



A Priori Error Analysis for Time-Stepping Discontinuous Galerkin Finite Element Approximation of Time Fractional Optimal Control Problem

Chenyang Zhang¹ · Huipo Liu² · Zhaojie Zhou¹

Received: 10 February 2019 / Revised: 17 April 2019 / Accepted: 23 April 2019 / Published online: 2 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In this paper a priori error analysis for time-stepping discontinuous Galerkin finite element approximation of optimal control problem governed by time fractional diffusion equation is presented. A time-stepping discontinuous Galerkin finite element method and variational discretization approach are used to approximate the state and control variables respectively. Regularity of the optimal control problem is discussed. Since the time fractional derivative is nonlocal, in order to reduce the computational cost a fast gradient projection algorithm is designed for the control problem based on the block triangular Toeplitz structure of the discretized state equation and adjoint state equation. Numerical examples are carried out to illustrate the theoretical findings and fast algorithm.

Keywords Time-stepping discontinuous Galerkin method · Time fractional diffusion equation · Optimal control problem · A priori error estimate · Fast algorithm

Mathematics Subject Classification 49J20 · 49K20 · 65N15 · 65N30

1 Introduction

PDE-constrained optimal control problem plays an important role in many real world applications. There has been extensive research on numerical methods or algorithms of optimal control problems in the literatures (see, [6,9,11,18]), most of which focused on control problems governed by integer order PDEs. In recent years, many researchers have started to study numerical approximation of optimal control problems governed by fractional differ-

The first and third author would like to thank the support of Natural Science Foundation of Shandong Province (No. ZR2016JL004) and National Natural Science Foundation of China (No. 11301311). The second author is supported by National Natural Science Foundation of China (No. 11771053).

✉ Zhaojie Zhou
zhouzhaojie@sdu.edu.cn

¹ School of Mathematics and Statistics, Shandong Normal University, Jinan, China

² Institute of Applied Physics and Computational Mathematics, Beijing, China

ential equations, which are frequently met in applications, for example, anomalous process ([3,14,20]).

Based on an optimal control framework an initial value inverse problem for time fractional diffusion equation was studied in [23] by spectral Galerkin method. Spectral Galerkin approximation of state integral constrained time fractional optimal control problem was discussed in [24]. In [17], Legendre pseudo-spectral method combining with $L1$ scheme was utilized to discretize control constrained optimal control problem governed by a time-fractional diffusion equation. In [25] finite element method combined with $L1$ -scheme was applied to discretize time fractional optimal control problem with Caputo derivative and a priori error estimate for semidiscrete case was proved. Regularity of time fractional optimal control problem and a fully discrete error estimate for $L1$ and back Euler scheme were presented in [13]. A time adaptive algorithm was developed in [27] for time-stepping discontinuous Galerkin approximation of time fractional optimal control problem. In [1,2], finite element approximation of optimal control problems governed by fractional Laplacian were investigated. Finite element approximation of optimal control problem governed by one dimensional space fractional diffusion equation was investigated in [26], where a priori error estimate was derived and a fast primal-dual active set algorithm was developed. A fast gradient projection algorithm for finite difference approximation of space fractional optimal control problem was proposed in [7].

This paper is devoted to the error analysis and fast algorithm for time-stepping discontinuous Galerkin finite element approximation of control constrained optimal control problem governed by time fractional diffusion equation. A piecewise linear finite element method combined with a piecewise constant time-stepping discontinuous Galerkin method is applied to discretize the state equation. For the discretization of control variable variational discretization approach (see, [10]) is adopted. Regularity of the optimal control problem is discussed. A priori error analysis for state variable, adjoint state variable and control variable is deduced. Due to the nonlocal property of time fractional derivative the numerical schemes for time fractional equation give rise to long tails in time, which leads to higher memory requirements and increasing computational time comparing with integer differential equations. For optimal control problem this problem becomes more serious, since we usually solve the discrete first order optimality condition in an iterative manner, which consists of discrete state equation, adjoint state equation and variational inequality. To reduce the computational cost, a fast gradient projection algorithm is designed for the control problem based on the block triangular Toeplitz structure of the coefficient matrix of the discretized system. The total computational cost for solving the state and adjoint state equation is of $O(MJ \log J)$ operations, where M and J denote the number of freedom for space and time, respectively. Finally, numerical examples are carried out to illustrate the theoretical findings and fast algorithm.

The paper is organized as follows. In Sect. 2, we recall some knowledge about fractional calculus. Regularity of the optimality system is discussed in Sect. 3. In Sect. 4, time-stepping discontinuous Galerkin approximation of optimal control problem is given and a priori error estimate for state, adjoint state and control variable is derived in Sect. 5. A fast gradient projection algorithm is developed in Sect. 6 based on the block triangular Toeplitz structure of the coefficient matrix of the discretized state and adjoint state equation. Numerical results are given in Sect. 7 to verify the theoretical findings and fast algorithm.

2 Preliminary

In this section we begin with recalling some definitions and properties of fractional derivatives as well as Sobolev spaces.

For $0 < \beta < 1$, the left Caputo and Riemann-Liouville fractional derivative are defined as follows ([21])

$${}_0^C \partial_t^\beta v = \frac{1}{\Gamma(1-\beta)} \int_0^t \frac{v'(s)}{(t-s)^\beta} ds$$

and

$${}_0^R \partial_t^\beta v = \frac{1}{\Gamma(1-\beta)} \frac{d}{dt} \int_0^t \frac{v(s)}{(t-s)^\beta} ds.$$

Here $\Gamma(\cdot)$ denotes the Gamma function. Similarly, the right Caputo and Riemann-Liouville fractional derivative of order β are defined by

$${}_t^C \partial_T^\beta v = -\frac{1}{\Gamma(1-\beta)} \int_t^T \frac{v'(s)}{(s-t)^\beta} ds$$

and

$${}_t^R \partial_T^\beta v = -\frac{1}{\Gamma(1-\beta)} \frac{d}{dt} \int_t^T \frac{v(s)}{(s-t)^\beta} ds.$$

Following [21], the following relations hold

$${}_0^R \partial_t^\beta v(t) = {}_0^C \partial_t^\beta v + \frac{v(0)t^{-\beta}}{\Gamma(1-\beta)}$$

and

$${}_t^R \partial_T^\beta v(t) = {}_t^C \partial_T^\beta v + \frac{v(T)(T-t)^{-\beta}}{\Gamma(1-\beta)}.$$

We can observe that the Riemann-Liouville and Caputo fractional derivatives are equal for the homogenous initial condition or terminal condition.

For a Lebesgue measurable subset ω of R^σ ($1 \leq \sigma \leq 4$), the symbol $(p, q)_\omega$ means $\int_\omega pq$. According [8] the following properties hold.

Lemma 2.1 Assume $\beta \in (0, 1) \setminus \{0.5\}$ and $\Lambda = (a, b)$. If $v \in H_0^\beta(\Lambda)$, we have

$$({}_a^R \partial_t^\beta v, {}_t^R \partial_b^\beta v)_\Lambda = \cos(\beta\pi) |v|_{H^\beta(\Lambda)}^2.$$

Further, if $0 < \beta < 1/2$ and $v, w \in H^\beta(\Lambda)$, we have

$$\begin{aligned} ({}_a^R \partial_t^\beta v, {}_t^R \partial_b^\beta w)_\Lambda &\leq |v|_{H^\beta(\Lambda)} |w|_{H^\beta(\Lambda)}, \\ ({}_a^R \partial_t^{2\beta} v, w)_\Lambda &= ({}_a^R \partial_t^\beta v, {}_t^R \partial_b^\beta w)_\Lambda = (w, {}_t^R \partial_b^{2\beta} v)_\Lambda. \end{aligned}$$

According to [22], we introduce the space of $\dot{H}^s(\Omega)$. Let $\{\lambda_k\}_{k=1}^\infty$ and $\{\phi_k\}_{k=1}^\infty$ denote the eigenvalues (ordered nondecreasingly with multiplicity counted) and the $L^2(\Omega)$ -orthonormal eigenfunctions of $-\Delta$ operator on the domain Ω with a zero Dirichlet boundary condition. For $s \geq 0$, let $\dot{H}^s(\Omega) \subset L^2(\Omega)$ defined by

$$\dot{H}^s(\Omega) := \left\{ v = \sum_{k=0}^\infty c_k \phi_k : \left(\sum_{k=0}^\infty c_k^2 \lambda_k^s \right)^{\frac{1}{2}} < \infty \right\}.$$

By Lemma 3.1(see, [22], Chap. 3, Page 38) we have $\dot{H}^1(\Omega) = H_0^1(\Omega)$ and $\dot{H}^2(\Omega) = H^2(\Omega) \cap H_0^1(\Omega)$.

3 Optimal Control Problem

Let Ω be a bounded domain of $R^d (1 \leq d \leq 3)$ with sufficiently smooth boundary $\partial\Omega$, and $\Omega_T = \Omega \times (0, T)$, $\Gamma_T = \partial\Omega \times (0, T)$. We study the following distributed optimal control problem governed by time fractional diffusion equation

$$\min_{q \in U_{ad}} \mathcal{J}(u, q) := \frac{1}{2} \|u(x, t) - u_d(x, t)\|_{L^2(\Omega_T)}^2 + \frac{\gamma}{2} \|q(x, t)\|_{L^2(\Omega_T)}^2 \tag{3.1}$$

subject to

$$\begin{cases} \frac{\partial u}{\partial t} - {}_0^R\partial_t^\beta \Delta u = f(x, t) + q(x, t), & (x, t) \in \Omega_T, \\ u(x, t) = 0, & (x, t) \in \Gamma_T, \\ u(x, 0) = 0, & x \in \Omega. \end{cases} \tag{3.2}$$

Here U_{ad} is the admissible set defined by

$$U_{ad} = \{q \in L^2(\Omega_T) : a \leq q(x, t) \leq b \text{ a.e. in } \Omega_T \text{ with } a, b \in \mathbb{R} \text{ and } a \leq b\}.$$

Since the functional \mathcal{J} is strictly convex and the admissible set U_{ad} is bounded, closed and convex, the control problem (3.1)–(3.2) admits a unique solution.

Assumption 3.1 In the following we suppose that $\gamma > 0$, $f \in L^2(0, T; L^2(\Omega))$ and $u_d \in L^2(0, T; L^2(\Omega))$ are fixed data.

According [16] the state variable satisfies the following estimate.

Theorem 3.2 Suppose that u is the solution of state equation with right hand term f . Then we have

$$(u_t, v)_{\Omega_T} + ({}_0^R\partial_t^\beta \nabla u, \nabla v)_{\Omega_T} = (f, v)_{\Omega_T}, \forall v \in L^2(0, T; \dot{H}^1(\Omega)) \tag{3.3}$$

and

$$\|u\|_{H^1(0, T; L^2(\Omega))} + \|u\|_{H^{(1+\beta)/2}(0, T; \dot{H}^1(\Omega))} + \|u\|_{H^\beta(0, T; \dot{H}^2(\Omega))} \leq C \|f\|_{L^2(0, T; L^2(\Omega))}.$$

For above optimal control problem we can derive the following first order optimality conditions ([27]).

