Numer. Math. Theor. Meth. Appl. Vol. **6**, No. 3, pp. 499-519
doi: 10.4208/nmtma.2013.1207nm August 2013

Spectral Optimization Methods for the Time Fractional Diffusion Inverse Problem

Xingyang Ye[1,2] and Chuanju Xu[1,*]

[1] *School of Mathematical Sciences, Xiamen University, Xiamen 361005, Fujian, China.*
[2] *School of Science, Jimei University, Xiamen 361021, Fujian, China.*

Received 23 April 2012; Accepted (in revised version) 24 September 2012

Available online 14 June 2013

Abstract. An inverse problem of reconstructing the initial condition for a time fractional diffusion equation is investigated. On the basis of the optimal control framework, the uniqueness and first order necessary optimality condition of the minimizer for the objective functional are established, and a time-space spectral method is proposed to numerically solve the resulting minimization problem. The contribution of the paper is threefold: 1) a priori error estimate for the spectral approximation is derived; 2) a conjugate gradient optimization algorithm is designed to efficiently solve the inverse problem; 3) some numerical experiments are carried out to show that the proposed method is capable to find out the optimal initial condition, and that the convergence rate of the method is exponential if the optimal initial condition is smooth.

AMS subject classifications: 65M12, 65M32, 65M70, 35S10, 49J20

Key words: Time fractional diffusion equation, inverse problem, spectral method, error estimate, conjugate gradient method.

1. Introduction

Optimal control problems can be found in many scientific and engineering applications, and it has become a very active and successful research area in recent years. Extensive research has been carried out on various theoretical aspects of control problems such as existence of optimal control, optimality conditions, regularity of the optimal solutions, and so on. The literature on this field is huge, and it is impossible to give even a very brief review here. However, to the best of the authors' knowledge, most research concerning control problems has been performed using partial differential equations of integer order, and there are not many published works related to the differential equations of fractional order.

*Corresponding author. *Email addresses:* `xingyangye@163.com` (X. Y. Ye), `cjxu@xmu.edu.cn` (C. Xu)

 ©2013 Global-Science Press

In this paper we are interested in control problems based on partial differential equations of fractional order. This work is motivated by the fact that the fractional partial differential equations are novel extensions of the traditional models, based on fractional calculus. They are now winning more and more scientific applications cross a variety of fields including control theory, biology, electrochemical processes, viscoelastic materials, polymer, finance, and etc. We will consider an inverse problem associated to the time fractional diffusion equation (TFDE) of the form

$$\partial_t^\alpha u(x,t) - \partial_x^2 u(x,t) = 0, \quad 0 < x < l, \quad t > 0,$$

where ∂_t^α means the fractional derivative of order α, $0 < \alpha < 1$. Precisely, the problem to be investigated is as follows: suppose we are given the observed date $\bar{u}(x,t)$, the goal is to find out the optimal initial condition $u(x,0)$ such that the corresponding solution to TFDE matches the observed data as closely as possible.

For the initial boundary value problem of TFDE, some theoretical and numerical results have been obtained by a number of authors. For example, Schneider and Wyss [18] and Wyss [20] used the Green functions to construct the explicit solution in some simple cases. Luchko [12, 13] derived the maximum principle and proved the unique existence of the generalized solution. Sakamoto and Yamamoto [16] investigated weak solutions of TFDE in 2D. The existing numerical methods includes finite difference [7, 11, 19], Galerkin finite element [4, 5, 15], finite difference/spectral method [9], time-space spectral method [8], and so on.

For the inverse problem concerning TFDE, although the research is relatively sparse, several studies have been carried out, and we see increasing interest in this topic from both scientific and engineering communities. We mention, among others, the work [3] to determine the fractional order α and variable diffusion coefficient by means of additional boundary data. The uniqueness of the inverse problem was proved theoretically on the basis of the eigenfunction expansion of the weak solution and the Gel'fand-Levitan theory. Sakamoto and Yamamoto also considered in their above mentioned paper an inverse source problem. They analyzed the stability of determining time-dependent factor in the source by some observation. Zhang and Xu [21] established the uniqueness of an inverse problem which consists in identifying the time independent source term for TFDE with homogenous Neumann boundary condition, and some numerical examples were presented.

In this work, we will focus on the numerical method to find out the optimal initial condition for the TFDE with known observed data. Unlike the work [21], which uses the eigenfunction expansion of the solution as the main tool, we will adapt the optimal control framework [10] to treat the inverse problem. By introducing an objective function which measures the discrepancy of the solution given by the TFDE problem and the known observation data, the optimal initial condition is then defined as the state such that the objective function attains its minimum. Thanks to the weak formulation of TFDE proposed in [8], we are able to derive a space-time spectral method for the considered inverse problem.

The remainder of this paper comprises five sections. In Section 2, we first describe the inverse problem, and give some preliminary results on the initial value problem associated to the TFDE. The formulation of the inverse problem into the control problem is also given in this section. In Section 3, we derive the necessary and sufficient optimality condition for the optimal control problem. The space-time spectral discretization and the error analysis are presented in Section 4, where some error estimates are provided. In Section 5, we describe the overall algorithm and present some numerical examples to validate our method. Some concluding remarks are given at the end of the paper.

2. Inverse problem of the time fractional diffusion equation

2.1. Time fractional diffusion equation

Let $\Lambda=(-1,1)$, $I=(0,T)$ and $\Omega=\Lambda\times I$. We consider the following one-dimensional TFDE

$$ {}_0\partial_t^{\alpha} u(x,t)-\partial_x^2 u(x,t)=0, \quad \forall (x,t)\in\Omega, \tag{2.1}$$

subject to the boundary and initial condition:

$$u(-1,t)=u(1,t)=0, \qquad \forall t\in I, \tag{2.2a}$$

$$u(x,0)=q(x), \qquad \forall x\in\Lambda, \tag{2.2b}$$

where $0<\alpha<1$, ${}_0\partial_t^{\alpha} u(x,t)$ is the left Caputo fractional derivative of order α, defined by

$$ {}_0\partial_t^{\alpha} u(x,t)=\frac{1}{\Gamma(1-\alpha)}\int_0^t \partial_\tau u(x,\tau)\frac{d\tau}{(t-\tau)^{\alpha}}.$$

The inverse problem that we are concerned with in this paper is:

Given the observation date $\bar{u}(x,t)$, find the optimal initial condition $q(x)$ such that the corresponding solution to (2.1)-(2.2b) *matches $\bar{u}(x,t)$ as well as possible.*

In order to define well the inverse problem, we first introduce some notations that will be used to construct the weak problem of the time fractional diffusion equation (2.1)-(2.2b). We use the symbol $\mathscr{O}$ to denote a domain which may stand for Λ, I or Ω. $C_0^{\infty}(\mathscr{O})$ means the space of all functions having continuous derivatives of all orders and compactly supported in $\mathscr{O}$. The notations $L^2(\mathscr{O}), H^s(\mathscr{O})$, and $H_0^s(\mathscr{O})$ stand for the usual Sobolev spaces, whose norms are respectively denoted by $\|\cdot\|_{0,\mathscr{O}}$ and $\|\cdot\|_{s,\mathscr{O}}$. For the Sobolev space X with norm $\|\cdot\|_X$, let

$$H^s(I;X):=\{v\,|\,\|v(\cdot,t)\|_X\in H^s(I)\}, \quad s\geq 0,$$

endowed with the norm:

$$\|v\|_{H^s(I;X)}:=\Big\|\|v(\cdot,t)\|_X\Big\|_{s,I}.$$

Particularly, when X stands for $H^{\mu}(\Lambda)$ or $H_0^{\mu}(\Lambda)$, $\mu\geq 0$, the norm of the space $H^s(I;X)$ will be denoted by $\|\cdot\|_{\mu,s,\Omega}$. Hereafter, in cases where no confusion would arise, the domain symbols I,Λ and Ω may be dropped from the notations.

We also need some definitions and properties regarding fractional derivatives.

