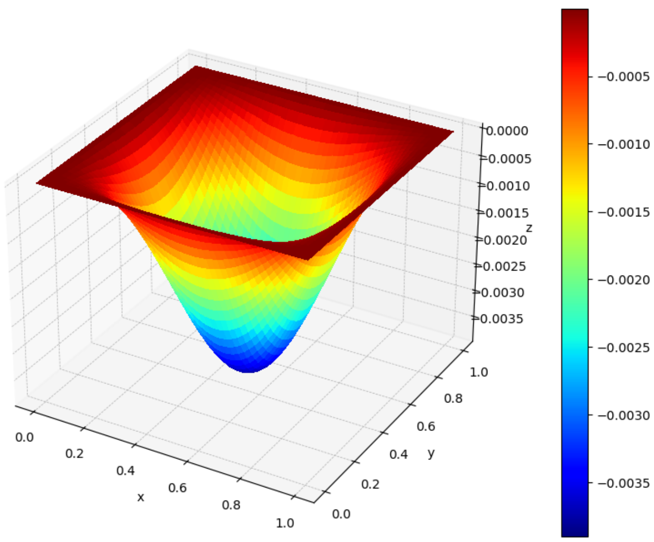

Figure 4. The contour of the numerical solution for z over a 160×160 triangle mesh.

Example 2. When setting $A = I$, $c = 40$, $\theta = 0.02$, we pick the exact solutions as shown below:

$$
\begin{aligned}
&y(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2),\\
&u(x_1, x_2) = \max\{0, -z(x_1, x_2)\}/0.02,\\
&\boldsymbol{p}_d = \boldsymbol{p}(x_1, x_2) = -\nabla y(x_1, x_2) = \begin{pmatrix} -\pi\cos(\pi x_1)\sin(\pi x_2)\\ -\pi\sin(\pi x_1)\cos(\pi x_2)\end{pmatrix},\\
&z(x_1, x_2) = \frac{1}{40+2\pi^2}\sin(\pi x_1)\cdot\sin(\pi x_2) - \frac{1}{40+8\pi^2}\sin(2\pi x_1)\cdot\sin(2\pi x_2),\\
&\boldsymbol{q}(x_1, x_2) = -\nabla z(x_1, x_2) = \begin{pmatrix} -\frac{\pi}{40+2\pi^2}\cos(\pi x_1)\cdot\sin(\pi x_2) + \frac{2\pi}{40+8\pi^2}\cos(2\pi x_1)\cdot\sin(2\pi x_2)\\ -\frac{\pi}{40+2\pi^2}\sin(\pi x_1)\cdot\cos(\pi x_2) + \frac{2\pi}{40+8\pi^2}\sin(2\pi x_1)\cdot\cos(2\pi x_2)\end{pmatrix},\\
&y_d = \sin(2\pi x_1)\cdot\sin(2\pi x_2),\\
&f(x_1, x_2) = \left[2\pi^2 + 40\right]\sin(\pi x_1)\cdot\sin(\pi x_2) - u(x_1, x_2).
\end{aligned}
\tag{5.3}
$$

Table 4. The error outcomes corresponding to state variables.

h	$\|y-y_h\|$	Rate	$\|P_h y - y_h\|$	Rate	$\|\boldsymbol{p}-\boldsymbol{p}_h\|$	Rate
1 / 10	5.21994e-02	–	3.27372e-03	–	2.01340e-01	–
1 / 20	2.61599e-02	0.9967	8.24801e-04	1.9888	1.00716e-01	0.9993
1 / 40	1.30875e-02	0.9992	2.06505e-04	1.9979	5.03638e-02	0.9998
1 / 80	6.54467e-03	0.9998	5.16401e-05	1.9996	2.51826e-02	1.0000
1 / 160	3.27245e-03	1.0000	1.29104e-05	2.0000	1.25914e-02	1.0000
1 / 320	1.63624e-03	1.0000	3.22732e-06	2.0001	6.29570e-03	1.0000
1 / 640	8.18122e-04	1.0000	8.06864e-07	1.9999	3.14785e-03	1.0000

Table 5. The error outcomes corresponding to costate variables.