Theorem 3.3 Assume that $q \in U_{ad}$ is the solution to optimal control problem (3.1)–(3.2) and u is the corresponding state variable given by (3.2). Then there exists an adjoint state such that (u, z, q) satisfies the following optimality conditions:

$$\begin{cases} \frac{\partial u}{\partial t} - {}_0^R\partial_t^\beta \Delta u = f + q, & (x, t) \in \Omega_T, \\ u(x, t) = 0, & (x, t) \in \Gamma_T, \\ u(x, 0) = 0, & x \in \Omega, \end{cases} \tag{3.4}$$

$$\begin{cases} -\frac{\partial z}{\partial t} - {}_t^R\partial_T^\beta \Delta z = u - u_d, & (x, t) \in \Omega_T, \\ z(x, t) = 0, & (x, t) \in \Gamma_T, \\ z(x, T) = 0, & x \in \Omega \end{cases} \tag{3.5}$$

and

$$\int_{\Omega_T} (\gamma q + z)(v - q) \geq 0, \quad \forall v \in U_{ad}. \tag{3.6}$$

Let

$$P_{U_{ad}}(q(x, t)) = \max\{a, \min(q(x, t), b)\}$$

denote the pointwise projection onto the admissible set U_{ad} . Then (3.6) is equivalent to

$$q = P_{U_{ad}}\left(-\frac{1}{\gamma}z\right).$$

Similar to state equation, the adjoint state variable admits the following estimate ([16]).

Theorem 3.4 *Let z be the solution of adjoint state equation with right hand term g . Then we have*

$$-(z_t, v)_{\Omega_T} + ({}^R_0\partial_T^\beta \nabla z, \nabla v)_{\Omega_T} = (g, v)_{\Omega_T}, \quad \forall v \in L^2(0, T; \dot{H}^1(\Omega)) \tag{3.7}$$

and

$$\|z\|_{H^1(0,T;L^2(\Omega))} + \|z\|_{H^{(1+\beta)/2}(0,T;\dot{H}^1(\Omega))} + \|z\|_{H^\beta(0,T;\dot{H}^2(\Omega))} \leq C\|g\|_{L^2(0,T;L^2(\Omega))}.$$

For the optimality system (3.4)–(3.6) we have the following regularity result.

Theorem 3.5 *Supposed that (u, z, q) is the solution of optimality system (3.4)–(3.6). Then we have*

$$u(x), z(x) \in H^1(0, T; L^2(\Omega)) \cap H^\beta(0, T; \dot{H}^2(\Omega)) \cap H^{(1+\beta)/2}(0, T; \dot{H}^1(\Omega))$$

and

$$q(x) \in H^1(0, T; L^2(\Omega)).$$

Proof Note that $f \in L^2(0, T; L^2(\Omega))$, $y_d \in L^2(0, T; L^2(\Omega))$ and $q \in U_{ad}$. Then by Theorems 3.3 and 3.4 we have

$$u(x), z(x) \in H^1(0, T; L^2(\Omega)) \cap H^\beta(0, T; \dot{H}^2(\Omega)) \cap H^{(1+\beta)/2}(0, T; \dot{H}^1(\Omega)).$$

For the projection operator $P_{U_{ad}}$ the following property holds (see, [15])

$$\|P_{U_{ad}}(v)\|_{W^{s,p}(0,T;L^2(\Omega))} \leq C\|v\|_{W^{s,p}(0,T;L^2(\Omega))}, \quad 0 \leq s \leq 1, \quad 1 \leq p \leq \infty.$$

This implies that

$$q(x) \in H^1(0, T; L^2(\Omega)).$$

□

4 Time Stepping Discontinuous Galerkin Discrete Scheme

Let V_h be the finite element space consisting of continuous piecewise linear functions over the triangulation T_h :

$$V_h = \{v_h \in H_0^1(\Omega) \cap C(\Omega); v_h \text{ is a linear function over } K, \forall K \in T_h\}.$$

Let $\Delta_\tau : 0 = t_0 < t_1 < \dots < t_{J-1} < t_J = T$ be a time grid with $\tau = T/J$. Set $I_j = (t_{j-1}, t_j)$ for each $1 \leq j \leq J$. Define the fully discrete finite element space

$$V_{hk} = \{\phi : \Omega \times [0, T] \rightarrow R; \phi(x, \cdot) \in V_h, \phi(\cdot, t)|_{I_j} \in P_0, j = 1, 2, \dots, J\}.$$

This implies that ϕ is a piecewise constant with respect to time. For $\phi \in V_{hk}, t \in I_j$, we define

$$\phi_j^\pm = \lim_{t \rightarrow t_j^\pm} \phi(t),$$

and

$$[[\phi_j]] = \phi_j^+ - \phi_j^-, \quad [[\phi^0]] = \phi_0^+.$$

Then the time-stepping discontinuous Galerkin approximation of the state equation can be written as

$$\sum_{j=0}^{J-1} ([[U_j]], w_j^+)_{\Omega} + ({}^R_0\partial_t^\beta \nabla U, \nabla w)_{\Omega_T} = (f + q, w)_{\Omega_T}, \forall w \in V_{hk}.$$

The time-stepping discontinuous Galerkin discrete scheme for the control problem (3.1)–(3.2) is to find $(U, Q) \in V_{hk} \times U_{ad}$, such that

$$\min_{Q \in U_{ad}} \mathcal{J}(U, Q)$$

subject to

$$\sum_{j=0}^{J-1} ([[U_j]], w_j^+)_{\Omega} + ({}^R_0\partial_t^\beta \nabla U, \nabla w)_{\Omega_T} = (f + Q, w)_{\Omega_T}, \forall w \in V_{hk}. \tag{4.1}$$

Here the control variable was implicitly discretized by variational discretization concept(see, [10]).

In order to obtain the discrete first-order optimal control conditions for the optimal control problem, we define the Lagrange functional

$$\begin{aligned} \mathcal{L}(U, Z, Q) &= \mathcal{J}(U, Q) + (f + Q, Z)_{\Omega_T} \\ &\quad - \sum_{j=0}^{J-1} ([[U_j]], Z_j^+)_{\Omega} - ({}^R_0\partial_t^\beta \nabla U, \nabla Z)_{\Omega_T}, Z \in V_{hk}. \end{aligned} \tag{4.2}$$

Then the discrete first order optimality condition can be deduced by computing

$$\frac{\partial \mathcal{L}(U, Z, Q)}{\partial U}(w) = 0, \frac{\partial \mathcal{L}(U, Z, Q)}{\partial Z}(w) = 0, \frac{\partial \mathcal{L}(U, Z, Q)}{\partial Q}(v - Q) \geq 0.$$

– Discrete adjoint state equation: Note that

$$\begin{aligned} &\frac{\partial \mathcal{L}(U, Z, Q)}{\partial U}(w) \\ &= \lim_{t \rightarrow 0^+} \frac{1}{t} (\mathcal{L}(U + tw, Z, Q) - \mathcal{L}(U, Z, Q)) \\ &= \lim_{t \rightarrow 0^+} \frac{1}{t} \left(- \sum_{j=0}^{J-1} ([[U + tw]_j], Z_j^+)_{\Omega} - ({}^R_0\partial_t^\beta \nabla(U + tw), \nabla Z)_{\Omega_T} + \frac{1}{2} \int_{\Omega_T} (U + tw - u_d)^2 dx dt \right) \end{aligned}$$

$$\begin{aligned}
 & + \sum_{j=0}^{J-1} ([U_j], Z_j^+)_{\Omega} + ({}_0^R \partial_t^\beta \nabla U, \nabla Z)_{\Omega_T} - \frac{1}{2} \int_{\Omega_T} (U - u_d)^2 dxdt \Big) \\
 = & \lim_{t \rightarrow 0^+} \frac{1}{t} \left(- \sum_{j=0}^{J-1} ([U_j], Z_j^+)_{\Omega} - ({}_0^R \partial_t^\beta \nabla(U), \nabla Z)_{\Omega_T} - \frac{1}{2} \int_{\Omega_T} (2U + t w - 2u_d) t w dxdt \right) \\
 = & - \sum_{j=0}^{J-1} ([w_j], Z_j^+)_{\Omega} - ({}_t^R \partial_T^\beta \nabla Z, \nabla w)_{\Omega_T} + \int_{\Omega_T} (U - u_d) w dxdt.
 \end{aligned}$$

Then we have

$$\sum_{j=0}^{J-1} ([w_j], Z_j^+)_{\Omega} + ({}_t^R \partial_T^\beta \nabla Z, \nabla w)_{\Omega_T} = (U - u_d, w)_{\Omega_T}, \forall w \in V_{hk}. \tag{4.3}$$

– Discrete variational inequality: Note that

$$\begin{aligned}
 \frac{\partial \mathcal{L}(U, Z, Q)}{\partial Q}(v - Q) & = \lim_{t \rightarrow 0^+} \frac{1}{t} \left(\mathcal{L}(U, Z, Q + t(v - Q)) - \mathcal{L}(U, Z, Q) \right) \\
 & = \lim_{t \rightarrow 0^+} \frac{1}{t} \left((f + Q + t(v - Q), Z)_{\Omega_T} + \frac{\gamma}{2} \int_{\Omega_T} (Q + t(v - Q))^2 dxdt \right. \\
 & \quad \left. - (f + Q, Z)_{\Omega_T} - \frac{\gamma}{2} \int_{\Omega_T} Q^2 dxdt \right) \\
 & = \lim_{t \rightarrow 0^+} \frac{1}{t} \left((t(v - Q), Z)_{\Omega_T} + \frac{\gamma}{2} \int_{\Omega_T} (2Q + t(v - Q)) t(v - Q) dxdt \right) \\
 & = \int_{\Omega_T} (v - Q) Z dxdt + \gamma \int_{\Omega_T} Q(v - Q) dxdt.
 \end{aligned}$$

This gives

$$(\gamma Q + Z, v - Q)_{\Omega_T} \geq 0. \tag{4.4}$$

Then the first order optimal condition for the discrete optimal control problem can be written as

$$\begin{cases} \sum_{j=0}^{J-1} ([U_j], w_j^+)_{\Omega} + ({}_0^R \partial_t^\beta \nabla U, \nabla w)_{\Omega_T} = (f + Q, w)_{\Omega_T}, \forall w \in V_{hk}, \\ \sum_{j=0}^{J-1} ([w_j], Z_j^+)_{\Omega} + ({}_t^R \partial_T^\beta \nabla Z, \nabla w)_{\Omega_T} = (U - u_d, w)_{\Omega_T}, \forall w \in V_{hk}, \\ (\gamma Q + Z, v - Q)_{\Omega_T} \geq 0, \forall v \in U_{ad}. \end{cases} \tag{4.5}$$

5 Error Analysis

In this section we are going to derive the error estimate of $\|u - U\|_{L^2(0,T;L^2(\Omega))}$, $\|z - Z\|_{L^2(0,T;L^2(\Omega))}$ and $\|q - Q\|_{L^2(0,T;L^2(\Omega))}$. Note the the state variable and control variable are coupled together. In the following we firstly sketch the overall proof strategy.