- *Right Caputo fractional derivative* [14]

$$ {}_t\partial_T^\alpha u(t) = -\frac{1}{\Gamma(1-\alpha)} \int_t^T \frac{u'(\tau)d\tau}{(\tau-t)^\alpha}, \quad 0<\alpha<1. \tag{2.3}$$

- *Left Riemann-Liouville fractional derivative*

$$ {}_0^R\partial_t^\alpha u(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{u(\tau)}{(t-\tau)^\alpha} \mathrm{d}\tau, \quad 0<\alpha<1. \tag{2.4}$$

- *Right Riemann-Liouville fractional derivative*

$$ {}_t^R\partial_T^\alpha u(t) = -\frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_t^T \frac{u(\tau)}{(\tau-t)^\alpha} \mathrm{d}\tau, \quad 0<\alpha<1. \tag{2.5}$$

The definitions of Riemann-Liouville and Caputo fractional derivative are linked by the following relationship, which can be verified by a direct calculation:

$$ {}_0^R\partial_t^\alpha v(t) = \frac{v(0)t^{-\alpha}}{\Gamma(1-\alpha)} + {}_0\partial_t^\alpha v(t), \tag{2.6a}$$

$$ {}_t^R\partial_T^\alpha v(t) = \frac{v(T)(T-t)^{-\alpha}}{\Gamma(1-\alpha)} + {}_t\partial_T^\alpha v(t). \tag{2.6b}$$

We employ the space introduced in [8]:

$$ B^s(\Omega) = H^s(I, L^2(\Lambda)) \cap L^2(I, H_0^1(\Lambda)), \quad \forall s>0, $$

equipped with the norm:

$$ \|v\|_{B^s(\Omega)} = \Big(\|v\|^2_{H^s(I,L^2(\Lambda))} + \|v\|^2_{L^2(I,H_0^1(\Lambda))} \Big)^{\frac{1}{2}}. $$

In this setting, the weak formulation of the problem (2.1)-(2.2b) reads: given $q(x) \in L^2(\Lambda)$, find $u \in B^{\alpha/2}(\Omega)$, such that

$$ \mathscr{A}(u,v) = \Big(\frac{q(x)t^{-\alpha}}{\Gamma(1-\alpha)}, v \Big)_\Omega, \quad \forall v \in B^{\frac{\alpha}{2}}(\Omega), \tag{2.7}$$

where the bilinear form $\mathscr{A}(\cdot,\cdot)$ is defined by

$$ \mathscr{A}(u,v) := ({}_0^R\partial_t^{\frac{\alpha}{2}} u, {}_t^R\partial_T^{\frac{\alpha}{2}} v)_\Omega + (\partial_x u, \partial_x v)_\Omega. $$

It has been proved [8] that the problem (2.7) is well-posed.

2.2. Inverse problem and optimal control

We now define the objective function as follows:

$$J(q) = \frac{1}{2}\|u-\bar{u}\|_{0,\Omega}^2 + \frac{\lambda}{2}\left\|q\right\|_{0,\Lambda}^2, \quad \forall q \in L^2(\Lambda), \tag{2.8}$$

where u is the solution of the problem (2.7) associated to the initial condition $q(x)$, $\bar{u}$ is the observation data, and $\lambda > 0$ is the regularization parameter.

Statement of the inverse problem: *Given the observed data $\bar{u}(x,t)$, $x \in \Lambda$, $t > 0$, determine the unknown initial condition $q(x)$, such that the objective function $J(q)$ attains its minimum.*

Precisely, the above inverse problem can be formulated into the following optimal control problem: given $\bar{u} \in L^2(I, L^2(\Lambda))$, find $q^* \in L^2(\Lambda)$, such that

$$J(q^*) = \min_{q \in L^2(\Lambda)} J(q). \tag{2.9}$$

From now on, we regard q as control variable, and u as state variable satisfying the state equation (2.7).

3. Optimality condition

The first order necessary optimality condition for the problem (2.9) takes the form

$$J'(q^*)(\eta) = 0, \quad \text{for all } \ \eta \in L^2(\Lambda), \tag{3.1}$$

where $J'(q^*)(\eta)$ is usually called the gradient of $J(q)$, which is defined through the Gâteaux differential of $J(q)$ at q^* along the "direction" η. Note that (3.1) is also the sufficient condition because the quadratic functional $J(q)$ is convex [10].

Now the key point is how to efficiently compute the gradient of the objective functional $J(q)$. To this end, we introduce the adjoint state equation of (2.1)-(2.2b) as follows:

$$\begin{cases} {}_t\partial_T^\alpha z - \partial_x^2 z = u - \bar{u}, & \forall (x,t) \in \Omega, \\ z(-1,t) = z(1,t) = 0, & \forall t \in I, \\ z(x,T) = 0, & \forall x \in \Lambda. \end{cases} \tag{3.2}$$

Its associated weak formulation reads: find $z \in B^{\alpha/2}(\Omega)$, such that

$$\mathscr{A}(\varphi, z) = (u - \bar{u}, \varphi)_\Omega, \quad \forall \varphi \in B^{\frac{\alpha}{2}}(\Omega). \tag{3.3}$$

Following the same idea as for the problem (2.7), it can be proved that (3.3) admits a unique solution $z \in B^{\alpha/2}(\Omega)$ for any given $u \in B^{\alpha/2}(\Omega)$. The solution z of the above problem is hereafter referred to the adjoint state variable.

Theorem 3.1. *Let $z \in B^{\alpha/2}(\Omega)$ be the solution of the adjoint state equation* (3.2)*, then there exists a unique solution $q^* \in L^2(\Lambda)$ to the optimal control problem* (2.9)*. Furthermore, the gradient of J can be obtained through*

$$J'(q)(\eta) = \frac{1}{\Gamma(1-\alpha)} \int_{\Omega} z(x,t)\eta(x)t^{-\alpha} dxdt + \lambda \int_{\Lambda} q(x)\eta(x)dx. \tag{3.4}$$

Proof. Let us consider the perturbation of the initial condition:

$$q(x) \to \tilde{q}(x) := q(x) + \varepsilon\eta(x),$$

where ε is a real parameter tending to 0, and $\eta \in L^2(\Lambda)$. Let $\tilde{u}$ be the solution of (2.1)-(2.2a) subject to the above perturbed initial condition $\tilde{q}(x)$. We define $\hat{u}$ as

$$\hat{u} = \lim_{\varepsilon \to 0} \frac{\tilde{u} - u}{\varepsilon}.$$

Then it is readily seen that $\hat{u}$ is the solution of the following problem:

$$\begin{cases} {}_0\partial_t^\alpha \hat{u} - \partial_x^2 \hat{u} = 0, & \forall (x,t) \in \Omega, \\ \hat{u}(-1,t) = \hat{u}(1,t) = 0, & \forall t \in I, \\ \hat{u}(x,0) = \eta(x), & \forall x \in \Lambda. \end{cases} \tag{3.5}$$

By virtue of (2.8), we have

$$\begin{aligned} J'(q)(\eta) &= \lim_{\varepsilon \to 0} \frac{J(q+\varepsilon\eta) - J(q)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{\int_\Omega [(\tilde{u}-\bar{u})^2 - (u-\bar{u})^2] dxdt + \lambda \int_\Lambda (\tilde{q}^2 - q^2) dx}{2\varepsilon} \\ &= \int_\Omega \lim_{\varepsilon \to 0} \frac{(\tilde{u}+u-2\bar{u})(\tilde{u}-u)}{2\varepsilon} dxdt + \lambda \int_\Lambda \lim_{\varepsilon \to 0} \frac{(\tilde{q}+q)\eta}{2} dx \\ &= \int_\Omega (u-\bar{u})\hat{u} dxdt + \lambda \int_\Lambda q\eta dx. \end{aligned} \tag{3.6}$$

Similarly, the second order Gâteaux derivative of $J(q)$ is given by

$$J''(q)(\eta,\eta) = \lim_{\varepsilon \to 0} \frac{J'(q+\varepsilon\eta)(\eta) - J'(q)(\eta)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{J'(\tilde{q})(\eta) - J'(q)(\eta)}{\varepsilon}.$$

By using (3.6) with q replaced by $\tilde{q}$, we obtain

$$\begin{aligned} J''(q)(\eta,\eta) &= \lim_{\varepsilon \to 0} \left(\int_\Omega \frac{\tilde{u}-u}{\varepsilon} \hat{u} dxdt + \lambda \int_\Lambda \frac{\tilde{q}-q}{\varepsilon} \eta dx \right) \\ &= \int_\Omega \hat{u}^2 dxdt + \lambda \int_\Lambda \eta^2 dx \geq 0. \end{aligned} \tag{3.7}$$

This means that the functional J is uniformly convex, and therefore problem (2.9) admits a unique solution q^*.

