Numerical Estimation of a Diffusion Coefficient in Subdiffusion

Bangti Jin*  Zhi Zhou†

Abstract

In this work, we consider the numerical recovery of a spatially dependent diffusion coefficient in a subdiffusion model from distributed observations. The subdiffusion model involves a Caputo fractional derivative of order \( \alpha \in (0,1) \) in time. The numerical estimation is based on the regularized output least-squares formulation, with an \( H^1(\Omega) \) penalty. We prove the well-posedness of the continuous formulation, e.g., existence and stability. Next, we develop a fully discrete scheme based on the Galerkin finite element method in space and backward Euler convolution quadrature in time. We prove the subsequential convergence of the sequence of discrete solutions to a solution of the continuous problem as the discretization parameters (mesh size and time step size) tend to zero. Further, under an additional regularity condition on the exact coefficient, we derive convergence rates in a weighted norm for the discrete approximations to the exact coefficient. The analysis relies heavily on suitable nonstandard nonsmooth data error estimates for the direct problem. We provide illustrative numerical results to support the theoretical study.

Keywords: parameter identification, subdiffusion, fully discrete scheme, convergence, error estimate

1 Introduction

Let \( \Omega \subset \mathbb{R}^d (d = 1, 2, 3) \) be a convex polyhedral domain with a boundary \( \partial \Omega \). Consider the following initial-boundary value problem of the subdiffusion equation:

\[
\begin{aligned}
\left\{ \begin{array}{l}
\partial_t^\alpha u(x,t) - \nabla \cdot (q(x)\nabla u(x,t)) = f(x,t), \quad (x,t) \in \Omega \times (0,T], \\
u(x,0) = u_0(x), \quad x \in \Omega, \\
u(x,t) = 0, \quad (x,t) \in \partial \Omega \times (0,T],
\end{array} \right.
\end{aligned}
\]  

(1.1)

where \( T > 0 \) is the final time. The functions \( f \) and \( u_0 \) are the given source term and initial condition, respectively, and their precise regularity will be specified below. The notation \( \partial_t^\alpha u \), denotes the Caputo fractional derivative in time of order \( \alpha \in (0,1) \), defined by

\[
\partial_t^\alpha u(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t - s)^{-\alpha} \partial_s u(s) \, ds.
\]

The fractional derivative \( \partial_t^\alpha u \) recovers the usual first order derivative as the order \( \alpha \to 1^- \) for sufficiently smooth functions \( u \). Thus the model is a fractional analogue of the classical diffusion model. The model (1.1) has received enormous attention in recent years, due to their extraordinary capability for describing anomalously slow diffusion processes, also known as subdiffusion, which displays local motion occasionally interrupted by long sojourns and trapping effects. These processes are characterized by sublinear growth of the particle mean squared displace with the time. The model has found many successful applications in physics, biology and finance etc; see the reviews [31, 30] for physical modeling and a long list of applications.

This work is concerned with numerically identifying the diffusion coefficient \( q \in L^\infty(\Omega) \) the model (1.1) from the (noisy) distributed observation

\[
u(x,t) = z^\delta(x,t), \quad (x,t) \in \Omega \times [0,T].
\]

(1.2)

*Department of Computer Science, University College London, Gower Street, London, WC1E 6BT, UK. (b.jin@ucl.ac.uk, bangti.jin@gmail.com)

†Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (zhizhou@polyu.edu.hk)
The inverse problem is a fractional analogue of the inverse conductivity problem for standard parabolic problems, which has been extensively studied both numerically and theoretically; see the monograph [13, Chapter 9] for relevant mathematical theory and the references [12, 34, 21, 22, 9, 29, 7, 32, 35] for a rather incomplete list of works on numerical identification of a diffusion coefficient in standard parabolic problems. Numerically, most of these existing works formulate the inverse problem into an output least-squares formulation, with a proper penalty, e.g., Sobolev smoothness or total variation.

In this work, we develop a rigorous numerical procedure for recovering a spatially dependent diffusion coefficient. We formulate an output least-squares formulation with an $H^1(\Omega)$ penalty, and provide a complete analysis of both continuous and discrete formulations, including well-posedness and convergence of discrete approximations, for weak regularity assumption on the problem data, in Sections 2 and 3, respectively. Furthermore, in Section 4 we derive some error estimates on the discrete approximation under a mild regularity assumption on the exact diffusion coefficient; see Theorem 4.1 and Corollary 4.1. The results extend the corresponding results for the standard parabolic case [12, 22, 35]. Due to the nonlocality of the Caputo derivative $\partial_t^\alpha u$, many powerful tools from PDE theory and classical numerical analysis, e.g., energy argument, are not directly applicable, and the solution operator has only limited smoothing properties, which represent the main technical challenges in the convergence analysis. Hence, the analysis differs significantly from the standard parabolic counterpart. The error analysis is complicated by the nonlinearity of the forward map, and thus standard techniques from optimal control theory also do not apply. We shall employ the positivity of the fractional derivative operators (in Theorem 2.1 and 3.1), nonsmooth data estimates (in Lemma 4.1) and novel test function $\varphi$ (in Theorem 4.1), to overcome these challenges, which represent the main technical novelties of the work.

Now we briefly review relevant works from the inverse problem literature. Inverse problems for fractional diffusion has started to attract much interest, and there has already been a vast literature (see, e.g., the review [19]). There are a number of interesting works on recovering the diffusion coefficient [4, 26, 27, 37, 23]. In an influential piece of work, Cheng et al [4] proved the unique recovery of both diffusion coefficient and fractional order from the lateral Cauchy data for the model (1.1) with a Dirac source in the one spatial dimensional case. The proof employs Laplace transform and the classical Sturm-Liouville theory. Very recently, Kian et al [23] proved uniqueness for the recovery of two coefficients from the Dirichlet-to-Neumann map [23]. Li et al [26, 27] discussed the numerical recovery of the diffusion coefficient (simultaneously with the fractional order), and showed various continuity results of the parameter to state map. However, the numerical discretization was not analyzed in [27]. Zhang [37] proved the unique recovery for the case of a time-dependent $q \equiv q(t)$, and devised a numerical scheme for its recovery. See also the work [30] for further numerical results on recovering the diffusion coefficient from boundary data in the one-dimensional case, using a space-time variational formulation, which allows only a zero initial condition. However, there is no analysis of the discretized problem, which is the focus of the present work. In sum, none of these existing works has rigorously studied the discretization schemes in a proper functional analytic setting, and it is precisely this gap that this work aims to fill in. We refer interested readers also to the works [38, 15, 20, 39] and references therein for further numerical methods on related nonlinear inverse problems.

The rest of the paper is organized as follows. In Section 2 we formulate the continuous problem, and analyze its well-posedness, e.g., existence and stability. Then in Section 3 we describe a fully discrete scheme, and show the convergence of the discrete approximations to a solution of the continuous problem as the discretization parameters tend to zero. In Section 4 we provide detailed error estimates for the discrete approximations under suitable regularity assumption on the exact coefficient. Finally, in Section 5 we present illustrative one- and two-dimensional numerical results to complement the theoretical analysis.

We end this section with some useful notation. Throughout, the notation $c$ denotes a generic constant, which may change at each occurrence, but is always independent of $q$, the mesh size $h$ and time stepsize $\tau$ etc. We shall employ standard notation for Sobolev spaces [11]. The spaces $L^p(\Omega)$ and $H^1(\Omega)$ are endowed with the norms $\|\cdot\|_{L^p(\Omega)}$ and $\|\cdot\|_{H^1(\Omega)}$, respectively, and the notation $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. We denote by $H^{-1}(\Omega)$ the dual space of $H^1_0(\Omega)$. For a Banach space $B$ (endowed with the norm $\|\cdot\|_B$), we define

$$L^2(0, T; B) = \{u(t) \in B \text{ for a.e. } t \in (0, T) \text{ and } \|u\|_{L^2(0, T; B)} < \infty\},$$

and the norm is given by $\|u\|_{L^2(0, T; B)} = \left( \int_0^T \|u(t)\|_B^2 \, dt \right)^{\frac{1}{2}}$. Similarly, the space $H^1(0, T; B)$ denotes

$$H^1(0, T; B) = \{u \in L^2(0, T; B) : u'(t) \in L^2(0, T; B)\},$$

as the discretization parameters tend to zero. In Section 4, we provide detailed error estimates for the discrete approximations under suitable regularity assumption on the exact coefficient. Finally, in Section 5 we present illustrative one- and two-dimensional numerical results to complement the theoretical analysis.

We end this section with some useful notation. Throughout, the notation $c$ denotes a generic constant, which may change at each occurrence, but is always independent of $q$, the mesh size $h$ and time stepsize $\tau$ etc. We shall employ standard notation for Sobolev spaces [11]. The spaces $L^p(\Omega)$ and $H^1(\Omega)$ are endowed with the norms $\|\cdot\|_{L^p(\Omega)}$ and $\|\cdot\|_{H^1(\Omega)}$, respectively, and the notation $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. We denote by $H^{-1}(\Omega)$ the dual space of $H^1_0(\Omega)$. For a Banach space $B$ (endowed with the norm $\|\cdot\|_B$), we define

$$L^2(0, T; B) = \{u(t) \in B \text{ for a.e. } t \in (0, T) \text{ and } \|u\|_{L^2(0, T; B)} < \infty\},$$

and the norm is given by $\|u\|_{L^2(0, T; B)} = \left( \int_0^T \|u(t)\|_B^2 \, dt \right)^{\frac{1}{2}}$. Similarly, the space $H^1(0, T; B)$ denotes

$$H^1(0, T; B) = \{u \in L^2(0, T; B) : u'(t) \in L^2(0, T; B)\},$$

2
with its norm given by \( \| u \|_{H^1(0,T;B)} = (\| u \|_{L^2(0,T;B)}^2 + \| u' \|_{L^2(0,T;B)}^2)^{1/2} \). Further, for any \( s \geq 0 \), we denote by \( \dot{H}^s(\Omega) = \{ v : (-\Delta)^s v \in L^2 \} \) (with \( \Delta \) being the Laplacian with a zero Dirichlet boundary condition, equipped with the norm \( \| v \|_{\dot{H}^s(\Omega)} = (\| v \|_{L^2(\Omega)}^2 + \| (-\Delta)^s v \|_{L^2(\Omega)}^2)^{1/2} \). Then \( \dot{H}^0(\Omega) = L^2(\Omega), \dot{H}^1(\Omega) = H^1_0(\Omega) \) and \( \dot{H}^2(\Omega) = H^2(\Omega) \cap H^1_0(\Omega) \).

## 2 Well-posedness of the continuous problem

In this section, we analyze the continuous formulation of the reconstruction approach. To recover the diffusion coefficient \( q \), we employ the following output least-squares formulation with an \( H^1(\Omega) \)-penalty:

\[
\min_{q \in A} J_\gamma(u; z^\delta) = \frac{1}{2} \int_0^T \int_{\Omega} |u(q) - z^\delta|^2 \, dx \, dt + \frac{\gamma}{2} \| \nabla q \|_{L^2(\Omega)}^2,
\]

subject to \( q \in A \) and \( u(q) \) satisfying the variational problem

\[
(\partial_t^\alpha u(q), v) + (q \nabla u(q), \nabla v) = (f, v), \quad \forall v \in \dot{H}^1(\Omega), \quad t \in (0, T], \quad \text{with} \quad u(0) = u_0.
\]

The admissible set \( A \) for the diffusion coefficient \( q(x) \) is given by

\[
A = \{ q \in H^1(\Omega) : c_0 \leq q \leq c_1 \ \text{a.e. in} \ \Omega \},
\]

with constants \( c_0, c_1 \in \mathbb{R} \) and \( 0 < c_0 < c_1 \). The \( H^1(\Omega) \) seminorm penalty is suitable for recovering a Sobolev smooth diffusion coefficient. The scalar \( \gamma > 0 \) is the regularization parameter, controlling the strength of the penalty [14]. The dependence of the functional \( J_\gamma \) on \( z^\delta \) will be suppressed whenever there is no confusion.

For the analysis in Sections 2 and 3, we make the following assumption on problem data. It is sufficient to ensure the existence of a unique solution \( u(q) \in L^2(0,T;H^1(\Omega)) \) for any \( q \in A \) [17].

**Assumption 2.1.** \( u_0 \in H^1_0(\Omega), \ f \in L^2(0,T;L^2(\Omega)), \ \text{and} \ z^\delta \in L^2(0,T;L^2(\Omega)). \)

First we show the well-posedness of problem [21] [22], which relies on a continuity result for the parameter-to-state map \( u(q) \). First, we recall a stability result on the solution operator. Below, for any \( q \in A \), the operator \( A(q) : H^1_0(\Omega) \to H^{-1}(\Omega) \) is defined by

\[
-\langle A(q) \varphi, \psi \rangle = (q \nabla \varphi, \nabla \psi), \quad \forall u, v \in H^1_0(\Omega).
\]

For any \( \varphi \in H^1_0(\Omega) \cap H^2(\Omega) \), there holds \( A(q) \varphi = \nabla \cdot (q \nabla \varphi) \in L^2(\Omega) \).

**Lemma 2.1.** For any \( q \in A \), let \( v \) solve

\[
\partial_t^\alpha v - A(q)v = f, \quad \forall t \in (0, T], \quad \text{with} \ v(0) = 0.
\]

Then there holds

\[
\| v \|_{L^2(0,T;H^1(\Omega))}^2 \leq c \| f \|_{L^2(0,T;H^{-1}(\Omega))}^2.
\]

**Proof.** Taking the test function \( \phi = v \) in the weak formulation, and then integrating from 0 to \( T \) gives

\[
\int_0^T (\partial_t^\alpha v(t), v(t)) \, dt + \int_0^T (q \nabla v, \nabla v) \, dt = \int_0^T (f, v) \, dt.
\]

Since \( v(0) = 0 \), there holds \( \int_0^T (\partial_t^\alpha v(t), v(t)) \, dt \geq 0 \) [28, Lemma 2.3], and by Poincaré’s inequality and Cauchy-Schwarz inequality, we obtain the desired estimate.