To decouple the estimate we firstly decompose $u - U$ and $z - Z$ as follows

$$\begin{aligned}
 u - U & = u - U(q) + U(q) - U, \\
 z - Z & = z - Z(u) + Z(u) - Z,
 \end{aligned}$$

where $U(q)$ and $Z(u)$ are auxiliary variables defined by

$$\sum_{j=0}^{J-1} ([[U(q)]_j], w_j^+)_{\Omega} + ({}^R_0\partial_t^\beta \nabla U(q), \nabla w)_{\Omega_T} = (f + q, w)_{\Omega_T}, \quad \forall w \in V_{hk} \quad (5.1)$$

and

$$\sum_{j=0}^{J-1} ([[w_j]], Z(u)_j^+)_{\Omega} + ({}^R_t\partial_T^\beta \nabla Z(u), \nabla w)_{\Omega_T} = (u - u_d, w)_{\Omega_T}, \quad \forall w \in V_{hk}. \quad (5.2)$$

In Sects. 5.1 and 5.2 we will prove that the estimate of $\|u - U\|_{L^2(0,T;L^2(\Omega))}$ and $\|z - Z\|_{L^2(0,T;L^2(\Omega))}$ are controlled by the estimate of $\|q - Q\|_{L^2(0,T;L^2(\Omega))}$. In order to finish the whole estimate we just need to derive the error estimate of $\|q - Q\|_{L^2(0,T;L^2(\Omega))}$. By introducing another auxiliary variable $Z(q)$, we obtain the error estimate of $\|q - Q\|_{L^2(0,T;L^2(\Omega))}$ in Sect. 5.3. Then combining the results obtained in Sects. 5.1–5.3 yields the final error estimate of $\|u - U\|_{L^2(0,T;L^2(\Omega))}$, $\|z - Z\|_{L^2(0,T;L^2(\Omega))}$ and $\|q - Q\|_{L^2(0,T;L^2(\Omega))}$, i.e., Theorem 5.13.

To achieve the final results we need to introduce some interpolation and projection operators. Suppose that X is a Banach space. For $v \in C((0, T]; X)$ and $v \in C([0, T]; X)$ we respectively define

$$(P_\tau v)|_{I_j} = v(t_j), \quad \forall 1 \leq j \leq J,$$

and

$$(G_\tau v)|_{I_j} = v(t_{j-1}), \quad \forall 1 \leq j \leq J.$$

According [16] the following estimates hold.

Lemma 5.1 *If $v \in H^\sigma(0, T)$, $0 \leq \alpha < 1/2 < \sigma < 1$, we have*

$$\begin{aligned} |(I - P_\tau)v|_{H^\alpha(0,T)} &\leq C\tau^{\sigma-\alpha} \|v\|_{H^\sigma(0,T)}, \\ |(I - G_\tau)v|_{H^\alpha(0,T)} &\leq C\tau^{\sigma-\alpha} \|v\|_{H^\sigma(0,T)}. \end{aligned}$$

Let $P_h : L^2(\Omega) \rightarrow V_h$ denote the $L^2(\Omega)$ –orthogonal projection operator defined by

$$(P_h\varphi, \chi) = (\varphi, \chi), \quad \forall \chi \in V_h,$$

which satisfies the following estimates ([12]): if $v \in \dot{H}^\sigma(\Omega)$ with $\sigma = 1, 2$,

$$\|v - P_h v\|_{L^2(\Omega)} + h \|v - P_h v\|_{\dot{H}^1(\Omega)} \leq h^\sigma \|v\|_{\dot{H}^\sigma(\Omega)}.$$

The following lemmas can be found in [22], which will be used in the following error analysis.

Lemma 5.2 *Assume that $v \in L^2(0, T; L^2(\Omega))$, and $v' \in L^2(0, T; L^2(\Omega))$, then we have*

$$\begin{aligned} (v', V)_{\Omega_T} &= (v(T), V_J^-)_{\Omega} - \sum_{j=0}^{J-1} ([[V_j]], (G_\tau P_h v)_i^+)_{\Omega}, \\ (v', V)_{\Omega \times (0,t_j)} &= \sum_{i=0}^{j-1} ([[(P_\tau P_h v)_i]], V_i^+)_{\Omega} - (v(0), V_0^+)_{\Omega}, \end{aligned}$$

for all $V \in V_{hk}$, and $1 \leq j \leq J$.

Lemma 5.3 *If $V \in V_{hk}$, then the following estimate holds*

$$\frac{1}{2} \left(\|V_j^-\|_{L^2(\Omega)}^2 + \|V_0^+\|_{L^2(\Omega)}^2 \right) \leq \sum_{i=0}^{j-1} (\llbracket V_i \rrbracket, V_i^+)_{\Omega},$$

for all $1 \leq j \leq J$.

5.1 Error Estimate for u-U

It is easy to see that $U(q)$ is the finite element approximation of u . Then according to [16], we can get the following error estimate.

Theorem 5.4 *If $f + q \in L^2(0, T; L^2(\Omega))$, then we have*

$$\|u - U(q)\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + \tau)\|f + q\|_{L^2(0,T;L^2(\Omega))}.$$

Nextly we turn to estimate $\|U(q) - U\|_{L^2(0,T;L^2(\Omega))}$.

Theorem 5.5 *Assume that $U(q)$ and U are the solution of the Eqs. (5.1) and (4.1), respectively. Then we have*

$$\|U(q) - U\|_{L^2(0,T;L^2(\Omega))} \leq C\|q - Q\|_{L^2(0,T;L^2(\Omega))}.$$

Proof Choosing $w = U(q) - U$ in (4.1) and (5.1), we deduce

$$\begin{aligned} & \sum_{j=0}^{J-1} (\llbracket (U(q) - U)_j \rrbracket, (U(q) - U)_j^+)_{\Omega} + ({}^R_0\partial_t^\beta \nabla(U(q) - U), \nabla(U(q) - U))_{\Omega_T} \\ &= (q - Q, U(q) - U)_{\Omega_T}. \end{aligned}$$

By Lemma 5.3, we deduce

$$\frac{1}{2} \|(U(q) - U)_J^-\|_{L^2(\Omega)}^2 \leq \sum_{j=0}^{J-1} (\llbracket (U(q) - U)_j \rrbracket, (U(q) - U)_j^+)_{\Omega}.$$

By Lemma 2.1, we obtain

$$({}^R_0\partial_t^\beta \nabla(U(q) - U), \nabla(U(q) - U))_{\Omega_T} = \cos\left(\frac{\beta}{2}\pi\right) |U(q) - U|_{H^{\beta/2}(0,T;H^1(\Omega))}^2.$$

Note that $0 < \beta < 1$ and $U(q) - U$ vanishes on $\partial\Omega$. Then we obtain

$$|U(q) - U|_{H^{\beta/2}(0,T;H^1(\Omega))}^2 \sim \|U(q) - U\|_{H^{\beta/2}(0,T;H^1(\Omega))}^2.$$

This implies that

$$\|U(q) - U\|_{L^2(0,T;L^2(\Omega))} \leq C|U(q) - U|_{H^{\beta/2}(0,T;H^1(\Omega))}.$$

Collecting above inequalities and using Young inequality leads to

$$\begin{aligned} & \|(U(q) - U)_J^-\|_{L^2(\Omega)}^2 + \cos\left(\frac{\beta}{2}\pi\right) |U(q) - U|_{H^{\beta/2}(0,T;H^1(\Omega))}^2 \\ & \leq C\|q - Q\|_{L^2(0,T;L^2(\Omega))}^2. \end{aligned}$$

Theorem result follows above estimate. □

By Theorems 5.4 and 5.5, we obtain the estimate of $\|u - U\|_{L^2(0,T;L^2(\Omega))}$.

Theorem 5.6 Assume that u and U are the solution of the Eqs. (3.4) and (4.1), respectively. Then the following estimate

$$\|u - U\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + \tau)\|f + q\|_{L^2(0,T;L^2(\Omega))} + C\|q - Q\|_{L^2(0,T;L^2(\Omega))}$$

holds.

5.2 Error Estimates for z-Z

Note that $Z(u)$ is the finite element approximation of z . Then in the following we are going to prove the estimate of $\|z - Z(u)\|_{L^2(0,T;L^2(\Omega))}$ by dual argument. To this end we firstly prove some auxiliary results.

Lemma 5.7 Assume that z and $Z(u)$ are solutions of (3.5) and (5.2). Then we have

$$\|(Z(u) - G_\tau P_h z)_j^+ \|_{L^2(\Omega)} + \|Z(u) - G_\tau P_h z\|_{H^{\beta/2}(0,t_j; \dot{H}^1(\Omega))} \leq C\|(I - G_\tau P_h)z\|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))}.$$

Proof Let $\theta = Z(u) - G_\tau P_h z$. Choosing $v = \theta_{\chi(0,t_j)}$, (3.7) reduce to

$$-(z', \theta)_{\Omega \times (0,t_j)} + ({}^R_t \partial_T^\beta \nabla z, \nabla \theta)_{\Omega \times (0,t_j)} = (u - u_d, \theta)_{\Omega \times (0,t_j)}.$$

Setting $w = \theta_{\chi(0,t_j)}$ in (5.2) we obtain

$$\sum_{i=0}^{j-1} ([[\theta_i]], Z(u)_i^+)_{\Omega} + ({}^R_t \partial_T^\beta \nabla Z(u), \nabla \theta)_{\Omega \times (0,t_j)} = (u - u_d, \theta)_{\Omega \times (0,t_j)}.$$

Subtracting above two formulas leads to

$$-(z', \theta)_{\Omega \times (0,t_j)} + ({}^R_t \partial_T^\beta \nabla(z - Z(u)), \nabla \theta)_{\Omega \times (0,t_j)} - \sum_{i=0}^{j-1} ([[\theta_i]], Z(u)_i^+)_{\Omega} = 0.$$

By Lemma 5.2 and $z(T) = 0$ we have

$$-(z', \theta)_{\Omega \times (0,t_j)} = \sum_{i=0}^{j-1} ([[\theta_i]], (G_\tau P_h z)_i^+)_{\Omega}.$$

A simple calculation then yields

$$({}^R_t \partial_T^\beta \nabla(z - Z(u)), \nabla \theta)_{\Omega \times (0,t_j)} - \sum_{i=0}^{j-1} ([[\theta_i]], \theta_i^+)_{\Omega} = 0.$$

Note that $z - Z(u) = z - G_\tau P_h z + G_\tau P_h z - Z(u)$. Then we obtain

$$({}^R_t \partial_T^\beta \nabla(z - G_\tau P_h z), \nabla \theta)_{\Omega \times (0,t_j)} = ({}^R_t \partial_T^\beta \nabla \theta, \nabla \theta)_{\Omega \times (0,t_j)} + \sum_{i=0}^{j-1} ([[\theta_i]], \theta_i^+)_{\Omega}.$$

