Numerische Mathematik (2025) 157:2323–2355
https://doi.org/10.1007/s00211-025-01495-2

Numerische
Mathematik

Numerical recovery of the diffusion coefficient in diffusion equations from terminal measurement

Bangti Jin[1] · Xiliang Lu[2] · Qimeng Quan[3] · Zhi Zhou[4]

Received: 19 September 2023 / Revised: 23 August 2025 / Accepted: 4 September 2025 /
Published online: 14 October 2025

Abstract

In this work, we investigate a numerical procedure for recovering a space-dependent diffusion coefficient in a (sub)diffusion model from the given terminal data, and provide a rigorous numerical analysis of the procedure. By exploiting decay behavior of the observation in time, we establish a novel Hölder type stability estimate for a large terminal time T. This is achieved by novel decay estimates of the (fractional) time derivative of the solution. To numerically recover the diffusion coefficient, we employ the standard output least-squares formulation with an $H^1(\Omega)$-seminorm penalty, and discretize the regularized problem by the Galerkin finite element method with continuous piecewise linear finite elements in space and backward Euler convolution quadrature in time. Further, we provide an error analysis of discrete approximations, and prove a convergence rate that matches the stability estimate. The derived $L^2(\Omega)$ error bound depends explicitly on the noise level, regularization parameter and discretization parameters, which gives a useful guideline of the a priori choice of discretization parameters with respect to the noise level in practical implementation. The error analysis is achieved using the conditional stability argument and discrete maximum-norm resolvent estimates. Several numerical experiments are also given to illustrate and complement the theoretical analysis.

Mathematics Subject Classification 65M30 · 65M15 · 65M60

1 Introduction

In this work, we study the inverse problem of recovering a space-dependent diffusion coefficient in (sub)diffusion equation from a terminal observation. Let $\Omega \subset \mathbb{R}^d$ ($d =$

The work of B. Jin is supported by Hong Kong Research Council grant (No. 14306423), and a start-up fund from The Chinese University of Hong Kong, X. Lu by the National Science Foundation of China (No. 12371424, No. U24A2002) and the Natural Science Foundation of Hubei Province (No. 2024AFE006), Q. Quan by UM Postdoctoral Fellow scheme from University of Macau, and that of Z. Zhou by Hong Kong Research Grants Council grant (No. 15303021) and an internal grant of The Hong Kong Polytechnic University (Project ID: P0045708, Work Programme: 4-ZZP7).

Extended author information available on the last page of the article

1, 2, 3) be a simply connected convex bounded domain with a smooth boundary $\partial\Omega$. The governing equation is given by

$$\begin{cases} \partial_t^\alpha u - \nabla\cdot(q\nabla u) = f, & \text{in } \Omega\times(0,T), \\ u = 0, & \text{on } \partial\Omega\times(0,T), \\ u(0) = u_0, & \text{in } \Omega, \end{cases} \tag{1.1}$$

where $T > 0$ is the final time and the notation $\partial_t^\alpha u$ denotes the left-sided Djrbashian–Caputo fractional derivative of order $\alpha \in (0, 1]$ in the time variable t defined by [27, p. 92]:

$$\partial_t^\alpha u(t) := \begin{cases} \dfrac{1}{\Gamma(1-\alpha)}\displaystyle\int_0^t (t-s)^{-\alpha}\partial_s u(s)\,\mathrm{d}s, & \text{for } \alpha\in(0,1), \\ \partial_t u(t), & \text{for } \alpha = 1, \end{cases}$$

with $\Gamma(z) = \int_0^\infty e^{-s}s^{z-1}\,\mathrm{d}s$, $\Re(z) > 0$, being Euler's Gamma function. The functions f and u_0 in (1.1) are given time-independent source and initial data, respectively. Due to its extraordinary modeling capability for describing the dynamics of subdiffusion processes (in which the mean square variance grows sublinearly with the time t), the model (1.1) has attracted much attention in physics, biology and finance etc. It has been successfully applied to many important research fields, e.g., subsurface flow [17, 34], thermal diffusion in media with fractal geometry [35], transport in column experiments [18] and highly heterogeneous aquifer [1]. The classical diffusion model (i.e., $\alpha = 1$) represents the most popular mathematical model to describe transport phenomena found in the nature.

In this work, the concerned inverse problem of the model (1.1) is to recover the unknown diffusion coefficient $q^\dagger$ from a noisy terminal observation z^δ:

$$z^\delta(x) = u(q^\dagger)(x,T) + \xi(x), \quad x\in\Omega,$$

where the exact data $z^\dagger := u(q^\dagger)(T)$ denotes the solution of problem (1.1) (corresponding to the exact potential $q^\dagger$) and ξ denotes the pointwise measurement noise. The accuracy of the data z^δ is measured by the noise level $\delta = \|u(q^\dagger)(T) - z^\delta\|_{L^2(\Omega)}$. The admissible set $\mathcal{A}$ is defined by $\mathcal{A} = \{q \in H^1(\Omega) : c_0 \le q(x) \le c_1 \text{ a.e. in } \Omega\}$, with $0 < c_0 < c_1 < \infty$. Due to the ill-posedness and highly nonlinearity, the numerical recovery of the diffusion coefficient is challenging.

The study of diffusion coefficient identification in anomalous diffusion has a notable history, dating back to at least the work [10]. In the one-dimensional case, Cheng et al. [10] proved the uniqueness for determining a spatially-dependent diffusion coefficient and the fractional order α, given the lateral Cauchy data and the Dirac delta function as initial condition. The proof makes use of the Laplace transform and Gel'fand–Leviton theory for inverse Sturm–Liouville problems. Zhang [42] proved the unique recovery of a time dependent diffusion coefficient from lateral Cauchy data. The study with terminal data, despite being highly practical, is still not well understood. Indeed, analyzing terminal data in the standard parabolic case where $\alpha = 1$ [3] has long been a challenging task, and has only been scarcely studied [4, 19]. In the one-dimensional case, Isakov [19] analyzed this inverse problem under some special assumptions on the boundary data. More recently, for normal diffusion with a zero source ($f = 0$), Triki [39] established a Lipschitz stability result

$$\|q_1 - q_2\|_{L^2(\Omega)} \le c_T \|(u(q_1) - u(q_2))(T)\|_{H^2(\Omega)}. \tag{1.2}$$

This result holds for sufficiently large T under certain positivity conditions on the initial data u_0, achieved by using careful spectral perturbation estimates (see Remark 1 for further details). However, this spectral perturbation argument in [39] is not directly applicable in the error analysis of fully discrete schemes. Moreover, this analysis relies heavily on the exponential decay property of the parabolic problem's solution operator, making it unsuitable for the subdiffusion model with $\alpha \in (0, 1)$, where the solution operator decays only polynomially [21, p. 196]. In this paper, we aim to address this gap by proposing a novel conditional stability result for the inverse problem. This result leverages a weighted energy estimate and is applicable for both normal diffusion ($\alpha = 1$) and subdiffusion ($\alpha \in (0, 1)$). Additionally, the strategy is amenable with the numerical analysis of discrete schemes.

Numerically, Li et al [30, 31] presented the first numerical recovery of the diffusion coefficient in the fractional case, including both smooth and nonsmooth data, but without an error analysis of the discrete scheme. Note that in practical computation, the regularized formulation is often discretized with the Galerkin FEM. The convergence of discrete approximations as the discretization parameters tend to zero has been analyzed; See [26, 41] for the standard parabolic case. However, deriving a convergence rate is far more challenging, due to the high degree of nonlinearity of the forward map and strong nonconvexity of the regularized functional. Thus there have been only very few error bounds on discrete approximations in the existing literature [22, 25, 40, 43], even though such a priori estimates can provide useful guidelines for the proper choice of discretization parameters. The analysis techniques in all these existing works require that the observational data is available over a time interval (for $\alpha = 1$) or whole space-time interval (for $\alpha \in (0, 1)$). The main technical tools include conditional stability and smoothing properties of solution operator. The current work aims to significantly extend the argument to cover terminal data, which is more practical.

In this work, we develop a numerical procedure for recovering the diffusion coefficient q using a regularized formulation [15, 20] and establish error bounds on the approximation. We make two new contributions in the work. First, under mild conditions on the problem data (u_0, f, T, q_1, q_2 and Ω), we prove a Hölder type conditional stability in Theorem 1:

$$\|q_1 - q_2\|_{L^2(\Omega)} \le c\|\nabla\big(u(q_1) - u(q_2)\big)(T)\|^{\frac{1}{2}}_{L^2(\Omega)}. \tag{1.3}$$

The overall proof relies only on a weighted energy argument (inspired by Bonito et al. [8]), some nonstandard smoothing properties and asymptotics of solution operators [21], and maximum-norm resolvent estimates [6, 7, 37]. To the best of our knowledge, this is the first stability result addressing the inverse problem for both the integer-order and fractional-order cases. Moreover, the analysis strategy also plays an essential role in the error analysis of the inversion scheme.

Second, we employ the standard output least square formulation to identify the diffusion coefficient. Motivated by the conditional stability analysis, an $H^1(\Omega)$-seminorm penalty is used in the formulation. Numerically, both Tikhonov functional and PDE constraint, i.e., problem (1.1), are discretized using the standard Galerkin finite element method (FEM) with continuous piecewise linear finite elements in space and backward Euler convolution quadrature in time; see e.g., [33] and [24, Chapter 4]. In particular, let h be the spatial mesh size, τ the time step size, γ the regularization parameter, and q_h^* denote the numerical reconstruction of the diffusion coefficient $q^\dagger$. We derive the following error estimate for the numerical approximation in Theorem 2:

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \le c(h\gamma^{-1}\eta^2 + \min(1, h^{-1}\eta)\gamma^{-\frac{1}{2}}\eta + h^2 + \tau)^{\frac{1}{2}}.$$

with $\eta = \delta + h^2 + \tau + \gamma^{\frac{1}{2}}$. The technical proof heavily relies on the conditional stability estimate (1.3) and several new smoothing properties and asymptotics of semi- and fully-discrete solution operators. Note that the analysis does not involve standard source type conditions, as is commonly done for nonlinear inverse problems [15, 20]. The derived $L^2(\Omega)$ error bound is given explicitly in terms of the discretization parameters h and τ, the noise level δ and the regularization parameter γ when the fixed value T is relatively large. Compared with existing works [22, 25, 40, 43], the present work requires overcoming new technical challenges. The key techniques for deriving conditional stability (1.3) include decay estimate in Lemma 3 and decay Lipschitz stability in Lemma 4 of $\partial_t^\alpha u$. Moreover, in the error analysis of the fully discrete scheme, the crucial discrete decay Lipschitz stability estimate does not follow as the continuous case, e.g., maximum-norm resolvent estimates. We develop innovative techniques to overcome the challenge, e.g., the decay estimates of the semi- and fully discrete solution operators (and their derivatives). The argument is applicable to both normal diffusion ($\alpha = 1$) and subdiffusion ($0 < \alpha < 1$), thereby significantly broadening the scope of existing works. Numerical experiments indicate that the conditional stability does not hold for small T, cf. Table 1, confirming the sharpness of the theoretical result.