To prove (3.4), we multiply each side of the first equation in (3.2) by $\hat{u}$, then integrate the resulted equation on the domain Ω to yield

$$\int_\Omega (u-\bar{u})\hat{u}dxdt = \int_\Omega \left({}_t\partial_T^\alpha z - \partial_x^2 z\right)\hat{u}dxdt. \tag{3.8}$$

In one side, taking into account the boundary conditions in (3.5) and (3.2), it holds

$$\int_\Omega \partial_x^2 z\hat{u}dxdt = \int_\Omega z\partial_x^2\hat{u}dxdt. \tag{3.9}$$

On the other side, by means of (2.6a), (2.6b), the terminal condition in (3.2), and the fractional integration by parts demonstrated in [8], we have

$$\begin{aligned}\int_\Omega {}_t\partial_T^\alpha z\hat{u}dxdt =& \int_\Omega \left({}_t^R\partial_T^\alpha z - \frac{z(x,T)(T-t)^{-\alpha}}{\Gamma(1-\alpha)}\right)\hat{u}dxdt\\ =& \int_\Omega {}_t^R\partial_T^\alpha z\hat{u}dxdt = \int_\Omega z\,{}_0^R\partial_t^\alpha\hat{u}dxdt\\ =& \int_\Omega z\,{}_0\partial_t^\alpha\hat{u}dxdt + \int_\Omega \frac{z\hat{u}(x,0)}{\Gamma(1-\alpha)t^\alpha}dxdt. \end{aligned}\tag{3.10}$$

Finally, combining (3.5), (3.8), (3.9), and (3.10), we obtain

$$\begin{aligned}\int_\Omega (u-\bar{u})\hat{u}dxdt =& \int_\Omega \left({}_0\partial_t^\alpha\hat{u} - \partial_x^2\hat{u}\right)zdxdt + \int_\Omega \frac{z\eta(x)}{\Gamma(1-\alpha)t^\alpha}dxdt\\ =& \frac{1}{\Gamma(1-\alpha)}\int_\Omega z(x,t)\eta(x)t^{-\alpha}dxdt. \end{aligned}\tag{3.11}$$

This, together with (3.6), leads to (3.4). □

4. Spectral approximation and a priori error estimates

In this section we consider a space-time spectral approximation to the optimal control problem and carry out an error analysis for the numerical solution.

We first define the space

$$P_M^0(\Lambda) := P_M(\Lambda)\cap H_0^1(\Lambda),$$

where P_M denotes the space of all polynomials degree less than or equal to M. The space-time spectral approximation space S_L is then defined as

$$S_L := P_M^0(\Lambda)\otimes P_N(I)\subset B^{\frac{\alpha}{2}}(\Omega),$$

where L stands for the parameter pair (M,N).

We then define the discrete objective function, which is an approximation to J, as follows:

$$J_L(q_L) := \frac{1}{2}\left\|u_L - \bar{u}\right\|_{0,\Omega}^2 + \frac{\lambda}{2}\left\|q_L\right\|_{0,\Lambda}^2, \quad \forall q_L \in P_M(\Lambda), \tag{4.1}$$

where $u_L = u_L(q_L) \in S_L$ is the solution of the following problem:

$$\mathscr{A}(u_L, v_L) = \left(\frac{q_L t^{-\alpha}}{\Gamma(1-\alpha)}, v_L\right)_\Omega, \quad \forall v_L \in S_L. \tag{4.2}$$

We propose the following spectral approximation to the optimal control problem (2.9): find $q_L^* \in P_M(\Lambda)$ such that

$$J_L(q_L^*) = \min_{q_L \in P_M(\Lambda)} J_L(q_L). \tag{4.3}$$

Similar to the continuous problem, it can be proved that the discrete optimal control problem (4.3) admits a unique solution $q_L^* \in P_M(\Lambda)$, which fulfills the first order optimality condition:

$$J_L'(q_L^*)(\eta) = 0, \quad \forall \eta \in P_M(\Lambda), \tag{4.4}$$

where

$$J_L'(q_L)(\eta) = \frac{1}{\Gamma(1-\alpha)}\int_\Omega z_L t^{-\alpha}\eta dx dt + \lambda\int_\Lambda q_L \eta dx \tag{4.5}$$

with $z_L \in S_L$, the solution of the discrete adjoint state equation:

$$\mathscr{A}(\varphi_L, z_L) = (u_L - \bar{u}, \varphi_L)_\Omega, \quad \forall \varphi_L \in S_L. \tag{4.6}$$

We now carry out an error analysis for the spectral approximation (4.3). To simplify the notations, we let c be a generic positive constant independent of any functions and of any discretization parameters. We use the expression $A \lesssim B$ to mean that $A \le cB$, and $A \cong B$ to mean that $A \lesssim B \lesssim A$.

We first introduce some approximation operators that will be used in the following context. We define the orthogonal projector $\Pi_M^{1,0} : H_0^1(\Lambda) \to P_M^0(\Lambda)$ by: $\forall v \in H_0^1(\Lambda)$, $\Pi_M^{1,0} v \in P_M^0(\Lambda)$, such that

$$\left(\left(\Pi_M^{1,0} v - v\right)', \phi_M'\right)_\Lambda = 0, \quad \forall \phi_M \in P_M^0(\Lambda).$$

Then, for all $v \in H^m(\Lambda) \cap H_0^1(\Lambda)$, $m \ge 1$, the following optimal error estimates hold (see [2]):

$$\left|\Pi_M^{1,0} v - v\right|_{1,\Lambda} \lesssim M^{1-m}\left\|v\right\|_{m,\Lambda}, \tag{4.7a}$$

$$\left\|\Pi_M^{1,0} v - v_{0,\Lambda}\right\| \lesssim M^{-m} \|v\|_{m,\Lambda}. \tag{4.7b}$$

Now, we construct the projection operator $\Pi_N^1 : H^1(I) \to P_N(I)$ by:

$$\int_I \left[(\Pi_N^1 v - v)' w_N' + (\Pi_N^1 v - v) w_N\right] dt = 0, \quad \forall w_N \in P_N(I).$$

The following error estimate is also known (see [2]):

$$\left\|\Pi_N^1 v - v\right\|_{k,I} \lesssim N^{k-m} \|v\|_{m,I}, \quad \forall v \in H^m(I), \quad m \geq 1, \quad k = 0,1. \tag{4.8}$$

For $0 < s < 1$, we can derive, by applying the standard space interpolation technique [1], the H^s-error estimate as follows:

$$\left\|\Pi_N^1 v - v\right\|_{s,I} \lesssim N^{s-m} \|v\|_{m,I}, \quad \forall v \in H^m(I), \quad m \geq 1. \tag{4.9}$$

Similar to the proof of Lemma 3.2 in [8], we obtain an error estimate for the composite projection operator $\Pi_N^1 \Pi_M^{1,0}$, which is stated below.

Lemma 4.1. *If $v \in H^s(I; H^\mu(\Lambda)) \cap H^\gamma(I; H_0^1(\Lambda))$, $0 < s < 1, \gamma \geq 1$, $\mu \geq 1$, then we have*

$$\left\|\partial_x (v - \Pi_N^1 \Pi_M^{1,0} v)\right\|_{0,0} \lesssim M^{1-\mu} \|v\|_{\mu,0} + N^{-\gamma} \|v\|_{1,\gamma}, \tag{4.10a}$$

$$\left\|{}_0^R\partial_t^s (v - \Pi_N^1 \Pi_M^{1,0} v)\right\|_{0,0} \lesssim N^{s-\gamma} \|v\|_{0,\gamma} + N^{s-\gamma} M^{-\mu} \|v\|_{\mu,\gamma} + M^{-\mu} \|v\|_{\mu,s}. \tag{4.10b}$$

We now introduce the auxiliary problem:

$$\mathscr{A}(u_L(q), v_L) = \left(\frac{q t^{-\alpha}}{\Gamma(1-\alpha)}, v_L\right)_\Omega, \qquad \forall v_L \in S_L, \tag{4.11a}$$

$$\mathscr{A}(\varphi_L, z_L(q)) = (u_L(q) - \bar{u}, \varphi_L)_\Omega, \qquad \forall \varphi_L \in S_L, \tag{4.11b}$$

where $q \in L^2(\Lambda)$ and $u_L(q), z_L(q) \in S_L$. Then it can be verified by a direct calculation that

$$J_L'(q)(\eta) = \frac{1}{\Gamma(1-\alpha)} \int_\Omega z_L(q) t^{-\alpha} \eta \, dx dt + \lambda \int_\Lambda q \eta \, dx, \quad \eta \in L^2(\Lambda), \tag{4.12}$$

where

$$J_L(q) = \frac{1}{2} \left\|u_L(q) - \bar{u}\right\|_{0,\Omega}^2 + \frac{\lambda}{2} \|q\|_{0,\Lambda}^2. \tag{4.13}$$

Following [8], the error between the solution of (2.7) and the solution of (4.11a) can be estimated as follows:

Lemma 4.2. *For any $q \in L^2(\Lambda)$, let $u(q)$ be the solution of* (2.7), *$u_L(q)$ be the solution of* (4.11a). *Suppose $u \in H^{\alpha/2}(I; H^\mu(\Lambda)) \cap H^\gamma(I; H_0^1(\Lambda))$, $0 < \alpha < 1$, $\gamma \geq 1$, $\mu \geq 1$, then we have*

$$\begin{aligned}\left\|u(q) - u_L(q)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \lesssim & N^{\frac{\alpha}{2}-\gamma} \|u\|_{0,\gamma} + N^{-\gamma} \|u\|_{1,\gamma} + N^{\frac{\alpha}{2}-\gamma} M^{-\mu} \|u\|_{\mu,\gamma} \\ & + M^{-\mu} \|u\|_{\mu,\frac{\alpha}{2}} + M^{1-\mu} \|u\|_{\mu,0}. \end{aligned} \tag{4.14}$$

We now analyze the approximation error of the proposed spectral method, and derive some error estimates for the control and state variables. The proof of the main result will be accomplished with a series of lemmas which we present below.