According Lemma 5.3, we have

$$\frac{1}{2}(\|\theta_j^-\|_{L^2(\Omega)}^2 + \|\theta_0^+\|_{L^2(\Omega)}^2) \leq \sum_{i=0}^{j-1} ([[\theta_i]], \theta_i^+)_{\Omega}.$$

This implies

$$\begin{aligned} & \frac{1}{2} \left(\|\theta_j^-\|_{L^2(\Omega)}^2 + \|\theta_0^+\|_{L^2(\Omega)}^2 \right) + |\theta|_{H^{\beta/2}(0,t_j; \dot{H}^1(\Omega))}^2 \\ & \leq C|z - G_\tau P_h z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} |\theta|_{H^{\beta/2}(0,t_j; \dot{H}^1(\Omega))}. \end{aligned}$$

By the Young inequality, we can deduce the theorem result. □

Lemma 5.8 *Assume that z and $Z(u)$ are solutions of (3.5) and (5.2). Then we derive*

$$\begin{aligned} & \max_{0 \leq j \leq J-1} \|z(t_j) - Z(u)_j^+\|_{L^2(\Omega)} + |z - Z(u)|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq C(h + \tau^{1/2}) \|u - u_d\|_{L^2(0,T; L^2(\Omega))}. \end{aligned} \tag{5.3}$$

Proof By the estimate of L^2 projection and the definition of G_τ , we have

$$\begin{aligned} & \|z(t_j) - (G_\tau P_h z)_j^+\|_{L^2(\Omega)} = \|(I - P_h)z(t_j)\|_{L^2(\Omega)} \\ & \leq h \|z(t_j)\|_{\dot{H}^1(\Omega)} \leq h \|z\|_{H^{(1+\beta)/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq h \|u - u_d\|_{L^2(0,T; L^2(\Omega))}. \end{aligned} \tag{5.4}$$

Let $z - G_\tau P_h z = z - P_h z + P_h z - G_\tau P_h z$. Using Lemma 5.1 gives

$$\begin{aligned} & |z - G_\tau P_h z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq |z - P_h z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} + |(I - G_\tau)P_h z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq h |z|_{H^{\beta/2}(0,T; \dot{H}^2(\Omega))} + |(I - G_\tau)z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq h |z|_{H^{\beta/2}(0,T; \dot{H}^2(\Omega))} + \tau^{1/2} |z|_{H^{(1+\beta)/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq C(h + \tau^{1/2}) \|u - u_d\|_{L^2(0,T; L^2(\Omega))}. \end{aligned} \tag{5.5}$$

Combining (5.4) and (5.5) as well as Lemma 5.7 leads to

$$\begin{aligned} & \|z(t_j) - Z(u)_j^+\|_{L^2(\Omega)} + |z - Z(u)|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq \|z(t_j) - (G_\tau P_h z)_j^+\|_{L^2(\Omega)} + \|(G_\tau P_h z)_j^+ - Z(u)_j^+\|_{L^2(\Omega)} \\ & \quad + |z - G_\tau P_h z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} + |G_\tau P_h z - Z(u)|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq \|z(t_j) - (G_\tau P_h z)_j^+\|_{L^2(\Omega)} + |z - G_\tau P_h z|_{H^{\beta/2}(0,T; \dot{H}^1(\Omega))} \\ & \leq C(h + \tau^{1/2}) \|u - u_d\|_{L^2(0,T; L^2(\Omega))}. \end{aligned} \tag{5.6}$$

Theorem 5.9 *Let z and $Z(u)$ be the solution of (3.5) and (5.2), respectively. Then the following estimate holds*

$$\|z - Z(u)\|_{L^2(0,T; L^2(\Omega))} \leq C(h^2 + \tau) \|u - u_d\|_{L^2(0,T; L^2(\Omega))}.$$

Proof To drive the estimate of $\|z - Z(u)\|_{L^2(0,T; L^2(\Omega))}$, we introduce the following auxiliary problem:

$$\begin{cases} (y', w)_{\Omega_T} + ({}^R_0 \partial_t^\beta \nabla y, \nabla w)_{\Omega_T} = (z - Z(u), w)_{\Omega_T}, & w \in L^2(0, T; \dot{H}^1(\Omega)), \\ y(\cdot, 0) = 0. \end{cases}$$

It is easy to see that y satisfies the stability estimate presented in Theorem 3.2.

Choosing $w = z - Z(u)$ yields

$$\|z - Z(u)\|_{L^2(0,T; L^2(\Omega))}^2 = (y', z - Z(u))_{\Omega_T} + ({}^R_0 \partial_t^\beta \nabla y, \nabla(z - Z(u)))_{\Omega_T}.$$

Using integration by parts we obtain

$$\|z - Z(u)\|_{L^2(0,T;L^2(\Omega))}^2 = -(z', y)_{\Omega_T} - (y', Z(u))_{\Omega_T} + ({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla y)_{\Omega_T}.$$

Choosing $w = Y = P_\tau P_h y$, (5.2) reduce to

$$\sum_{j=0}^{J-1} ([Y_j], Z(u)_j^+)_{\Omega} + ({}_t^R\partial_T^\beta \nabla Z(u), \nabla Y)_{\Omega_T} = (u - u_d, Y)_{\Omega_T}.$$

By Theorem 3.4 and $v = Y$, we have

$$-(z', Y)_{\Omega_T} + ({}_t^R\partial_T^\beta \nabla z, \nabla Y)_{\Omega_T} = (u - u_d, Y)_{\Omega_T}.$$

Subtracting above two equations leads to

$$({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla Y)_{\Omega_T} = (z', Y)_{\Omega_T} + \sum_{j=0}^{J-1} ([Y_j], Z(u)_j^+)_{\Omega}.$$

Note that

$$\begin{aligned} &({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla y)_{\Omega_T} \\ &= ({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla(y - Y))_{\Omega_T} + ({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla Y)_{\Omega_T} \\ &= ({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla(y - Y))_{\Omega_T} + (z', Y)_{\Omega_T} + \sum_{j=0}^{J-1} ([Y_j], Z(u)_j^+)_{\Omega}. \end{aligned}$$

By Lemma 5.2 and $y(\cdot, 0)$, we have

$$\sum_{j=0}^{J-1} ([Y_j], Z(u)_j^+)_{\Omega} = (y', Z(u))_{\Omega_T}.$$

Therefore we further have

$$\begin{aligned} \|z - Z(u)\|_{L^2(0,T;L^2(\Omega))}^2 &= ({}_t^R\partial_T^\beta \nabla(z - Z(u)), \nabla(y - Y))_{\Omega_T} - (z', y - Y)_{\Omega_T} \\ &\leq |z - Z(u)|_{H^{\beta/2}(0,T;\dot{H}^1(\Omega))} |y - Y|_{H^{\beta/2}(0,T;\dot{H}^1(\Omega))} \\ &\quad + \|z'\|_{L^2(0,T;L^2(\Omega))} \|y - Y\|_{L^2(0,T;L^2(\Omega))}. \end{aligned} \tag{5.6}$$

Using similar proof of Lemma 5.8, we have

$$|y - Y|_{H^{\beta/2}(0,T;\dot{H}^1(\Omega))} \leq C(h + \tau^{1/2}) \|z - Z(u)\|_{L^2(0,T;L^2(\Omega))} \tag{5.7}$$

and according [5], we have

$$\|y - Y\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + \tau) \|z - Z(u)\|_{L^2(0,T;L^2(\Omega))}. \tag{5.8}$$

Inserting (5.3), (5.7) and (5.8) into (5.6) yields the theorem result. □

Now it remains to estimate $\|Z(u) - Z\|_{L^2(0,T;L^2(\Omega))}$.

Theorem 5.10 *Assume that $Z(u)$ and Z be the solution of (5.2) and (4.3). Then the following estimate holds*

$$\|Z(u) - Z\|_{L^2(0,T;L^2(\Omega))} \leq C\|u - U\|_{L^2(0,T;L^2(\Omega))}.$$

Proof By (4.3) and (5.2), we obtain

$$\sum_{j=0}^{J-1} ([w_j], (Z(u) - Z)_j^+)_{\Omega} + ({}^R_0\partial_T^\beta \nabla(Z(u) - Z), \nabla w)_{\Omega_T} = (u - U, w)_{\Omega_T}.$$

Setting $w = Z(u) - Z$ and using Lemma 5.3, we have

$$\frac{1}{2} \|(Z(u) - Z)_0^+\|_{L^2(\Omega)}^2 \leq \sum_{j=0}^{J-1} ([Z(u) - Z]_j, (Z(u) - Z)_j^+)_{\Omega}. \tag{5.9}$$

By Lemma 2.1, we have

$$({}^R_0\partial_T^\beta \nabla(Z(u) - Z), \nabla(Z(u) - Z))_{\Omega_T} = \cos\left(\frac{\beta}{2}\pi\right) |Z(u) - Z|_{H^{\beta/2}(0,T;H^1(\Omega))}^2.$$

Note that

$$\|Z(u) - Z\|_{L^2(0,T;L^2(\Omega))} \leq C|Z(u) - Z|_{H^{\beta/2}(0,T;H^1(\Omega))}. \tag{5.10}$$

In analogous to the proof of Theorem 5.5, using (5.9) and (5.10) as well as Young inequality results in

$$\begin{aligned} & \|(Z(u) - Z)_0^+\|_{L^2(\Omega)}^2 + \cos\left(\frac{\beta}{2}\pi\right) |Z(u) - Z|_{H^{\beta/2}(0,T;H^1(\Omega))}^2 \\ & \leq C\|u - U\|_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

Again by (5.10) we obtain

$$\|(Z(u) - Z)_0^+\|_{L^2(\Omega)}^2 + \|Z(u) - Z\|_{L^2(0,T;L^2(\Omega))}^2 \leq C\|u - U\|_{L^2(0,T;L^2(\Omega))}^2.$$

□

Combining Theorems 5.9 and 5.10 yields the error estimate of the adjoint state variable $\|z - Z\|_{L^2(0,T;L^2(\Omega))}$.

Theorem 5.11 Assume that z and Z be the solution of the Eqs. (3.7) and (4.3). Then we have

$$\|z - Z\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + \tau)\|u - u_d\|_{L^2(0,T;L^2(\Omega))} + C\|u - U\|_{L^2(0,T;L^2(\Omega))}.$$

5.3 Error Estimates for q - Q

Note that the estimates of $u - U$ and $z - Z$ depend on $q - Q$. In the following analysis we will derive the estimate of $q - Q$ to finish the whole error analysis.