h	$\|z-z_h\|$	Rate	$\|P_h z - z_h\|$	Rate	$\|\boldsymbol{q}-\boldsymbol{q}_h\|$	Rate
1 / 10	2.40957e-03	–	2.10238e-03	–	1.30367e-02	–
1 / 20	8.43589e-04	1.5142	5.32606e-04	1.9809	4.62634e-03	1.4946
1 / 40	3.60456e-04	1.2267	1.33474e-04	1.9965	2.00549e-03	1.2059
1 / 80	1.71650e-04	1.0704	3.33944e-05	1.9989	9.60433e-04	1.0622
1 / 160	8.47163e-05	1.0188	8.35117e-06	1.9996	4.74782e-04	1.0164
1 / 320	4.22183e-05	1.0048	2.08802e-06	1.9998	2.36707e-04	1.0042
1 / 640	2.10917e-05	1.0012	5.22011e-07	2.0000	1.18268e-04	1.0010

In Tables 4–6, the errors linked to the control as well as the state and costate variables–coming from a uniformly refined sequence of meshes–are provided herein. It can be seen that first-order error estimation results are also exhibited by $\|u-u_h\|$, $\|y-y_h\|$, $\|z-z_h\|$ together with $\|\boldsymbol{p}-\boldsymbol{p}_h\|$ and $\|\boldsymbol{q}-\boldsymbol{q}_h\|$, while a convergence result of order $O(h^2)$ is demonstrated by the discretization errors of $\|P_h y - y_h\|$, $\|P_h z - z_h\|$ and $\|u_I - u_h\|$. The numerical solution profiles of u, $\boldsymbol{p}$, z, y, and $\boldsymbol{q}$ are illustrated in Figure 5–Figure 9, respectively. In Table 6, the convergence outcomes of the differences between u and u_h, as well as between u_I and u_h, exhibit a gradual trend towards $O(h)$ and $O(h^2)$ convergence, respectively. It is discernible that the convergence results from $h = 1/10$ to $h = 1/80$ are not satisfactory and less favorable than those of Example 1. The convergence results from $h = 1/160$ to $h = 1/640$ become increasingly better. In finite element analysis, the "climbing" of convergence orders is a common phenomenon. The asymptotic convergence order only emerges after mesh refinement, which is consistent with the asymptotic property of MFEMs.

In the two illustrative examples provided, we demonstrate convergence behavior of the proposed algorithm.

Table 6. The error results for the control u.

h	$\|u - u_h\|$	Rate	$\|u_I - u_h\|$	Rate
1 / 10	3.30062e-02	–	3.29616e-02	–
1 / 20	1.06085e-02	1.6375	9.94640e-03	1.7285
1 / 40	3.65511e-03	1.5372	2.96218e-03	1.7475
1 / 80	1.36897e-03	1.4168	7.98279e-04	1.8917
1 / 160	5.97736e-04	1.1955	2.03089e-04	1.9748
1 / 320	2.86589e-04	1.0605	5.10182e-05	1.9930
1 / 640	1.41804e-04	1.0151	1.27679e-05	1.9985

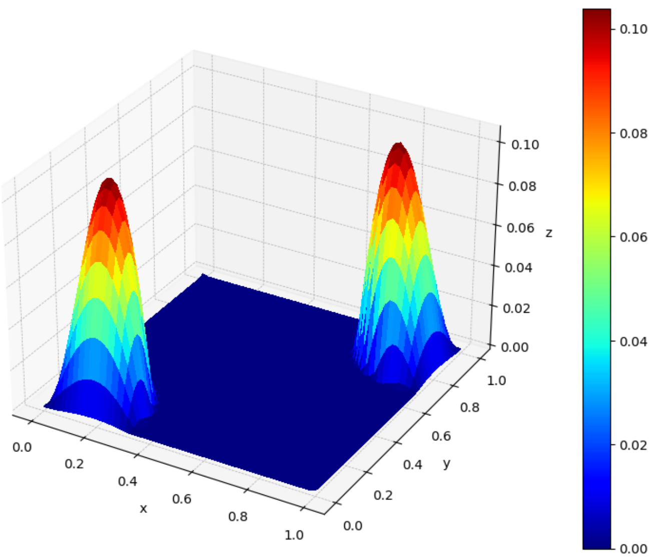

Figure 5. The contour of the numerical solution for u over a 160 × 160 triangle mesh.