Lemma 4.3. *For all $p,q \in L^2(\Lambda)$, we have*

$$J_L'(p)(p-q) - J_L'(q)(p-q) \geq \lambda \left\|p-q\right\|_{0,\Lambda}^2. \tag{4.15}$$

Proof. A direct calculation by using (4.12) gives

$$\begin{aligned} J_L'(p)(p-q) - J_L'(q)(p-q) &= \int_\Omega \frac{(z_L(p)-z_L(q))(p-q)}{\Gamma(1-\alpha)t^\alpha} dx dt + \lambda \int_\Lambda (p-q)^2 dx \\ &= \Big(\frac{(p-q)t^{-\alpha}}{\Gamma(1-\alpha)}, z_L(p)-z_L(q)\Big)_\Omega + \lambda \left\|p-q\right\|_{0,\Lambda}^2. \end{aligned} \tag{4.16}$$

For the first term in the right hand side, it follows from (4.11a) and (4.11b) that

$$\begin{aligned} \Big(\frac{(p-q)t^{-\alpha}}{\Gamma(1-\alpha)}, z_L(p)-z_L(q)\Big)_\Omega &= \mathscr{A}\left(u_L(p)-u_L(q), z_L(p)-z_L(q)\right) \\ &= \big(u_L(p)-u_L(q), u_L(p)-u_L(q)\big)_\Omega \\ &\geq 0. \end{aligned} \tag{4.17}$$

Then, combining (4.16) and (4.17) gives (4.15). □

Lemma 4.4. *Let q^* be the solution of the continuous optimization problem* (2.9)*, q_L^* be the solution of the discrete optimization problem* (4.3)*. Suppose $q^* \in H^\mu(\Lambda)$, $\mu \geq 1$, then it holds*

$$\left\|q^* - q_L^*\right\|_{0,\Lambda} \lesssim M^{-\mu} \left\|q^*\right\|_{\mu,\Lambda} + \left\|z(q^*) - z_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)}, \tag{4.18}$$

where $z(q^*)$ and $z_L(q^*)$ are respectively the solutions of (3.3) *and* (4.11b) *associated to q^*.*

Proof. To prove the asserted result, we split the error to be estimated in the following way:

$$\left\|q^* - q_L^*\right\|_{0,\Lambda} \leq \left\|q^* - p_L\right\|_{0,\Lambda} + \left\|p_L - q_L^*\right\|_{0,\Lambda}, \quad \forall p_L \in P_M(\Lambda). \tag{4.19}$$

First it follows from Lemma 4.3:

$$\lambda \left\|p_L - q_L^*\right\|_{0,\Omega}^2 \leq J_L'(p_L)(p_L - q_L^*) - J_L'(q_L^*)(p_L - q_L^*), \quad \forall p_L \in P_M(\Lambda). \tag{4.20}$$

In virtue of (3.1) and (4.4), we have

$$J'(q^*)(p_L - q_L^*) = J_L'(q_L^*)(p_L - q_L^*) = 0, \quad \forall p_L \in P_M(\Lambda).$$

Combining these equalities with (3.4), (4.5) and (4.12), and using the Hölder inequality, we obtain

$$\lambda \left\|p_L - q_L^*\right\|_{0,\Lambda}^2 \leq J_L'(p_L)(p_L - q_L^*) - J'(q^*)(p_L - q_L^*)$$

$$
\begin{aligned}
&=J_L'(p_L)(p_L-q_L^*)-J_L'(q^*)(p_L-q_L^*)+J_L'(q^*)(p_L-q_L^*)-J'(q^*)(p_L-q_L^*)\\
&=\lambda\int_\Lambda (p_L-q^*)(p_L-q_L^*)dx+\frac{1}{\Gamma(1-\alpha)}\int_\Omega \big(z_L(p_L)-z_L(q^*)\big)t^{-\alpha}(p_L-q_L^*)dxdt\\
&\quad+\frac{1}{\Gamma(1-\alpha)}\int_\Omega \big(z(q^*)-z_L(q^*)\big)t^{-\alpha}(p_L-q_L^*)dxdt\\
&=\lambda\int_\Lambda (p_L-q^*)(p_L-q_L^*)dx+\frac{1}{\Gamma(1-\alpha)}\int_I t^{-\alpha}\int_\Lambda \big(z_L(p_L)-z_L(q^*)\big)(p_L-q_L^*)dxdt\\
&\quad+\frac{1}{\Gamma(1-\alpha)}\int_I t^{-\alpha}\int_\Lambda \big(z(q^*)-z_L(q^*)\big)(p_L-q_L^*)dxdt\\
&\lesssim\lambda\left\|p_L-q^*\right\|_{0,\Lambda}\left\|p_L-q_L^*\right\|_{0,\Lambda}+\int_I t^{-\alpha}\left\|z_L(p_L)(\cdot,t)-z_L(q^*)(\cdot,t)\right\|_{0,\Lambda}\left\|p_L-q_L^*\right\|_{0,\Lambda}dt\\
&\quad+\int_I t^{-\alpha}\left\|z(q^*)(\cdot,t)-z_L(q^*)(\cdot,t)\right\|_{0,\Lambda}\left\|p_L-q_L^*\right\|_{0,\Lambda}dt\\
&\lesssim\lambda\left\|p_L-q^*\right\|_{0,\Lambda}\left\|p_L-q_L^*\right\|_{0,\Lambda}+\left\|p_L-q_L^*\right\|_{0,\Lambda}\left\|t^{-\alpha}\right\|_{L^s(I)}\left\|z_L(p_L)-z_L(q^*)\right\|_{L^{s'}(I;L^2(\Lambda))}\\
&\quad+\left\|p_L-q_L^*\right\|_{0,\Lambda}\left\|t^{-\alpha}\right\|_{L^s(I)}\left\|z(q^*)-z_L(q^*)\right\|_{L^{s'}(I;L^2(\Lambda))},
\end{aligned}
$$

where $s=2/(1+\alpha)$, $s'=2/(1-\alpha)$. By simplifying the both sides, we obtain

$$
\begin{aligned}
\lambda\left\|p_L-q_L^*\right\|_{0,\Lambda}\lesssim&\lambda\left\|p_L-q^*\right\|_{0,\Lambda}+\left\|t^{-\alpha}\right\|_{L^s(I)}\left\|z_L(p_L)-z_L(q^*)\right\|_{L^{s'}(I;L^2(\Lambda))}\\
&+\left\|t^{-\alpha}\right\|_{L^s(I)}\left\|z(q^*)-z_L(q^*)\right\|_{L^{s'}(I;L^2(\Lambda))}.
\end{aligned}
\tag{4.21}
$$

Furthermore, by the Embedding Theorem [1], we have

$$
H^{\frac{\alpha}{2}}(I)\hookrightarrow L^{s'}(I).
$$

As a result, it holds

$$
\left\|z_L(p_L)-z_L(q^*)\right\|_{L^{s'}(I;L^2(\Lambda))}\lesssim\left\|z_L(p_L)-z_L(q^*)\right\|_{H^{\frac{\alpha}{2}}(I;L^2(\Lambda))}\lesssim\left\|z_L(p_L)-z_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)},
$$

and

$$
\left\|z(q^*)-z_L(q^*)\right\|_{L^{s'}(I;L^2(\Lambda))}\lesssim\left\|z(q^*)-z_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)}.
$$