Theorem 5.12 Let q and Q is the solution to the problem (3.6) and (4.4). Then the following estimate holds

$$\|q - Q\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + \tau).$$

Proof It follows from (3.6) and (4.4)

$$\begin{aligned} \gamma \|q - Q\|_{L^2(0,T;L^2(\Omega))}^2 &= \int_{\Omega_T} \gamma q(q - Q) - \int_{\Omega_T} \gamma Q(q - Q) \\ &\leq \int_{\Omega_T} z(Q - q) + \int_{\Omega_T} Z(q - Q) \\ &= \int_{\Omega_T} (z - Z(q))(Q - q) + \int_{\Omega_T} (Z(q) - Z)(Q - q). \end{aligned}$$

Here $Z(q)$ is an auxiliary variable defined by

$$\sum_{j=0}^{J-1} ([[w_j]], Z(q)_j^+)_{\Omega} + ({}_t^R \partial_t^\beta \nabla Z(q), \nabla w)_{\Omega_T} = (U(q) - u_d, w)_{\Omega_T}, \forall w \in V_{hk}. \tag{5.11}$$

By (4.1) and (5.1), we have

$$(Q - q, w)_{\Omega_T} = \sum_{j=0}^{J-1} ([[U - U(q)]_j], w_j^+)_{\Omega} + ({}_0^R \partial_t^\beta \nabla(U - U(q)), \nabla w)_{\Omega_T}.$$

Choosing $w = Z(q) - Z$, and using integration by parts yields

$$\begin{aligned} & \int_{\Omega_T} (Q - q)(Z(q) - Z) \\ &= \sum_{j=0}^{J-1} ([[U - U(q)]_j], (Z(q) - Z)_j^+)_{\Omega} + ({}_0^R \partial_t^\beta \nabla(U - U(q)), \nabla(Z(q) - Z))_{\Omega_T} \\ &= \sum_{j=0}^{J-1} ([[U - U(q)]_j], (Z(q) - Z)_j^+)_{\Omega} + ({}_t^R \partial_t^\beta \nabla(Z(q) - Z), \nabla(U - U(q)))_{\Omega_T}. \end{aligned}$$

By (4.3) and (5.11), we obtain

$$\sum_{j=0}^{J-1} ([[w_j]], (Z(q) - Z)_j^+)_{\Omega} + ({}_t^R \partial_t^\beta \nabla(Z(q) - Z), \nabla w)_{\Omega_T} = (U(q) - U, w)_{\Omega_T}.$$

Setting $w = U - U(q)$ gives

$$\int_{\Omega_T} (Q - q)(Z(q) - Z) = - \int_{\Omega_T} (U(q) - U)^2 \leq 0.$$

Then we deduce

$$\begin{aligned} \gamma \|q - Q\|_{L^2(0,T;L^2(\Omega))}^2 &\leq \int_{\Omega_T} (z - Z(q))(Q - q) \\ &\leq \|z - Z(q)\|_{L^2(0,T;L^2(\Omega))} \|q - Q\|_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

Further we have

$$\|q - Q\|_{L^2(0,T;L^2(\Omega))} \leq C \|z - Z(q)\|_{L^2(0,T;L^2(\Omega))}.$$

Now it remains to estimate $\|z - Z(q)\|_{L^2(0,T;L^2(\Omega))}$. Note that $z - Z(q) = z - Z(u) + Z(u) - Z(q)$, thus we only need to estimate $\|Z(u) - Z(q)\|_{L^2(0,T;L^2(\Omega))}$. By (5.2) and (5.11), we have

$$\sum_{j=0}^{J-1} ([[w_j]], (Z(u) - Z(q))_j^+)_{\Omega} + ({}_t^R \partial_t^\beta \nabla(Z(u) - Z(q)), \nabla w)_{\Omega_T} = (u - U(q), w)_{\Omega_T}.$$

Setting $w = Z(u) - Z(q)$ and according to Lemma 5.3, we have

$$\frac{1}{2} \|(Z(u) - Z(q))_0^+\|_{L^2(\Omega)}^2 \leq \sum_{j=0}^{J-1} ([[Z(u) - Z(q)]_j], (Z(u) - Z(q))_j^+)_{\Omega}, 1 \leq j \leq J.$$

By Lemma 2.1, we have

$$(\int_t^R \partial_T^\beta \nabla(Z(u) - Z(q)), \nabla(Z(u) - Z(q)))_{\Omega_T} = \cos(\frac{\beta}{2}\pi) |Z(u) - Z(q)|_{H^{\beta/2}(0,T;H^1(\Omega))}^2.$$

Collecting above equations leads to

$$\begin{aligned} & \frac{1}{2} \|(Z(u) - Z(q))_0^+\|_{L^2(\Omega)}^2 + \cos(\frac{\beta}{2}\pi) |Z(u) - Z(q)|_{H^{\beta/2}(0,T;H^1(\Omega))}^2 \\ & \leq C \|u - U(q)\|_{L^2(0,T;L^2(\Omega))} \|Z(u) - Z(q)\|_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

Since

$$\|Z(u) - Z(q)\|_{L^2(0,T;L^2(\Omega))} \leq C |Z(u) - Z(q)|_{H^{\beta/2}(0,T;H^1(\Omega))},$$

by Young inequality we obtain

$$\|(Z(u) - Z(q))_0^+\|_{L^2(\Omega)}^2 + \|Z(u) - Z(q)\|_{L^2(0,T;L^2(\Omega))}^2 \leq C \|u - U(q)\|_{L^2(0,T;L^2(\Omega))}^2. \tag{5.12}$$

By Theorems 5.4, 5.9 and (5.12), we can get the error estimate of the control variable

$$\begin{aligned} \|q - Q\|_{L^2(0,T;L^2(\Omega))} & \leq C \|z - Z(q)\|_{L^2(0,T;L^2(\Omega))} \\ & \leq C (\|z - Z(u)\|_{L^2(0,T;L^2(\Omega))} + \|Z(u) - Z(q)\|_{L^2(0,T;L^2(\Omega))}) \\ & \leq C(h^2 + \tau) \|u - u_d\|_{L^2(0,T;L^2(\Omega))} + C \|u - U(q)\|_{L^2(0,T;L^2(\Omega))} \\ & \leq C(h^2 + \tau). \end{aligned} \quad \square$$

5.4 Main Result

Combining Theorems 5.6, 5.11 and 5.12, we obtain the following main result.

Theorem 5.13 *Let (u, z, q) and (U, Z, Q) respectively are the solution of the Eqs. (3.4–3.6) and (4.5). Then the following error estimate is established*

$$\|u - U\|_{L^2(0,T;L^2(\Omega))} + \|z - Z\|_{L^2(0,T;L^2(\Omega))} + \|q - Q\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + \tau).$$

Proof By Theorems 5.6 and 5.12, we obtain

$$\begin{aligned} \|u - U\|_{L^2(0,T;L^2(\Omega))} & \leq C(h^2 + \tau) \|f + q\|_{L^2(0,T;L^2(\Omega))} + C \|q - Q\|_{L^2(0,T;L^2(\Omega))} \\ & \leq C(h^2 + \tau). \end{aligned}$$

Hence Theorem 5.11 implies

$$\begin{aligned} \|z - Z\|_{L^2(0,T;L^2(\Omega))} & \leq C(h^2 + \tau) \|u - u_d\|_{L^2(0,T;L^2(\Omega))} + C \|u - U\|_{L^2(0,T;L^2(\Omega))} \\ & \leq C(h^2 + \tau). \end{aligned} \quad \square$$

6 Fast Algorithm

In this section, a fast gradient projection algorithm is designed to solve the optimal control problem. In the following F_J and F_J^* denote the Fourier matrix and the inverse Fourier matrix. The symbol \otimes denotes the Kronecker tensor product.

6.1 Gradient Projection Algorithm

The following gradient projection algorithm is used to solve the discrete optimal control problem.

Algorithm 1: gradient projection algorithm

Require:

Given the time step τ , space step h and tolerance η .

Ensure:

- 1: Given the initial value \mathbf{Q} , and setting $error = 1$.
- 2: **while** $error > \eta$ **do**
- 3: Solving the state equation (4.1) in the discrete optimal system to get state variable \mathbf{U} ;
- 4: Solving the adjoint state equation (4.3) in the discrete optimal system to get adjoint state variable \mathbf{Z} ;
- 5: Using the pointwise projection $P_{U_{ad}}$ onto the admissible U_{ad} to compute the control variable:

$$\mathbf{Q}_{new} = P_{U_{ad}}\left(-\frac{1}{\gamma}\mathbf{Z}\right).$$

- 6: Calculate the error

$$error = norm(\mathbf{Q} - \mathbf{Q}_{new}, inf) / norm(\mathbf{Q}, inf).$$

- 7: Update the control variable $\mathbf{Q} = \mathbf{Q}_{new}$,

- 8: **end while**
-

In above algorithm the discrete control system is solved in an iterative manner. The main computational cost comes from solving the discrete state and adjoint equation. The discrete state equation and adjoint state equation can be solved in a time-marching fashion. When the time division and space division are relatively large, due to nonlocal property of the time fractional derivative, the computation amount is very large. Therefore, fast algorithm is necessary.

6.2 The Coefficient Matrix of the Optimal Control Problem

In this part we are going to investigate the structure of the coefficient matrix of discrete state and adjoint state equation. Let V_h be the finite element space consisting of continuous piecewise linear functions

$$V_h = span\{\varphi_1, \varphi_2, \dots, \varphi_M\}.$$

According to [27], by choosing w to vanish outside I_j the discrete first order optimality system is reduced to

$$\left\{ \begin{aligned} (U^j - U^{j-1}, w_h) + \sum_{k=1}^j C_{j,k}(\nabla U^k, \nabla w_h) &= \tau(\bar{f}^j + Q^j, w_h), \quad j = 1, 2, \dots, J, \\ (Z^{j-1} - Z^j, w_h) + \sum_{k=j}^J C_{k,j}(\nabla Z^{k-1}, \nabla w_h) &= \tau(U^j - \bar{u}_d^j, w_h), \quad j = J, J-1, \dots, 1, \\ (\gamma Q^j + Z^{j-1}, v_h - Q^j) &\geq 0, \quad \forall v_h \in U_{ad}, \quad j = 1, 2, \dots, J, \\ U^0 &= 0, Z^J = 0. \end{aligned} \right.$$

Here $\bar{f}^j = \frac{1}{\tau} \int_{I_j} f dt$, $\bar{u}_d^j = \frac{1}{\tau} \int_{I_j} u_d dt$ and $C_{j,k}$ are defined as follows

$$C_{j,k} = \begin{cases} \frac{\tau^{1-\beta}}{\Gamma(2-\beta)} \left((j-k+1)^{1-\beta} - 2(j-k)^{1-\beta} + (j-k-1)^{1-\beta} \right), & k = 1, 2, \dots, j-1, \\ \frac{\tau^{1-\beta}}{\Gamma(2-\beta)}, & k = j. \end{cases}$$

We can observe that the time-stepping discontinuous Galerkin discrete scheme can be viewed as a modified backward Euler scheme, since the discrete state and adjoint state are piecewise constant in time.