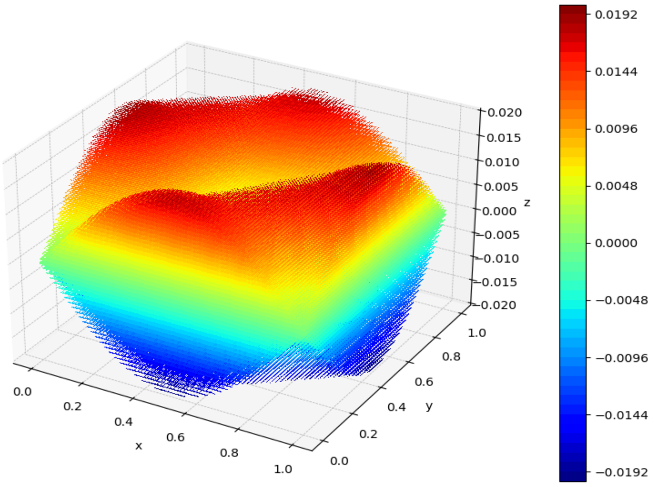

Figure 6. The contour of the numerical solution for $\boldsymbol{p}$ over a 160×160 triangle mesh.

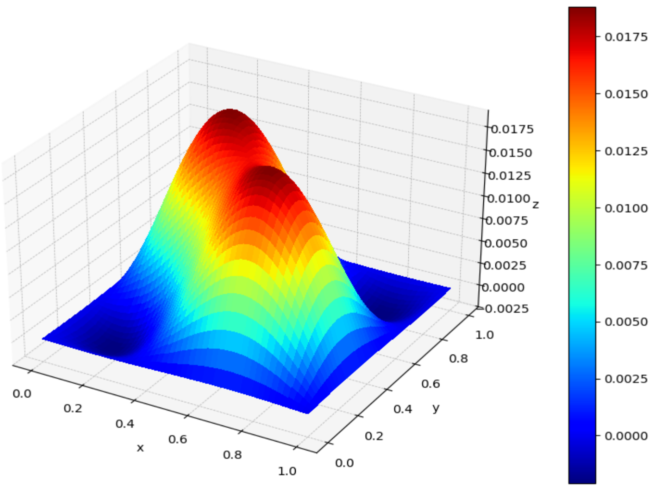

Figure 7. The contour of the numerical solution for z over a 160×160 triangle mesh.

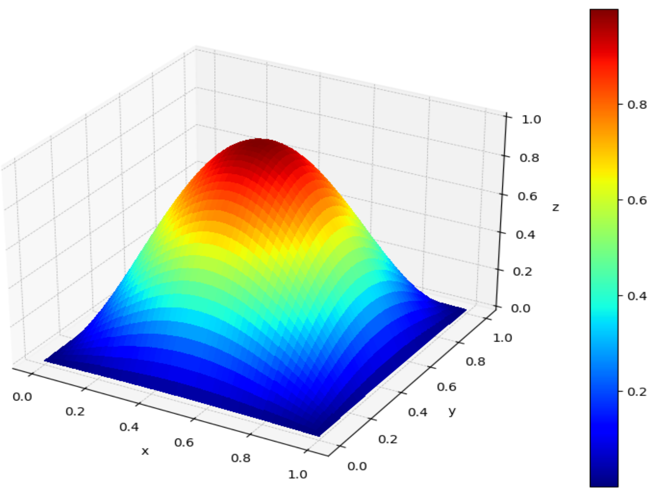

Figure 8. The contour of the numerical solution for y over a 160×160 triangle mesh.