The next result gives the continuity of the parameter-to-state map.

**Lemma 2.2.** If the sequence \( \{ q_n \} \subset A \) that converges to \( q \in A \) in \( L^1(\Omega) \) and a.e., then

\[
\lim_{n \to \infty} \| u(q) - u(q^n) \|_{L^2(0,T;H^1(\Omega))} = 0.
\]

3
Proof. Let \( v^n = u(q^n) - u(q^n) \). Then it satisfies \( v^n(0) = 0 \) and
\[
\partial_t^2 v^n - \nabla \cdot ((q^n \nabla v^n)) = \nabla \cdot ((q - q^n) \nabla u(q)), \quad \forall t \in (0,T).
\]

Then by Lemma \[2.1\] and the definition of the \( H^{-1}(\Omega) \), we obtain
\[
\| v^n \|_{L^2(0,T;H^1(\Omega))} \leq c \| \nabla \cdot ((q - q^n) \nabla u(q)) \|_{L^2(0,T;H^{-1}(\Omega))} \\
\leq c \| (q - q^n) \nabla u(q) \|_{L^2(0,T;L^2(\Omega))}.
\]

Since \( q^n \in \mathcal{A} \) and \( q^n \) converges to \( q \) in \( L^1(\Omega) \) and almost everywhere, by Lebesgue’s dominated convergence theorem \[11\],
\[
\lim_{n \to \infty} \| (q - q^n) \nabla u(q) \|_{L^2(0,T;L^2(\Omega))} = 0,
\]
which shows the desired estimate.

The next result gives the existence of a minimizer.

**Theorem 2.1.** Under Assumption \[2.1\] there exists at least one minimizer to problem \( (2.1) \)–\( (2.2) \).

**Proof.** Since the functional \( J_\gamma \) is bounded from below by zero, there exists a minimizing sequence \( \{ q^n \}_{n \geq 1} \subset \mathcal{A} \) such that \( \lim_{n \to \infty} J_\gamma(q^n) = \inf_{q \in \mathcal{A}} J_\gamma(q) \). Thus, the sequence \( \{ q^n \}_{n \geq 1} \) is uniformly bounded in \( H^1(\Omega) \) seminorm, which together with the box constraint \( q^n \in \mathcal{A} \), implies that it is also uniformly bounded in \( H^1(\Omega) \). Thus there exists a subsequence, still denoted by \( \{ q^n \}_{n \geq 1} \) that converges to some \( q^* \in \mathcal{A} \) weakly in \( H^1(\Omega) \), and by compact Sobolev embedding theorem \[11\], converges also in \( L^1(\Omega) \). Further, by standard measure theory, convergence in \( L^1(\Omega) \) implies almost everywhere convergence up to a subsequence \[11\] Theorem 1.21, p. 29. Thus, we may assume that the subsequence \( \{ q^n \}_{n \geq 1} \) converges to \( q^* \) in \( L^1(\Omega) \) and almost everywhere.

Then by Lemma \[2.2\]
\[
\lim_{n \to \infty} \int_0^T \int_\Omega |u(q^n) - z^\delta|^2 \, dx \, dt = \int_0^T \int_\Omega |u(q^*) - z^\delta|^2 \, dx \, dt.
\]

This and weak lower semi-continuity of semi-norms imply that \( q^* \) is a minimizer to \( (2.1) \). \( \square \)

The following continuous dependence results follow from a standard argument \[8\ \[13\].

**Theorem 2.2.** Under Assumption \[2.1\] the following statements hold.

(i) Let the sequence \( \{ z_j \}_{j \geq 1} \) be convergent to \( z^* \) in \( L^2(0,T;L^2(\Omega)) \), and \( q_j^* \in \mathcal{A} \) the corresponding minimizer to \( J_\gamma(\cdot; z_j) \). Then \( \{ q_j^* \}_{j \geq 1} \) contains a subsequence convergent to a minimizer of \( J_\gamma(\cdot; z^*) \) over \( \mathcal{A} \) in \( H^1(\Omega) \).

(ii) Let \( \{ \delta_j \}_{j \geq 1} \subset \mathbb{R}_+ \) with \( \delta_j \to 0 \), \( \{ z^\delta_j \}_{j \geq 1} \subset L^2(0,T;L^2(\Omega)) \) be a sequence satisfying \( \| z^\delta_j - z^* \|_{L^2(0,T;L^2(\Omega))} = \delta_j \) for some exact data \( z^* \), and \( q_j^* \) be a minimizer to \( J_\gamma(\cdot; z^\delta_j) \) over \( \mathcal{A} \). If the sequence \( \{ \gamma_j \}_{j \geq 1} \) satisfies
\[
\lim_{j \to \infty} \gamma_j = 0 \quad \text{and} \quad \lim_{j \to \infty} \frac{\delta_j^2}{\gamma_j} = 0,
\]
then \( \{ q_j^* \}_{j \geq 1} \) contains a subsequence converging to a minimum-\( H^1(\Omega) \) seminorm solution in \( H^1(\Omega) \).

### 3 Numerical approximation and convergence analysis

Now we describe the discretization of problem \( (2.1) \)–\( (2.2) \) and show the convergence of the approximations.

#### 3.1 Numerical approximation

First, we describe a spatially semidiscrete scheme for problem \( (1.1) \) based on the Galerkin FEM; see \[17\] for a recent overview on the numerical approximation of the subdiffusion model. Let \( T_h \) be a shape regular quasi-uniform triangulation of the domain \( \Omega \) into \( d \)-simplexes, denoted by \( T \), with a mesh size \( h \). Over \( T_h \), we define a continuous piecewise linear finite element space \( X_h \) by
\[
X_h = \{ v_h \in H^1_0(\Omega) : v_h|_T \text{ is a linear function } \forall T \in T_h \}.
\]
and similarly the space $V_h$ by

$$V_h = \{ v_h \in H^1(\Omega) : v_h|_T \text{ is a linear function } \forall T \in T_h \}.$$  

The spaces $X_h$ and $V_h$ will be employed to approximate the state $u$ and the diffusion coefficient $q$, respectively. We define the $L^2(\Omega)$ projection $P_h : L^2(\Omega) \rightarrow X_h$ by

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

Note that the operator $P_h$ satisfies the following error estimate: for any $s \in [1, 2]$,

$$\|P_h \varphi - \varphi\|_{L^2(\Omega)} + h\|\nabla(P_h \varphi - \varphi)\|_{L^2(\Omega)} \leq h^s \|\varphi\|_{H^s(\Omega)}, \quad \forall \varphi \in H^s(\Omega).$$  

Let $I_h$ be the interpolation operator associated with the finite element space $V_h$. Then it has the following error estimates for $s = 1, 2$ (see e.g., [10, Theorem 1.103]):

$$\|v - I_h v\|_{L^2(\Omega)} + h\|v - I_h v\|_{H^1(\Omega)} \leq c h^s \|v\|_{H^s(\Omega)}, \quad \forall v \in H^s(\Omega),$$  

$$\|v - I_h v\|_{L^\infty(\Omega)} + h\|v - I_h v\|_{W^{1, \infty}(\Omega)} \leq c h^s \|v\|_{W^{2, \infty}(\Omega)}, \quad \forall v \in W^{2, \infty}(\Omega).$$  

Now we partition the time interval $[0, T]$ uniformly, with grid points $t_n = n \tau$, $n = 0, \ldots, N$, and a time step size $\tau = T/N$. The fully discrete scheme for problem (1.1) reads: Given $U^0_h = P_h u_0 \in X_h$, find $U^n_h \in X_h$ such that

$$(\partial_t^\tau (U^n_h - U^0_h), \chi) + (q \nabla U^n_h, \nabla \chi) = (f^n, \chi), \quad \forall \chi \in X_h, \quad n = 1, 2, \ldots, N,$$  

where $f^n = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} f(s) \, ds$ and $\partial_t^\tau \varphi^n$ denotes the backward Euler convolution quadrature (CQ) approximation (with $\varphi^j = \varphi(t_j)$):

$$\partial_t^\tau \varphi^n = \tau^{-\alpha} \sum_{j=0}^n b_j^{(\alpha)} \varphi^{n-j}, \quad \text{with } (1 - \xi)^\alpha = \sum_{j=0}^\infty b_j^{(\alpha)} \xi^j. \quad (3.4)$$  

Upon letting the discrete operator $A_h(q) : X_h \rightarrow X_h$ by $- (A_h(q)v_h, \chi) = (q \nabla v_h, \nabla \chi)$ for all $v_h, \chi \in X_h$, the fully discrete scheme (3.3) can be rewritten as

$$\partial_t^\tau (U^n_h - U^0_h) - A_h(q) U^n_h = P_h f^n, \quad n = 1, 2, \ldots, N.$$  

Now we can formulate the finite element discretization of problem (2.1)–(2.2):

$$\min_{q_h \in A_h} J_{\gamma, h, \tau}(q_h) = \frac{T}{2} \sum_{n=1}^N \int_{\Omega} \left| U^n_h(q_h) - z^n_h \right|^2 \, dx + \frac{\gamma}{2} \|\nabla q_h\|_{L^2(\Omega)}^2, \quad (3.5)$$  

with $z^n_h = \tau^{-1} \int_{t_{n-1}}^{t_n} z^\delta \, dt$, subject to $q_h \in A_h$ and $U^n_h(q_h)$ satisfying

$$\partial_t^\tau (U^n_h(q_h) - U^0_h) + A_h(q_h) \nabla U^n_h(q_h) = P_h f^n, \quad n = 1, 2, \ldots, N, \quad (3.6)$$  

with $U^0_h = P_h u_0$. The discrete admissible set $A_h$ is taken to be

$$A_h = \{ q_h \in V_h : a \leq q_h(x) \leq b \quad \text{in} \quad \Omega \}.$$  

Clearly, $A_h = A \cap V_h$. Problem (3.5)–(3.6) is a finite-dimensional nonlinear optimization problem with PDE and box constraint, and can be solved efficiently. The analysis of problem (3.5)–(3.6) is the main focus of Sections 3.2 and 4.
3.2 Existence and convergence

This part is devoted to the convergence analysis of the discrete approximations given by the scheme (3.5)–(3.6) to the continuous formulation (2.1)–(2.2). We begin with some a priori estimate on the solutions of the time-stepping scheme (3.3). The proof relies on positivity of CQ.

Lemma 3.1. Let $V^n_h \in X_h$ solve
\[
(\partial_t^\alpha V^n_h, \chi) + (q_h \nabla V^n_h, \nabla \chi) = (f^n_h, \chi), \quad \forall \chi \in X_h, \quad n = 1, 2, \ldots, N, \quad \text{with } V^0_h = 0.
\]
Then there holds
\[
\tau \sum_{n=1}^{N} (\nabla V^n_h, \nabla V^n_h) \leq c \tau \sum_{n=1}^{N} (f^n_h, V^n_h).
\]

Proof. Upon letting $\chi = V^n_h \in X_h$ and then summing over $n$ leads to
\[
\tau \sum_{n=1}^{N} (\partial_t^\alpha V^n_h, V^n_h) + \tau \sum_{n=1}^{N} (q_h \nabla V^n_h, V^n_h) = \tau \sum_{n=1}^{N} (f^n_h, V^n_h).
\]
Now we shall show that the first term on the left hand side is nonnegative. To this end, we extend \(\{V^n_h\}_{n=0}^{N}\) to \(\{V^n_h\}_{n=-\infty}^{\infty}\) and \(\{b^n(\alpha)\}_{n=0}^{\infty}\) to \(\{b^n(\alpha)\}_{n=-\infty}^{\infty}\) by zero. Then \(\partial_t^\alpha V^n_h\) can be written as \(\partial_t^\alpha V^n_h = \tau^{-\alpha} \sum_{j=-\infty}^{\infty} b^n(\alpha) j V^j_h\). Next we denote the discrete Fourier transform \(\tilde{V}^n_h(\xi)\) by \(\tilde{V}^n_h(\xi) = \sum_{n=-\infty}^{\infty} V^n_h e^{-i n \xi}\). By Parseval’s theorem, since $V^0_h = 0$, we have
\[
\sum_{j=1}^{N} (\partial_t^\alpha V^n_h, V^n_h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\tilde{\partial_t^\alpha V^n_h}(\xi), [\tilde{V}^n_h(\xi)]^*) \, d\xi.
\]
By the property of discrete Fourier transform, we have
\[
\sum_{j=1}^{N} (\partial_t^\alpha V^n_h, V^n_h) = \frac{\tau^{-\alpha}}{2\pi} \int_{-\pi}^{\pi} \left| b^n(\alpha) \right| \left| \tilde{V}^n_h(\xi) \right|^2 \, d\xi
\]
\[
= \frac{\tau^{-\alpha}}{2\pi} \int_{-\pi}^{\pi} \left( 1 - e^{-i \xi} \right)^\alpha \left| \tilde{y}_n(\xi) \right|^2 \, d\xi
\]
\[
= \frac{\tau^{-\alpha}}{\pi} \int_{0}^{\pi} \left[ \Re \left( 1 - e^{-i \xi} \right) \right] \left| \tilde{y}_n(\xi) \right|^2 \, d\xi \geq 0.
\]
Then Cauchy-Schwarz inequality and Poincaré’s inequality imply the desired estimate.