Using these results to (4.21) yields

$$
\begin{aligned}
\lambda\left\|p_L-q_L^*\right\|_{0,\Lambda}\lesssim&\lambda\left\|p_L-q^*\right\|_{0,\Lambda}+\left\|t^{-\alpha}\right\|_{L^s(I)}\left\|z_L(p_L)-z_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)}\\
&+\left\|t^{-\alpha}\right\|_{L^s(I)}\left\|z(q^*)-z_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)}.
\end{aligned}
\tag{4.22}
$$

Note that $z_L(p_L)-z_L(q^*)$ solves

$$
\mathscr{A}\big(\varphi_L,z_L(p_L)-z_L(q^*)\big)=\big(u_L(p_L)-u_L(q^*),\varphi_L\big)_\Omega,\quad\forall\varphi_L\in S_L,
\tag{4.23}
$$

and $u_L(p_L)-u_L(q^*)$ satisfies

$$\mathscr{A}(u_L(p_L)-u_L(q^*),v_L)=\Big(\frac{(p_L-q^*)t^{-\alpha}}{\Gamma(1-\alpha)},v_L\Big)_\Omega,\quad \forall v_L\in S_L. \tag{4.24}$$

On the other hand, the bilinear form $\mathscr{A}(\cdot,\cdot)$ satisfies the following continuity and coercivity [8]:

$$\mathscr{A}(u,v)\lesssim \|u\|_{B^{\frac{\alpha}{2}}(\Omega)}\|v\|_{B^{\frac{\alpha}{2}}(\Omega)},\quad \mathscr{A}(v,v)\gtrsim \|v\|^2_{B^{\frac{\alpha}{2}}(\Omega)},\quad \forall u,v\in B^{\frac{\alpha}{2}}(\Omega).$$

Thus, taking $v_L=u_L(p_L)-u_L(q^*)$ in (4.24) gives

$$\big\|u_L(p_L)-u_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}\lesssim \big\|t^{-\alpha}\big\|_{L^s(I)}\big\|p_L-q^*\big\|_{0,\Lambda}. \tag{4.25}$$

Similarly, taking $\varphi_L=z_L(p_L)-z_L(q^*)$ in (4.23) yields

$$\big\|z_L(p_L)-z_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}\lesssim \big\|u_L(p_L)-u_L(q^*)\big\|_{0,\Omega}\lesssim \big\|u_L(p_L)-u_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}.$$

Bringing (4.25) into above inequality, we obtain

$$\big\|z_L(p_L)-z_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}\lesssim \big\|t^{-\alpha}\big\|_{L^s(I)}\big\|p_L-q^*\big\|_{0,\Lambda}. \tag{4.26}$$

Then by combining (4.22) and (4.26), we get

$$\begin{aligned}\lambda\big\|p_L-q_L^*\big\|_{0,\Lambda}\lesssim&\lambda\big\|p_L-q^*\big\|_{0,\Lambda}+\big\|t^{-\alpha}\big\|^2_{L^s(I)}\big\|p_L-q^*\big\|_{0,\Lambda}\\&+\big\|t^{-\alpha}\big\|_{L^s(I)}\big\|z(q^*)-z_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}.\end{aligned} \tag{4.27}$$

Plugging (4.27) into (4.19) yields

$$\begin{aligned}\big\|q^*-q_L^*\big\|_{0,\Lambda}\lesssim&\Big(2+\frac{1}{\lambda}\big\|t^{-\alpha}\big\|^2_{L^s(I)}\Big)\big\|p_L-q^*\big\|_{0,\Lambda}\\&+\frac{1}{\lambda}\big\|t^{-\alpha}\big\|_{L^s(I)}\big\|z(q^*)-z_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}.\end{aligned} \tag{4.28}$$

Since the above estimate is true for all $p_L\in P_M(\Lambda)$, we take $p_L=\Pi_M q^*$ in (4.28), with Π_M standing for the standard L^2-projector, to obtain

$$\begin{aligned}\big\|q^*-q_L^*\big\|_{0,\Lambda}\lesssim&\Big(2+\frac{1}{\lambda}\big\|t^{-\alpha}\big\|^2_{L^s(I)}\Big)M^{-\mu}\big\|q^*\big\|_{\mu,\Lambda}\\&+\frac{1}{\lambda}\big\|t^{-\alpha}\big\|_{L^s(I)}\big\|z(q^*)-z_L(q^*)\big\|_{B^{\frac{\alpha}{2}}(\Omega)}.\end{aligned} \tag{4.29}$$

Finally, a simple calculation shows

$$\big\|t^{-\alpha}\big\|_{L^s(I)}=\Big(\frac{1+\alpha}{1-\alpha}\Big)^{\frac{1+\alpha}{2}}T^{\frac{1-\alpha}{2}},\quad s=\frac{2}{1+\alpha}.$$

This quantity is bounded for any fixed T and $\alpha\in(0,1)$. Thus we obtain (4.18). □

Lemma 4.5. *Given $q \in L^2(\Lambda)$, let $z(q) \in B^{\alpha/2}(\Omega)$ be the solution of the adjoint state problem* (3.3)*, $z_L(q)$ be the solution of its approximation problem* (4.11b)*, then we have*

$$\left\|z(q)-z_L(q)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \lesssim \left\|u(q)-u_L(q)\right\|_{0,\Omega}+\inf_{\forall \varphi_L \in S_L}\left\|z-\varphi_L\right\|_{B^{\frac{\alpha}{2}}(\Omega)}. \tag{4.30}$$

Proof. We denote by $\tilde{z}_L \in S_L$ the solution of the Galerkin approximation to (3.3) as follows:

$$\mathscr{A}(\varphi_L,\tilde{z}_L)=(u-\bar{u},\varphi_L)_\Omega, \quad \forall \varphi_L \in S_L. \tag{4.31}$$

Then it holds

$$\mathscr{A}(\varphi_L, z-\tilde{z}_L)=0, \quad \forall \varphi_L \in S_L,$$

and therefore

$$\left\|z-\tilde{z}_L\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \lesssim \inf_{\varphi_L \in S_L}\left\|z-\varphi_L\right\|_{B^{\frac{\alpha}{2}}(\Omega)}. \tag{4.32}$$

Subtracting (4.11b) from (4.31) gives

$$\mathscr{A}(\varphi_L,\tilde{z}_L-z_L(q))=(u-u_L(q),\varphi_L)_\Omega, \quad \forall \varphi_L \in S_L.$$

By taking $\varphi_L=\tilde{z}_L-z_L(q)$ in the above equation and using the coercivity of $\mathscr{A}(\cdot,\cdot)$, we obtain

$$\left\|\tilde{z}_L-z_L(q)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \lesssim \left\|u-u_L(q)\right\|_{0,\Omega}. \tag{4.33}$$

Finally, (4.30) results from binding (4.32), (4.33), and the triangle inequality. $\square$

We are now in a position to prove the main result concerning the approximation error for our optimal control problem.

Theorem 4.1. *Suppose q^* and q_L^* are respectively the solutions of the continuous optimization problem* (2.9) *and its discrete counterpart* (4.3)*, u^* and u_L^* are the state solutions of* (2.7) *and* (4.2) *associated to q^* and q_L^* respectively, z^* and z_L^* are the associated solutions of* (3.3) *and* (4.6) *respectively. If $q^* \in H^\mu(\Lambda), u^*, z^* \in H^{\alpha/2}(I;H^\mu(\Lambda))\cap H^\gamma(I;H_0^1(\Lambda))$, $0<\alpha<1$, $\gamma \geq 1$, $\mu \geq 1$, then the following estimate holds:*

$$\begin{aligned}
&\left\|q^*-q_L^*\right\|_{0,\Lambda}+\left\|u^*-u_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)}+\left\|z^*-z_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \\
\lesssim & M^{-\mu}\left\|q^*\right\|_{\mu,\Lambda}+M^{1-\mu}\left(\left\|u^*\right\|_{\mu,0}+\left\|z^*\right\|_{\mu,0}\right)+N^{-\gamma}\left(\left\|u^*\right\|_{1,\gamma}+\left\|z^*\right\|_{1,\gamma}\right) \\
&+N^{\frac{\alpha}{2}-\gamma}\left(\left\|u^*\right\|_{0,\gamma}+\left\|z^*\right\|_{0,\gamma}\right)+N^{\frac{\alpha}{2}-\gamma}M^{-\mu}\left(\left\|u^*\right\|_{\mu,\gamma}+\left\|z^*\right\|_{\mu,\gamma}\right) \\
&+M^{-\mu}\left(\left\|z^*\right\|_{\mu,\frac{\alpha}{2}}+\left\|u^*\right\|_{\mu,\frac{\alpha}{2}}\right).
\end{aligned} \tag{4.34}$$