The rest of the paper is organized as follows. In Sect. 2, we show the conditional stability of the inverse problem. Then in Sect. 3, we describe the numerical reconstruction scheme, and provide a complete error analysis for discrete approximations. Finally, in Sect. 4, we present one- and two-dimensional numerical experiments to complement the theoretical results. Last, we give some useful notations. For any $m \ge 0$ and $p \ge 1$, we denote by $W^{m,p}(\Omega)$ and $W_0^{m,p}(\Omega)$ the standard Sobolev spaces of order m, equipped with the norm $\|\cdot\|_{W^{m,p}(\Omega)}$ [2]. We denote by $W^{-m,p'}(\Omega)$ the dual space of $W_0^{m,p}(\Omega)$, with p' being the conjugate exponent of p. Further, we write $H^m(\Omega)$ and $H_0^m(\Omega)$ with the norm $\|\cdot\|_{H^m(\Omega)}$ if $p = 2$ and write $L^p(\Omega)$ with the norm $\|\cdot\|_{L^p(\Omega)}$ if $m = 0$. The notation $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. We also make use of Bochner spaces. For a Banach space B, the space $W^{m,p}(0, T; B)$ is defined as

$$W^{m,p}(0, T; B) = \{v : v(t) \in B \text{ a.e. } t \in (0, T) \text{ with } \|v\|_{W^{m,p}(0,T;B)} < \infty\},$$

equipped with the norm

$$\|v\|_{W^{m,p}(0,T;B)} := \left(\sum_{k=0}^{m} \int_0^T \|\partial_t^k u(t)\|_B^p \, \mathrm{d}t \right)^{\frac{1}{p}}.$$

The space $L^\infty(0, T; B)$ is defined analogously. Throughout the paper, we use c, with or without a subscript, to denote a generic constant that may vary from one occurrence to another, but is always independent of the discretization parameters h and τ, the noise level δ, the regularization parameter γ and the terminal time T.

2 Conditional stability

In this section, we establish a novel conditional stability estimate for the inverse conductivity problem with the terminal data. Let the operator $A(q)$ be the realization of $-\nabla\cdot(q\nabla\cdot)$ with a zero Dirichlet boundary condition, with its domain $\mathrm{Dom}(A(q)) := H^2(\Omega) \cap H_0^1(\Omega)$. Then the solution $u(q)$ to problem (1.1) is given by

$$u(q) = F(t; q)u_0 + \int_0^t E(s; q) f \, \mathrm{d}s, \tag{2.1}$$

where the solution operators $F(t; q)$ and $E(t; q)$ are given respectively by [21, Section 6.2.1]

$$F(t; q) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\sigma}} e^{zt} z^{\alpha-1} (z^\alpha + A(q))^{-1} \, \mathrm{d}z \tag{2.2}$$

$$E(t; q) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\sigma}} e^{zt} (z^\alpha + A(q))^{-1} \, \mathrm{d}z, \tag{2.3}$$

with the contour $\Gamma_{\theta,\sigma} \subset \mathbb{C}$ (oriented with an increasing imaginary part) given by $\Gamma_{\theta,\sigma} = \{z \in \mathbb{C} : |z| = \sigma, |\arg(z)| \le \theta\} \cup \{z \in \mathbb{C} : z = \rho e^{\pm i\theta}, \rho \ge \sigma\}$. Throughout, we fix $\theta \in (\frac{\pi}{2}, \pi)$ so that $z^\alpha \in \Sigma_{\alpha\theta} \subset \Sigma_\theta$ for all $z \in \Sigma_\theta := \{z \in \mathbb{C}\backslash\{0\} : |\arg(z)| \le \theta\}$. Moreover, the following identity holds: $\partial_t F(t; q) = -A(q)E(t; q)$ [21, Lemma 6.2].

The next lemma gives useful smoothing properties of the operators $F(t; q)$ and $E(t; q)$. For any $s \in \mathbb{R}$, the notation $A(q)^s$ denotes the fractional power of $A(q)$, defined by spectral decomposition. The cases $s = 0, 1$ are known [21, Theorem 6.4], and the case $0 < s < 1$ follows from the standard interpolation theory [32, Proposition 2.3].

Lemma 1 *For any $q \in \mathcal{A}$, and any $s \in [0, 1]$, there exists $c > 0$ independent of q and t such that*

$$t^{s\alpha} \|A(q)^s F(t; q)\|_{L^2(\Omega)\to L^2(\Omega)} + t^{1-(1-s)\alpha} \|A(q)^s E(t; q)\|_{L^2(\Omega)\to L^2(\Omega)} \le c.$$

The following maximum norm resolvent estimate will be used extensively. See [37, Theorem 1], [6, Theorem 1.1] and [7, Theorem 2.1] for the proof. This estimate involves the elliptic regularity pickup in $W^{1,\infty}(\Omega)$, for which $q \in \mathcal{A}$ is insufficient, and the condition $q \in \mathcal{A} \cap W^{1,\infty}(\Omega)$ is sufficient. Thus, the uniform bound depends on q via its $W^{1,\infty}(\Omega)$ norm.

Lemma 2 *For $q \in W^{1,\infty}(\Omega) \cap \mathcal{A}$ with $\|q\|_{W^{1,\infty}(\Omega)} \le c_q$, there exists $c > 0$ dependent of c_q such that for any $z \in \Sigma_\theta$ and $\theta \in (\frac{\pi}{2}, \pi)$*

$$|z| \| \big(z + A(q)\big)^{-1} \|_{L^\infty(\Omega) \to L^\infty(\Omega)} + |z|^{\frac{1}{2}} \| \big(z + A(q)\big)^{-1} \|_{L^\infty(\Omega) \to W^{1,\infty}(\Omega)} \le c. \quad (2.4)$$

Using Lemma 2, we can derive an a priori estimate for the time-fractional derivative $\partial_t^\alpha u$.

Lemma 3 *Let $u_0 \in W^{2,\infty}(\Omega) \cap H_0^1(\Omega)$, $f \in L^\infty(\Omega)$, and $q \in W^{1,\infty}(\Omega) \cap \mathcal{A}$, and let $u(q)$ be the solution to problem (1.1). Then there exists $c > 0$, independent of t and q, such that*

$$\|\partial_t^\alpha u(q)\|_{W^{1,\infty}(\Omega)} \le c t^{-\frac{\alpha}{2}}.$$

Proof By Lemma 2 and letting $\sigma = t^{-1}$ in the contour $\Gamma_{\theta,\sigma}$, we have

$$\begin{aligned} \|F(t;q)\|_{L^\infty(\Omega) \to W^{1,\infty}(\Omega)} &\le c \int_{\Gamma_{\theta,\sigma}} |e^{zt}| |z|^{\alpha-1} \| \big(z^\alpha + A(q)\big)^{-1} \|_{L^\infty(\Omega) \to W^{1,\infty}(\Omega)} \, |\mathrm{d}z| \\ &\le c \int_{\Gamma_{\theta,\sigma}} |e^{zt}| |z|^{\frac{\alpha}{2}-1} \, |\mathrm{d}z| \le c t^{-\frac{\alpha}{2}}. \end{aligned} \quad (2.5)$$

Let $w(t) = \partial_t^\alpha u(q)$. Then the function $w(t)$ satisfies $w(0) = f - A(q)u_0 \in L^\infty(\Omega)$ and for every $t \in (0, T)$,

$$\partial_t^\alpha w + A(q) w = 0, \quad \text{in } H^{-1}(\Omega).$$

It follows from the representation (2.1) that $\partial_t^\alpha u(q) = w(t) = F(t;q)\big(f - A(q)u_0\big)$. This, the estimate (2.5) and the assumption on q, u_0 and f lead to

$$\|\partial_t^\alpha u(q)\|_{W^{1,\infty}(\Omega)} \le c \|F(t;q)\|_{L^\infty(\Omega) \to W^{1,\infty}(\Omega)} \|f - A(q)u_0\|_{L^\infty(\Omega)} \le c t^{-\frac{\alpha}{2}}.$$

This completes the proof of the lemma. □

Next we provide a crucial Lipschitz stability estimate of the time (fractional) derivative with respect to the $L^2(\Omega)$ norm of the diffusion coefficient.

Lemma 4 *Let $u_0 \in W^{2,\infty}(\Omega) \cap H_0^1(\Omega)$ and $f \in L^\infty(\Omega)$, and let $u(q_1)$ and $u(q_2)$ be the solutions of problem (1.1) with $q_1 \in \mathcal{A}$ and $q_2 \in W^{1,\infty}(\Omega) \cap \mathcal{A}$, respectively. Then for any small $\epsilon > 0$, there exists $c = c(q_2) > 0$, independent of q_1, t and T, such that*

$$\|\partial_t^\alpha (u(q_1) - u(q_2))\|_{L^2(\Omega)} \le c \max(t^{-\alpha}, t^{-\frac{\alpha}{2}(1-2\epsilon)}) \|q_1 - q_2\|_{L^2(\Omega)},$$

Proof Let $u_i = u(q_i)$ and $w(t) = \partial_t^\alpha (u_1 - u_2)$. Then w satisfies

$$\begin{cases} \partial_t^\alpha w - \nabla \cdot (q_1 \nabla w) = \nabla \cdot \big((q_1 - q_2)\nabla \partial_t^\alpha u_2\big), & \text{in } \Omega \times (0, T), \\ w = 0, & \text{on } \partial\Omega \times (0, T), \\ w(0) = \nabla \cdot \big((q_1 - q_2)\nabla u_0\big), & \text{in } \Omega. \end{cases}$$

The solution representation (2.1) leads to

$$w(t) = F(t; q_1)\big(\nabla \cdot \big((q_1 - q_2)\nabla u_0\big)\big) + \int_0^t E(t-s; q_1)\big(\nabla \cdot \big((q_1 - q_2)\nabla \partial_s^\alpha u_2(s)\big)\big) \, \mathrm{d}s := \mathrm{I}_1 + \mathrm{I}_2.$$

Since $q_1 \in \mathcal{A}$, there exist constants c and c' independent of q_1 such that

$$c\|A(q_1)^{\frac{1}{2}} v\|_{L^2(\Omega)} \le \|\nabla v\|_{L^2(\Omega)} \le c'\|A(q_1)^{\frac{1}{2}} v\|_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega). \tag{2.6}$$