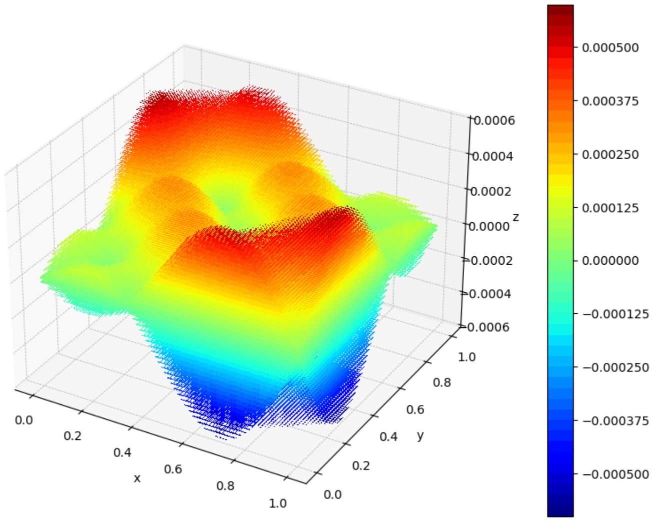

Figure 9. The contour of the numerical solution for $\boldsymbol{q}$ over a 160×160 triangle mesh.

6. Conclusions

This paper addresses our discussion on error estimates and the superconvergence of mixed covolume approximation in relation to a linear elliptic optimal control problem with pointwise control constraints. The discretized optimality condition was derived by the discretize-then-optimize approach. We demonstrated the superconvergence result as $\|u_I - u_h\| \leq Ch^2$. In future research, we will extend the framework of this paper to three-dimensional situations and conduct a systematic investigation into the a posteriori error estimates for the mixed covolume method as applied to the problem governed by equations (1.1)–(1.4). Furthermore, we shall consider optimal control problems belonging to parabolic and hyperbolic types.

Author contributions

All authors contributed to the study conception and design. All authors read and approved the final manuscript.

Use of Generative-AI tools declaration

The authors declare that they did not utilize any artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by National Natural Science Foundation of China (12571388), Natural Science Research Start-up Foundation of Recruiting Talents of Nanjing University of Posts and Telecommunications (NY223127), Natural Science Foundation of Nanjing University of Posts and Telecommunications (NY224169), National Natural Science Foundation of China Key Program (12431014), and the Science and Technology Innovation Program of Hunan Province (2024RC3159).

Conflict of interest

The authors declare that they have no conflicts of interest relevant to this study.

References

1. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Berlin: Springer, 1971.
2. J. Lin, Y. Chen, Y. Huang, A finite element method for elliptic optimal control problem in the unbounded domain, *J. Appl. Math. Comput.,* **71** (2025), 4375–4396. https://doi.org/10.1007/s12190-025-02381-8
3. R. Li, W. B. Liu, N. Yan, A posteriori error estimates of recovery type for distributed convex optimal control problems, *J. Sci. Comput.,* **33** (2007), 155–182. https://doi.org/10.1007/s10915-007-9147-7
4. X. Lin, Y. Chen, Y. Huang, A priori and a posteriori error analysis of hp spectral element discretization for optimal control problems with elliptic equations, *J. Comput. Appl. Math.,* **423** (2023), 114960. https://doi.org/10.1016/j.cam.2022.114960
5. Y. Chen, Y. Dai, Superconvergence for optimal control problems governed by semi-linear elliptic equations, *J. Sci. Comput.,* **39** (2009), 206–221. https://doi.org/10.1007/s10915-008-9258-9
6. C. Meyer, A. Rösch, Superconvergence properties of optimal control problems, *SIAM J. Control Optim.,* **43** (2004), 970–985. https://doi.org/10.1137/S0363012903431608
7. D. Yang, Y. Chang, W. Liu, A priori error estimates and superconvergence analysis for an optimal control problems of bilinear type, *J. Comput. Math.,* **4** (2008), 471–487. https://www.jstor.org/stable/43693457
8. N. Yan, Superconvergence analysis and a posteriori error estimation of a finite element method for an optimal control problem governed by integral equations, *Appl. Math.,* **54** (2009), 267–283. https://doi.org/10.1007/s10492-009-0017-5
9. Y. Chen, Superconvergence of mixed finite element methods for optimal control problems, *Math. Comp.,* **77** (2008), 1269–1291. https://doi.org/10.1090/S0025-5718-08-02104-2
10. Y. Chen, Superconvergence of quadratic optimal control problems by triangular mixed finite element methods, *Inter. J. Numer. Meths. Eng.,* **75** (2008), 881–898. https://doi.org/10.1002/nme.2272
11. Y. Chen, Y. Huang, W. B. Liu, N. Yan, Error estimates and superconvergence of mixed finite element methods for convex optimal control problems, *J. Sci. Comput.,* **42** (2010), 382–403. https://doi.org/10.1007/s10915-009-9327-8
12. T. Hou, C. Liu, Y. Yang, Error estimates and superconvergence of a mixed finite element method for elliptic optimal control problems, *Comp. Math. Appl.,* **74** (2017), 714–726. https://doi.org/10.1016/j.camwa.2017.05.021
13. T. Hou, H. Leng, T. Luan, Two-grid methods for P_0^2-P_1 mixed finite element approximation of general elliptic optimal control problems with low regularity, *Numer. Meth. Part. Differ. Equ.,* **36** (2020), 1184–1202. https://doi.org/10.1002/num.22471