We denote the mass matrix and stiff matrix by \mathcal{M} and S , respectively, whose entries are calculated by $\mathcal{M} = (\varphi_i, \varphi_j)$ and $S = (\nabla \varphi_i, \nabla \varphi_j)$, $i, j = 1, 2, \dots, M$. Note that

$$\begin{aligned} C_{11}(\nabla U^1, \nabla w_h) + (U^1, w_h) &= \tau(\bar{f}^1 + Q^1, w_h), \\ C_{21}(\nabla U^1, \nabla w_h) - (U^1, w_h) + C_{22}(\nabla U^2, \nabla w_h) + (U^2, w_h) &= \tau(\bar{f}^2 + Q^2, w_h), \\ C_{31}(\nabla U^1, \nabla w_h) + C_{32}(\nabla U^2, \nabla w_h) - (U^2, w_h) + C_{33}(\nabla U^3, \nabla w_h) + (U^3, w_h) &= \tau(\bar{f}^3 + Q^3, w_h), \\ &\vdots \\ C_{J1}(\nabla U^1, \nabla w_h) + C_{J2}(\nabla U^2, \nabla w_h) + \dots + C_{J,J-1}(\nabla U^{J-1}, \nabla w_h) - (U^{J-1}, w_h) \\ &\quad + C_{JJ}(\nabla U^J, \nabla w_h) + (U^J, w_h) = \tau(\bar{f}^J + Q^J, w_h). \end{aligned}$$

Then the discrete state equation can be rewritten as follows

$$AU = \mathfrak{b}. \tag{6.1}$$

Here

$$\begin{aligned} \mathbb{U} &= (\mathbf{U}^1, \mathbf{U}^2, \dots, \mathbf{U}^J)^T, \\ \mathfrak{b} &= (\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^J)^T, \end{aligned}$$

with $\mathbf{b}^j = (\tau(\bar{f}^j + Q^j, \varphi_i))_{M \times 1}$, $j = 1, \dots, J$, $i = 1, 2, \dots, M$ and $\mathbf{U}^j \in R^M$. Since the time subdivision is uniform, by the definition of C_{jk} we have

$$\left\{ \begin{aligned} C_{11} &= C_{22} = \dots = C_{JJ}, \\ C_{21} &= C_{32} = \dots = C_{J,J-1}, \\ C_{31} &= C_{42} = \dots = C_{J,J-2}, \\ &\vdots \\ C_{J-2,1} &= C_{J-1,2} = C_{J,3}, \\ C_{J-1,1} &= C_{J,2}. \end{aligned} \right.$$

For convenience, we define $C_1 = C_{11}, C_2 = C_{21}, \dots, C_J = C_{J1}$. Then the coefficient matrix A takes the following form

$$A = \begin{pmatrix} \mathcal{M} + C_1 \mathcal{S} & & & & & \\ C_2 \mathcal{S} - \mathcal{M} & \mathcal{M} + C_1 \mathcal{S} & & & & \\ C_3 \mathcal{S} & C_2 \mathcal{S} - \mathcal{M} & \mathcal{M} + C_1 \mathcal{S} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ C_{J-1} \mathcal{S} & C_{J-2} \mathcal{S} & C_{J-3} \mathcal{S} & \cdots & \mathcal{M} + C_1 \mathcal{S} & \\ C_J \mathcal{S} & C_{J-1} \mathcal{S} & C_{J-2} \mathcal{S} & \cdots & C_2 \mathcal{S} - \mathcal{M} & \mathcal{M} + C_1 \mathcal{S} \end{pmatrix}.$$

In an analogous way, we can rewrite the discrete adjoint state equation as follows

$$AZ = \mathbf{c}, \quad (6.2)$$

among

$$\begin{aligned} Z &= (\mathbf{Z}^{J-1}, \mathbf{Z}^{J-2}, \dots, \mathbf{Z}^0)^T, \\ \mathbf{c} &= (\mathbf{c}^1, \mathbf{c}^2, \dots, \mathbf{c}^J)^T, \\ \mathbf{c}^j &= (\tau(U^{J-j+1} - \bar{u}_d^{J-j+1}, \varphi_i))_{M \times 1}, j = 1, \dots, J, i = 1, \dots, M. \end{aligned}$$

It is easy to see that the coefficient matrix for state equation and adjoint state equation are both block lower triangular Toeplitz-like with tri-diagonal block matrix (BL3TB).

6.3 Fast Algorithm

Note that the computational effort to solve the optimal control problem mainly focuses on how to solve the state equation and the adjoint state equation. Then the remaining problem is how to solve the following linear system in a fast manner

$$A\mathcal{X} = \mathbf{y}. \quad (6.3)$$

According [19], the main idea is to replace the exact A with an approximate A_ϵ , where A_ϵ satisfies the following form

$$A_\epsilon = \begin{bmatrix} A_0 & \epsilon A_{J-1} & \cdots & \epsilon A_2 & \epsilon A_1 \\ A_1 & A_0 & \epsilon A_{J-1} & \cdots & \epsilon A_2 \\ \vdots & A_1 & A_0 & \ddots & \vdots \\ A_{J-2} & \cdots & \ddots & \ddots & \epsilon A_{J-1} \\ A_{J-1} & A_{J-2} & \cdots & A_1 & A_0 \end{bmatrix}.$$

Therefore the solution of the Eq. (6.3) can be represented as

$$\mathcal{X} = A^{-1} \mathbf{y} \approx A_\epsilon^{-1} \mathbf{y} = \mathcal{X}_\epsilon.$$

Note that A_ϵ is a Block ϵ -circulant matrix (see [4]). This allows us to use fast Fourier transform to reduce the computational cost.

Theorem 6.1 (see [4], Theorem 2.10) *Let A_ϵ is the Block ϵ -circulant matrix with the first column $[A_0, A_1, \dots, A_{J-1}]^T$. Then we have*

$$A_\epsilon^{-1} = [(D_\delta^{-1} F_J^*) \otimes I_M] \text{diag}(A_0^{-1}, A_1^{-1}, \dots, A_{J-1}^{-1}) [(F_J D_\delta \otimes I_M)], \quad (6.4)$$

Table 1 Analysis of computation

Step	Operation
1) compute δ and D_δ	$O(J)$
2) compute A_j	$O(J + N \log J)$
3) Compute y_ϵ	$O(MJ + MJ \log J)$
4) Compute \tilde{y}_ϵ	$O(MJ + MJ \log J)$
5) Compute \mathcal{X}_ϵ	$O(MJ + MJ \log J)$

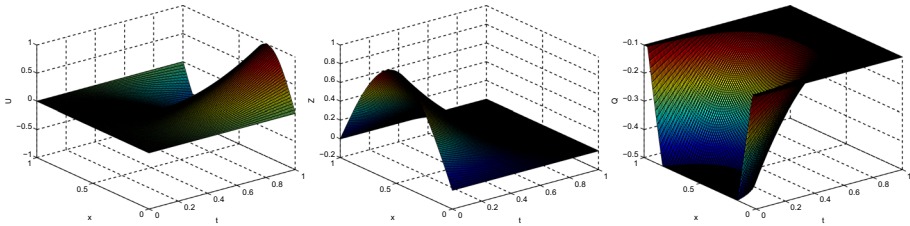


Fig. 1 discrete state U, adjoint state Z and control Q

in which, $D_\delta = \text{diag}(1, \delta, \dots, \delta^{J-1})$, $\delta = \sqrt[J]{\epsilon}$ is a diagonal matrix. The matrix A_k , $k = 0, 1, \dots, J - 1$, are $M \times M$ dimensionality and satisfy

$$\begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_{J-1} \end{bmatrix} = [(\sqrt[J]{F} F_J D_\delta) \otimes I_M] \begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_{J-1} \end{bmatrix}.$$

According to [19], the Eq. (6.4) can be implemented by the following algorithm, which is called approximate inversion method (AIM).

Algorithm 2: AIM for solving the *BLTB* system

1. Input ϵ , y and tri-diagonal matrices A_0, A_1, \dots, A_{J-1} ,
2. Calculate $\delta = \sqrt[J]{\epsilon}$ and $D_\delta = \text{diag}(1, \delta, \dots, \delta^{J-1})$,
3. Calculate tri-diagonal matrices A_j , $j = 0, 1, \dots, J - 1$,
4. Calculate $y_\epsilon = [(F_J D_\delta) \otimes I_M]y$,
5. Calculate $\tilde{y}_\epsilon = \text{diag}(A_0^{-1}, A_1^{-1}, \dots, A_{J-1}^{-1})y_\epsilon$,
6. Calculate $\mathcal{X}_\epsilon = [(D_\delta^{-1} F_J^*) \otimes I_M]\tilde{y}_\epsilon$.

Applying Algorithm 2 to Algorithm 1 yields a fast projection gradient method for the optimal control problem. The Table 1 displays the computational cost required for each step, which implies that the total computational cost is $O(MJ \log J)$.

7 Numerical Examples

In this section numerical experiments will be carried out to illustrate the error analysis and algorithm presented in Sects. 5 and 6.

Table 2 Errors of state, adjoint state and control in $L^2(0, T; L^2(\Omega))$ for different β and fixed time partition $J=M^2$

	M=10	M=20	M=30	M=40	M=50	M=60	M=70	Order
$\beta=0.4$								
$\ u-U\ _{L^2(0,T;L^2(\Omega))}$	5.94e-3	1.46e-3	6.46e-4	3.63e-4	2.33e-4	1.62e-4	1.19e-4	2.00
$\ z-Z\ _{L^2(0,T;L^2(\Omega))}$	3.28e-3	8.31e-4	3.71e-4	2.10e-4	1.34e-4	9.36e-5	6.89e-5	1.98
$\ q-Q\ _{L^2(0,T;L^2(\Omega))}$	3.45e-3	8.54e-4	3.78e-4	2.12e-4	1.35e-4	9.40e-5	6.89e-5	2.01
$\beta=0.8$								
$\ u-U\ _{L^2(0,T;L^2(\Omega))}$	5.41e-3	1.35e-3	6.04e-4	3.43e-4	2.21e-4	1.54e-4	1.14e-4	1.97
$\ z-Z\ _{L^2(0,T;L^2(\Omega))}$	2.44e-3	6.73e-4	3.13e-4	1.81e-4	1.19e-4	8.36e-5	6.22e-5	1.89
$\ q-Q\ _{L^2(0,T;L^2(\Omega))}$	4.05e-3	9.63e-4	4.18e-4	2.31e-4	1.46e-4	1.01e-4	7.33e-5	2.05

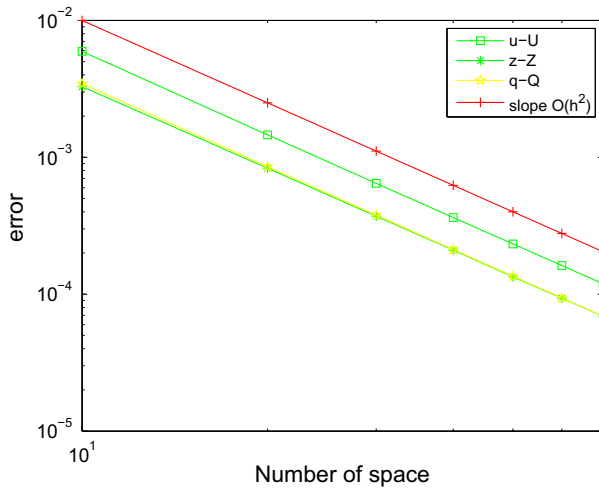


Fig. 2 Space convergence rates of state, adjoint state and control for $\beta = 0.4$