Then the self-adjointness of the operator $F(t; q_1)$, integration by parts (using the fact that $F(t; q_1)\varphi \in H_0^1(\Omega) \cap H^2(\Omega)$ for $\varphi \in L^2(\Omega)$), and the commutativity of $F(t; q_1)$ and $A(q_1)$ imply

$$\begin{aligned} \|\mathrm{I}_1\|_{L^2(\Omega)} &= \sup_{\|\varphi\|_{L^2(\Omega)}=1} \big(F(t; q_1)\nabla \cdot \big((q_1 - q_2)\nabla u_0, \varphi\big) \\ &= \sup_{\|\varphi\|_{L^2(\Omega)}=1} \big(\nabla \cdot \big((q_1 - q_2)\nabla u_0, F(t; q_1)\varphi\big) \\ &= \sup_{\|\varphi\|_{L^2(\Omega)}=1} \big((q_2 - q_1)\nabla u_0, \nabla F(t; q_1)\varphi\big) \\ &= \sup_{\|\varphi\|_{L^2(\Omega)}=1} \big((q_2 - q_1)\nabla u_0, \nabla\big[A(q_1)^{\frac{1}{2}} F(t; q_1)\big] A(q_1)^{-\frac{1}{2}}\varphi\big). \end{aligned}$$

Then Lemma 1, the equivalence property (2.6), and the boundedness of the operator $A(q_1)^{-\frac{1}{2}}$ in $L^2(\Omega)$ (uniform in q_1) imply

$$\begin{aligned} \|\mathrm{I}_1\|_{L^2(\Omega)} \le{}& c\|q_1 - q_2\|_{L^2(\Omega)} \|\nabla u_0\|_{L^\infty(\Omega)} \|A(q_1)F(t; q_1)\|_{L^2(\Omega)\to L^2(\Omega)} \\ &\times \sup_{\|\varphi\|_{L^2(\Omega)}=1} \|A(q_1)^{-\frac{1}{2}}\varphi\|_{L^2(\Omega)} \le ct^{-\alpha}\|q_1 - q_2\|_{L^2(\Omega)}. \end{aligned}$$

Similarly, by the boundedness of the operator $A(q_1)^{-\frac{1}{2}+\epsilon}$ in $L^2(\Omega)$ (uniform in q_1) for small $\epsilon > 0$ and Lemmas 1–3, we have

$$\begin{aligned}
&\|E(t-s;q_1)\big(\nabla\cdot\big((q_1-q_2)\nabla\partial_s^\alpha u_2(s)\big)\big)\|_{L^2(\Omega)}\\
&\quad= \sup_{\|\varphi\|_{L^2(\Omega)}=1} \big((q_2-q_1)\nabla\partial_s^\alpha u_2(s), \nabla A(q_1)^{\frac{1}{2}-\epsilon}E(t-s;q_1)A(q_1)^{-\frac{1}{2}+\epsilon}\varphi\big)\\
&\quad\le c\|q_1-q_2\|_{L^2(\Omega)}\|\nabla\partial_s^\alpha u_2(s)\|_{L^\infty(\Omega)}\|A(q_1)^{1-\epsilon}E(t-s;q_1)\|_{L^2(\Omega)\to L^2(\Omega)}\\
&\qquad\times \sup_{\|\varphi\|_{L^2(\Omega)}=1}\|A(q_1)^{-\frac{1}{2}+\epsilon}\varphi\|_{L^2(\Omega)} \le c(t-s)^{\epsilon\alpha-1}s^{-\frac{\alpha}{2}}\|q_1-q_2\|_{L^2(\Omega)}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\|\mathrm{I}_2\|_{L^2(\Omega)} &\le \int_0^t \|E(t-s;q_1)\big(\nabla\cdot((q_1-q_2)\nabla\partial_s^\alpha u_2(s))\big)\|_{L^2(\Omega)}\,\mathrm{d}s\\
&\le c\|q_1-q_2\|_{L^2(\Omega)}\int_0^t (t-s)^{\epsilon\alpha-1}s^{-\frac{\alpha}{2}}\,\mathrm{d}s \le ct^{-\frac{\alpha}{2}(1-2\epsilon)}\|q_1-q_2\|_{L^2(\Omega)}.
\end{aligned}$$

The bounds on I_1 and I_2 and the triangle inequality complete the proof of the lemma. □

Next we give a novel conditional stability estimate. First we state the standing assumption.

Assumption 1 $q^\dagger \in W^{1,\infty}(\Omega)\cap\mathcal{A}$, $u_0\in W^{2,\infty}(\Omega)\cap H_0^1(\Omega)$, and $f\in L^\infty(\Omega)$.

Under Assumption 1, the following regularity results hold. Let $p>\max(d,2)$. Then for any $\theta<\frac{1}{2}-\frac{d}{2p}$ and $r>\frac{1}{\alpha\theta}$, the solution $u=u(q^\dagger)$ to problem (1.1) satisfies [25, (2.5)–(2.6)]

(i) $u\in W^{\alpha\theta,r}(0,T;W^{2(1-\theta),p})\hookrightarrow L^\infty(0,T;W^{1,\infty}(\Omega))$;
(ii) $\|u(t)\|_{H^2(\Omega)}+\|\partial_t^\alpha u(t)\|_{L^2(\Omega)}+t^{1-\alpha}\|\partial_t u(t)\|_{L^2(\Omega)}+t\|\partial_t u(t)\|_{L^2(\Omega)}\le c$, for a.e. $t\in(0,T]$.

Additionally, we assume that the following positivity condition holds:

$$(q^\dagger|\nabla u(q^\dagger)|^2+(f-\partial_t^\alpha u(q^\dagger))u(q^\dagger))(T)\ge c>0. \tag{2.7}$$

This condition was proved in [22] for parabolic equations (i.e., $\alpha=1$) and in [25] for time-fractional diffusion (i.e., $0<\alpha<1$). For example, if Ω is a $C^{2,\mu}$ domain with $\mu\in(0,1)$, $q^\dagger\in C^{1,\mu}(\Omega)\cap\mathcal{A}$, $f\in C^\mu(\Omega)$ with $f\ge c_f>0$ and $u_0\in C^{2,\mu}(\Omega)\cap H_0^1(\Omega)$ with $u_0\ge0$ in Ω, and $f+\nabla\cdot(q^\dagger\nabla u_0)\le0$ in Ω, then condition (2.7) holds. See [22, Section 4.3] and [25, Proposition 3.5] for detailed discussions.

Now we give a Hölder type conditional stability estimate for the inverse problem. The estimate is conditional since the coefficients are required to have extra regularity. To the best of our knowledge, this appears to be the first result of the kind for the time-fractional model (1.1) with a terminal observation. The analysis strategy will also guide the error analysis of the fully discrete scheme in Sect. 3 below.

Theorem 1 *Let Assumption 1 hold, and $q \in \mathcal{A}$ with $\|\nabla q\|_{L^2(\Omega)} \le c$. Then for small $\epsilon > 0$, the following conditional stability estimate holds*

$$\begin{aligned}&\int_\Omega \Big(\frac{q^\dagger - q}{q^\dagger}\Big)^2 (q^\dagger |\nabla u(q^\dagger)|^2 + (f - \partial_t^\alpha u(q^\dagger))u(q^\dagger))(T)\,\mathrm{d}x\\ &\quad \le c\|\nabla\big(u(q) - u(q^\dagger)\big)(T)\|_{L^2(\Omega)} + c\max(T^{-\alpha}, T^{-\frac{\alpha}{2}(1-2\epsilon)})\|q^\dagger - q\|^2_{L^2(\Omega)},\end{aligned}$$

with $c > 0$ independent of q and T. Moreover, if condition (2.7) holds, then there exist $T_0 > 0$ and $c > 0$, independent of q, such that for all $T \ge T_0$,

$$\|q - q^\dagger\|_{L^2(\Omega)} \le c\|\nabla\big(u(q) - u(q^\dagger)\big)(T)\|^{\frac{1}{2}}_{L^2(\Omega)}.$$

Proof Let $u^\dagger = u(q^\dagger)$ and $u = u(q)$. Then by the weak formulations of $u^\dagger$ and u, there holds for any $\varphi \in H_0^1(\Omega)$,

$$\big((q^\dagger - q)\nabla u^\dagger(T), \nabla\varphi\big) = -\big(q\nabla(u^\dagger - u)(T), \nabla\varphi\big) - \big(\partial_t^\alpha(u^\dagger - u)(T), \varphi\big) =: \mathrm{I}_1 + \mathrm{I}_2.$$

Let $\varphi \equiv \frac{q^\dagger - q}{q^\dagger} u^\dagger(T) \in H_0^1(\Omega)$. We claim the following the crucial identity

$$\big((q^\dagger - q)\nabla u^\dagger(T), \nabla\varphi\big) = \frac{1}{2}\int_\Omega \left(\frac{q^\dagger - q}{q^\dagger}\right)^2 \big(q^\dagger|\nabla u^\dagger|^2 + (f - \partial_t^\alpha u^\dagger)u^\dagger\big)(T)\,\mathrm{d}x. \tag{2.8}$$

Indeed, following the derivation in [8, 22], direct computation with integration by parts and product rule leads to

$$\begin{aligned}&\big((q^\dagger - q)\nabla u^\dagger(T), \nabla\varphi\big)\\ &= \int_\Omega \left(\Big|\frac{q^\dagger - q}{q^\dagger}\Big|^2 q^\dagger|\nabla u^\dagger|^2 + \frac{q^\dagger - q}{q^\dagger}\nabla\frac{q^\dagger - q}{q^\dagger}\cdot\big(q^\dagger\nabla u^\dagger\big)u^\dagger\right)(T)\,\mathrm{d}x\\ &= \int_\Omega \left(\Big|\frac{q^\dagger - q}{q^\dagger}\Big|^2 q^\dagger|\nabla u^\dagger|^2 + \frac{1}{2}\nabla\Big(\frac{q^\dagger - q}{q^\dagger}\Big)^2\cdot\big(q^\dagger\nabla u^\dagger\big)u^\dagger\big)\right)(T)\,\mathrm{d}x\\ &= \frac{1}{2}\int_\Omega \left(\Big|\frac{q^\dagger - q}{q^\dagger}\Big|^2\big(q^\dagger|\nabla u^\dagger|^2 + \big(-\nabla\cdot(q^\dagger\nabla u^\dagger)u^\dagger\big)\right)(T)\,\mathrm{d}x.\end{aligned}$$

Then using the identity $-\nabla\cdot(q^\dagger\nabla u^\dagger) = f - \partial_t^\alpha u^\dagger$ in $\Omega \times \{t = T\}$ gives the claim (2.8). Meanwhile, direct computation yields

$$\nabla\varphi = \nabla\Big(\frac{q^\dagger - q}{q^\dagger}\Big)u^\dagger(T) + \Big(\frac{q^\dagger - q}{q^\dagger}\Big)\nabla u^\dagger(T).$$

Now Assumption 1 implies the a priori estimate

$$\|u^\dagger(T)\|_{L^\infty(\Omega)} + \|\nabla u^\dagger(T)\|_{L^2(\Omega)} \le c_T.$$

This, the box constraint on $q^\dagger, q \in \mathcal{A}$ and the a priori bound $\|\nabla q\|_{L^2(\Omega)} \le c$ yield

$$\|\varphi\|_{L^2(\Omega)} \le c\|q - q^\dagger\|_{L^2(\Omega)} \quad \text{and} \quad \|\nabla\varphi\|_{L^2(\Omega)} \le c.$$