Lemma 3.2. The following statements hold
\[
\sum_{n=0}^{m} b^n(\alpha) = b^{(\alpha-1)}_m \quad \text{and} \quad \tau^{-\alpha} \sum_{n=0}^{m} b^n(\alpha) \leq ct^{-\alpha}_{m+1}.
\]

Proof. Let $\sum_{n=0}^{m} b^n(\alpha) = v_m$. Then
\[
\sum_{m=0}^{\infty} v_m \xi^m = \sum_{m=0}^{\infty} \xi^m \sum_{n=0}^{m} b^n(\alpha) = \left( \sum_{n=0}^{\infty} b^n(\alpha) \xi^n \right) \left( \sum_{m=0}^{\infty} \xi^m \right)
\]
\[
= (1 - \xi)^\alpha (1 - \xi)^{-1} = (1 - \xi)^{\alpha-1}.
\]
Therefore, $v_m = b^{(\alpha-1)}_m \leq c(m + 1)^{-\alpha}$ [18, Lemma 2.3], which shows the second assertion.

The next result gives a discrete analogue of the following inequality: $\varphi(t)\partial_t^\alpha (\varphi(t) - \varphi(0)) \geq \frac{1}{2} \partial_t^\alpha (|\varphi(t)|^2 - \varphi(0)^2)$. 

6
Lemma 3.3. Let \( \bar{\partial}_t^* \varphi^n \) be the backward Euler CQ defined as (3.4). Then there holds

\[
(\bar{\partial}_t^* (\varphi^n - \varphi^0)) \varphi^n \geq \frac{1}{2} \bar{\partial}_t^* (|\varphi^n|^2 - |\varphi^0|^2)
\]

Proof. By the definition of the backward Euler CQ in (3.4), we deduce

\[
(\bar{\partial}_t^* (\varphi^n - \varphi^0)) \varphi^n = \varphi^n \left( \tau^{-\alpha} \sum_{j=0}^{n} b_{n-j}^{(\alpha)} (\varphi^j - \varphi^0) \right)
\]

Now by the definition, the binomial coefficient \( b_j^{(\alpha)} < 0 \) for \( j \geq 1 \), and thus

\[
\sum_{j=0}^{n-1} b_{n-j}^{(\alpha)} \varphi^j \varphi^0 \geq \frac{1}{2} \sum_{j=0}^{n-1} b_{n-j}^{(\alpha)} |\varphi^j|^2 + \frac{1}{2} \sum_{j=0}^{n-1} b_{n-j}^{(\alpha)} |\varphi|^2
\]

and

\[
(\sum_{j=0}^{n} b_{n-j}^{(\alpha)} \varphi^j \varphi^0) \leq \frac{1}{2} (\sum_{j=0}^{n} b_{n-j}^{(\alpha)} |\varphi|^2 + \frac{1}{2} (\sum_{j=0}^{n} b_{n-j}^{(\alpha)} |\varphi^0|^2).
\]

Then the desired result follows immediately.

The next result gives a discrete continuity result.

Lemma 3.4. Let the sequence \( \{q_h^k \} \subset \mathcal{A}_h \) be convergent to \( q_h^* \in \mathcal{A}_h \) in \( L^1(\Omega) \). Then

\[
\lim_{j \to \infty} \tau \sum_{n=1}^{N} \int_{\Omega} |U_{h,\tau}^n(q_h^j) - z^\delta(t_n)|^2 \, dx = \tau \sum_{n=1}^{N} \int_{\Omega} |U_{h,\tau}^n(q_h^j) - z^\delta(t_n)|^2 \, dx.
\]

Proof. Using Lemma 3.3, the proof is similar to that of Lemma 2.2, upon noting the fact that in a finite-dimensional space \( V_h \), all norms are equivalent, and the convergence in \( L^1(\Omega) \) implies almost every convergence. Thus the proof is omitted.

Then we can obtain the existence of a discrete minimizer \( q_h^* \in \mathcal{A}_h \). The proof is identical with that in Theorem 2.1, and hence omitted. Note that the discrete minimizer \( q_h^* \) depends implicitly also on the time step size \( \tau \), through the weak formulation (3.6).

Theorem 3.1. Under Assumption 2.1 there exists at least one minimizer \( q_h^* \in \mathcal{A}_h \) to problem (3.5)-(3.6).

Below we analyze the convergence of the sequence \( \{q_h^k \}_{h>0} \) as \( h, \tau \to 0 \). The next result is an analogue of Lemma 2.2 and plays an important role in the convergence analysis. For the sequence of discrete solutions \( U_{h,\tau}^n = U_{h,\tau}^n(q_h) \) to problem (3.6), we define a piecewise constant in time interpolation \( u_{h,\tau}(t) \) by

\[
u_{h,\tau}(t) = U_{h,\tau}^n, \quad t \in [t_n, t_{n+1}], \quad n = 0, \ldots, N - 1.
\]

Lemma 3.5. Let \( U_{h,\tau}^n = U_{h,\tau}^n(q_h) \) be the discrete solutions to problem (3.6) with \( q_h \in \mathcal{A}_h \), and suppose that the sequence \( \{q_h \in \mathcal{A}_h \}_{h>0} \) converges to some \( q^* \in \mathcal{A} \) in \( L^1(\Omega) \) and almost everywhere. Then under Assumption 2.1, the piecewise constant interpolation \( u_{h,\tau} \) satisfies

\[
u_{h,\tau}(q_h) \to u(q^*) \text{ weakly in } L^2(0, T; H^1(\Omega)), \quad h, \tau \to 0.
\]

Proof. Taking the test function \( \chi = U_{h}^n - U_{h}^0 \) in (3.6) and summing over \( n \) yield

\[
\tau \sum_{n=0}^{N} (\bar{\partial}_t^*(U_{h}^n - U_{h}^0), U_{h}^n - U_{h}^0) + \tau \sum_{n=1}^{N} (q_h \nabla U_{h}^n, \nabla(U_{h}^n - U_{h}^0)) = \tau \sum_{n=1}^{N} (f_h^n, U_{h}^n - U_{h}^0),
\]
This identity, the nonnegativity of the discrete convolution $\tilde{\partial}_t^\alpha$ (see the proof of Lemma 3.1), Poincaré inequality and Young’s inequality, and the $L^2(\Omega)$ stability of $P_h$ lead to

$$\tau \sum_{n=1}^N \| \nabla U_h^n \|_2^2 \leq c \tau \sum_{n=1}^N \left( \| \nabla U_h^0 \|_2^2 + \| f_h^n \|_{H^{-1}(\Omega)}^2 \right)$$

$$\leq c \left( \| \nabla u_0 \|_2^2 + \| f \|_{L^2(0,T;H^{-1}(\Omega))} \right).$$

Thus, the sequence $\{u_{h,\tau}\}_{h,\tau > 0}$ is uniformly bounded in $L^2(0,T;H^1(\Omega))$, and thus there exists a subsequence, still denoted by $\{u_{h,\tau}\}_{h,\tau > 0}$, such that

$$u_{h,\tau} \text{ converges weakly to some } u^* \text{ in } L^2(0,T;H^1(\Omega)). \quad (3.7)$$

Meanwhile, by taking the test function $\chi = \tilde{\partial}_t^\alpha (U_h^n - U_h^0)$ in (3.6),

$$\tau \sum_{n=0}^N (\tilde{\partial}_t^\alpha (U_h^n - U_h^0), \tilde{\partial}_t^\alpha (U_h^n - U_h^0)) + \tau \sum_{n=1}^N (q_h \nabla U_h^n, \tilde{\partial}_t^\alpha \nabla (U_h^n - U_h^0))$$

$$= \tau \sum_{n=1}^N (f_h^n, \tilde{\partial}_t^\alpha (U_h^n - U_h^0)). \quad (3.8)$$

Then Lemma 3.3 allows bounding the second term by

$$\tau \sum_{n=1}^N (q_h \nabla U_h^n, \tilde{\partial}_t^\alpha \nabla (U_h^n - U_h^0)) \geq \frac{1}{2} \int_\Omega q_h(x) \left[ \tau \sum_{n=1}^N \tilde{\partial}_t^\alpha \left( |\nabla U_h^n|^2 - |\nabla U_h^0|^2 \right) \right] dx$$

$$= \frac{1}{2} \int_\Omega q_h(x) \left[ \tau^{1-\alpha} \sum_{n=0}^N \sum_{j=0}^n b_{n-j}^{(\alpha)} (|\nabla U_h^j|^2 - |\nabla U_h^0|^2) \right] dx$$

$$= \frac{1}{2} \int_\Omega q_h(x) \left[ \tau^{1-\alpha} \sum_{j=0}^N (|\nabla U_h^j|^2 - |\nabla U_h^0|^2) \sum_{n=j}^N b_{n-j}^{(\alpha)} \right] dx$$

$$= \frac{1}{2} \int_\Omega q_h(x) \left[ \tau \sum_{j=0}^N \left( |\nabla U_h^j|^2 - |\nabla U_h^0|^2 \right) b_{N-j}^{(\alpha-1)} \right] dx$$

$$\geq \frac{1}{2} \int_\Omega q_h(x) \left[ \tau \sum_{j=0}^N (|\nabla U_h^j|^2 - |\nabla U_h^0|^2) b_{N-j}^{(\alpha-1)} \right] dx$$

$$\geq -c \int_\Omega q_h(x) \left[ \tau \sum_{n=0}^N t_{N-j+1}^{\alpha} \right] |\nabla U_h^0|^2 dx \geq -c \| U_h^0 \|^2_{L^2(\Omega)},$$

where the second last line follows from the fact that $b_{j}^{(\alpha-1)} > 0$ for all $j \geq 0$, and the last line from Lemma 3.2. This and Young’s inequality imply

$$\tau \sum_{n=1}^N \| \tilde{\partial}_t^\alpha (U_h^n - U_h^0) \|^2_{L^2(\Omega)} \leq c (\| \nabla u_0 \|_{L^2(\Omega)} + \| f \|_{L^2(0,T;L^2(\Omega))}).$$

Thus, the sequence of piecewise constant interpolation, denoted by $\{\tilde{\partial}_t^\alpha (u_{h,\tau} - v_h^0)\}_{h,\tau > 0}$, is uniformly bounded in $L^2(0,T;L^2(\Omega))$, and there exists a subsequence, still denoted by $\{\tilde{\partial}_t^\alpha (u_{h,\tau} - u_h^0)\}_{h,\tau > 0}$, and some $v^* \in L^2(0,T;L^2(\Omega))$ such that it converges to $v^*$ weakly in $L^2(0,T;L^2(\Omega))$. Next we claim that $u^*$ satisfies the weak formulation of $u(q^*)$, cf. (2.2). Take a smooth test function $\phi \in C^1([0,T];H^1_0(\Omega))$ with $\phi(T) = 0$, and define an approximation $\phi_{h,\tau}$ by $\phi_{h,\tau}(t) = \tau^{-1} \int_{t_{n-1}}^{t_n} P_h \phi(t) dt$, $t \in (t_{n-1}, t_n]$. Clearly,