14. A. Rösch, B. Vexler, Optimal control of the Stokes equations: A priori error analysis for finite element discretization with postprocessing, *SIAM J. Numer. Anal.,* **44** (2006), 1903–1920. https://doi.org/10.1137/050637364
15. X. Yang, W. Wang, Z. Zhou, H. X. Zhang, An efficient compact difference method for the fourth-order nonlocal subdiffusion problem, *Taiwan J. Math.,* **29** (2025), 35–66. https://doi.org/10.11650/tjm/240906
16. T. Liu, H. Zhang, X. Yang, The BDF2-ADI compact difference scheme on graded meshes for the three-dimensional PIDEs with multi-term weakly singular kernels, *Numer. Algorithms,* 2025, 1–34. https://doi.org/10.1007/s11075-025-02246-y
17. J. Zhang, X. Yang, S. Wang, A three-layer FDM for the Neumann initial-boundary value problem of 2D Kuramoto-Tsuzuki complex equation with strong nonlinear effects, *Commun. Nonlinear Sci. Numer. Simul.,* (2025), 109255. https://doi.org/10.1016/j.cnsns.2025.109255
18. Z. Zhang, X. Yang, S. Wang, The alternating direction implicit difference scheme and extrapolation method for a class of three dimensional hyperbolic equations with constant coefficients, *Electron. Res. Arch.,* **33** (2025), 3348–3377. https://doi.org/10.3934/era.2025148
19. T. F. Russell, Time stepping along characteristics with incomplete iteration for a Galerkin approximation of miscible displacement in porous media, *SIAM J. Numer. Anal.,* **22** (1985), 970–1013. https://doi.org/10.1137/0722059
20. X. Yang, Z. Zhang, Analysis of a new NFV scheme preserving DMP for two-dimensional sub-diffusion equation on distorted meshes, *J. Sci. Comput.,* **99** (2024), 80. https://doi.org/10.1007/s10915-024-02511-7
21. X. Yang, Z. Zhang, Superconvergence analysis of a robust orthogonal Gauss collocation method for 2D fourth-order subdiffusion equations, *J. Sci. Comput.,* **100** (2024), 62. https://doi.org/10.1007/s10915-024-02616-z
22. X. Yang, Z. Zhang, On conservative, positivity preserving, nonlinear FV scheme on distorted meshes for the multi-term nonlocal Nagumo-type equations, *Appl. Math. Lett.,* **150** (2024), 108972. https://doi.org/10.1016/j.aml.2023.108972
23. J. Douglas, E. J. Roberts, Global estimates for mixed finite element methods for second order elliptic equations, *Math. Comp.,* **44** (1985), 39–52. https://doi.org/10.1090/S0025-5718-1985-0771029-9
24. M. Fortin, F. Brezzi, *Mixed and Hybrid Finite Element Methods*, New York: Springer-Verlag, 1991. https://doi.org/10.1007/978-1-4612-3172-1
25. S. Yang, Z. W. Jiang, Mixed covolume method for parabolic problems on triangular grids, *Appl. Math. Comput.*, **215** (2009), 1251–1265. https://doi.org/10.1016/j.amc.2009.06.068
26. P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, Boston-London-Melbourne, 1985. https://doi.org/10.1137/1.9781611972030
27. J. H. Brandts, Superconvergence and a posteriori error estimation for triangular mixed finite elements, *Numer. Math.,* **68** (1994), 311–324. https://doi.org/10.1007/s002110050064