Hence, by Lemma 4 and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
|\mathrm{I}_1| &\le c_T \|\nabla(u - u^\dagger)(T)\|_{L^2(\Omega)}, \\
|\mathrm{I}_2| &\le c\|\partial_t^\alpha(u - u^\dagger)(T)\|_{L^2(\Omega)}\|q - q^\dagger\|_{L^2(\Omega)} \\
&\le c\max(T^{-\alpha}, T^{-\frac{\alpha}{2}(1-2\epsilon)})\|q - q^\dagger\|^2_{L^2(\Omega)}.
\end{aligned}$$

Note that the constant c_T depends on $u^\dagger(T)$ (and thus also on T), but in view of Lemma 4, the constant c is independent of T. Moreover, under condition (2.7), by applying the box constraint on $q, q^\dagger \in \mathcal{A}$, we derive

$$\|q - q^\dagger\|^2_{L^2(\Omega)} \le c_T\|\nabla\big(u - u^\dagger\big)(T)\|_{L^2(\Omega)} + c\max(T^{-\alpha}, T^{-\frac{\alpha}{2}(1-2\epsilon)})\|q - q^\dagger\|^2_{L^2(\Omega)}.$$

Let T_0 be sufficiently large such that $c\max(T^{-\alpha}, T^{-\frac{\alpha}{2}(1-2\epsilon)}) \le \frac{1}{2}$. Then for any $T \ge T_0$, the desired estimate follows. □

Remark 1 Theorem 1 extends several existing works. The weighted stability results of the type for the observation over a space-time domain were implicitly obtained in [22, 25]. The only result for the terminal data case was obtained by Triki [39, Theorem 1.1] for the standard parabolic problem, who proved the following Lipschitz stability for a large T: for $q, q^\dagger \in C^1(\overline{\Omega})$, there holds

$$\|q - q^\dagger\|_{L^2(\Omega)} \le c\|u(q^\dagger)(T) - u(q)(T)\|_{H^2(\Omega)},$$

where the constant c depends on the terminal time T (exponentially) and the domain Ω. This result was shown for the case $f \equiv 0$ and u_0 satisfying a mild positivity condition, and the proof relies on the decay estimate on $\partial_t u$, which itself was proved using refined spectral perturbation estimates. Theorem 1 provides a novel Hölder stability estimate using an energy estimate and can be adapted to the error analysis of numerical approximations in Sect. 3.

3 Numerical approximation and error analysis

Now we develop a numerical procedure based on the regularized output least-squares formulation, and discretize the regularized problem using backward Euler convolution quadrature (BECQ) in time and Galerkin FEM with continuous piecewise linear elements in space. Furthermore, we provide a complete error analysis of the fully discrete scheme.

3.1 Regularized problem and numerical approximation

To identify the diffusion coefficient q, we employ the standard Tikhonov regularization with an $H^1(\Omega)$ seminorm penalty [15, 20], which gives the following minimization problem:

$$\min_{q\in\mathcal{A}} J_\gamma(q) = \frac{1}{2}\|u(q)(T) - z^\delta\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|\nabla q\|^2_{L^2(\Omega)}, \tag{3.1}$$

where $\gamma > 0$ is the regularization parameter and $u(t) \equiv u(q)(t) \in H^1_0(\Omega)$ with $u(0) = u_0$ satisfies

$$(\partial_t^\alpha u(t), \varphi) + (q\nabla u(t), \nabla\varphi) = (f, \varphi), \quad \forall\varphi \in H^1_0(\Omega), \text{ a.e. } t \in (0, T). \tag{3.2}$$

By a standard argument [15, 20], it can be proved that problem (3.1)–(3.2) has at least one global minimizer q_γ^δ, which is continuous with respect to the perturbations in the data z^δ. Moreover, as the noise level $\delta \to 0^+$, the sequence $\{q_\gamma^\delta\}_{\delta>0}$ of minimizers contains a subsequence that converges to the exact coefficient $q^\dagger$ in $H^1(\Omega)$ if γ is chosen properly.

In practice, one needs to discretize the regularized formulation (3.1)–(3.2) suitably. For time discretization, we divide the interval $[0, T]$ uniformly into N subintervals, with a time step size $\tau := T/N$ and grid $t_n := n\tau$, $n = 0, 1 \ldots N$. To approximate the fractional derivative $\partial_t^\alpha v(t_n)$, we employ backward Euler convolution quadrature (BECQ) defined by

$$\bar\partial_\tau^\alpha v^n := \tau^{-\alpha}\sum_{j=0}^{n} b_j^{(\alpha)}(v^{n-j} - v^0), \quad \text{with } v^j = v(t_j),$$

where the weights $b_j^{(\alpha)}$ are generated by the power series expansion $(1-\zeta)^\alpha = \sum_{j=0}^\infty b_j^{(\alpha)}\zeta^j$. The weights $b_j^{(\alpha)}$ are given by $b_j^{(\alpha)} = (-1)^j\alpha(\alpha-1)\cdots(\alpha-j+1)/j!$, with $b_0^{(\alpha)} = 1$ and $b_j^{(\alpha)} < 0$ for $j \geq 1$. When $\alpha = 1$, it reduces to the standard backward Euler scheme.

For the spatial discretization, we employ the standard Galerkin FEM. Let $h \in (0, h_0]$ for some $h_0 > 0$ and $\mathcal{T}_h := \cup\{T_j\}_{j=1}^{N_h}$ be a shape regular quasi-uniform simplicial triangulation of the domain Ω into mutually disjoint open face-to-face subdomains T_j, such that $\Omega_h := \mathrm{Int}(\cup_j\{\bar T_j\}) \subset \Omega$ with all the boundary vertices of the domain Ω_h lying on $\partial\Omega$ and $\mathrm{dist}(x, \partial\Omega) \leq ch^2$ for $x \in \partial\Omega_h$ [28, Section 5.3]. On the triangulation $\mathcal{T}_h$, we define the space V_h of continuous piecewise linear finite element functions by

$$V_h := \{v_h \in H^1(\Omega_h) : v_h|_T \text{ is a linear polynomial, } \forall T \in \mathcal{T}_h\}.$$

Note that the functions in V_h can be naturally extended to the entire domain Ω by linear polynomials, and we denote the space of extended functions also by V_h. Moreover, we define the space X_h (that vanish outside Ω_h) by

$$X_h := \{v_h \in H_0^1(\Omega_h) : v_h|_T \text{ is a linear polynomial } \forall T \in \mathcal{T}_h \text{ and } v_h|_{\Omega\setminus\Omega_h} = 0\}$$

The spaces V_h and X_h are then used to discretize the diffusion coefficient q and the state u, respectively. Note that if Ω is a convex polygon, then $X_h = V_h \cap H_0^1(\Omega)$. Next we recall several useful estimates. We denote by Π_h the Lagrange nodal interpolation operator on V_h. Since $\mathrm{dist}(x, \partial\Omega) \le ch^2$ for $x \in \partial\Omega_h$, we have the following error estimates:

$$\|v - \Pi_h v\|_{L^2(\Omega)} + h\|\nabla(v - \Pi_h v)\|_{L^2(\Omega)} \le ch^2\|v\|_{H^2(\Omega)}, \quad \forall v \in H^2(\Omega), \tag{3.3}$$

$$\|v - \Pi_h v\|_{L^\infty(\Omega)} + h\|\nabla(v - \Pi_h v)\|_{L^\infty(\Omega)} \le ch\|v\|_{W^{1,\infty}(\Omega)}, \quad \forall v \in W^{1,\infty}(\Omega). \tag{3.4}$$

For the proof, see [9, Theorem 4.4.20] for a convex polyhedral domain and Lemma 11 for a convex domain with a curved boundary. Moreover, we define the standard $L^2(\Omega)$-projection operator $P_h : L^2(\Omega) \to X_h$ by

$$(P_h v, \varphi_h) = (v, \varphi_h), \quad \forall v \in L^2(\Omega), \varphi_h \in X_h,$$

Then for $1 \le p \le \infty$ and $s = 0, 1, 2$, $k = 0, 1$ with $k \le s$ [6, 13]:

$$\|v - P_h v\|_{W^{k,p}(\Omega)} \le Ch^{s-k}\|v\|_{W^{s,p}(\Omega)}, \quad \forall v \in W^{s,p}(\Omega) \cap H_0^1(\Omega). \tag{3.5}$$

Now we can formulate a fully discrete scheme for the regularized problem (3.1)–(3.2) as

$$\min_{q_h \in \mathcal{A}_h} J_{\gamma,h,\tau}(q_h) = \frac{1}{2}\|U_h^N(q_h) - z^\delta\|_{L^2(\Omega)}^2 + \frac{\gamma}{2}\|\nabla q_h\|_{L^2(\Omega)}^2, \tag{3.6}$$

with $\mathcal{A}_h = \mathcal{A} \cap V_h$, where $U_h^n \equiv U_h^n(q_h) \in X_h$ satisfies $U_h^0 = P_h u_0$ and

$$(\bar\partial_\tau^\alpha U_h^n, \varphi_h) + (q_h \nabla U_h^n, \nabla\varphi_h) = (f, \varphi_h), \quad \forall \varphi_h \in X_h,\ n = 1, 2, \ldots, N. \tag{3.7}$$

Note that problem (3.6)–(3.7) is finite-dimensional, and using the norm equivalence in a finite-dimensional space, the continuity of the objective functional $J_{\gamma,h,\tau}(q_h)$, and a standard compactness argument in calculus of variation, one can prove that the discrete problem (3.6)–(3.7) is well-posed: there exists at least one global minimizer $q_h^* \in \mathcal{A}_h$, and it depends continuously on the data. Further, as the discretization parameters h and τ tend to zero, the numerical approximation q_h^* converges to the regularized solution to problem (3.1)–(3.2). We aim to establish a bound on the error $q_h^* - q^\dagger$ in terms of the noise level δ, discretization parameters h and τ and regularization parameter γ. For the error analysis, we need the following assumption on the problem data.

Assumption 2 $q^\dagger \in H^2(\Omega) \cap W^{1,\infty}(\Omega) \cap \mathcal{A}$, $u_0 \in W^{2,\infty}(\Omega) \cap H_0^1(\Omega)$, and $f \in L^\infty(\Omega)$.

Now we give the main result in this section, i.e., a weighted $L^2(\Omega)$ error bound on the approximation q_h^*. The proof heavily relies on some technical estimates, whose proofs are deferred to Sect. 3.2.