$$\lim_{h,\tau \to 0} \| \phi_{h,\tau} - \phi \|_{L^2(0,T;H^1(\Omega))} = 0.$$
Then by discrete summation by parts and straightforward computation, there holds
\[
\tau \sum_{n=0}^{N} (\tilde{\partial}_t^\alpha(U^n_h - U^0_h), \phi_{h,\tau}(t_n)) = \int_0^T (\tilde{\partial}_t^\alpha(u_{h,\tau}^0 - u^0_h), P_h(\phi(t))dt
\]
\[
= \int_0^T (u_{h,\tau}^0 - u^0_h, R\tilde{\partial}_t^\alpha P_h(\phi(t))dt,
\]
where the notation \( R\tilde{\partial}_t^\alpha P_h(\phi(t)) \) denotes
\[
R\tilde{\partial}_t^\alpha P_h(\phi(t)) = \sum_{i=n}^{N} b_n^{(\alpha)} R\tilde{\partial}_t^\alpha P_h(\phi(t + (i-n)\tau)), \quad t \in (t_{n-1}, t_n], \quad n = 1, 2, \ldots, N.
\]
By the approximation property of \( R\tilde{\partial}_t^\alpha P_h \) and \( P_h \), since \( \phi \in C^1([0,T]; H_0^1(\Omega)) \), \( R\tilde{\partial}_t^\alpha P_h(\phi(t)) \) converges to \( \tilde{\partial}_t^\alpha \phi(t) \) in \( L^2(0,T; L^2(\Omega)) \) as \( h, \tau \to 0^+ \), and thus
\[
\lim_{h,\tau \to 0} \int_0^T (u_{h,\tau}^0 - u^0_h, R\tilde{\partial}_t^\alpha P_h(\phi(t))dt = \int_0^T (u^* - u_0, R\tilde{\partial}_t^\alpha \phi(t))dt
\]
and meanwhile, by the weak convergence of \( \tilde{\partial}_t^\alpha(u_{h,\tau}^0 - u^0_h) \) to \( v^* \) in \( L^2(0,T; L^2(\Omega)) \) and the approximation property of \( P_h \),
\[
\lim_{h,\tau \to 0} \int_0^T (\tilde{\partial}_t^\alpha(u_{h,\tau}^0 - u^0_h), P_h(\phi(t))dt = \int_0^T (v^*, \phi(t))dt.
\]
Comparing these two identities shows that \( v^* = \partial_t^\alpha(u^* - u_0) \), i.e., \( v^* \) is the weak fractional order derivative of \( u^* - u_0 \). Now taking the test function \( \chi = \phi_{h,\tau}(t_n) \) in (3.6) and summing over \( n \), we obtain
\[
\tau \sum_{n=0}^{N} (\tilde{\partial}_t^\alpha(U^n_h - U^0_h), \phi_{h,\tau}(t_n)) + \tau \sum_{n=1}^{N} (q_h \nabla U^n_h, \nabla \phi_{h,\tau}(t_n)) = \tau \sum_{n=1}^{N} (f^n_h, \phi_{h,\tau}(t_n)),
\]
and by the definition of piecewise constant interpolations \( \tilde{\partial}_t(U^n_h - U^0_h) \) and \( u_{h,\tau}(t) \) and the construction of the test function \( \phi_{h,\tau}(t_n) \), it is equivalent to
\[
\int_0^T (\tilde{\partial}_t^\alpha(u_{h,\tau}^0 - u^0_h), P_h(\phi(t))dt + \int_0^T (q_h \nabla u_{h,\tau}, \nabla P_h(\phi(t))dt = \int_0^T (f_{h,\tau}, P_h(\phi(t))dt,
\]
where \( f_{h,\tau}(t) = \tau^{-1} \sum_{t_{n-1}}^{t_n} P_h f(t)dt, \) for \( t \in (t_{n-1}, t_n], \) \( n = 1, \ldots, N. \) Upon passing limit on both sides and noting the construction of the approximation \( f_{h,\tau}(t) \),
\[
\lim_{h,\tau \to 0} \|f_{h,\tau} - f\|_{L^2(0,T; L^2(\Omega))} = 0,
\]
we deduce
\[
\lim_{h,\tau \to 0} \int_0^T (\tilde{\partial}_t^\alpha(U^n_h - U^0_h), P_h(\phi(t))dt = \int_0^T (\tilde{\partial}_t^\alpha(u^* - u_0), \phi(t))dt,
\]
\[
\lim_{h,\tau \to 0} \int_0^T (f_{h,\tau}, P_h(\phi(t))dt = \int_0^T (f, \phi(t))dt.
\]
Further, to analyze the term \( \int_0^T (q_h \nabla u_{h,\tau}, \nabla P_h(\phi(t))dt \), we employ the following splitting
\[
|\int_0^T (q_h \nabla u_{h,\tau}, \nabla P_h(\phi(t))dt| \leq |\int_0^T (q_h \nabla u_{h,\tau}, \nabla P_h(\phi(t))dt| - |\int_0^T (q_h \nabla u_{h,\tau}, \nabla \phi(t))dt|
\]
We bound the three terms separately. By the approximation property of $P_h$ and uniform boundedness of $u_{h,\tau}$ in $L^2(0, T; H^1(\Omega))$ due to (3.7), we deduce
\[
\lim_{h, \tau \to 0^+} I \leq \lim_{h, \tau \to 0^+} c\|u_{h,\tau}\|_{L^2(0, T; H^1(\Omega))}\|P_h \phi - \phi\|_{L^2(0, T; H^1(\Omega))} = 0.
\]
Next, since $q_h$ converges to $q^*$ in $L^1(\Omega)$ and almost everywhere and (3.7), by Lebesgue’s dominated convergence theorem, we have
\[
\lim_{h, \tau \to 0^+} II \leq \lim_{h, \tau \to 0^+} \|u_{h,\tau}\|_{L^2(0, T; H^1(\Omega))}\|(q_h - q^*)\phi\|_{L^2(0, T; H^1(\Omega))} = 0.
\]
The third term III tends to zero as $h, \tau \to 0^+$, in view of the weak convergence in (3.7). Consequently, combining the three assertions together yields
\[
\lim_{h, \tau \to 0} \int_0^T (q_h \nabla u_{h,\tau}, \nabla P_h \phi(t)) \, dt = \int_0^T (q^* \nabla u^*, \nabla \phi(t)) \, dt.
\]
In sum, the limit $u^*$ satisfies
\[
\int_0^T (f^* - u_0, \phi) \, dt + \int_0^T (q^* \nabla u^*, \nabla \phi) \, dt = \int_0^T (f, \phi) \, dt, \quad \forall \phi \in C([0, T]; H^1(\Omega)).
\]
By the density of the space $C^1([0, T]; H^1_0(\Omega))$ in $L^2(0, T; H^1_0(\Omega))$, the identity holds also for any $\phi \in L^2(0, T; H^1_0(\Omega))$. This immediately shows that $u^*$ is a weak solution to problem (1.1) with $q^*$, i.e., $u^* = u(q^*)$. Since every subsequence contains a convergent sub-subsequence, the whole sequence converges to $u(q^*)$. This completes the proof of the lemma.

Now we can state the main result of this part, i.e., the convergence of the discrete solutions $\{q_h^*\}_{h>0}$ to the continuous optimization problem (2.1)–(2.2).

**Theorem 3.2.** Let $\{q_h^*\}_{h>0}$ be a sequence of minimizers to problem (3.5)–(3.6). Then under Assumption 2.1 it contains a subsequence convergent to a minimizer of problem 2.1–2.2 in $H^1(\Omega)$.

**Proof.** Since the constant function $q_h = c_0$ belongs to the admissible set $A_h$ for any $h$, there holds $J_{\gamma, h, \tau}(q_h^*) \leq J_{\gamma, h, \tau}(c_0) < \infty$, from which it directly follows that the sequence $\{q_h^*\}_{h>0}$ is uniformly bounded in the $H^1(\Omega)$-seminorm. This and the box constraint in $A_h$ imply that the sequence $\{q_h^* \in A_h\}_{h>0}$ is uniformly bounded in the $H^1(\Omega)$ norm. Thus there exists a subsequence, still denoted by $\{q_h^*\}_{h>0}$ such that it converges weakly in the $H^1(\Omega)$ to some $q^* \in A$. We claim that $q^*$ is a minimizer to problem (2.1)–(2.2). For any $q \in A$, by the density of $W^{1,\infty}(\Omega)$ in $H^1(\Omega)$ [11] (e.g., by means of mollifier), there exists a sequence $\{q^*_\} \subset A \cap W^{1,\infty}(\Omega)$ such that $\lim_{\tau \to 0^+} \|q^* - q^\tau\|_{H^1(\Omega)} = 0$ and almost everywhere. Now let $q_h^* = \mathcal{I}_h q^\tau \in V_h$. By the minimizing property of $q_h^*$, there holds
\[
J_{\gamma, h, \tau}(q_h^*) \leq J_{\gamma, h, \tau}(q_h^*).
\]
By the weak lower semi-continuity of norms, weak convergence of $u_{h,\tau}(q_h^*)$ to $u(q^*)$ in $L^2(0, T; H^1(\Omega))$ from Lemma 3.5 and the continuous embedding from $H^1(\Omega)$ into $L^2(\Omega)$, and the construction of the piecewise constant function $z^\delta(t) = \tau^{-1} \int_{t_n-1}^{t_n} z^\delta(t) \, dt$, for $t \in (t_{n-1}, t_n]$, $n = 1, \ldots, N$, $\lim_{\tau \to 0} \|z^\delta - z^\tau\|_{L^2(0,T;L^2(\Omega))} = 0$, we have
\[
\|\nabla q^*\|_{L^2(\Omega)} \leq \lim_{h \to 0} \inf \|\nabla q_h^*\|_{L^2(\Omega)}
\]
\[
\|u(q^*) - z^\delta\|_{L^2(0,T;L^2(\Omega))} \leq \lim_{h, \tau \to 0} \|u_{h,\tau}(q_h^*) - z^\tau\|_{L^2(0,T;L^2(\Omega))}
\]
\[ J_0(q^*) \leq \lim_{h, \tau \to 0^+} J_{\gamma, h, \tau}(q_h^*). \] (3.10)

and consequently

\[ J_j(q^*) \leq \lim_{h, \tau \to 0^+} J_{\gamma, h, \tau}(q_h^*). \] (3.11)

Thus, taking limit as \( h, \tau \to 0^+ \) in the inequality (3.9) yields \( J_j(q^*) \leq J_j(q') \). Further, since \( q' \to q \) in \( H^1(\Omega) \) and almost everywhere as \( \epsilon \to 0^+ \), by Lemma 2.2 there holds

\[ \lim \limits_{\epsilon \to 0^+} J_j(q') = J_j(q). \] (3.12)

Combining the three relations (3.10)–(3.12) yields \( J_j(q^*) \leq J_j(q) \) for any \( q \in A \). This shows the weak convergence to a minimizer \( q^* \) in \( H^1(\Omega) \). Meanwhile, by the weak lower semi-continuity of the norms and a standard argument by contradiction, we deduce

\[ \lim_{h, \tau \to 0^+} \| \nabla q_h^* \|^2_{L^2(\Omega)} = \| \nabla q^* \|^2_{L^2(\Omega)}. \]

Therefore, the subsequence \( \{q_h^*\}_{h>0} \) converges to \( q^* \) in \( H^1(\Omega) \), completing the proof. \( \square \)

**Remark 3.1.** Since the continuity results (at both discrete and continuous levels) are stated with respect to \( L^1(\Omega) \), the results in Sections 2 and 3 extend straightforwardly to closely related regularized formulations, e.g., total variation penalty or \( L^1(\Omega) \) extend straightforwardly to closely related regularized formulations, e.g., total variation penalty or \( \| u \|_{L^2(T_0,T;L^2(\Omega))} \) (with \( T_0 \in [0,T] \)) fidelity.

### 4 Error estimates

Now we derive error estimates of approximations \( q_h^* \) under the following regularity on the problem data.

**Assumption 4.1.** The following conditions hold.

(i) There exists some \( \beta > \max\left(\frac{d}{2} - 1, 0\right) \) such that

\[ u_0 \in H^2(\Omega) \cap H_0^1(\Omega), \quad f \in L^\infty(0,T;\dot{H}^\beta(\Omega)) \cap C^2([0,T];L^2(\Omega)), \quad q^1 \in W^{2,\infty}(\Omega). \]

(ii) \( \delta^0 \in H^2([0,T];L^2(\Omega)) \)

Under Assumption 4.1 (i), there exists a solution \( u \in C([0,T];\dot{H}^2(\Omega)) \cap C^2([0,T];L^2(\Omega)) \) and for any \( s \in [0, \beta) \) and \( r \in [0,2] \), there holds

\[ \| u(t) \|_{\dot{H}^2(\Omega)} + t^{\frac{\beta}{2}} \| u(t) \|_{\dot{H}^2_r(\Omega)} + t^{1-(1-\frac{s}{2})\alpha} \| \partial_t u(t) \|_{\dot{H}^r(\Omega)} + t^{2-\alpha} \| \partial_{tt} u(t) \|_{L^2(\Omega)} \leq c. \] (4.1)

See [33] [17] for a proof of the regularity estimate.

The better regularity on the observation \( z^0 \) and \( u(q) \) enables slightly modifying the discrete optimization problem \( J_{h,\gamma, \tau} \) instead of using \( z_h^0 := \tau^{-1} \int_{t_{n-1}}^{t_n} \delta^0(t) \mathrm{d}t \). In particular, we can employ the trapezoid rule: with \( a_0 = a_N = 1/2 \) and \( a_i = 1, i = 1, \ldots, N-1 \),

\[ \min_{q_h \in A_h} J_{\gamma, h, \tau}(q_h) = \frac{\tau}{2} \sum_{n=0}^{N} a_i \int_{\Omega} |U_h^n(q_h) - z^\delta(t_n)|^2 \mathrm{d}x + \frac{\gamma}{2} \| \nabla q_h \|^2_{L^2(\Omega)}, \] (4.2)

subject to \( q_h \in A_h \) and \( U_h^n(q_h) \) satisfying \( U_h^0 = P_h u_0 \) and

\[ \partial_{tt} U_h^n(q_h) - A_h(q_h) \nabla U_h^n(q_h) = P_h f(t_n), \quad n = 1, 2, \ldots, N. \] (4.3)

This change allows deriving a better rate in \( \tau \) in Theorem 4.1 below. Under Assumption 4.1, Theorem 3.1 and 3.2 in Section 3 remain valid for problem 4.2–4.3. The goal of this part is to derive error estimates for the approximation constructed by (4.2)–(4.3).