Proof. First we have the following triangle inequality:

$$\left\|u^*-u_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \leq \left\|u(q^*)-u_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} + \left\|u_L(q^*)-u_L(q_L^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)}.$$

By means of (4.25), we obtain

$$\left\|u^*-u_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \lesssim \left\|u(q^*)-u_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} + \left\|q^*-q_L^*\right\|_{0,\Lambda}. \tag{4.35}$$

Similarly, we have

$$\left\|z^*-z_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \lesssim \left\|z(q^*)-z_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} + \left\|q^*-q_L^*\right\|_{0,\Lambda}. \tag{4.36}$$

Thus, by combining (4.18), (4.30), (4.35) and (4.36), we obtain

$$\begin{aligned}&\left\|q^*-q_L^*\right\|_{0,\Lambda} + \left\|u^*-u_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} + \left\|z^*-z_L^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \\ \lesssim& M^{-\mu}\left\|q^*\right\|_{\mu,\Lambda} + \left\|u(q^*)-u_L(q^*)\right\|_{B^{\frac{\alpha}{2}}(\Omega)} + \inf_{\varphi_L\in S_L}\left\|z^*-\varphi_L\right\|_{B^{\frac{\alpha}{2}}(\Omega)}.\end{aligned} \tag{4.37}$$

The last term in the right hand side can be estimated by using Lemma 4.1 as follows

$$\begin{aligned}\inf_{\varphi_L\in S_L}\left\|z^*-\varphi_L\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \leq& \left\|z^*-\Pi_N^1\Pi_M^{1,0}z^*\right\|_{B^{\frac{\alpha}{2}}(\Omega)} \\ \lesssim& M^{1-\mu}\left\|z^*\right\|_{\mu,0} + N^{-\gamma}\left\|z^*\right\|_{1,\gamma} \\ &+ N^{\frac{\alpha}{2}-\gamma}\left\|z^*\right\|_{0,\gamma} + N^{\frac{\alpha}{2}-\gamma}M^{-\mu}\left\|z^*\right\|_{\mu,\gamma} + M^{-\mu}\left\|z^*\right\|_{\mu,\frac{\alpha}{2}}.\end{aligned} \tag{4.38}$$

Finally, plugging the estimates (4.14) with $q=q^*$ and (4.38) into (4.37), we obtain the sought result (4.34). ☐

5. Conjugate gradient optimization algorithm and numerical results

We will first derive the linear system of the spectral approximation based on Gauss numerical quadratures, then describe the conjugate gradient algorithm for the associated discrete optimization problem.

5.1. Implementation with Gauss numerical quadratures

We start with defining the temporal and spatial sampling points.

Let $\{\hat{\xi}_k^M\}_{k=0}^M$ and $\{\hat{\rho}_k^M\}_{k=0}^M$ be, respectively, the Gauss-Lobatto-Legendre (GLL) points and weights, such that

$$\int_{-1}^{1}\phi(x)dx = \sum_{k=0}^{M}\phi(\hat{\xi}_k^M)\hat{\rho}_k^M, \quad \forall\phi(x)\in P_{2M-1}(\Lambda).$$

The $(M+1)\times(N+1)$ GLL points in Ω are then defined by

$$(x_k,t_l) := \Big(\hat{\xi}_k^M, (\hat{\xi}_l^N+1)\frac{T}{2}\Big), \quad k=0,1,\cdots,M; \quad l=0,1,\cdots,N.$$

The corresponding weights are

$$\rho_k^M \rho_l^N \text{ with } \rho_k^M := \hat{\rho}_k^M,\ \rho_l^N := \frac{T\hat{\rho}_l^N}{2}, \quad k = 0,1,\cdots,M; \quad l = 0,1,\cdots,N.$$

The discrete scalar product $(\cdot,\cdot)_M$ and $(\cdot,\cdot)_L$ are defined by

$$(u,v)_M := \sum_{k=0}^{M} \int_I u(x_k,t)v(x_k,t)\rho_k^M dt, \tag{5.1a}$$

$$(u,v)_L := \sum_{k=0}^{M}\sum_{l=0}^{N} u(x_k,t_l)v(x_k,t_l)\rho_k^M \rho_l^N. \tag{5.1b}$$

Let $\{h_i^x : i = 0,\cdots,M\}$ and $\{h_j^t : j = 0,\cdots,N\}$ be the Lagrangian polynomials respectively associated with GLL points $\{x_i : i = 0,\cdots,M\}$ and $\{t_j : j = 0,\cdots,N\}$. It is readily seen that the set $\{h_i^x h_j^t,\ i = 1,\cdots,M-1;\ j = 0,\cdots,N\}$ serves as a basis of $P_M^0(\Lambda)\otimes P_N(I)$, i.e.,

$$P_M^0(\Lambda)\otimes P_N(I) = \text{span}\left\{h_i^x(x)h_j^t(t),\ i = 1,\cdots,M-1;\ j = 0,\cdots,N\right\}.$$

Similarly, the set $\{h_i^x, i = 0,\cdots,M\}$ forms a basis of $P_M(\Lambda)$:

$$P_M(\Lambda) = \text{span}\{h_i^x(x),\ i = 0,\cdots,M\},$$

which is the spectral approximate space for the initial condition. Expressing u_L and z_L in the Lagrangian basis $\{h_i^x h_j^t,\ i = 1,\cdots,M-1;\ j = 0,\cdots,N\}$, and choosing each test function v_L, φ_L to be these basis functions, we arrive at the matrix statement of (4.2) and (4.6):

$$\mathbf{Au} = \mathbf{f}, \tag{5.2a}$$

$$\mathbf{Bz} = \mathbf{F}, \tag{5.2b}$$

where $\mathbf{u} = (u_{ij})_{(M-1)(N+1)}$ and $\mathbf{z} = (z_{ij})_{(M-1)(N+1)}$ are the state and costate unknown vectors, with u_{ij} and z_{ij} approximations to $u(x_i,t_j)$ and $z(x_i,t_j)$ respectively, $\mathbf{A} = (a_{mn,ij})_{((M-1)(N+1))^2}$, $\mathbf{B} = (b_{mn,ij})_{((M-1)(N+1))^2}$ with

$$a_{mn,ij} = \delta_{im}\rho_m^M \int_I {}_0^R\partial_t^{\frac{\alpha}{2}} h_j^t(t)\, {}_t^R\partial_T^{\frac{\alpha}{2}} h_n^t(t)dt + \sum_{k=0}^{M} D_{ki}D_{km}\delta_{jn}\, \rho_k^M \rho_n^N, \tag{5.3a}$$

$$b_{mn,ij} = \delta_{im}\rho_m^M \int_I {}_0^R\partial_t^{\frac{\alpha}{2}} h_n^t(t)\, {}_t^R\partial_T^{\frac{\alpha}{2}} h_j^t(t)dt + \sum_{k=0}^{M} D_{ki}D_{km}\delta_{jn}\, \rho_k^M \rho_n^N, \tag{5.3b}$$

$$m,i = 1,\cdots,M-1; \quad n,j = 0,\cdots,N.$$

In (5.3a) and (5.3b), δ denotes the Kronecker symbol, $D_{ij} = \partial_x h_j^x(x_i)$, and $D = (D_{ij})$ is usually called the derivative matrix in space. The right hand side vectors $\mathbf{f} = (f_{mn})_{(M-1)(N+1)}$ and $\mathbf{F} = (F_{mn})_{(M-1)(N+1)}$ are given by

$$f_{mn} = \frac{q_m \rho_m^M}{\Gamma(1-\alpha)} \int_I h_n^t(t)t^{-\alpha}dt, \quad F_{mn} = (u_{mn} - \bar{u}(x_m,t_n))\rho_m^M \rho_n^N,$$

with $m=1,\cdots,M-1$; $n=0,\cdots,N$, where $q_m=q_L(x_m)$.

The discrete gradient direction, denoted by $\mathbf{g}=(g_m)_{M+1}$, can be obtained by

$$g_m=\frac{1}{\Gamma(1-\alpha)}\sum_{j=0}^{N}z_{mj}\int_I h_j^t(t)t^{-\alpha}dt+\lambda q_m,\quad m=0,\cdots,M.$$

5.2. Conjugate gradient optimization algorithm

We propose below a conjugate gradient algorithm for the overall discrete optimization problem. The algorithm is an iterative process to construct vector sequences $\mathbf{q^{(k)}}$, which will get closer and closer to the desired solutions.