Theorem 2 *Let Assumption 2 hold, and* $\{(q_h^*, u_h^n(q_h^*))\}_{n=0}^N$ *be the solutions of problem* (3.6)–(3.7). *Then with* $\eta_T = T^{\alpha-1}\tau + \max(1, T^{-\alpha})h^2 + \delta + \gamma^{\frac{1}{2}}$*, there holds*

$$\begin{aligned}\int_\Omega \Big(\frac{q^\dagger - q_h^*}{q^\dagger}\Big)^2 &(q^\dagger|\nabla u(q^\dagger)|^2 + (f - \partial_t^\alpha u(q^\dagger))u(q^\dagger))(T)\,\mathrm{d}x\\ &\le c\big(h\gamma^{-1}\eta_T^2 + \min(1, h^{-1}\eta_T)\gamma^{-\frac{1}{2}}\eta_T + h^2\max(T^{-\alpha}, T^{-2\alpha}) + \tau T^{-\alpha-1}\big)\\ &\quad + c\max(T^{-\alpha}, T^{-\frac{\alpha}{2}(1-2\epsilon)}, T^{-\alpha(1-\epsilon)}, T^{-\alpha(2-\epsilon)})\|q^\dagger - q_h^*\|_{L^2(\Omega)}^2,\end{aligned}$$

where the generic constants are independent of h, τ, δ, γ *and* T. *Moreover, under condition* (2.7), *with* $\eta := \tau + h^2 + \delta + \gamma^{\frac{1}{2}}$, *there exits* $T_0 > 0$ *such that for any* $T \ge T_0$,

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)}^2 \le c(h\gamma^{-1}\eta^2 + \min(1, h^{-1}\eta)\gamma^{-\frac{1}{2}}\eta + h^2 + \tau).$$

Proof The proof proceeds similarly to the conditional stability estimate in Theorem 1 and requires several (new) technical estimates in the propositions below. Let $u^\dagger \equiv u(q^\dagger)$. For any test function $\varphi \in H_0^1(\Omega)$, using the weak formulations of $u^\dagger$ and $U_h^N(q_h^*)$, we have

$$\begin{aligned}&\big((q^\dagger - q_h^*)\nabla u^\dagger(T), \nabla\varphi\big)\\ &\quad = \big((q^\dagger - q_h^*)\nabla u^\dagger(T), \nabla(\varphi - P_h\varphi)\big) + \big((q^\dagger - q_h^*)\nabla u^\dagger(T), \nabla P_h\varphi\big)\\ &\quad = \big((q^\dagger - q_h^*)\nabla u^\dagger(T), \nabla(\varphi - P_h\varphi)\big) + (q_h^*\nabla(U_h^N(q_h^*) - u^\dagger(T)), \nabla P_h\varphi\big)\\ &\qquad + \big(q^\dagger\nabla u^\dagger(T) - q_h^*\nabla U_h^N(q_h^*), \nabla P_h\varphi\big)\\ &\quad = -\big(\nabla\cdot((q^\dagger - q_h^*)\nabla u^\dagger(T)), \varphi - P_h\varphi\big) + (q_h^*\nabla(U_h^N(q_h^*) - u^\dagger(T)), \nabla P_h\varphi\big)\\ &\qquad + \big(\bar\partial_\tau^\alpha U_h^N(q_h^*) - \partial_t^\alpha u^\dagger(T), P_h\varphi\big) =: \mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3.\end{aligned}$$

Now we bound the three terms I_i, $i = 1, 2, 3$, separately. Let $\varphi = \frac{q^\dagger - q_h^*}{q^\dagger}u^\dagger(T)$. Then by the box constraint $q^\dagger, q_h^* \in \mathcal{A}$ and Proposition 1 below, φ satisfies

$$\|\varphi\|_{L^2(\Omega)} \le c, \quad \|P_h\varphi\|_{L^2(\Omega)} \le c \quad \text{and} \quad \|\nabla\varphi\|_{L^2(\Omega)} \le c(1 + \|\nabla q_h^*\|_{L^2(\Omega)}). \quad (3.8)$$

By Assumption 2, we have

$$\|\Delta u^\dagger(T)\|_{L^2(\Omega)} + \|\nabla u^\dagger\|_{L^\infty(\Omega)} \le c.$$

Thus, direct computation yields

$$\begin{aligned}&\|\nabla\cdot\big((q^\dagger-q_h^*)\nabla u^\dagger(T)\big)\|_{L^2(\Omega)}\\&\quad\le\|q^\dagger-q_h^*\|_{L^\infty(\Omega)}\|\Delta u^\dagger(T)\|_{L^2(\Omega)}+\|\nabla q^\dagger\|_{L^\infty(\Omega)}\|\nabla u^\dagger(T)\|_{L^2(\Omega)}\\&\qquad+\|\nabla q_h^*\|_{L^2(\Omega)}\|\nabla u^\dagger(T)\|_{L^\infty(\Omega)}\le c(1+\|\nabla q_h^*\|_{L^2(\Omega)}).\end{aligned}$$

This inequality, the estimates in (3.8) and Proposition 1 below imply

$$|\mathrm{I}_1|\le ch(1+\|\nabla q_h^*\|_{L^2(\Omega)})\|\nabla\varphi\|_{L^2(\Omega)}\le ch(1+\|\nabla q_h^*\|_{L^2(\Omega)}^2)\le ch\gamma^{-1}\eta_T^2.$$

Next, by the triangle inequality, the inverse inequality in the space X_h [38, equation (1.12)], Proposition 1 below and the approximation property of P_h in (3.5), we have

$$\begin{aligned}|\mathrm{I}_2|&\le\big(\|\nabla(U_h^N(q_h^*)-P_hu^\dagger(T))\|_{L^2(\Omega)}+\|\nabla(u^\dagger(T)-P_hu^\dagger(T))\|_{L^2(\Omega)}\big)\|\nabla\varphi\|_{L^2(\Omega)}\\&\le c\big(h^{-1}\|U_h^N(q_h^*)-P_hu^\dagger(T)\|_{L^2(\Omega)}+\|\nabla(u^\dagger(T)-P_hu^\dagger(T))\|_{L^2(\Omega)}\big)\|\nabla\varphi\|_{L^2(\Omega)}\\&\le c(h+h^{-1}\eta_T)\gamma^{-\frac12}\eta_T.\end{aligned}$$

Upon setting $\varphi_h\equiv U_h^N(q_h^*)$ in (3.7), the box constraint of $\mathcal{A}_h$, Cauchy–Schwarz inequality and Poincaré inequality, Assumption 2 and Remark 4 below give

$$\|\nabla U_h^N(q_h^*)\|_{L^2(\Omega)}\le c(\|\bar\partial_\tau^\alpha U_h^N(q_h^*)\|_{L^2(\Omega)}+\|f\|_{L^2(\Omega)})\le c_T.$$

Thus, we obtain the a priori estimate $\|\nabla\big(U_h^N(q_h^*)-u^\dagger(T)\big)\|_{L^2(\Omega)}\le c$, and consequently,

$$|\mathrm{I}_2|\le c\min(1,h^{-1}\eta_T)\gamma^{-\frac12}\eta_T.$$

Next, to bound the term I_3, we employ the splitting

$$\begin{aligned}&\big(\bar\partial_\tau^\alpha U_h^N(q_h^*)-\partial_t^\alpha u^\dagger(T),P_h\varphi\big)\\&\quad=\big(\bar\partial_\tau^\alpha U_h^N(q_h^*)-\bar\partial_\tau^\alpha U_h^N(q^\dagger),P_h\varphi\big)+\big(\bar\partial_\tau^\alpha U_h^N(q^\dagger)-\bar\partial_\tau^\alpha U^N(q^\dagger),P_h\varphi\big)\\&\qquad+\big(\bar\partial_\tau^\alpha U^N(q^\dagger)-\partial_t^\alpha u^\dagger(T),P_h\varphi\big)=:\mathrm{I}_3^1+\mathrm{I}_3^2+\mathrm{I}_3^3,\end{aligned}$$

where $U^N(q^\dagger)$ denotes the time semi-discrete solution, cf. (3.17) below. and then bound the three terms separately. It follows from Lemmas 8–9 below and the estimate (3.8) that

$$|\mathrm{I}_3^2|\le ch^2\max(T^{-\alpha},T^{-2\alpha})\quad\text{and}\quad|\mathrm{I}_3^3|\le c\tau T^{-\alpha-1}.$$

Meanwhile, Proposition 2 below and the estimate (3.8) lead to

$$|\mathrm{I}_3^1|\le c\max(T^{-\alpha},T^{-\frac{\alpha}{2}(1-2\epsilon)},T^{-\alpha(1-\epsilon)},T^{-\alpha(2-\epsilon)})\|q^\dagger-q_h^*\|_{L^2(\Omega)}^2.$$

Upon combing these estimates with the following identity [8, 22]

$$\big((q^\dagger - q_h^*)\nabla u^\dagger(T), \nabla\varphi\big) = \frac{1}{2}\int_\Omega \Big(\frac{q^\dagger - q_h^*}{q^\dagger}\Big)^2 (q^\dagger|\nabla u^\dagger|^2 + (f - \partial_t^\alpha u^\dagger)u^\dagger)(T)\,\mathrm{d}x,$$

(the derivation is analogous to that of the identity (2.8)), we prove the first assertion. Since $\eta_T \le c\eta$ for large T, the second assertion follows exactly as Theorem 1. □

Remark 2 The estimate in Theorem 2 provides useful guidelines for choosing the algorithmic parameters: Given the noise level δ, we may choose $\gamma \sim \delta^2$ and $h \sim \delta^{\frac{1}{2}}$. The choice $\gamma \sim \delta^2$ differs from the usual condition for Tikhonov regularization, i.e., $\lim_{\delta\to 0^+} \frac{\delta^2}{\gamma(\delta)} = 0$, but it agrees with that with conditional stability (see, e.g., [14, Theorems 1.1 and 1.2]). It is noteworthy that the error bound in Theorem 2 is comparable with that for the standard parabolic case [22, Theorem 4.5] and the time fractional case [23, 25]. Theorem 2 requires only the terminal data, whereas previous results [23, 25] in the fractional case require full space-time data. Thus it represents a substantial improvement for the concerned inverse problem.