We begin with some preliminary estimates under Assumption 4.1 (i).
Lemma 4.1. Let \( q^1 \) be the exact diffusion coefficient and \( u \equiv u(q^1) \) be the solution to problem \( (2.2) \), and \( \{ U_h^n(q^1) \} \) and \( \{ I_h^n(q^1) \} \) be the solutions to the scheme \( (3.3) \) corresponding to \( q^1 \) and \( I_h q^1 \), respectively. Then under Assumption \( (4.1(i)) \), with \( \ell_h = |\log h| \),

\[
\| u(t_n) - U_h^n(q^1) \|_{L^2(\Omega)} \leq c(\tau t_n^{-1} + h^2 \ell_h), \\
\| u(t_n) - U_h^n(I_h q^1) \|_{L^2(\Omega)} \leq c(\tau t_n^{-1} + h^2 \ell_h).
\]

Proof. The first estimate is immediate from \[16\]

\[
\| u(t_n) - U_h^n(q^1) \|_{L^2(\Omega)} \leq c h^2 \ell_h \left( \| A(q^1) u_0 \|_{L^2(\Omega)} + \| f \|_{L^\infty(0,T;H^s(\Omega))} \right) \\
+ c \tau \left( t_n^{-1} \| A(q^1) u_0 + f(0) \|_{L^2(\Omega)} + \int_0^{t_n} (t_n - s)^{\alpha - 1} \| f'(s) \|_{L^2(\Omega)} \, ds \right).
\]

To show the second estimate, we bound \( \rho_h^n := U_h^n(q^1) - U_h^n(I_h q^1) \), which satisfies

\[
\frac{\partial \rho_h^n}{\partial t^n} + A_h(q^1) \rho_h^n = [A_h(I_h q^1) - A_h(q^1)]U_h^n(I_h q^1), \quad \text{with} \quad \rho_h^n = 0, \quad n = 1, 2, \ldots, N,
\]

where \( A_h(q^1), A_h(I_h q^1) : X_h \to X_h \) are the discrete analogues of the elliptic operators \( A(q^1) \) and \( A(q_h) \) associated with \( q^1 \) and \( I_h q^1 \), respectively. Thus, it can be written as

\[
\rho_h^n = \tau \sum_{i=1}^n E_{h,\tau}^n [A_h(I_h q^1) - A_h(q^1)]U_h^n(I_h q^1), \quad (4.4)
\]

where \( E_{h,\tau}^n \) is the fully discrete solution operator, which satisfies that for all \( v_h \in X_h \),

\[
\| E_{h,\tau}^n v_h \|_{L^2(\Omega)} = \| A_h(q^1)^{\frac{1}{2}} E_{h,\tau}^n (A_h(q^1)^{-\frac{1}{2}} v_h) \|_{L^2(\Omega)} \\
\leq c \tau_n^{1 + \frac{\alpha}{2}} \| (A_h(q^1)^{-\frac{1}{2}} v_h) \|_{L^2(\Omega)} \leq c \tau_n^{1 + \frac{\alpha}{2}} \| v_h \|_{H^{-\frac{\alpha}{2}}(\Omega)}.
\]

It follows from this estimate and the solution representation \( (4.4) \) that

\[
\| \rho_h^n \|_{L^2(\Omega)} \leq c \tau_n \sum_{i=1}^n \tau_n^{-\frac{\alpha}{2}} \| [A_h(I_h q^1) - A_h(q^1)]U_h^n(I_h q^1) \|_{H^{-\frac{\alpha}{2}}(\Omega)}.
\]

Further, the definitions of \( P_h \) and \( A_h \) and the \( H^1(\Omega) \)-stability of \( P_h \) yield

\[
\| [A_h(I_h q^1) - A_h(q^1)]U_h^n(I_h q^1) \|_{H^{-\frac{\alpha}{2}}(\Omega)} = \sup_{v \in H^1} \frac{\langle [A_h(I_h q^1) - A_h(q^1)]U_h^n(I_h q^1), v \rangle}{\| v \|_{H^1(\Omega)}} \\
= \sup_{v \in H^1} \frac{\langle [A_h(I_h q^1) - A_h(q^1)]U_h^n(I_h q^1), P_h v \rangle}{\| v \|_{H^1(\Omega)}} = \sup_{v \in H^1} \frac{\langle (q^1 - I_h q^1) \nabla U_h^n(I_h q^1), \nabla P_h v \rangle}{\| v \|_{H^1(\Omega)}} \\
\leq c \tau_n \| q^1 \|_{W^{2,\infty}(\Omega)} \| \nabla U_h^n(I_h q^1) \|_{L^2(\Omega)},
\]

since \( q \in W^{2,\infty}(\Omega) \) by Assumption \( (4.1(i)) \) and \( (3.2) \). Thus, we deduce

\[
\| \rho_h^n \|_{L^2(\Omega)} \leq c \tau_n \sum_{i=1}^n \tau_n^{-\frac{\alpha}{2}} \leq c \tau_n \int_0^T t^{-1 + \frac{\alpha}{2}} \, dt \leq c h^2.
\]

This and the triangle inequality completes the proof of the lemma. \hfill \( \Box \)

Next we give an estimate on the CQ approximation of the fractional derivative. The detailed proof is deferred to the appendix.

Lemma 4.2. Let \( q^1 \) be the exact diffusion coefficient and \( u \equiv u(q^1) \) be the solution to problem \( (2.2) \). Then under Assumption \( (4.1) \) there holds

\[
\| \partial_t^\alpha (u(t_n) - u_0) - \partial_t^\alpha (u(t_n) - u_0) \|_{L^2(\Omega)} \leq c \tau_n t_n^{-1}.
\]
The next lemma gives the quadrature error estimate.

**Lemma 4.3.** Let $q^\dagger$ be the exact diffusion coefficient and $u \equiv u(q^\dagger)$ the solution to problem \((2.2)\). Then under Assumption 4.3
\[
\sum_{n=0}^{N} a_i \|u(t_n) - z^\dagger(t_n)\|_{L^2(\Omega)}^2 \leq c(\delta^2 + \tau^{1+\alpha}).
\]

**Proof.** Let $g(t) = z^\dagger(t) - u(t)$. By the regularity estimate (4.1) and Assumption 4.1
\[
\|g\|_{C([0,T];L^2(\Omega))} \leq c, \quad \|g(t)\|_{L^2(\Omega)} \leq c\tau^{-1} \quad \text{and} \quad \|g''(t)\|_{L^2(\Omega)} \leq c\tau^{-2}.
\]

By the triangle inequality, we have
\[
|\int_{t}^{\tau} \sum_{n=0}^{N} a_i \|g(t_n)\|_{L^2(\Omega)}^2 \|g(t)\|_{L^2(\Omega)}^2 \mathrm{d}t|
\leq \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|g(t)\|_{L^2(\Omega)}^2 \mathrm{d}t - \frac{\tau}{2} \left(\|g(t_{n-1})\|_{L^2(\Omega)}^2 + \|g(t_n)\|_{L^2(\Omega)}^2\right) \leq \sum_{n=1}^{N} I_n.
\]

Next we analyze the two cases $n = 1$ and $n > 1$ separately. First, for the case $n = 1$,
\[
I_1 \leq \left|\int_{0}^{\tau} (\|g(t)\|_{L^2(\Omega)}^2 - \|g(t_0)\|_{L^2(\Omega)}^2) \mathrm{d}t\right| + \left|\int_{0}^{\tau} (\|g(t)\|_{L^2(\Omega)}^2 - \|g(\tau)\|_{L^2(\Omega)}^2) \mathrm{d}t\right| := I_{1,0} + I_{1,1}.
\]

Using (4.5), the term $I_{1,0}$ can be bounded by
\[
I_{1,0} \leq c\|g(t)\|_{C([0,T];L^2(\Omega))} \int_{0}^{\tau} \|g(\tau) - g(t)\|_{L^2(\Omega)} \mathrm{d}t
\leq c\tau \int_{0}^{\tau} \|g(t)\|_{L^2(\Omega)} \mathrm{d}s \leq c\tau \int_{0}^{\tau} s^{\alpha - 1} \mathrm{d}s \leq c\tau^{1+\alpha}.
\]

Similarly, we can deduce $I_{1,1} \leq c\tau^{1+\alpha}$. Further, for the case $n > 1$, $g(t)$ is smooth, and thus by standard interpolation error estimates, for some $\xi_n \in [t_{n-1}, t_n]$
\[
I_n \leq c\tau^2 \int_{t_{n-1}}^{t_n} \frac{d^2}{dt^2} \|g(t)\|_{L^2(\Omega)}^2 |t = \xi_n| \mathrm{d}t.
\]

In view of the bounds in (4.5),
\[
\left|\frac{d^2}{dt^2} \|g(\xi_n)\|_{L^2(\Omega)}^2 \right| \leq 2(\|g'(\xi_n)\|_{L^2(\Omega)}^2 + \|g(\xi_n)\|_{L^2(\Omega)}^2) \|g''(\xi_n)\|_{L^2(\Omega)}^2 \leq c\tau^{\alpha - 2}.
\]

The last two estimates together imply
\[
\sum_{n=2}^{N} I_n \leq c\tau^3 \sum_{n=2}^{N} \tau^{\alpha - 2} \leq c\tau^{1+\alpha}.
\]

Then the assertion follows from the triangle inequality and the definition of the noise level.

**Remark 4.1.** One can only obtain an $O(\tau + \delta^2)$ rate the discrete objective function $J_{\gamma, h, \tau}$ in (3.5). The $\alpha$ exponent in Lemma 4.3 reflects the limited temporal smoothing property of the solution $u(t)$: the larger the fractional order $\alpha$ is, the smoother is the solution $u(t)$ in time and the faster is the decay of quadrature error.

The next result gives a priori bounds on the bound $q^\dagger$ and error estimates on the approximation $U^n_h(q^n_h)$. This result will play a crucial role in the proof of Theorem 4.1.
Lemma 4.4. Let $q^I$ be the exact coefficient and $u \equiv u(q^I)$ the solution to problem (2.2). Let $q_h^* \in \mathcal{A}_h$ be the solution to problem (4.2)–(4.3), and $\{U_h^N(q_h^*)\}_{n=1}^N$ the fully discrete solution to problem (3.6). Then under Assumption 4.1 there holds

$$
\tau \sum_{n=1}^N ||U_h^N(q_h^*) - u(t_n)||^2_{L^2(\Omega)} + \gamma \|\nabla q_h^*\|_{L^2(\Omega)}^2 \leq c(\tau^{1+\alpha} + h^4 t_h^2 + \delta^2 + \gamma).
$$

Proof. By the minimizing property of $q_h^* \in \mathcal{A}_h$ and $I_h q^I \in \mathcal{A}_h$, we deduce

$$
J_{\gamma,h,\tau}(q_h^*) \leq J_{\gamma,h,\tau}(I_h q^I).
$$

By the triangle inequality, we derive

$$
\tau \sum_{n=1}^N ||U_h^N(q_h^*) - u(t_n)||^2_{L^2(\Omega)} \leq c t \sum_{n=1}^N ||U_h^N(q_h^*) - z(t_n)||^2_{L^2(\Omega)} + c t \sum_{n=0}^N a_n \|z(t_n) - u(t_n)||^2_{L^2(\Omega)}.
$$

These two inequalities and Lemma 4.3 imply

$$
\tau \sum_{n=1}^N ||U_h^N(q_h^*) - u(t_n)||^2_{L^2(\Omega)} + \gamma \|\nabla q_h^*\|_{L^2(\Omega)}^2 
\leq c t \sum_{n=1}^N ||U_h^N(I_h q^I) - z(t_n)||^2_{L^2(\Omega)} + c \gamma \|\nabla I_h q^I\|_{L^2(\Omega)}^2 + c(\delta^2 + \tau^{1+\alpha}).
$$

Since $q^I \in W^{1,\infty}(\Omega)$ by Assumption 4.1, $\|\nabla I_h q^I\|_{L^2(\Omega)} \leq c$, cf. (3.2). Meanwhile, by Lemma 4.1 we have

$$
||U_h^N(I_h q^I) - z(t_n)||^2_{L^2(\Omega)} \leq 2 ||U_h^N(I_h q^I) - u(t_n)||^2_{L^2(\Omega)} + 2 ||u(t_n) - z(t_n)||^2_{L^2(\Omega)}
\leq c(\tau a_n + h^2 \ell_h)^2 + c ||u(t_n) - z(t_n)||^2_{L^2(\Omega)}.
$$

Consequently,

$$
\tau \sum_{n=1}^N ||(U_h^N(I_h q^I) - z(t_n))||^2_{L^2(\Omega)} \leq c t \sum_{n=1}^N (t_n^{\delta} + h^2 \ell_h)^2 + c t \sum_{n=0}^N a_n ||u(t_n) - z(t_n)||^2_{L^2(\Omega)}
\leq c t \sum_{n=1}^N t_n^{\delta - 2} + c h^4 \ell_h^2 + c(\tau^{1+\alpha} + \delta^2) \leq c(\tau^{1+\alpha} + h^4 \ell_h^2 + \delta^2).
$$

Combining the preceding estimates completes the proof of the lemma.

We shall also need the following lemma on backward Euler CQ.

Lemma 4.5. Let $q^I$ be the exact coefficient, and $u \equiv u(q^I)$ the solution to problem (1.1). Then for $\phi^m = \frac{q^I - q_h^*}{q^I - q_h^*} u(t_m)$, and $\epsilon \in (0, \min(\frac{1}{2}, 1 - \alpha))$, there hold

$$
|| \tau^{-\alpha} \sum_{n=j}^m b_{n-j}^{(\alpha)} P_h(\phi^n - \phi^m) ||_{L^2(\Omega)} \leq c \epsilon t_j^{1-\epsilon}.
$$

Proof. By the associativity of CQ, i.e., $\partial_t^\alpha \phi^n = \partial_t^{\alpha - 1} \partial_t \phi^n$, if $\phi^0 = 0,$

$$
I := \tau^{-\alpha} \sum_{n=j}^m b_{n-j}^{(\alpha)} P_h(\phi^n - \phi^m) = \tau^{1-\alpha} \sum_{n=j}^m b_{n-j}^{(\alpha-1)} \frac{P_h(\phi^n - \phi^{n+1})}{\tau}.
$$

Thus, the $L^2(\Omega)$-stability of $P_h$, the bound on $|b_{n-j}^{(\alpha-1)}| \leq c(j+1)^{-\alpha}$ and (4.1) imply

$$
||I||_{L^2(\Omega)} \leq \tau^{1-\alpha} \sum_{n=j}^m |b_{n-j}^{(\alpha-1)}| \frac{\phi^n - \phi^{n+1}}{\tau} \|\phi^n - \phi^{n+1}\|_{L^2(\Omega)}
$$

14
Consequently, this completes the proof of the lemma.

Next we bound the three terms. Direct computation with the triangle inequality gives the solution to problem

\[ 2 + h \]

Then under Assumption 4.1, for \( n \in \mathbb{N} \), the solution to problem

\[ \sum_{m=1}^{N} \sum_{n=1}^{m} \int_{\Omega} \left( \frac{q^1 - q^h_n}{q^1} \right)^2 \left( q^1 \| \nabla u(t_n) \|^2 + (f(t_n) - \partial_t u(t_n))u(t_n) \right) \, dx \]

\[ \leq c(h \gamma^{-1} - \eta + h^{-1} + h^{-1} \gamma^{-2} \eta) \eta. \]

Proof. For any test function \( \varphi \) to be specified below, we have the splitting

\[ ((q^1 - q^h_n) \nabla u(t_n), \nabla \varphi) = ((q^1 - q^h_n) \nabla u(t_n), \nabla (\varphi - P_h \varphi)) + (q^1 \nabla u(t_n) - q^h_n \nabla u(t_n), \nabla P_h \varphi). \]

Thus, applying integration by parts to the first term leads to

\[ ((q^1 - q^h_n) \nabla u(t_n), \nabla \varphi) = -((\nabla \cdot (q^1 - q^h_n) \nabla u(t_n)), \varphi - P_h \varphi) + (q^h_n \nabla (U_{h_n}^n(q^h_n) - u(t_n)), \nabla P_h \varphi) \]

\[ + (q^1 \nabla u(t_n) - q^h_n \nabla U_{h_n}^n(q^h_n), \nabla P_h \varphi) = \sum_{i=1}^{3} \Gamma_i. \quad (4.6) \]

Next we bound the three terms. Direct computation with the triangle inequality gives

\[ \| \nabla \cdot ((q^1 - q^h_n) \nabla u(t_n)) \|_{L^2(\Omega)} \leq \| \nabla q^1 \|_{L^\infty(\Omega)} \| \nabla u(t_n) \|_{L^2(\Omega)} + \| q^1 - q^h_n \|_{L^\infty(\Omega)} \| \Delta u(t_n) \|_{L^2(\Omega)} \]

We claim that the integral \( g(t_j) \) is decreasing in \( t_j \in [\tau, t_m] \). Indeed, for any \( 0 < \tilde{t}_1 < \tilde{t}_2 \leq t_m \), by changing variables, there holds

\[ g(\tilde{t}_2) = \int_{t_2}^{t_m} (s - \tilde{t}_2 + \tau)^{-\alpha}s^{\alpha+\epsilon-1} ds \geq g(\tilde{t}_2). \]

Consequently,

\[ ||\varphi||_{L^2(\Omega)} \leq c t_j^{-\epsilon} \int_{\tau}^{t_j} (s + \tau)^{-\alpha}s^{\alpha+\epsilon-1} ds \leq c t_j^{-\epsilon} \int_{\tau}^{t_j} s^{\epsilon-1} ds = c t_j^{-\epsilon}. \]

This completes the proof of the lemma.