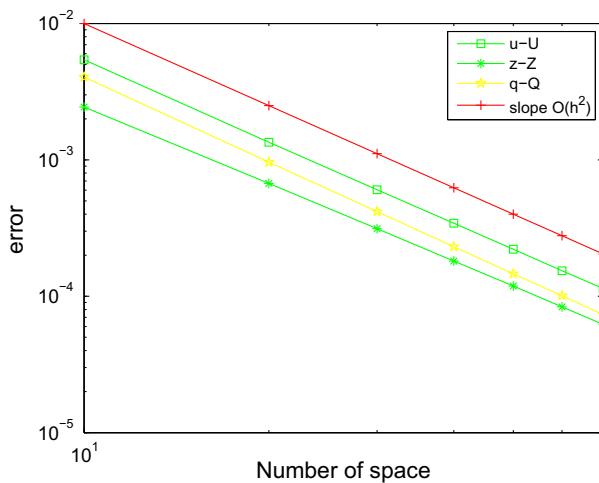


Fig. 3 Space convergence rates of state, adjoint state and control for $\beta = 0.8$

Table 3 Errors of state, adjoint state and control for different β and $M=J$

	$J=400$	$J=500$	$J=600$	$J=700$	$J=800$	$J=900$	$J=1000$	Order	
$\beta=0.4$	$\ u-U\ _{L^2(0,T;L^2(\Omega))}$	9.72e-4	7.79e-4	6.50e-4	5.58e-4	4.88e-4	4.34e-4	3.91e-4	0.99
	$\ z-Z\ _{L^2(0,T;L^2(\Omega))}$	8.78e-4	7.04e-4	5.87e-4	5.04e-4	4.41e-4	3.93e-4	3.54e-4	0.98
	$\ q-Q\ _{L^2(0,T;L^2(\Omega))}$	7.76e-4	6.20e-4	5.16e-4	4.42e-4	3.86e-4	3.43e-4	3.09e-4	1.00
$\beta=0.8$	$\ u-U\ _{L^2(0,T;L^2(\Omega))}$	7.31e-4	5.94e-4	5.01e-4	4.34e-4	3.82e-4	3.43e-4	3.10e-4	0.94
	$\ z-Z\ _{L^2(0,T;L^2(\Omega))}$	6.74e-4	5.48e-4	4.62e-4	4.00e-4	3.52e-4	3.16e-4	2.85e-4	0.94
	$\ q-Q\ _{L^2(0,T;L^2(\Omega))}$	9.09e-4	7.22e-4	5.98e-4	5.10e-4	4.44e-4	3.93e-4	3.53e-4	1.03

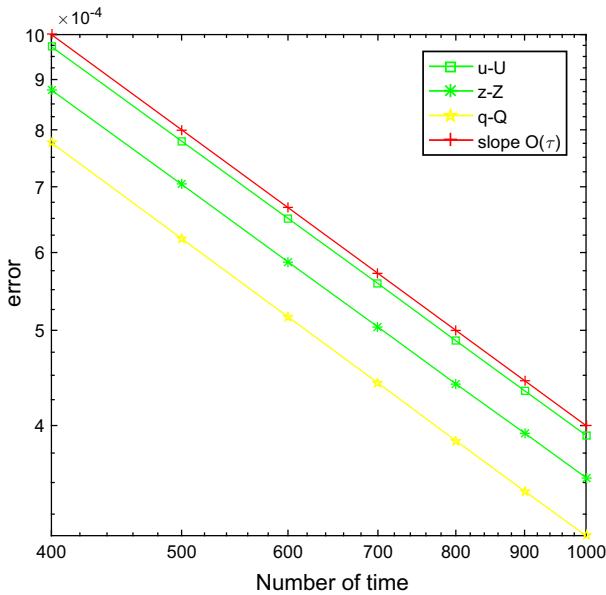


Fig. 4 Time convergence rates of state, adjoint state and control for $\beta = 0.4$

7.1 Example 1

We consider the control problem with $\Omega = [0, 1]$, $\gamma = 1$, $T = 1$, $\eta = 1.0 \times 10^{-5}$ and $\epsilon = 0.5 \times 10^{-8}$. The exact solutions are given by

$$\begin{aligned}
 u &= t^2 \sin(2\pi x), \\
 z &= (1 - t)^2 \sin(\pi x), \\
 q &= \max(-0.5, \min(-z, -0.1)).
 \end{aligned}$$

The right hand term f and the desired state u_d can be calculated by the exact solutions and governing equation.

The space-time surfaces of discrete state variable, adjoint state variable and the control variable are displayed in Fig. 1 .

In order to achieve the convergence rate of space we fix time step as $J = M^2$. The errors of state, adjoint state and control variable are presented in Table 2. The corresponding

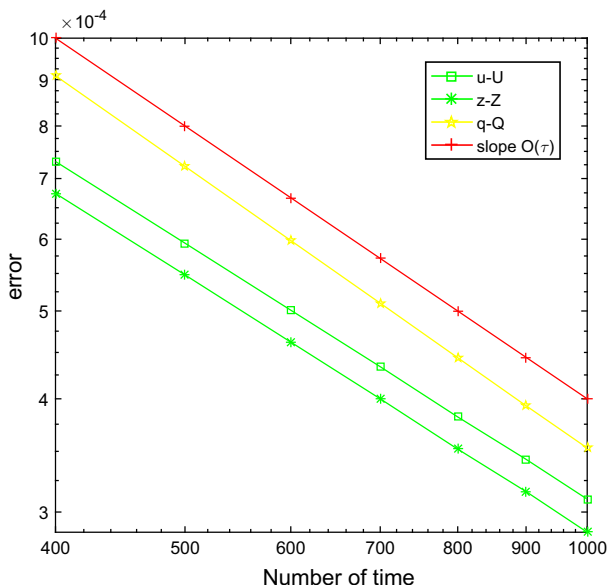


Fig. 5 Time convergence rates of state, adjoint state and control for $\beta = 0.8$

Table 4 Comparison of computational time

	$J = 2^7$	$J = 2^8$	$J = 2^9$	$J = 2^{10}$	$J = 2^{11}$	$J = 2^{12}$	Order
AIM	0.978	1.935	3.664	7.122	14.007	29.375	0.98
BSFM	9.058	33.7	144.9	679.3	2856.4	15796.1	2.15

convergence rate for space is shown in Figs. 2 and 3, which is in agreement with the theoretical result.

Similarly, in order to test the convergence rate of time we set $M = J$. The errors of state, adjoint state and control variable are list in Table 3 and the corresponding convergence rate is given in Figs. 4 and 5. It is easy to see that the convergence rate is consistent with the theoretical prediction.

The comparison of computational time for fast algorithm (AIM) and the traditional method with the block forward substitution (BSFM) is given in Table 4 with $M = 2^7$. We can observe that fast algorithm can effectively reduce the computational time. From Figs. 6 and 7 we can find that the computational time for fast algorithm (AIM) increases linearly, which is in contrast to the traditional method with the block forward substitution (BSFM).

7.2 Example 2

We consider the control problem with $\Omega = [0, 1]$, $\gamma = 1$, $T = 1$, and $\eta = 1.0 \times 10^{-5}$. f and u_d are given by

$$f = 10,$$

$$u_d = -20x(1 - x)e^t.$$

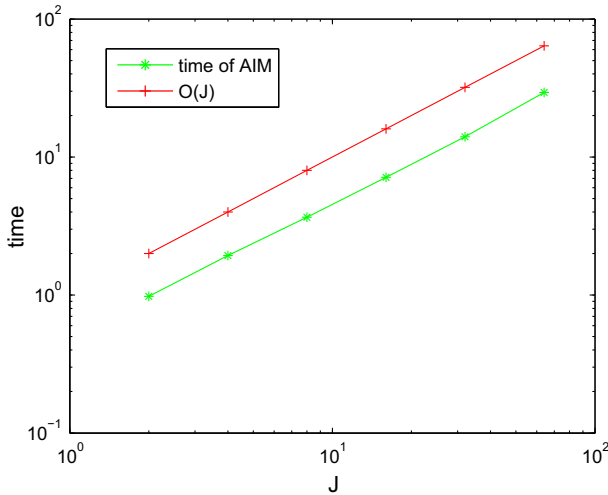


Fig. 6 Computational time of fast algorithm (AIM)

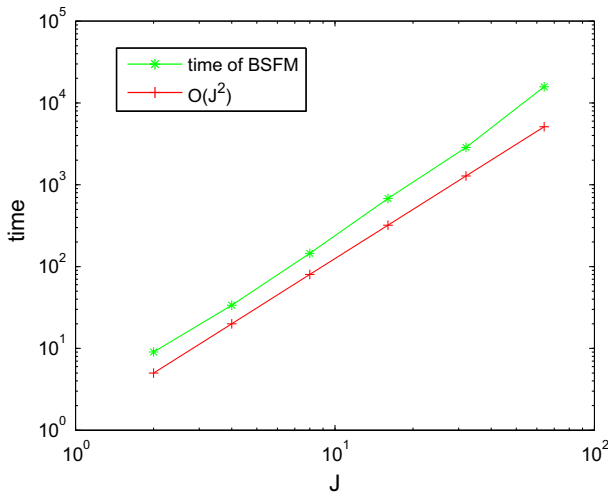


Fig. 7 Computational time of BFSM

In this example, no exact solutions can be achieved. The control variable is chosen as $q = \max(-0.5, \min(-z, -0.1))$. We use the numerical solutions on a much finer grid as the reference solutions. In order to achieve the convergence rate of space, we fix $J = 2^{10}$ and $M = 2^{10}$. The errors of state, adjoint state and control variable are given in Table 5. The corresponding convergence rate for space is shown in Figs. 8 and 9 for different β , which is in agreement with the theoretical result.

Similarly, we set $J = 2^{11}$ and $M = 2^{11}$ to verify the convergence rate of time. The errors of state, adjoint state and control variable are listed in Table 6 and the corresponding convergence rate is given in Figs. 10 and 11 for different β . It is easy to observe that the convergence rate is consistent with the theoretical prediction.