3.2 Preliminary technical estimates

In this part, we derive crucial a priori bounds on $u(q^\dagger)(T) - U_h^N(q_h^*)$ and ∇q_h^*; see Proposition 1 for the precise statement. For any $q \in \mathcal{A}$, we define a discrete elliptic operator $A_h(q) : X_h \to X_h$ by

$$(A_h(q)v_h, \varphi_h) = (q\nabla v_h, \nabla\varphi_h), \quad \forall v_h, \varphi_h \in X_h.$$

Then problem (3.7) is equivalent to an operator equation in X_h: with $U_h^0 = P_h u_0$,

$$\bar\partial_\tau^\alpha U_h^n + A_h(q_h)U_h^n = P_h f, \quad n = 1, 2, \ldots, N.$$

Using the discrete Laplace transform, the solution $U_h^n(q_h)$ is given by

$$\begin{aligned} U_h^n(q_h) &= F_{h,\tau}^n(q_h)U_h^0 + \tau\sum_{j=1}^n E_{h,\tau}^j(q_h)P_h f \\ &= F_{h,\tau}^n(q_h)P_h u_0 + (I - F_{h,\tau}^n(q_h))A_h(q_h)^{-1}P_h f, \end{aligned}$$

where the fully discrete solution operators $F_{h,\tau}^n(q_h)$ and $E_{h,\tau}^n(q_h)$ are defined respectively by [24, Section 3.2]

$$F_{h,\tau}^n(q_h) = \frac{1}{2\pi i}\int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}}\delta_\tau(e^{-z\tau})^{\alpha-1}(\delta_\tau(e^{-z\tau})^\alpha + A_h(q_h))^{-1}\,\mathrm{d}z, \tag{3.9}$$

$$E_{h,\tau}^n(q_h) = \frac{1}{2\pi i}\int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}}(\delta_\tau(e^{-z\tau})^\alpha + A_h(q_h))^{-1}\,\mathrm{d}z, \tag{3.10}$$

with the kernel function $\delta_\tau(\zeta) := \tau^{-1}(1-\zeta)$ and the contour $\Gamma^\tau_{\theta,\sigma} := \{z \in \Gamma_{\theta,\sigma} : |\Im(z)| \le \frac{\pi}{\tau}\}$ (oriented with an increasing imaginary part). Like in the continuous case, we need suitable smoothing properties of the fully discrete solution operators $F^n_{h,\tau}(q)$ and $E^n_{h,\tau}(q)$. Recall that for any fixed $\theta \in (\frac{\pi}{2}, \pi)$, there exists $\theta' \in (\frac{\pi}{2}, \pi)$ such that for all $\alpha \in (0, 1]$ and $z \in \Gamma^\tau_{\theta,\sigma}$ [24, Lemma 3.1]:

$$c_1|z| \le |\delta_\tau(e^{-z\tau})| \le c_2|z|, \quad \delta_\tau(e^{-z\tau}) \in \Sigma_{\theta'}, \quad |\delta_\tau(e^{-z\tau})^\alpha - z^\alpha| \le c_3\tau|z|^{1+\alpha}, \tag{3.11}$$

where the constants c_1, c_2 and c_3 are independent of τ. Further, for any fixed $q \in \mathcal{A}$, let $\lambda(q)$ and $\lambda_h(q)$ be the smallest eigenvalues of operators $A(q)$ and $A_h(q)$, respectively. Then by Courant–Fischer–Weyl minmax theorem, there exists $c_0 > 0$, independent of q, such that $c_0 \le \lambda(q) \le \lambda_h(q)$. The following discrete resolvent estimate holds

$$\begin{aligned} \|\big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q)\big)^{-1}\|_{L^2(\Omega)\to L^2(\Omega)} &\le c\min(|z|^{-\alpha}, \lambda_h(q)^{-1}) \\ &\le c\min(|z|^{-\alpha}, 1), \quad \forall z \in \Gamma^\tau_{\theta,\sigma}. \end{aligned} \tag{3.12}$$

Now we can give smoothing properties of the solution operators $F^n_{h,\tau}(q)$ and $E^n_{h,\tau}(q)$. The proof of the cases $s = 0$ and $s = 1$ can be found in [43, Lemma 4.3], and the case $0 < s < 1$ follows from the standard interpolation theory.

Lemma 5 *For any $q \in \mathcal{A}$ and any $s \in [0, 1]$, there exists c, independent of h, τ, t_n and q, such that for all $v_h \in X_h$*

$$t_n^{s\alpha}\|A_h(q)^s F^n_{h,\tau}(q)v_h\|_{L^2(\Omega)} + t_n^{1-(1-s)\alpha}\|A_h(q)^s E^n_{h,\tau}(q)v_h\|_{L^2(\Omega)} \le c\|v_h\|_{L^2(\Omega)}.$$

Remark 3 Lemma 5 implies a sharper estimate

$$\|F^n_{h,\tau}(q)v_h\|_{L^2(\Omega)} \le c\min(1, t_n^{-\alpha})\|v_h\|_{L^2(\Omega)}.$$

Indeed, the estimate $\|F^n_{h,\tau}(q)v_h\|_{L^2(\Omega)} \le c\|v_h\|_{L^2(\Omega)}$ follows directly from $s = 0$. Moreover, the $L^2(\Omega)$ uniform boundedness of $A_h(q)^{-1}$ in X_h and the assertion for $s = 1$ give

$$\|F^n_{h,\tau}(q)v_h\|_{L^2(\Omega)} = \|A_h(q)F^n_{h,\tau}(q)A_h(q)^{-1}v_h\|_{L^2(\Omega)} \le ct_n^{-\alpha}\|v_h\|_{L^2(\Omega)}.$$

Combining these two estimates yields the desired estimate.

The analysis uses frequently the following discrete $L^p(\Omega)$ resolvent estimate. Note that the estimate (3.13) is different from that in Lemma 2 in that the latter allows also mappings to the space $W^{1,\infty}(\Omega)$. This difference has important consequences in the error analysis, and it has to be overcome alternatively.

Lemma 6 *Let Ω be a smooth domain or a convex polygon and $q \in W^{1,\infty}(\Omega)$. Then for any $p \in [1, \infty]$ and $v_h \in X_h$, there holds*

$$(1 + |z|)\|(z + A_h(q))^{-1}v_h\|_{L^p(\Omega)} \le c\|v_h\|_{L^p(\Omega)}, \quad \forall z \in \Sigma_\theta,\ \theta \in (\tfrac{\pi}{2}, \pi). \tag{3.13}$$

Proof See [12, Theorem 1.1] for the case $d = 1$, and [29, Theorem 1.1] for the case $d = 2, 3$. Note that for $d = 2, 3$, the work [29] discussed only the case $p = \infty$. The case $p = 1$ follows by a duality argument

$$\begin{aligned}\|(z + A_h(q))^{-1}v_h\|_{L^1(\Omega)} &= \sup_{\|w\|_{L^\infty(\Omega)}=1} \big((z + A_h(q))^{-1}v_h, w\big)\\ &= \sup_{\|w\|_{L^\infty(\Omega)}=1} \big(v_h, (z + A_h(q))^{-1}P_h w\big)\\ &\le \sup_{\|w\|_{L^\infty(\Omega)}=1} \|v_h\|_{L^1(\Omega)}\|(z + A_h(q))^{-1}P_h w\|_{L^\infty(\Omega)}\\ &\le c(1 + |z|)^{-1}\|v_h\|_{L^1(\Omega)}.\end{aligned}$$

The intermediate case $p \in (1, \infty)$ follows from Riesz–Thorin interpolation theorem. □

Now we can give an error bound on the FEM approximation $U_h^N(\Pi_h q^\dagger)$ by the fully discrete scheme (3.7) with $q_h = \Pi_h q^\dagger$. It plays a central role in proving Proposition 1.

Lemma 7 *If Assumption 2 holds, then there exists $c > 0$, independent of τ, h and t_n, such that*

$$\|u(q^\dagger)(t_n) - U_h^n(\Pi_h q^\dagger)\|_{L^2(\Omega)} \le c(t_n^{\alpha-1}\tau + \max(1, t_n^{-\alpha})h^2), \quad n = 1, 2, \ldots, N.$$

Proof Let $u^n \equiv u(q^\dagger)(t_n)$ and $U_h^n \equiv U_h^n(q^\dagger)$. Then the following a priori estimate holds [25, Lemma A.1]

$$\|u^n - U_h^n\|_{L^2(\Omega)} \le c(\tau t_n^{\alpha-1} + h^2), \quad n = 1, 2, \ldots, N.$$

Next we prove

$$\|U_h^n - U_h^n(\Pi_h q^\dagger)\|_{L^2(\Omega)} \le c\max(1, t_n^{-\alpha})h^2, \quad n = 1, 2, \ldots, N.$$

Using the discrete solution operator $F_{h,\tau}^n(q)$ in (3.9), U_h^n and $U_h^n(\Pi_h q^\dagger)$ can be represented respectively by

$$\begin{aligned}U_h^n &= F_{h,\tau}^n(q^\dagger)U_h^0 + (I - F_{h,\tau}^n(q^\dagger))A_h(q^\dagger)^{-1}P_h f,\\ U_h^n(\Pi_h q^\dagger) &= F_{h,\tau}^n(\Pi_h q^\dagger)U_h^0 + (I - F_{h,\tau}^n(\Pi_h q^\dagger))A_h(\Pi_h q^\dagger)^{-1}P_h f.\end{aligned}$$

Thus the error $e_h^n := U_h^n - U_h^n(\Pi_h q^\dagger)$ satisfies $e_h^0 = 0$ and for $n = 1, 2, \ldots, N$

$$\begin{aligned}e_h^n = {}&(F_{h,\tau}^n(q^\dagger) - F_{h,\tau}^n(\Pi_h q^\dagger))U_h^0 + (A_h(q^\dagger)^{-1} - A_h(\Pi_h q^\dagger)^{-1})P_h f\\ &+ (-F_{h,\tau}^n(q^\dagger)A_h(q^\dagger)^{-1} + F_{h,\tau}^n(\Pi_h q^\dagger)A_h(\Pi_h q^\dagger)^{-1})P_h f := \mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3.\end{aligned}$$

It remains to bound the three terms separately. For the term I_1, by the definition of $F_{h,\tau}^n(q^\dagger)$ (with the contour $\sigma = t_n^{-1}$ in $\Gamma_{\theta,\sigma}^\tau$), we have

$$\mathrm{I}_1 = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}} \delta_\tau(e^{-z\tau})^{\alpha-1} K_{h,\tau}(q^\dagger) U_h^0 \,\mathrm{d}z.$$

with the operator $K_{h,\tau}(q^\dagger) : X_h \to X_h$ given by

$$K_{h,\tau}(q^\dagger) := \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1} - \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(\Pi_h q^\dagger)\big)^{-1}.$$

It follows directly from the identity

$$B_1^{-1} - B_2^{-1} = B_1^{-1}(B_2 - B_1)B_2^{-1} \tag{3.14}$$

that

$$\begin{aligned} K_{h,\tau}(q^\dagger) =& \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}\big(A_h(\Pi_h q^\dagger) - A_h(q^\dagger)\big)\big(\delta_\tau(e^{-z\tau})^\alpha \\ &+ A_h(\Pi_h q^\dagger)\big)^{-1}. \end{aligned}$$

Consequently,

$$\begin{aligned} \|K_{h,\tau}(q^\dagger)\|_{L^p(\Omega)\to L^2(\Omega)} \le & \left\| \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1} A_h(q^\dagger) \right\|_{L^2(\Omega)\to L^2(\Omega)} \\ & \times \left\| A_h(\Pi_h q^\dagger)^{-1} - A_h(q^\dagger)^{-1} \right\|_{L^p(\Omega)\to L^2(\Omega)} \\ & \times \left\| \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(\Pi_h q^\dagger)\big)^{-1} A_h(\Pi_h q^\dagger) \right\|_{L^p(\Omega)\to L^p(\Omega)}. \end{aligned}$$

Now we recall the following a priori estimate:

$$\|A_h(\Pi_h q^\dagger)^{-1} - A_h(q^\dagger)^{-1}\|_{L^p(\Omega)\to L^2(\Omega)} \le ch^2, \quad \text{with } p > \max(d+\epsilon, 2). \tag{3.15}$$