**Theorem 4.1.** Let \( q^1 \) be the exact diffusion coefficient, \( u \equiv u(q^1) \) the solution to problem 2.2, and \( q^h_n \in \mathcal{A}_h \) the solution to problem 4.2-4.3. Then under Assumption 4.1 for \( d = 1, 2 \), there holds (with \( \eta = \tau^{\frac{d}{2}} + h^2 \ell_h + \delta + \gamma \frac{d}{2} \))
Now we choose the test function $\varphi$ for the term $I_1$ which together with the trivial inequality gives

$$\|
abla \cdot (q^1 - q^h_\ast) \nabla u(t_n)\|_{L^2(\Omega)} \leq c + \|
abla q^h_\ast\|_{L^2(\Omega)} \|
abla u(t_n)\|_{L^\infty(\Omega)} \leq c(1 + t_n^{\min(0, -\frac{d}{4} - \epsilon)}) \|
abla q^h_\ast\|_{L^2(\Omega)},$$

where the second line is due to Sobolev embedding $\|
abla u\|_{L^\infty(\Omega)} \leq c\|u\|_{H^{s}(\Omega)}$ with $s > \frac{d}{4} + 1$ (by the convexity of the domain and elliptic regularity [3 Corollary 19.7, p. 166]). This and the Cauchy-Schwarz inequality imply that the first term $I_1^\ast$ is bounded by

$$\|I_1^\ast\| \leq c(1 + \|
abla q^h_\ast\|_{L^2(\Omega)})\|\varphi - P_h \varphi\|_{L^2(\Omega)}.$$ 

Now we choose the test function $\varphi$ to be $\varphi \equiv \varphi^n = \frac{q^l - q^h_{\ast}}{q^1} u(t_n) \in H^1_0(\Omega)$, and then straightforward computation gives

$$\nabla \varphi^n = (q^l - q^h_{\ast}) \nabla (q^1 - q^h_{\ast}) u(t_n) + q^1 - q^h_{\ast} \nabla u(t_n).$$

By the box constraint of $A$ and the regularity estimate (4.1), we have

$$\|
abla \varphi^n\|_{L^2(\Omega)} \leq c \left[1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)}\|u(t_n)\|_{L^\infty(\Omega)} + \|
abla u(t_n)\|_{L^2(\Omega)} \right] \leq c(1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)}),$$

and the approximation property of the projection operator $P_h$ implies

$$\|\varphi^n - P_h \varphi^n\|_{L^2(\Omega)} \leq c h \|
abla \varphi^n\|_{L^2(\Omega)} \leq c h(1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)}).$$

Thus, in view of Lemma 4.4, the term $I_1^\ast$ in the splitting (4.6) can be bounded by

$$\|I_1^\ast\| \leq c h t_n^{\min(0, -\frac{d}{4} - \epsilon)} (1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)})^2 \leq c h \|
abla \varphi^n\|_{L^2(\Omega)} \leq c h(1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)}),$$

which together with the trivial inequality $\tau \sum_{n=1}^N t_n^{\min(0, -\frac{d}{4} - \epsilon)} \leq c$ implies

$$\tau \sum_{n=1}^N I_1^\ast \leq c h \gamma^{-1} \eta^2. \tag{4.7}$$

For the term $I_2$, by the triangle inequality, inverse inequality, $H^1(\Omega)$ stability of $P_h$, we have

$$\|\nabla(u(t_n) - U_h^n(q^h_{\ast}))\|_{L^2(\Omega)} \leq \|\nabla(u(t_n) - P_h u(t_n))\|_{L^2(\Omega)} + h^{-1} \|P_h u(t_n) - U_h^n(q^h_{\ast})\|_{L^2(\Omega)} \leq c(h + h^{-1} \|u(t_n) - U_h^n(q^h_{\ast})\|_{L^2(\Omega)}),$$

and consequently, the Cauchy-Schwarz inequality and Lemma 4.4 imply

$$\tau \sum_{n=1}^N I_2^\ast \leq \tau \sum_{n=1}^N \|\nabla u(t_n) - U_h^n(q^h_{\ast})\|_{L^2(\Omega)} \|\nabla \varphi^n\|_{L^2(\Omega)} \leq c \left(1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)} \right) \left(1 + \|
abla q^h_{\ast}\|_{L^2(\Omega)} \right) \leq c(h \gamma^{-\frac{1}{2}} + h^{-1} \gamma^{-\frac{3}{2}} \eta). \tag{4.8}$$

Next we bound the third term $I_3$. It follows directly from (2.2) and (3.6) that

$$I_3^\ast = (q^1 \nabla u(t_n) - q^h_\ast \nabla U_h^n(q^h_{\ast}), \nabla P_h \varphi^n) = (\partial_t^n(u(t_n) - U_h^n(q^h_{\ast}) - \partial_t^n(u(t_n) - u_0), P_h \varphi^n)$$
\[
\begin{align*}
&= (\overline{\partial}_\nu^\alpha [(U_h^n(q_h^*) - U_h^0) - (u(t_n) - u_0)], P_h\phi^n) \\
&\quad + (\overline{\partial}_\nu^\alpha (u(t_n) - u_0) - \partial_\nu^\alpha (u(t_n) - u_0), P_h\phi^n) =: I_{3,1}^n + I_{3,2}^n.
\end{align*}
\]

It remains to bound the two terms \(I_{3,1}^n\) and \(I_{3,2}^n\) separately. By Lemma \[4.2\] there holds
\[
|I_{3,2}^n| = \|\overline{\partial}_\nu^\alpha (u(t_n) - u_0) - \partial_\nu^\alpha (u(t_n) - u_0)\|_{L^2(\Omega)} \|P_h\phi^n\|_{L^2(\Omega)} \leq c\tau t_n^{-1}, \quad n = 1, 2, \ldots, N.
\]

Consequently,
\[
|\tau^2 \sum_{m=1}^N \sum_{n=1}^m I_{3,2}^n| \leq c\tau^3 \sum_{m=1}^N \sum_{n=1}^m t_n^{-1} \leq c\tau \log(1 + t_N / \tau).
\]

It remains to bound the term \(I_{3,1}^n\). Since \(U_h^n(q_h^*) = U_h^0\) and \(u(0) = u_0\), straightforward computation with summation by parts yields
\[
\tau \sum_{m=1}^N I_{3,1}^n = \tau \sum_{m=1}^N \sum_{n=1}^m [(U_h^n(q_h^*) - U_h^0) - (u(t_n) - u_0)], P_h\phi^n) = \tau \sum_{j=0}^m [((U_h^n(q_h^*) - U_h^0) - (u(t_j) - u_0)], \tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} P_h\phi^n).\]

Next we appeal to the splitting
\[
\tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} P_h\phi^n = \tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} P_h(\phi^n - \phi_m^n) + \tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} P_h\phi_m^n := IV_{j,m}^1 + IV_{j,m}^2.
\]

By Lemma \[3.2\] the sum \(IV_{j,m}^2\) satisfies
\[
|IV_{j,m}^2| \leq c\|\phi_m^n\|_{L^2(\Omega)} \left(\tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} \right) \leq ct_{m-j+1}^{-\alpha}\|\phi_m^n\|_{L^2(\Omega)} \leq ct_{m-j+1}^{-\alpha},
\]

since \(\|\phi_m^n\|_{L^2(\Omega)} \leq c\). Then Lemma \[4.4\] and Cauchy-Schwarz inequality imply
\[
|IV_{j,m}^1| \leq c\|IV_{j,m}^1\|_{L^2(\Omega)} \left(\tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} \right) \leq ct_{m-j+1}^{-\alpha}\|IV_{j,m}^1\|_{L^2(\Omega)} \leq c t_{m-j+1}^{-\alpha},
\]

where the second inequality is due to \(\tau \sum_{m=j}^N t_{m-j+1}^{-\alpha} \leq c t_{N-j+1}^{-\alpha} \leq c t_{N-j+1}^{-\alpha}\). Similarly, by Lemma \[4.5\]
\[
|IV_{j,m}^1| \leq c\|IV_{j,m}^1\|_{L^2(\Omega)} \left(\tau^{-\alpha} \sum_{n=0}^m b_{n-j}^{(\alpha)} \right) \leq c\|IV_{j,m}^1\|_{L^2(\Omega)} \leq c t_{m-j+1}^{-\alpha},
\]

These two estimates and the triangle inequality lead to
\[
|\tau^2 \sum_{m=1}^N \sum_{n=1}^m (\overline{\partial}_\nu^\alpha [(U_h^n(q_h^*) - U_h^0) - (u(t_n) - u_0)], P_h\phi^n)| \leq c\eta. \tag{4.9}
\]

The three estimates \[4.7\], \[4.8\] and \[4.9\] together imply
\[
|\tau^2 \sum_{m=1}^N \sum_{n=1}^m (\overline{\partial}_\nu^\alpha [(U_h^n(q_h^*) - U_h^0) - (u(t_n) - u_0)], P_h\phi^n)| \leq c(h\gamma^{-1}\eta + \gamma^{-\frac{1}{2}}\eta + h^{-1}\gamma^{-\frac{1}{2}}\eta).\]
Finally, this and the identity
\[
((q^l - q^h) \nabla u(t_n), \nabla \varphi^n) = \int_{\Omega} \left( \frac{q^l - q^h}{q^l} \right)^2 q^l |\nabla u(t_n)|^2 + (f(t_n) - \partial^a u(t_n))u(t_n) \, dx
\]
lead immediately to the desired assertion. This completes the proof of the theorem.

**Remark 4.2.** The restriction on \(d = 1, 2\) is due to limited regularity pickup on general convex polyhedral domains. The result holds also for a polyhedral domain with suitable conditions [5, Theorem 4, p. 18].

The next result is an immediate corollary of Theorem 4.1.