Optimization algorithm Given initial guess $\mathbf{q^{(0)}}$

(a) Solve problems $\mathbf{Au^{(0)}}=\mathbf{f^{(0)}}$ and $\mathbf{Bz^{(0)}}=\mathbf{F^{(0)}}$ to determine $\mathbf{z^{(0)}}$. Compute $\mathbf{g^{(0)}}$ by

$$g_m^{(0)}=\frac{1}{\Gamma(1-\alpha)}\sum_{j=0}^{N}z_{mj}^{(0)}\int_I h_j^t(t)t^{-\alpha}dt+\lambda q_m^{(0)},\quad m=0,\cdots,M.$$

Set $\mathbf{s^{(0)}}=\mathbf{g^{(0)}}$, $k=0$.

(b) Compute the auxiliary vector $\mathbf{\tilde{z}^{(k)}}$ by solving $\mathbf{A\tilde{u}^{(k)}}=\mathbf{\tilde{f}^{(k)}}$ and $\mathbf{B\tilde{z}^{(k)}}=\mathbf{\tilde{F}^{(k)}}$, where

$$\mathbf{\tilde{f}^{(k)}}=(\tilde{f}_{mn}^{(k)})_{(M-1)(N+1)},\quad \mathbf{\tilde{F}^{(k)}}=(\tilde{F}_{mn}^{(k)})_{(M-1)(N+1)}$$

with

$$\tilde{f}_{mn}^{(k)}=\frac{s_m^{(k)}\rho_m^M}{\Gamma(1-\alpha)}\int_I h_n^t(t)t^{-\alpha}dt,\quad \tilde{F}_{mn}^{(k)}=\tilde{u}_{mn}^{(k)}\rho_m^M\rho_n^N.$$

Then compute $\mathbf{\tilde{g}^{(k)}}$ by

$$\tilde{g}_m^{(k)}=\frac{1}{\Gamma(1-\alpha)}\sum_{j=1}^{N}\tilde{z}_{mj}^{(k)}\int_I h_j^t(t)t^{-\alpha}dt+\lambda s_m^{(k)}.$$

Set $\rho_k=\frac{\sum_{m=0}^M g_m^{(k)}g_m^{(k)}\rho_m^M}{\sum_{m=0}^M \tilde{g}_m^{(k)}s_m^{(k)}\rho_m^M}$.

(c) Update: $\mathbf{q^{(k+1)}}=\mathbf{q^{(k)}}-\rho_k\mathbf{s^{(k)}}$, $\mathbf{g^{(k+1)}}=\mathbf{g^{(k)}}-\rho_k\mathbf{\tilde{g}^{(k)}}$.

(d) If $\frac{\sum_{m=0}^M(g_m^{(k+1)})^2\rho_m^M}{\sum_{m=0}^M(g_m^{(0)})^2\rho_m^M}\le$ tolerance, then take $\mathbf{q}^*=\mathbf{q^{(k+1)}}$, stop.

Otherwise, let $\beta_k=\frac{\sum_{m=0}^M(g_m^{(k+1)})^2\rho_m^M}{\sum_{m=0}^M(g_m^{(k)})^2\rho_m^M}$, $\mathbf{s^{(k+1)}}=\mathbf{g^{(k+1)}}+\beta_k\mathbf{s^{(k)}}$.

Set: $k=k+1$, go to **(b).**

Remark 5.1. In our calculation, the integrals in the above algorithm are evaluated by Gauss-Lobatto-Jacobi numerical quadrature formula, see [8] for details.

5.3. Numerical results

We carry out in this subsection a series of numerical experiments to demonstrate the efficiency of the proposed optimization algorithm. The main purpose is to numerically verify the a priori error estimates we obtained in the previous sections. In all the calculations, we take $T = 1$ and use the exact solution as the observation data $\bar{u}$.

Example 5.1. Consider the Eq. (2.1) with a right hand side function $f(x,t)$, chosen such that the exact solution $u(x,t) = \sin \pi t \sin \pi x$, and thus, the corresponding initial condition $q(x) = u(x,0) = 0$.

It is an easy matter to verify that the exact solution of the optimization problem (2.9) is $q^*(x) = 0$ and the corresponding state solution $u^*(x,t) = \sin \pi t \sin \pi x$. We thus expect that the proposed spectral optimization algorithm should be able to find the optimal initial condition and the corresponding state solution exactly up to spectral accuracy.

In the first test, we study the effect of the regularization parameter λ in term of the convergence rate of the iterative optimization algorithm. In Fig. 1 we present the convergence history as a function of the iteration number with $M = N = 18$, $\alpha = 0.5$ for several values of λ. It is observed that the algorithm has better convergence property for $\lambda = 1$. The convergence slows down slightly with decreasing λ. In particular, the algorithm fails to converge with $\lambda = 0$.

We then check the convergence behavior of numerical solutions with respect to the polynomial degrees M. In Fig. 2 we plot the errors as functions of the polynomial degrees M with $\lambda = 1, \alpha = 0.5, N = 20$. The initial guess $\mathbf{q}^{(0)}$ is taken to be the constant 1. As expected, the errors show an exponential decay, since in this semi-log representation one observes that the error variations are essentially linear versus the degrees of polynomial. It is worthwhile to mention that, although the initial guess has been taken far from the exact initial condition, the convergence of the optimization algorithm was attained within six iterations (not reported in the paper).

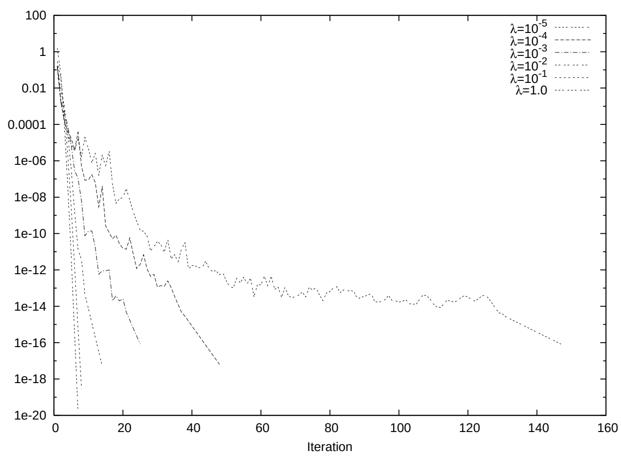

Figure 1: Impact of λ on the convergence rate of the gradient of the objective function.

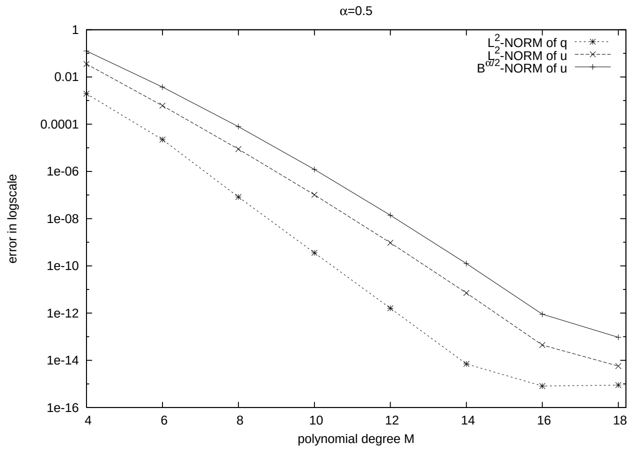

Figure 2: Errors of q and u versus M with $N=20$, $\alpha=0.5$.

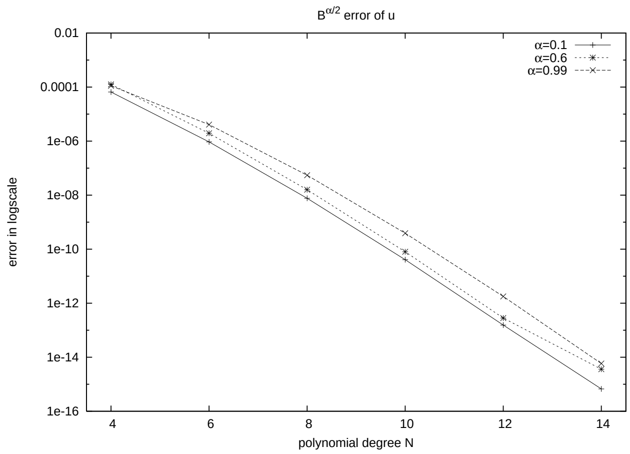

Figure 3: $B^{\alpha/2}$ errors of u versus N with $M=20$, $\alpha=0.1,0.6,0.99$.

We now investigate the temporal error about the state solution u. In Fig. 3, we plot the errors versus N with $M=20$ for several different α. The straight line of the error curves indicates that the convergence in time is also exponential.

Example 5.2. We choose f such that the exact solution u and the exact initial condition q are respectively $u(x,t)=\sin\pi x\cos\pi t$ and $q(x)=\sin\pi x$.