We provide a short proof, following the argument of [22, Lemma A.1], for the convenience of readers. Let $v_h(\Pi_h q^\dagger) = A_h(\Pi_h q^\dagger)^{-1} P_h f$ and $v_h(q^\dagger) = A_h(q^\dagger)^{-1} P_h f$ for any $f \in L^p(\Omega)$. Then the difference $w_h := v_h(q^\dagger) - v_h(\Pi_h q^\dagger)$ satisfies for every $\varphi_h \in X_h$,

$$\begin{aligned} (\Pi_h q^\dagger \nabla w_h, \nabla \varphi_h) &= ((\Pi_h q^\dagger - q^\dagger)\nabla v_h(q^\dagger), \nabla \varphi_h) \\ &= ((\Pi_h q^\dagger - q^\dagger)\nabla(v_h(q^\dagger) - v(q^\dagger)), \nabla \varphi_h) + ((\Pi_h q^\dagger - q^\dagger)\nabla v(q^\dagger), \nabla \varphi_h), \end{aligned}$$

with $v(q^\dagger) := A(q^\dagger)^{-1} f$. Let $\varphi_h = w_h \in X_h$. Then Hölder's inequality, Assumption 2 and the approximation properties (3.3)–(3.4) of the operator Π_h lead to

$$\begin{aligned}\|\nabla w_h\|_{L^2(\Omega)} &\leq c\|\Pi_h q^\dagger - q^\dagger\|_{L^\infty(\Omega)}\|\nabla(v_h(q^\dagger) - v(q^\dagger))\|_{L^2(\Omega)} \\ &\quad + c\|\Pi_h q^\dagger - q^\dagger\|_{L^2(\Omega)}\|\nabla v(q^\dagger)\|_{L^\infty(\Omega)} \\ &\leq ch^2\big(\|q^\dagger\|_{W^{1,\infty}(\Omega)} + \|q^\dagger\|_{H^2(\Omega)}\big)\|v(q^\dagger)\|_{W^{2,p}(\Omega)} \leq ch^2\|f\|_{L^p(\Omega)},\end{aligned}$$

in view of Sobolev embedding $W^{1,p}(\Omega) \hookrightarrow L^\infty(\Omega)$ and the full $W^{2,p}(\Omega)$ regularity pickup of the elliptic operator $A(q^\dagger)$. Last, Poincaré inequality yields

$$\|w_h\|_{L^2(\Omega)} \leq c\|\nabla w_h\|_{L^2(\Omega)} \leq ch^2\|f\|_{L^p(\Omega)}.$$

This directly implies the estimate (3.15). Meanwhile, the discrete resolvent estimate (3.13) implies

$$\|(\delta_\tau(e^{-z\tau})^\alpha + A_h(q))^{-1} A_h(q)\|_{L^p(\Omega)\to L^p(\Omega)} \leq c, \quad \forall q \in \mathcal{A}. \tag{3.16}$$

Combining the estimates (3.15) and (3.16) gives

$$\|K_{h,\tau}(q^\dagger)\|_{L^p(\Omega)\to L^2(\Omega)} \leq ch^2.$$

Thus the following bound on the term I_1 holds:

$$\|\mathrm{I}_1\|_{L^2(\Omega)} \leq ch^2\|u_0\|_{L^p(\Omega)} \int_{\Gamma^\tau_{\theta,\sigma}} |e^{zt_n}||z|^{\alpha-1}\,|\mathrm{d}z| \leq ct_n^{-\alpha}h^2.$$

For the term I_2, by the estimate (3.15) and the $L^p(\Omega)$ stability of P_h, cf. (3.5), we obtain

$$\|\mathrm{I}_2\|_{L^2(\Omega)} \leq \|A_h(q^\dagger)^{-1} - A_h(\Pi_h q^\dagger)^{-1}\|_{L^p(\Omega)\to L^2(\Omega)}\|P_h f\|_{L^p(\Omega)} \leq ch^2.$$

Last, for the term I_3, we use the splitting $\mathrm{I}_3 = \mathrm{I}_3^1 + \mathrm{I}_3^2$, with

$$\begin{aligned}\mathrm{I}_3^1 &= (F^n_{h,\tau}(\Pi_h q^\dagger) - F^n_{h,\tau}(q^\dagger))A_h(q^\dagger)^{-1}P_h f, \\ \mathrm{I}_3^2 &= -F^n_{h,\tau}(\Pi_h q^\dagger)(A_h(q^\dagger)^{-1} - A_h(\Pi_h q^\dagger)^{-1})P_h f.\end{aligned}$$

The argument for the term I_1 and the boundedness of the operator $A_h(q^\dagger)^{-1}P_h$ in $L^p(\Omega)$ [9, Section 8.5] imply

$$\|\mathrm{I}_3^1\|_{L^2(\Omega)} \leq ch^2 t_n^{-\alpha}\|A_h(q^\dagger)^{-1}P_h f\|_{L^p(\Omega)} \leq ch^2 t_n^{-\alpha}\|f\|_{L^p(\Omega)}.$$

Meanwhile, Lemma 5, Remark 3, the $L^p(\Omega)$ stability of P_h and the estimate (3.15) lead to

$$\begin{aligned}\|\mathrm{I}_3^2\|_{L^2(\Omega)} &= \|F_{h,\tau}^n(\Pi_h q^\dagger)(A_h(q^\dagger)^{-1} - A_h(\Pi_h q^\dagger)^{-1})P_h f\|_{L^2(\Omega)} \\ &\leq \|F_{h,\tau}^n(\Pi_h q^\dagger)\|_{L^2(\Omega)\to L^2(\Omega)}\|(A_h(\Pi_h q^\dagger)^{-1} - A_h(q^\dagger)^{-1})P_h f\|_{L^2(\Omega)} \\ &\leq c\min(1, t_n^{-\alpha})\|(A_h(\Pi_h q^\dagger)^{-1} - A_h(q^\dagger)^{-1})P_h f\|_{L^2(\Omega)} \\ &\leq c\min(1, t_n^{-\alpha})h^2\|P_h f\|_{L^p(\Omega)} \leq c\min(1, t_n^{-\alpha})h^2\|f\|_{L^p(\Omega)}.\end{aligned}$$

The desired estimate now follows by combining the preceding estimates. □

Next we provide a crucial a priori estimate of $u(q^\dagger)(T) - U_h^N(q_h^*)$ and ∇q_h^*.

Proposition 1 *Let Assumption 2 hold, and q_h^* be a minimizer of problem (3.6)–(3.7). Then there exists c, independent of τ, h, δ, γ and T, such that*

$$\|u(q^\dagger)(T) - U_h^N(q_h^*)\|_{L^2(\Omega)} + \gamma^{\frac{1}{2}}\|\nabla q_h^*\|_{L^2(\Omega)} \leq c(T^{\alpha-1}\tau + \max(1, T^{-\alpha})h^2 + \delta + \gamma^{\frac{1}{2}}).$$

Proof Let $u \equiv u(q^\dagger)$. Since q_h^* minimizes problem (3.6)–(3.7) and $\Pi_h q^\dagger \in \mathcal{A}_h$, we have

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

By the $H^1(\Omega)$ stability of the operator Π_h, cf. (3.3), and Lemma 7, we obtain

$$\begin{aligned}&\|U_h^N(q_h^*) - z^\delta\|_{L^2(\Omega)}^2 + \gamma\|\nabla q_h^*\|_{L^2(\Omega)}^2 \\ &\quad\leq \|U_h^N(\Pi_h q^\dagger) - z^\delta\|_{L^2(\Omega)}^2 + \gamma\|\nabla \Pi_h q^\dagger\|_{L^2(\Omega)}^2 \\ &\quad\leq c\big(\|U_h^N(\Pi_h q^\dagger) - u(T)\|_{L^2(\Omega)}^2 + \|u(T) - z^\delta\|_{L^2(\Omega)}^2 + \gamma\big) \\ &\quad\leq c(T^{2\alpha-2}\tau^2 + \max(1, T^{-2\alpha})h^4 + \delta^2 + \gamma).\end{aligned}$$

Then the triangle inequality yields

$$\begin{aligned}&\|u(T) - U_h^N(q_h^*)\|_{L^2(\Omega)}^2 + \gamma\|\nabla q_h^*\|_{L^2(\Omega)}^2 \\ &\quad\leq c\big(\|u(T) - z^\delta\|_{L^2(\Omega)}^2 + \|z^\delta - U_h^N(q_h^*)\|_{L^2(\Omega)}^2 + \gamma\|\nabla q_h^*\|_{L^2(\Omega)}^2\big) \\ &\quad\leq c(T^{2\alpha-2}\tau^2 + \max(1, T^{-2\alpha})h^4 + \delta^2 + \gamma).\end{aligned}$$

This completes the proof of the lemma. □

3.3 Bound on the error $\bar\partial_\tau^\alpha U_h^n(q_h^*) - \partial_t^\alpha u(q^\dagger)(t_n)$

Next, we estimate the decay of the discrete (fractional) derivative $\bar\partial_\tau^\alpha U_h^n(q_h^*)$ and bound the term $\bar\partial_\tau^\alpha U_h^n(q_h^*) - \partial_t^\alpha u(q^\dagger)(t_n)$ in terms of $\|q_h^* - q^\dagger\|_{L^2(\Omega)}$; see Lemma 10 and Proposition 2 for the precise statement. These estimates play a central role in establishing Theorem 2. We need the following time semidiscrete scheme for problem (1.1): Find $U^n \equiv U^n(q) \in H_0^1(\Omega)$ with $U^0 = u_0$ such that

$$\bar\partial_\tau^\alpha U^n(q) + A(q)U^n(q) = f, \quad n = 1, 2, \ldots, N. \tag{3.17}$$

The discrete Laplace transform gives

$$U^n = F_\tau^n(q)u_0 + \tau \sum_{j=1}^{n} E_\tau^j(q) f = F_\tau^n(q)u_0 + (I - F_\tau^n(q))A(q)^{-1} f, \tag{3.18}$$

where the time-semidiscrete solution operators $F_\tau^n\tau(q)$ and $E_\tau^n(q)$ are defined respectively by [24, Section 3.2]

$$F_\tau^n(q) = \frac{1}{2\pi i}\int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}}\delta_\tau(e^{-z\tau})^{\alpha-1}(\delta_\tau(e^{-z\tau})^\alpha + A(q))^{-1}\,\mathrm{d}z,$$
$$E_\tau^n(q) = \frac{1}{2\pi i}\int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}}(\delta_\tau(e^{-z\tau})^\alpha + A(q))^{-1}\,\mathrm{d}z.$$

The next lemma gives a temporal error estimate for the approximate time (fractional) derivative.