Table 5 Errors of state, adjoint state and control for $M = 2^{10}$

	$J = 2^{10}$	$M = 2^2$	$M = 2^3$	$M = 2^4$	$M = 2^5$	$M = 2^6$	Order
$\beta = 0.4$	$\ u - U\ _{L^2(0,T;L^2(\Omega))}$	0.0389	0.00971	0.00243	6.06e-4	1.52e-4	2.00
	$\ z - Z\ _{L^2(0,T;L^2(\Omega))}$	0.0306	0.00770	0.00193	4.81e-4	1.22e-4	1.99
	$\ q - Q\ _{L^2(0,T;L^2(\Omega))}$	0.0130	0.00305	7.06e-4	1.79e-4	4.52e-5	2.04
$\beta = 0.8$	$\ u - U\ _{L^2(0,T;L^2(\Omega))}$	0.0311	0.00777	0.00195	4.91e-4	1.28e-4	1.98
	$\ z - Z\ _{L^2(0,T;L^2(\Omega))}$	0.0245	0.00617	0.00155	3.94e-4	1.07e-4	1.96
	$\ q - Q\ _{L^2(0,T;L^2(\Omega))}$	0.0127	0.00319	7.76e-4	1.96e-4	5.07e-5	1.99

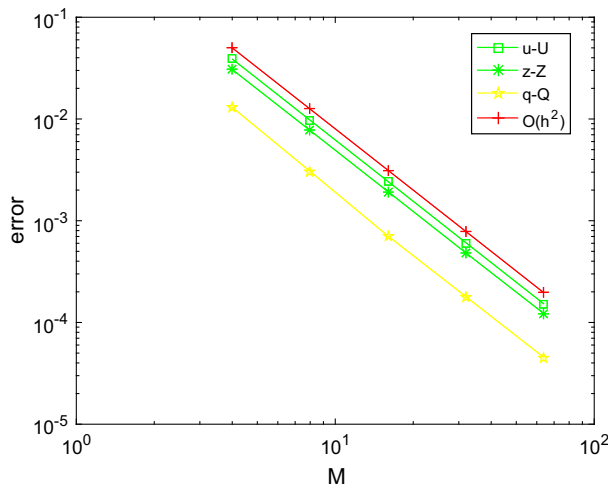


Fig. 8 Space convergence rates of state, adjoint state and control for $\beta = 0.4$

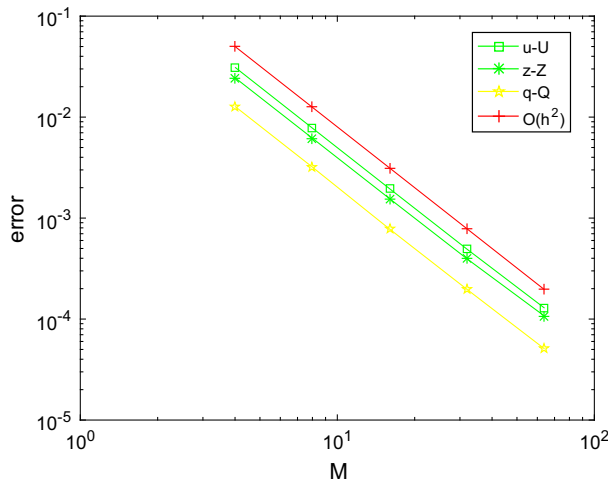


Fig. 9 Space convergence rates of state, adjoint state and control for $\beta = 0.8$

Table 6 Errors of state, adjoint state and control for $J = 2^{11}$

	$M = 2^{11}$	$J = 2^3$	$J = 2^4$	$J = 2^5$	$J = 2^6$	$J = 2^7$	$J = 2^7$	Order
$\beta = 0.4$	$\ u - U\ _{L^2(0,T;L^2(\Omega))}$	0.0436	0.0235	0.0124	0.00638	0.00322	0.00159	0.95
	$\ z - Z\ _{L^2(0,T;L^2(\Omega))}$	0.0449	0.0253	0.0137	0.00711	0.00360	0.00177	0.93
	$\ q - Q\ _{L^2(0,T;L^2(\Omega))}$	0.0392	0.0208	0.00997	0.00462	0.00239	0.00116	1.01
$\beta = 0.8$	$\ u - U\ _{L^2(0,T;L^2(\Omega))}$	0.0431	0.0209	0.0101	0.00487	0.00236	0.00114	1.04
	$\ z - Z\ _{L^2(0,T;L^2(\Omega))}$	0.0371	0.0182	0.00892	0.00432	0.00210	0.00103	1.03
	$\ q - Q\ _{L^2(0,T;L^2(\Omega))}$	0.0286	0.0131	0.00650	0.00310	0.00149	7.14e-4	1.06

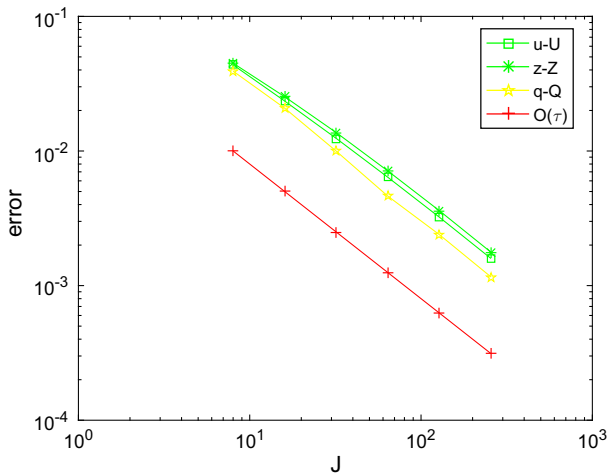


Fig. 10 Time convergence rates of state, adjoint state and control for $\beta = 0.4$

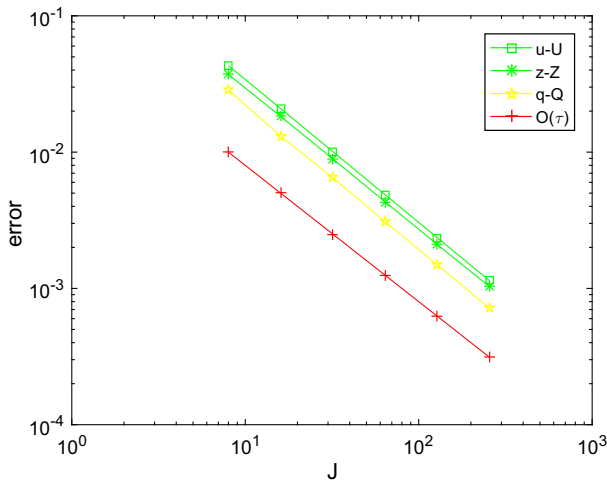


Fig. 11 Time convergence rates of state, adjoint state and control for for $\beta = 0.8$

References

1. Antil, H., Otárola, E.: A FEM for an optimal control problem of fractional powers of elliptic operators. *SIAM J. Control Optim.* **53**(6), 3432–3456 (2015)
2. Antil, H., Otárola, E., Salgado, A.J.: A space-time fractional optimal control problem: analysis and discretization. *SIAM J. Control Optim.* **54**(3), 1295–1328 (2016)
3. Benson, D.A., Wheatcraft, S.W., Meerschaert, M.M.: The fractional order governing equations of Levy motion. *Water Resour. Res.* **36**, 1413–1423 (2000)
4. Bini, D.A., Latouche, G., Meini, B.: *Numerical Methods for Structured Markov Chains*. Oxford University Press, Oxford (2005)
5. Casasy, E., Chrysafoinos, K.: A discontinuous Galerkin time-stepping scheme for the velocity tracking problem. *SIAM J. Numer. Anal.* **5**, 2281–2306 (2012)
6. Deckelnick, K., Hinze, M.: Convergence of a finite element approximation to a state-constrained elliptic control problem. *SIAM J. Numer. Anal.* **45**(4), 1937–1953 (2007)
7. Du, N., Wang, H., Liu, W.B.: A fast gradient projection method for a constrained fractional optimal control. *J. Sci. Comput.* **68**, 1–20 (2016)
8. Ervin, V.J., Roop, J.P.: Variational formulation for the stationary fractional advection dispersion equation. *Numer. Methods Partial Differ. Equ.* **22**(3), 558–576 (2006)
9. Gong, W., Hinze, M., Zhou, Z.J.: A priori error analysis for finite element approximation of parabolic optimal control problems with pointwise control. *SIAM J. Control Optim.* **52**(1), 97–119 (2014)
10. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.* **30**(1), 45–61 (2005)
11. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: *Optimization with PDE Constraints*. Springer, Dordrecht (2009)
12. Jin, B.T., Lazarov, R., Zhou, Z.: Error estimates for a semidiscrete finite element method for fractional order parabolic equations. *SIAM J. Numer. Anal.* **51**, 445–466 (2012)
13. Jin, B.T., Li, B.Y., Zhou, Z.: Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint. *IMA J. Numer. Anal.* (2018). <https://doi.org/10.1093/imanum/drxxxx>
14. Klafter, J., Sokolov, I.M.: Anomalous diffusion spreads its wings. *Phys. World* **18**(8), 29–32 (2005)
15. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in L^2 for a class of evolution equations. *SIAM J. Control. Optim.* **46**(5), 1726–1753 (2007)
16. Li, B.J., Luo, H., Xie, X.P.: Error estimates of a discontinuous Galerkin method for time fractional diffusion problems with nonsmooth data. [arXiv:1809.02015](https://arxiv.org/abs/1809.02015) (2018)
17. Li, S.Y., Zhou, Z.J.: Legendre pseudo-spectral method for optimal control problem governed by a time-fractional diffusion equation. *Int. J. Comput. Math.* **95**(6–7), 1308–1325 (2018)
18. Liu, W.B., Yan, N.N.: *Adaptive Finite Element Methods for Optimal Control Governed by PDEs*. Beijing Science Press, Beijing (2008)
19. Lu, X., Pang, H.K., Sun, H.W.: Fast approximate inversion of a block triangular Toeplitz matrix with applications to fractional sub-diffusion equation. *Numer. Linear Algebra Appl.* **22**(5), 866–883 (2015)
20. Meerschaert, M.M., Sikorskii, A.: *Stochastic Models for Fractional Calculus*. De Gruyter Studies in Mathematics, 43. De Gruyter, Berlin (2012)
21. Podlubny, I.: *Fractional Differential Equations*. Academic Press, New York (1999)
22. Thomée, V.: *Galerkin Finite Element Methods for Parabolic Problems*. Springer, Berlin (2006)
23. Ye, X.Y., Xu, C.J.: Spectral optimization methods for the time fractional diffusion inverse problem. *Numer. Math. Theor. Meth. Appl.* **6**(3), 499–519 (2013)
24. Ye, X.Y., Xu, C.J.: A spectral method for optimal control problem governed by the abnormal diffusion equation with integral constraint on the state. *Sci. Sin. Math.* **46**(7), 1053–1070 (2016)
25. Zhou, Z.J., Gong, W.: Finite element approximation of optimal control problems governed by time fractional diffusion equations. *Comput. Math. Appl.* **71**(1), 301–308 (2016)
26. Zhou, Z.J., Tan, Z.Y.: Finite element approximation of optimal control problem governed by space fractional equation. *J. Sci. Comput.* **78**, 1840–1861 (2019)
27. Zhou, Z.J., Zhang, C.Y.: Time-stepping discontinuous Galerkin approximation of optimal control problem governed by time fractional diffusion equation. *Numer. Algorithms* **79**(2), 437–455 (2018)