It is believed that a non trivial optimal condition is usually more difficult to capture than a trivial condition, therefore this example can be served to better demonstrate the ability of the algorithm. In this case, the solution of (2.9) differs from the one of the original inverse problem. The error of the regularized solution depends on the magnitude of the regularization parameter, and has been subject of many research, see e.g., [6,17].

Let us emphasize here that the choice of an appropriate regularization parameter λ is

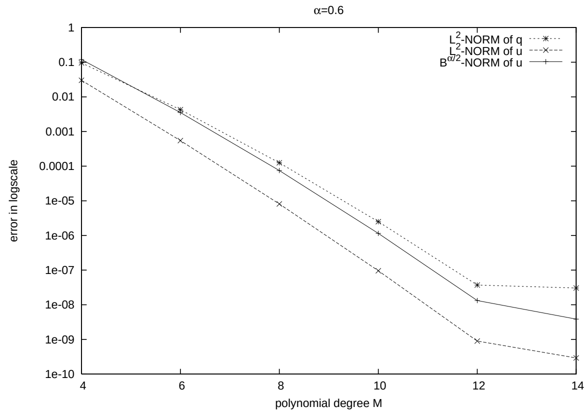

Figure 4: Errors of q and u versus M with $N = 14$, $\alpha = 0.6$.

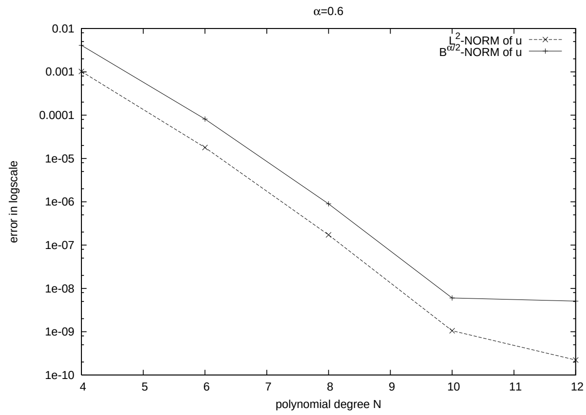

Figure 5: L^2 and $B^{\alpha/2}$ errors of u versus N with $M = 14$, $\alpha = 0.6$.

an important issue for inverse problems. In one side, the regularization parameter plays a role to stabilize the optimization algorithm, and it is required to be large enough to have good stabilization effect. In the other side, this stabilization parameter should not be too large to guarantee that the solution of the stabilized optimization problem approaches the exact solution as well as possible.

First we investigate the influence of the regularization parameter λ on the accuracy. Let e_q (resp: e_u) denote the L^2 error between the exact solution q (resp: u) and the numerical solution q_L^* (resp: u_L^*). In Table 1, we list e_q and e_u for a number of λ with $M = N = 14$. It is observed that the error decays quasi-linearly as λ decreases until the error stemming from the spectral approximation becomes dominant. It is seen that when the regularization parameter $\lambda = 10^{-12}$, the numerical solutions appear to be more satisfactory. Thus, in what follows, we take $\lambda = 10^{-12}$ and $\mathbf{q}^{(0)} = 10$.

We now investigate the errors of the numerical solution with respect to the temporal and spatial approximations. In Fig. 4 and Fig. 5 we report the errors in logarithmic scale

Table 1: L^2 errors of q_L^* and u_L^* for different λ with $M=N=14$, $\alpha=0.6$.

λ	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}	10^{-12}	10^{-14}
e_q	0.98	0.33	4.94E-3	4.96E-5	5.09E-7	1.15E-7	3.05E-8	1.10E-7
e_u	0.14	4.71E-2	7.01E-4	7.05E-6	7.05E-8	8.18E-10	2.91E-10	6.47E-10

as a function of the polynomial degrees M and N respectively. Clearly, all the errors show an exponential decay until the errors associated to the regularization become dominant.

6. Concluding remarks

We presented an efficient optimization algorithm for the inverse problem of the time fractional diffusion equation, which consists in finding the optimal initial condition such that the corresponding solution matches the observed data as closely as possible. The proposed algorithm is based on a time-space spectral approximation to the time fractional diffusion equation and a conjugate gradient iteration to solve the discrete optimal control problem. We established the well-posedness of the regularized optimal problem and derived an error estimate for both control and state variables. The error estimate and the numerical tests showed that the convergence of the spectral optimization algorithm was exponential for smooth solutions. The effect of the regularization parameter on the accuracy and the iteration number was investigated numerically.

There are several potential extensions of the present method. Firstly, the proposed algorithm can be adapted to two and three dimensional problems. Secondly, it can be extended to other inverse problems, such as inverse source, inverse boundary conditions and so on. Thirdly, some popular strategies, such as discrepancy principle, generalized cross-validation and so on, for choosing a more suitable regularization parameter can be considered.

Acknowledgments The work of X. Y was partially supported by the Natural Science Foundation of Fujian Province, China (Grant No. 2012J01013). The work of C. X. was partially supported by National NSF of China (Grants 11071203 and 91130002).

References

[1] R. A. Adams, Sobolev Spaces, Academic Press, New York, 1975.

[2] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, Spectral Methods Fundamentals in Single Domains, Springer-Verlag, Berlin, 2006.

[3] J. Cheng, J. Nakagawa, M. Yamamoto and T. Yamazaki, *Uniqueness in an inverse problem for a one-dimensional fractional diffusion equation*, Inverse Problems, 25(11) (2009), pp. 115002.

[4] V. J. Ervin and J. P. Roop, *Variational formulation for the stationary fractional advection dispersion equation*, Numer. Methods Partial Differential Equations, 22(3) (2006), pp. 558–576.

[5] V. J. Ervin and J. P. Roop, *Variational solution of fractional advection dispersion equations on bounded domains in $\mathbb{R}^d$*, Numer. Methods Partial Differential Equations, 23(2) (2007), pp. 256–281.

[6] V. ISAKOV, Inverse Problems for Partial Differential Equations, Springer Science, New York, 2006.

[7] T. A. M LANGLANDS AND B. I. HENRY, *The accuracy and stability of an implicit solution method for the fractional diffusion equation*, J. Comput. Phys., 205(2) (2005), pp. 719–736.

[8] XIANJUAN LI AND CHUANJU XU, *A space-time spectral method for the time fractional diffusion equation*, SIAM J. Numer. Anal., 47(3) (2009), pp. 2108–2131.

[9] YUMIN LIN AND CHUANJU XU, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225(2) (2007), pp. 1533–1552.

[10] J. L. LIONS, Optimal Control of Systems Governed by Partial Differential Equations, volume 170 of Grundlehren Math. Wiss., Springer-Verlag, Berlin, 1971.

[11] F. LIU, S. SHEN, V. ANH AND I. TURNER, *Analysis of a discrete non-markovian random walk approximation for the time fractional diffusion equation*, ANZIAM J, 46(4) (2005), pp. 488–504.

[12] Y. LUCHKO, *Maximum principle for the generalized time-fractional diffusion equation*, J. Math. Anal. Appl, 351(1) (2009), pp. 218–223.

[13] Y. LUCHKO, *Some uniqueness and existence results for the initial-boundary-value problems for the generalized time-fractional diffusion equation*, Comput. Math. Appl, 59(5) (2010), pp. 1766–1772.

[14] I. PODLUBNY, Fractional Differential Equations, Acad. Press, New York, 1999.

[15] J. P. ROOP, *Computational aspects of FEM approximation of fractional advection dispersion equations on bounded domains in* $\mathbb{R}^2$, J. Comput. Appl. Math., 193(1) (2006), pp. 243–268.

[16] K. SAKAMOTO AND M. YAMAMOTO, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, J. Math. Anal. Appl., 382(1) (2011), pp. 426–447.

[17] A. A. SAMARSKII AND P. N. VABISHCHEVICH, Numerical Methods for Solving Inverse Problems of Mathematical Physics, De Gruyter, Berlin, 2007.

[18] W. R SCHNEIDER AND W. WYSS, *Fractional diffusion and wave equations*, J. Math. Phys, 30(1) (1989), pp. 134–144.

[19] Z. SUN AND X. WU, *A fully discrete difference scheme for a diffusion-wave system*, Appl. Numer. Math., 56(2) (2006), pp. 193–209.

[20] W. WYSS, *The fractional diffusion equation*, J. Math. Phys., 27(11) (1986), pp. 2782–2785.

[21] Y. ZHANG AND X. XU, *Inverse source problem for a fractional diffusion equation*, Inverse Problems, 27(3) (2011), pp. 035010.