**Corollary 4.1.** Let \(q^l\) be the exact diffusion coefficient, \(u \equiv u(q^l)\) the solution to problem (4.2), and \(q^h \in A_h\) the solution to problem (4.2)-(4.3). Then under Assumption 4.1 for \(d = 1, 2\), there holds (with \(\eta = \tau + \frac{\gamma}{2} + h^2 \ell_h + \delta + \gamma \frac{\tau}{2}\))
\[
\int_0^T \int_0^t \left( \frac{q^l - q^h}{q^l} \right)^2 q^l |\nabla u(s)|^2 + (f(s) - \partial^a u(s))u(s) \right) \, dx \, ds \, dt \leq c \epsilon h^{-1} \eta + h^{-\frac{\gamma}{2}} + h^{-1} \gamma \frac{\tau}{2} \eta.
\]

**Proof.** In view of Theorem 4.1 it suffices to bound the quadrature error:
\[
\left| \int_0^T \int_0^t |\nabla u(s)|^2 \, ds \, dt - \tau^2 \sum_{m=1}^N \sum_{n=1}^m |\nabla u(t_n)|^2 \right| + \left| \int_0^T \int_0^t (f(s) - \partial^a u(s))u(s) \, ds \, dt - \tau^2 \sum_{m=1}^N \sum_{n=1}^m (f(t_n) - \partial^a u(t_n))u(t_n) \right| := I + II.
\]

It remains to bound the two terms I and II. For the first term,
\[
I \leq \left| \sum_{m=1}^N \left( \int_{t_{m-1}}^{t_m} \int_0^{t_m} |\nabla u(s)|^2 \, ds \, dt - \tau^2 \sum_{n=1}^m |\nabla u(t_n)|^2 \right) \right| + \left| \tau \sum_{m=1}^N \int_{t_{m-1}}^{t_m} |\nabla u(s)|^2 \, ds \right|
\]
\[
\leq \tau \sum_{m=1}^N \int_0^{t_m} |\nabla u(s)|^2 \, ds - \tau \sum_{n=1}^m |\nabla u(t_n)|^2 + \tau \sum_{m=1}^N \int_{t_{m-1}}^{t_m} |\nabla u(s)|^2 \, ds.
\]

By the regularity estimate 4.1, \(\|\nabla u'(s)\|_{L^2(\Omega)} \leq c s^{\frac{\gamma}{2}-1}\) and \(\|\nabla u(t)\|_{C([0,T];L^2(\Omega))} \leq c\). Clearly \(\tau \sum_{m=1}^N \int_{t_{m-1}}^{t_m} |\nabla u(s)|^2 \, ds \leq ct\). Further,
\[
\int_{\Omega} I_{m} \, dx \leq \sum_{n=1}^m \int_{t_{n-1}}^{t_m} |\nabla u(s) + u(t_n)|_{L^2(\Omega)} \|\nabla (u(s) - u(t_n))\|_{L^2(\Omega)} \, ds \leq c \|\nabla u\|_{C([0,T];L^2(\Omega))} \sum_{n=1}^m \int_{t_{n-1}}^{t_m} \|\nabla u(\zeta)\|_{L^2(\Omega)} \, ds \leq c \|\nabla u\|_{C([0,T];L^2(\Omega))} T \int_0^T s^{\frac{\gamma}{2}-1} \, ds \leq ct.
\]

Combining the preceding two estimates, we obtain \(\int_{\Omega} I_{m} \, dx \leq ct\). The term II can be bounded similarly as \(\int_{\Omega} I_{m} \, dx \leq ct \ln \tau\). Indeed, under Assumption 4.1, the regularity estimate 4.1, we have \(\|\partial^a u\|_{L^2(\Omega)} \leq c\) and \(\|\partial^a u'(t)\|_{L^2(\Omega)} \leq ct^{-1}\), and thus the function \(g(t) \equiv \partial^a u(t) - f(t)\) satisfies \(\|g(t)\|_{L^2(\Omega)} \leq c\) and \(\|g'(t)\|_{L^2(\Omega)} \leq ct^{-1}\). Then repeating the argument completes the proof.
Remark 4.3. The error estimate in Corollary 4.1 would provide the usual $L^2(\Omega)$ error estimate, if the following structural condition holds: For the exact diffusion coefficient $q^\dagger$ and the corresponding state variable $u \equiv u(q^\dagger)$, there holds
\begin{equation}
\int_0^T \int_0^t \left( q^\dagger |\nabla u(s)|^2 + (f(s) - \partial_t^\alpha u(s))u(s) \right) ds dt > c_0 \quad \text{a.e. } x \in \bar{\Omega}. \tag{4.10}
\end{equation}

In the classical parabolic case, similar structural conditions have been assumed in the literature, e.g., the following characteristic condition [34, 21]: 
\[ t^{-1} \int_0^t \nabla u(q)(x,s) \cdot \nu \geq \delta_0 > 0 \text{ for all } (x,t) \in \Omega \times (0,T), \]
where $\nu$ is a constant vector, or [35] Theorem 6.4 \[ a_0 \int_0^t |\nabla u(q)(s)|^2 + t \int_0^t (u_q(q)(s) - f(s)) ds \geq 0 \text{ a.e. in } \Omega \times (0,T). \]
Note that this latter condition is not positively homogeneous (with respect to problem data). Next we comment on the condition (4.10). If $f \geq 0$ in $Q \equiv \Omega \times (0,T]$, $u_0 > 0$ in $\Omega$, then maximum principle [29] implies $u > 0$ in $Q$. Further, $w = \partial_t^\alpha u$ satisfies
\[ \begin{cases}
\partial_t^\alpha w - \nabla \cdot (q \nabla w) = \partial_t^\alpha f, & \text{in } Q, \\
w(0) = \nabla \cdot (q \nabla u_0) + f(0), & \text{in } \Omega, \\
w(t) = 0, & \text{on } \partial \Omega \times (0,T].
\end{cases} \]
If $\partial_t^\alpha f(t) \leq 0$ and $\nabla q \cdot \nabla u_0 + q \Delta u_0 + f(0) \leq 0$, then maximum principle implies $\partial_t^\alpha u = w \leq 0$ in $Q$. Further, if $f > 0$ in $Q$, then $f - \partial_t^\alpha u > 0$ in $Q$, which implies $(f - \partial_t^\alpha u) > 0$ in $Q$. Thus condition (4.10) holds.

Remark 4.4. Theorem 4.1 and Corollary 4.1 show that the convergence rate is of order $O(\delta^{1/2})$ in the weighted norm, provided that $\gamma = O(h^2) = O(\delta^2) = O(\tau^{1+\alpha})$. The error estimate in Theorem 4.1 and Corollary 4.1 is expected to be sub-optimal, due to the presence of the factor $h^{-1}$, which arises from the use of inverse inequality in (4.8). It remains unclear how to achieve optimality, even in the standard parabolic case [35].

5 Numerical results and discussions

Now we present numerical results to illustrate the fully discrete scheme (3.5)–(3.6) with one- and two-dimensional examples, with the measurement $z^\delta$ over the time interval $[T_0, T]$ (by a straightforward adaptation of the formulation; see Remark 3.1), with $T$ fixed at 1. Throughout, the corresponding discrete problem is solved by the conjugate gradient (CG) method [2], with the gradient computed using the standard adjoint technique. The lower and upper bounds in the admissible set $A$ are taken to be 0.5 and 5, respectively, and are enforced by a projection step after each CG iteration (but it was never active in the numerical experiments). The minimization method converges generally within tens of iterations. The noisy data $z^\delta$ is generated by
\[ z^\delta(x,t) = u(q^\dagger)(x,t) + \varepsilon \sup_{(x,t) \in \Omega \times [T_0, T]} |u(x,t)| \xi(x,t), \quad (x,t) \in \Omega \times [T_0, T], \]
where $\xi(x,t)$ follows the standard Gaussian distribution, and $\varepsilon \geq 0$ denotes the (relative) noise level. The noisy data $z^\delta$ is first generated on a fine spatial-temporal mesh and then interpolated to a coarse spatial/temporal mesh for the inversion step. The scalar $\gamma$ in the functional $J_\gamma$ plays an important role in determining the accuracy of the reconstructions, but it is notoriously challenging to choose (see e.g., [14]). In our experiments, its value is determined by a trial and error manner, first for the fractional order $\alpha = 0.50$, and then employed for the other two cases $\alpha = 0.25$ and $\alpha = 0.75$, which might be suboptimal but works reasonably well in practice.

5.1 Numerical results in one spatial dimension

First we present numerical results for two examples on unit interval $\Omega = (0,1)$. The reference data $z(q^\dagger)$ is computed with a mesh size $h = 1/400$ and time step size $T = 1/2048$, and the inversion step is carried out with a mesh size $h = 1/200$ and time step size $\tau = 1/1024$, unless otherwise specified.

The first example has a smooth exact coefficient $q^\dagger$, and the problem is homogeneous.
Example 5.1. \( u_0 = x(1-x), \; f \equiv 0, \; q^1 = 2 + \sin(2\pi x) \).

First, we let \( T_0 = 0.75 \) and study how the reconstruction error changes according to different parameters. The numerical results for the example with different noise levels \( \varepsilon \), and fixed \( h \) and \( \tau \), are summarized in Table 1. The chosen \( \gamma \) is relatively small, since the magnitude of the exact data \( u(q^1) \) is actually very small: for example, upon convergence, the functional value \( J_{\h,\tau}(q^1_\h) \) is about \( O(10^{-12}) \) for exact data and about \( O(10^{-9}) \) for \( \varepsilon = 1.00e-2 \). Clearly, the \( L^2(\Omega) \) error of the reconstruction \( q^1_\h \), i.e., \( \| q^1 - q^1_\h \|_{L^2(\Omega)} \), decreases steadily as the noise level \( \varepsilon \) tends to zero (Note that even at \( \varepsilon = 0 \), the reconstruction error is nonzero due to the presence of discretization errors). The convergence is consistently observed for all three fractional orders. Interestingly, for a fixed noise level \( \varepsilon \), as the fractional order \( \alpha \) increases from \( 0.25 \) to \( 0.75 \), the reconstruction error tends to deteriorate slightly. It might be related to the fact that for homogeneous subdiffusion, the smaller is the fractional order \( \alpha \), the quicker the state \( u(t) \) approaches a “quasi”-steady state; Then the inverse problem reduces to the elliptic counterpart, i.e., \( -\nabla \cdot \alpha \frac{\partial u}{\partial x} = f \), which is known to be beneficial for numerical reconstruction \[19\]. However, the precise mechanism remains to be ascertained. We refer to Fig. [4] for reconstructions: the recoveries are qualitatively comparable with each other and all reasonably accurate for \( \varepsilon \) up to \( \varepsilon = 5.00e-2 \). These observations concur well with the numbers in Table 1.

| \( \varepsilon \) | 0 | 1.00e-3 | 5.00e-3 | 1.00e-2 | 3.00e-2 | 5.00e-2 |
|----|----|----|----|----|----|----|
| \( \gamma \) | 1.00e-14 | 1.00e-13 | 3.00e-13 | 5.00e-13 | 1.00e-12 | 3.00e-12 |
| \( \alpha = 0.25 \) | 7.75e-3 | 9.95e-3 | 1.33e-2 | 1.53e-2 | 2.50e-2 | 3.64e-2 |
| \( \alpha = 0.50 \) | 8.73e-3 | 1.00e-2 | 1.33e-2 | 1.50e-2 | 2.65e-2 | 4.11e-2 |
| \( \alpha = 0.75 \) | 9.92e-3 | 1.16e-2 | 1.80e-2 | 2.24e-2 | 3.30e-2 | 5.16e-2 |

Next we examine the convergence with respect to the mesh size \( h \) and time step size \( \tau \); see Tables 2 and 3 for the empirical convergence with respect to \( h \) and \( \tau \), respectively. The reference regularized solution \( q^* \) is computed with \( h = 1/800 \) and \( \tau = 1/2048 \), and it differs slightly from the exact diffusion coefficient \( q^1 \), due to the presence of data noise (\( \varepsilon = 1.0e-2 \)). Clearly, the \( L^2(\Omega) \) error \( \| q^* - q^1_\h \|_{L^2(\Omega)} \) of the reconstruction \( q^1_\h \) (which depends also implicitly on \( \tau \) via the optimization problem (3.5)–(3.6)) decreases as either the mesh size \( h \) or time step size \( \tau \) tends to zero, and the convergence is generally steady. These observations partially confirm the convergence result in Theorem 3.2.

| \( M \) | 10 | 20 | 40 | 80 | 160 | 320 |
|----|----|----|----|----|----|----|
| \( \alpha = 0.25 \) | 5.39e-2 | 2.74e-2 | 2.33e-2 | 1.46e-2 | 2.04e-2 | 1.15e-2 |
| \( \alpha = 0.50 \) | 5.38e-2 | 2.56e-2 | 2.51e-2 | 1.56e-2 | 1.16e-2 | 6.51e-3 |
| \( \alpha = 0.75 \) | 4.61e-2 | 2.57e-2 | 2.26e-2 | 2.41e-2 | 1.14e-2 | 8.00e-3 |

| \( \tau \) | \( 2^{-9} \) | \( 2^{-8} \) | \( 2^{-7} \) | \( 2^{-6} \) | \( 2^{-5} \) | \( 2^{-4} \) |
|----|----|----|----|----|----|----|
| \( \alpha = 0.25 \) | 3.78e-2 | 3.88e-2 | 2.03e-2 | 8.30e-3 | 2.38e-2 | 6.27e-3 |
| \( \alpha = 0.50 \) | 3.90e-2 | 3.80e-2 | 1.98e-2 | 1.92e-2 | 2.07e-2 | 8.46e-3 |
| \( \alpha = 0.75 \) | 9.31e-2 | 4.47e-2 | 2.64e-2 | 1.06e-2 | 1.45e-2 | 6.64e-3 |

Last, we take \( T_0 = 0 \) and examine the convergence of the errors

\[ e_u = \| q^1 - q^1_\h \|_{L^2(\Omega)} \quad \text{and} \quad e_u = \left( \tau \sum_{n=1}^{N} \| u(t_n) - U^n_h(q^1_\h) \|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}, \]
Figure 1: Numerical reconstructions for Example 5.1 (top: $\alpha = 0.25$; middle: $\alpha = 0.50$; bottom: $\alpha = 0.75$).

with respect to $\epsilon$. Motivated by the error estimates in Theorem 4.1 and Remark 4.4, we fix a small $\tau = 1/2048$ and let $h = \sqrt{\epsilon}$ and $\gamma = 10^{-4} \times \epsilon^2$. The errors $e_q$ and $e_u$ are plotted in Fig. 2: a first-order convergence $O(\epsilon)$ is observed. This shows sub-optimality of the theoretical convergence rate in Theorem 4.1. This remains an outstanding question for the analysis of the discrete problem, and seems open even for the standard parabolic case.

Figure 2: Plot of $e_u$ and $e_q$ versus $\epsilon$, with $h = \sqrt{\epsilon}$, $\gamma = 10^{-4} \times \epsilon^2$ and $\tau = 1/2048$. 
The second example has a nonsmooth exact coefficient $q^\dagger$, and the problem is inhomogeneous. The notation min denotes the pointwise minimum.

**Example 5.2.** $u_0(x) = x^2(1 - x)^2$, $f(x, t) = e^{x(1-x)}x(1 - x)t$, $q^\dagger = 2 + \min\left(\frac{1}{2}, \sin^4(2\pi x)\right)$, and $T_0 = 0.75$.

The numerical results for the example with different noise levels are given in Table 4 and Fig. 3. The observations from Example 5.1 remain largely valid: the error $\|q^\dagger - q^\ast h\|_{L^2(\Omega)}$ decreases as the noise level $\varepsilon$ decreases to zero. The results are mostly comparable for all three fractional orders. For high noise levels, e.g., $\varepsilon = 5.00e-2$, the error in the reconstruction is clearly dominated by the oscillations within the flat regions, which is reminiscent of the Gibbs phenomenon arising from the approximation of the kinks, and also the deviations in the valley. Nonetheless, all the results are fair and represent acceptable approximations.

| $\varepsilon$ | $\gamma$ | 0.00e-3 | 5.00e-3 | 1.00e-2 | 3.00e-2 | 5.00e-2 |
|--------------|----------|---------|---------|---------|---------|---------|
| $\alpha = 0.25$ | 1.00e-15 | 2.00e-13 | 4.00e-13 | 1.00e-12 | 4.00e-12 | 9.00e-12 |
| $\alpha = 0.50$ | 6.13e-3 | 6.95e-3 | 9.15e-3 | 1.57e-2 | 2.45e-2 | 2.87e-2 |
| $\alpha = 0.75$ | 6.23e-3 | 1.14e-2 | 1.03e-2 | 1.54e-2 | 2.18e-2 | 2.87e-2 |

Figure 3: Numerical reconstructions for Example **5.2** (Top: $\alpha = 0.25$; middle: $\alpha = 0.50$; bottom: $\alpha = 0.75$)
5.2 Numerical results in two spatial dimension

Now we present numerical results for the following example on the unit square \(\Omega = (0,1)^2\). The domain \(\Omega\) is first uniformly divided into \(M^2\) small squares, each with side length \(1/M\), and then a triangulation is obtained by connecting the low-left and upper-right vertices of each small square. The reference data is first computed on a finer mesh with \(M = 100\) and a time step size \(\tau = 1/2000\). The inversion is carried out with a mesh \(M = 40\) and \(\tau = 1/500\).

**Example 5.3.** \(u_0(x_1, x_2) = x_1(1-x_1)\sin(\pi x_2), f \equiv 0, q^l(x_1, x_2) = 1 + \sin(\pi x_1)x_2(1-x_2),\) and \(T_0 = 0.8\).

The numerical results for the example with different noise levels are presented in Table 5 and Figs. 4–6. The empirical observations are in excellent agreement with for Example 5.1 e.g., convergence with decreasing noise level \(\varepsilon\) and slightly improved reconstructions for increasing fractional orders \(\alpha\). Figs. 4–6 indicate that the error \(e_q = q_h^* - q^1\) lies mainly in recovering the peak, however, the overall shape is well recovered.

| \(\varepsilon\) | \(\gamma\) | \(\alpha = 0.25\) | \(\alpha = 0.50\) | \(\alpha = 0.75\) |
|---|---|---|---|---|
| 0 | 1.00e-3 | 1.51e-3 | 1.61e-3 | 1.59e-3 |
| 5e-3 | 1.00e-2 | 1.75e-3 | 1.86e-3 | 2.21e-3 |
| 1.00e-2 | 3.00e-2 | 2.87e-3 | 2.80e-3 | 3.38e-3 |
| 3.00e-2 | 5.00e-2 | 3.64e-3 | 3.62e-3 | 4.66e-3 |
| 5.00e-2 | 8.00e-2 | 5.82e-3 | 5.68e-3 | 1.13e-2 |

Figure 4: Numerical reconstructions for Example 5.3 with \(\alpha = 0.25\).

A Proof of Lemma 4.2

The proof relies on the discrete Laplace transform, and the following two well known estimates

\[
c_1|z| \leq |\delta_\tau(e^{-z\tau})| \leq c_2|z|, \quad \forall z \in \Gamma_{\theta, \delta}, \tag{A.1}
\]

\[
|\delta_\tau(e^{-z\tau})| \leq |z| \sum_{k=1}^{\infty} \frac{|z\tau|^{k-1}}{k!} \leq |z|e^{|z|\tau}, \quad \forall z \in \Sigma_\theta, \tag{A.2}
\]

and the resolvent estimate

\[
\|(z - A(q))^{-1}\| \leq c|z|^{-1}, \quad \forall z \in \Sigma_\theta. \tag{A.3}
\]
Proof. Let $y(t) = u(t) - u_0$. Then $y(t)$ satisfies
\[
\partial_t^\alpha y(t) - Ay(t) + Au_0 = f(t), \quad 0 < t \leq T.
\]
Taking Laplace transform gives
\[
z^\alpha \hat{y}(z) - A\hat{y}(z) + z^{-1}Au_0 = \hat{f}(z),
\]
i.e., $\hat{y}(z) = (z^\alpha - A)^{-1}(\hat{f}(z) - z^{-1}Au_0)$. Since $\partial_t^\alpha y(t) = z^\alpha \hat{y}(z)$ and $\partial_t^\alpha y = \delta_t(z^\alpha \hat{y}(z))$,
\[
\partial_t^\alpha y(t_n) - \partial_t^\alpha y(t_n) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt_n} K(z)(z^{-1}Au_0 - \hat{f}(z)) \, dz
\]
\[
+ \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta} \setminus \Gamma_{\theta,\delta}} e^{zt_n} K(z)(z^{-1}Au_0 - \hat{f}(z)) \, dz,
\]
with $K(z) = (\delta_t(e^{-zt})^\alpha - z^\alpha)(z^\alpha - A)^{-1}$. Recall the following estimate:
\[
|\delta_t(e^{-zt})^\alpha - z^\alpha| \leq cT z^{1+\alpha}, \quad \forall z \in \Gamma_{\theta,\delta}.
\]
(A.4)
By choosing $\delta = c/t_n$ and (A.3), $I = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt_n} K(z) z^{-1} (Au_0 - f(0)) dz$ is bounded by

$$||I||_{L^2(\Omega)} \leq c \tau \|Au_0 - f(0)\|_{L^2(\Omega)} \left( \int_0^{\pi/2} e^{-c\rho t_n} d\rho + \int_0^\theta ct_n^{-1} d\theta \right) \leq c\tau t_n^{-1} \|Au_0 - f(0)\|_{L^2(\Omega)}.$$  

Further, by (A.2), for any $\delta$ by choosing $\theta \in (\pi/2, \pi)$ close to $\pi$,

$$|e^{zt_n} (\delta_r (e^{-z\tau}) - z^\alpha) z^{-1}| \leq e^{t_n \rho \cos \theta} (c|z|^\alpha e^{\alpha \rho \tau} + |z|^\alpha) |z|^{-1} \leq c|z|^{\alpha - 1} e^{-c\rho t_n}.$$  

Then the term $I = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt_n} K(z) z^{-1} (Au_0 - f(0)) dz$ is bounded by

$$||I||_{L^2(\Omega)} \leq c \|Au_0 - f(0)\|_{L^2(\Omega)} \int_0^{\pi/2} e^{-c\rho t_n} d\rho \leq c\tau t_n^{-1} \|Au_0 - f(0)\|_{L^2(\Omega)}.$$  

In view of the splitting $f(t) = f(t) + tf'(0) + df_t^2 f''(t)$, it remains to bound the other two terms. Upon extending $f''(t)$ by zero to $\mathbb{R}_-$, straightforward computation gives

$$\partial_t^2 y(t_n) - \partial_z^2 y(t_n) = -\frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt_n} K(z) z^{-2} dz f'(0) ds - \frac{1}{2\pi i} \int_0^{t_n} \int_{\Gamma_{\theta,\delta}} e^{z(t_n - s)} z^{-2} K(z) dz f''(s) ds.$$  

Then repeating the preceding argument leads to

$$\|\partial_t^2 y(t_n) - \partial_z^2 y(t_n)\|_{L^2(\Omega)} \leq c \tau \left( \|f'(0)\|_{L^2(\Omega)} + \int_0^{t_n} \|f''(s)\|_{L^2(\Omega)} ds \right).$$  

Combining the preceding estimates shows the desired assertion.

\qed

References

[1] R. A. Adams and J. F. Fournier. *Sobolev Spaces*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] O. M. Alifanov, E. A. Artyukhin, and S. V. Rumyantsev. *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*. Begell House, New York, 1995.

[3] A. A. Alikhanov. A priori estimates for solutions of boundary value problems for equations of fractional order. *Differ. Uravn.*, 46(5):658–664, 2010.

[4] 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):115002, 16, 2009.

[5] M. Dauge. Regularity and singularities in polyhedral domains. Slide, available at https://perso.univ-rennes1.fr/monique.dauge/publis/Talk_Karlsruhe08.pdf, last accessed on August 1, 2019.

[6] M. Dauge. *Elliptic Boundary Value Problems on Corner Domains*. Springer-Verlag, Berlin, 1988.

[7] P. DuChateau, R. Thelwell, and G. Butters. Analysis of an adjoint problem approach to the identification of an unknown diffusion coefficient. *Inverse Problems*, 20(2):601–625, 2004.

[8] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic, Dordrecht, 1996.
[9] H. W. Engl and J. Zou. A new approach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction. *Inverse Problems*, 16(6):1907–1923, 2000.

[10] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer-Verlag, New York, 2004.

[11] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, Boca Raton, FL, 2015.

[12] S. Gutman. Identification of discontinuous parameters in flow equations. *SIAM J. Control Optim.*, 28(5):1049–1060, 1990.

[13] V. Isakov. *Inverse Problems for Partial Differential Equations*. Springer, New York, second edition, 2006.

[14] K. Ito and B. Jin. *Inverse Problems: Tikhonov Theory and Algorithms*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.

[15] D. Jiang, Z. Li, Y. Liu, and M. Yamamoto. Weak unique continuation property and a related inverse source problem for time-fractional diffusion-advection equations. *Inverse Problems*, 33(5):055013, 22, 2017.

[16] B. Jin, R. Lazarov, and Z. Zhou. Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. *SIAM J. Sci. Comput.*, 38(1):A146–A170, 2016.

[17] B. Jin, R. Lazarov, and Z. Zhou. Numerical methods for time-fractional evolution equations with nonsmooth data: a concise overview. *Comput. Methods Appl. Mech. Engrg.*, 346:332–358, 2019.

[18] B. Jin, B. Li, and Z. Zhou. Numerical analysis of nonlinear subdiffusion equations. *SIAM J. Numer. Anal.*, 56(1):1–23, 2018.

[19] B. Jin and W. Rundell. A tutorial on inverse problems for anomalous diffusion processes. *Inverse Problems*, 31(3):035003, 40, 2015.

[20] B. Kaltenbacher and W. Rundell. On an inverse potential problem for a fractional reaction-diffusion equation. *Inverse Problems*, 35(6):065004, 31 pp., 2019.

[21] T. Kärkkäinen. Error estimates for distributed parameter identification in parabolic problems with output least squares and Crank-Nicolson method. *Appl. Math.*, 42(4):259–277, 1997.

[22] Y. L. Keung and J. Zou. Numerical identifications of parameters in parabolic systems. *Inverse Problems*, 14(1):83–100, 1998.

[23] Y. Kian, L. Oksanen, E. Soccorsi, and M. Yamamoto. Global uniqueness in an inverse problem for time fractional diffusion equations. *J. Diff. Equations*, 264(2):1146–1170, 2018.

[24] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and Applications of Fractional Differential Equations*. Elsevier Science B.V., Amsterdam, 2006.

[25] P. Kügler. Identification of a temperature dependent heat conductivity from single boundary measurements. *SIAM J. Numer. Anal.*, 41(4):1543–1563, 2003.

[26] G. Li, W. Gu, and X. Jia. Numerical inversions for space-dependent diffusion coefficient in the time fractional diffusion equation. *J. Inverse Ill-Posed Probl.*, 20(3):339–366, 2012.

[27] G. Li, D. Zhang, X. Jia, and M. Yamamoto. Simultaneous inversion for the space-dependent diffusion coefficient and the fractional order in the time-fractional diffusion equation. *Inverse Problems*, 29(6):065014, 36, 2013.

[28] X. Li and C. Xu. A space-time spectral method for the time fractional diffusion equation. *SIAM J. Numer. Anal.*, 47(3):2108–2131, 2009.
[29] Y. Liu, W. Rundell, and M. Yamamoto. Strong maximum principle for fractional diffusion equations and an application to an inverse source problem. *Fract. Calc. Appl. Anal.*, 19(4):888–906, 2016.

[30] R. Metzler, J. H. Jeon, A. G. Cherstvy, and E. Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.*, 16(44):24128–24164, 2014.

[31] R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339(1):1–77, 2000.

[32] T. K. Nilssen and X. C. Tai. Parameter estimation with the augmented Lagrangian method for a parabolic equation. *J. Optim. Theory Appl.*, 124(2):435–453, 2005.

[33] 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):426–447, 2011.

[34] X.-C. Tai and T. Kärkkäinen. Identification of a nonlinear parameter in a parabolic equation from a linear equation. *Mat. Apl. Comput.*, 14(2):157–184, 1995.

[35] L. Wang and J. Zou. Error estimates of finite element methods for parameter identifications in elliptic and parabolic systems. *Discrete Contin. Dyn. Syst. Ser. B*, 14(4):1641–1670, 2010.

[36] T. Wei and Y. S. Li. Identifying a diffusion coefficient in a time-fractional diffusion equation. *Math. Comput. Simul.*, 151:77–95, 2018.

[37] Z. Zhang. An undetermined coefficient problem for a fractional diffusion equation. *Inverse Problems*, 32(1):015011, 21, 2016.

[38] Z. Zhang and Z. Zhou. Recovering the potential term in a fractional diffusion equation. *IMA J. Appl. Math.*, 82(3):579–600, 2017.

[39] G.-H. Zheng and M.-H. Ding. Identification of the degradation coefficient for an anomalous diffusion process in hydrology. *Inverse Problems*, pages in press, https://doi.org/10.1088/1361-6420/ab1ed7, 2019.