Lemma 8 *Let Assumption 2 hold, and $u(q^\dagger)$ and $\{U^n(q^\dagger)\}_{n=0}^N$ be the solutions of problems (1.1) and (3.17) for $q^\dagger$, respectively. Then there exists $c > 0$, independent of τ, t_n and $q^\dagger$, such that*

$$\|\partial_t^\alpha u(q^\dagger)(t_n) - \bar\partial_\tau^\alpha U^n(q^\dagger)\|_{L^2(\Omega)} \le c\tau t_n^{-\alpha-1}.$$

Proof Let $u \equiv u(q^\dagger)$ and $U^n \equiv U^n(q^\dagger)$. Then $W^n := \bar\partial_\tau^\alpha U^n$ satisfies $W^0 = f - A(q^\dagger)u_0$ and

$$\bar\partial_\tau^\alpha W^n + A(q^\dagger)W^n = 0, \quad n = 1, 2, \ldots, N.$$

It follows from the solution representations (2.1) and (3.18) that

$$\partial_t^\alpha u(t_n) = F(t_n; q^\dagger)(f - A(q^\dagger)u_0) \quad \text{and} \quad \bar\partial_\tau^\alpha U^n = F_\tau^n(q^\dagger)(f - A(q^\dagger)u_0).$$

It follows from the estimate $\|F(t_n; q) - F_\tau^n(q)\|_{L^2(\Omega)\to L^2(\Omega)} \le cn^{-1}t_n^{-\alpha}$ [24, Lemma 15.6] and Assumption 2 that

$$\begin{aligned}
&\|\partial_t^\alpha u(q^\dagger)(t_n) - \bar\partial_\tau^\alpha U^n(q^\dagger)\|_{L^2(\Omega)} \\
&\quad \le \|F(t_n; q^\dagger) - F_\tau^n(q^\dagger)\|_{L^2(\Omega)\to L^2(\Omega)}\|f - A(q^\dagger)u_0\|_{L^2(\Omega)} \\
&\quad \le c\tau t_n^{-\alpha-1}(\|f\|_{L^2(\Omega)} + \|u_0\|_{H^2(\Omega)}) \le c\tau t_n^{-\alpha-1}.
\end{aligned}$$

This completes the proof of the lemma. □

The next lemma bounds the error between the discrete (fractional) derivative due to the spatial discretization.

Lemma 9 *Let Assumption 2 hold, and* $\{U^n(q^\dagger)\}_{n=0}^N$ *and* $\{U_h^n(q^\dagger)\}_{n=0}^N$ *be the solutions of problem (3.17) with* $q^\dagger$ *and problem (3.7) with* $q^\dagger$*, respectively. Then there exists* $c > 0$*, independent of* τ*,* h*,* t_n *and* $q^\dagger$*, such that*

$$\|\bar\partial_\tau^\alpha\big(U^n(q^\dagger) - U_h^n(q^\dagger)\big)\|_{L^2(\Omega)} \le ch^2 \max(t_n^{-\alpha}, t_n^{-2\alpha}).$$

Moreover, if $u_0 \in W^{2,\infty}(\Omega)$*, then with* $\ell_h := |\log h|$*,*

$$\|\bar\partial_\tau^\alpha\big(U^n(q^\dagger) - U_h^n(q^\dagger)\big)\|_{L^\infty(\Omega)} \le cht_n^{-\frac{\alpha}{2}} + ch^2\ell_h^2 \max(t_n^{-\alpha}, t_n^{-2\alpha}).$$

Proof Let $U^n \equiv U^n(q^\dagger)$ and $U_h^n \equiv U_h^n(q^\dagger)$, and $e_{h,\tau}^n := \bar\partial_\tau^\alpha(U^n - U_h^n)$. Note that the sequences $W^n := \bar\partial_\tau^\alpha U^n$ and $W_h^n := \bar\partial_\tau^\alpha U_h^n$ satisfy

$$\begin{cases} W^0 = f - A(q^\dagger)u_0, \\ \bar\partial_\tau^\alpha W^n + A(q^\dagger)W^n = 0, \quad n = 1, \ldots, N, \end{cases}$$
$$\begin{cases} W_h^0 = P_h f - A_h(q^\dagger)P_h u_0, \\ \bar\partial_\tau^\alpha W_h^n + A_h(q^\dagger)W_h^n = 0, \quad n = 1, \ldots, N. \end{cases}$$

By the solution representation (3.18), there holds

$$\begin{aligned} e_{h,\tau}^n &= F_\tau^n(q^\dagger)(f - A(q^\dagger)u_0) - F_{h,\tau}^n(q^\dagger)(P_h f - A_h(q^\dagger)P_h u_0) \\ &= \big(F_\tau^n(q^\dagger) - F_{h,\tau}^n(q^\dagger)P_h\big)f + \big(F_{h,\tau}^n(q^\dagger)A_h(q^\dagger)P_h - F_\tau^n(q^\dagger)A(q^\dagger)\big)u_0 \\ &:= \mathrm{I}_1 + \mathrm{I}_2. \end{aligned} \tag{3.19}$$

Now we bound the two terms I_1 and I_2 separately. Let

$$B_{h,\tau} = \big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1} - \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}P_h.$$

It follows from the estimates in (3.11) that for all $z \in \Gamma_{\theta,\sigma}^\tau$ [16, Theorem 5.2 and Remark 7.4]

$$\|B_{h,\tau}\|_{L^2(\Omega)\to L^2(\Omega)} \le ch^2. \tag{3.20}$$

Then choosing $\sigma = t_n^{-1}$ in the contour $\Gamma_{\theta,\sigma}^\tau$ leads to

$$\begin{aligned} \|\mathrm{I}_1\|_{L^2(\Omega)} &\le c\|f\|_{L^2(\Omega)} \int_{\Gamma_{\theta,\sigma}^\tau} |e^{zt_n}||\delta_\tau(e^{-z\tau})|^{\alpha-1}\|B_{h,\tau}\|_{L^2(\Omega)\to L^2(\Omega)}\,|\mathrm{d}z| \\ &\le ch^2\|f\|_{L^2(\Omega)} \int_{\Gamma_{\theta,\sigma}^\tau} |e^{zt_n}||z|^{\alpha-1}\,|\mathrm{d}z| \le ch^2 t_n^{-\alpha}. \end{aligned}$$

To estimate the term I_2, by the identity

$$\begin{aligned}&\big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}A_h(q^\dagger)P_h - \big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1}A(q^\dagger)\\ &\quad = \big(P_h - I\big) + \delta_\tau(e^{-z\tau})^\alpha B_{h,\tau},\end{aligned}$$

we derive

$$\mathrm{I}_2 = \frac{1}{2\pi \mathrm{i}}\int_{\Gamma^\tau_{\theta,\sigma}} e^{zt_n}\delta_\tau(e^{-z\tau})^{\alpha-1}(P_h - I)u_0 \,\mathrm{d}z + \frac{1}{2\pi \mathrm{i}}\int_{\Gamma^\tau_{\theta,\sigma}} e^{zt_n}\delta_\tau(e^{-z\tau})^{2\alpha-1} B_{h,\tau}u_0 \mathrm{d}z.$$

Then with $\sigma = t_n^{-1}$ in the contour $\Gamma^\tau_{\theta,\sigma}$, by the estimates (3.20), (3.11), and (3.5), we derive

$$\begin{aligned}\|\mathrm{I}_2\|_{L^2(\Omega)} &\le ch^2 \int_{\Gamma^\tau_{\theta,\sigma}} |e^{zt_n}|(|z|^{\alpha-1}\|u_0\|_{H^2(\Omega)} + |z|^{2\alpha-1}\|u_0\|_{L^2(\Omega)})\, |\mathrm{d}z|\\ &\le ch^2 t_n^{-\alpha}\|u_0\|_{H^2(\Omega)} + ch^2 t_n^{-2\alpha}\|u_0\|_{L^2(\Omega)} \le ch^2 \max(t_n^{-\alpha}, t_n^{-2\alpha}).\end{aligned}$$

To bound $\|e^n_{h,\tau}\|_{L^\infty(\Omega)}$, we split $e^n_{h,\tau}$ into

$$e^n_{h,\tau} := \big(\bar\partial^\alpha_\tau U^n - P_h\bar\partial^\alpha_\tau U^n\big) + \big(P_h\bar\partial^\alpha_\tau U^n - \bar\partial^\alpha_\tau U^n_h\big) := \mathrm{I}_3 + \mathrm{I}_4.$$

It follows from the estimates in (3.11), (2.4), and the argument of Lemma 3 that

$$\|F^n_\tau(q^\dagger)\|_{L^\infty(\Omega)\to W^{1,\infty}(\Omega)} \le ct_n^{-\frac{\alpha}{2}}.$$

From the approximation property of P_h in (3.5) and the assumption $u_0 \in W^{2,\infty}(\Omega)$, we have

$$\begin{aligned}&\|\mathrm{I}_3\|_{L^\infty(\Omega)} \le ch\|\bar\partial^\alpha_\tau U^n\|_{W^{1,\infty}(\Omega)}\\ &\quad \le ch\|F^n_\tau(q^\dagger)\|_{L^\infty(\Omega)\to W^{1,\infty}(\Omega)}\big(\|f\|_{L^\infty(\Omega)} + \|A(q^\dagger)u_0\|_{L^\infty(\Omega)}\big) \le cht_n^{-\frac{\alpha}{2}}.\end{aligned}$$

It remains to bound the term I_4. From the error representation (3.19), we obtain

$$\mathrm{I}_4 = \big(P_hF^n_\tau(q^\dagger) - F^n_{h,\tau}(q^\dagger)P_h\big)f + \big(F^n_{h,\tau}(q^\dagger)A_h(q^\dagger)P_h - P_hF^n_\tau(q^\dagger)A(q^\dagger)\big)u_0.$$

Let R_h be the standard Ritz projection. Then direct computation with the identity (3.14) and the relation $A_h(q^\dagger)R_h = P_hA(q^\dagger)$ [38, p. 11] gives

$$\begin{aligned}&P_h\big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1} - \big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}P_h\\ &\quad = A_h(q^\dagger)\big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}(P_h - R_h)\big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1}.\end{aligned}$$

Hence, we obtain

$$P_h F_\tau^n(q^\dagger) - F_{h,\tau}^n(q^\dagger)P_h = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}}\delta_\tau(e^{-z\tau})^{\alpha-1} \times \left[A_h(q^\dagger)\big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}(P_h - R_h)\big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1}\right] \mathrm{d}z$$

and

$$F_{h,\tau}^n(q^\dagger)A_h(q^\dagger)P_h - P_h F_\tau^n(q^\dagger)A(q^\dagger) = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\sigma}^\tau} e^{zt_{n-1}}\delta_\tau(e^{-z\tau})^{2\alpha-1} \times \left[A_h(q^\dagger)\big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}(P_h - R_h)\big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1}\right] \mathrm{d}z.$$

The resolvent estimates (3.13) and (2.4) imply

$$\|A_h(q^\dagger)\big(\delta_\tau(e^{-z\tau})^\alpha + A_h(q^\dagger)\big)^{-1}\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \le c,$$
$$\|A(q^\dagger)\big(\delta_\tau(e^{-z\tau})^\alpha + A(q^\dagger)\big)^{-1}\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \le c.$$

Now the estimate $\|(P_h - R_h)A(q^\dagger)^{-1}\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \le ch^2\ell_h^2$ holds ([36, p. 1658] and [6, p. 220]). Then letting $\sigma = t_n^{-1}$ in the contour $\Gamma_{\theta,\sigma}^\tau$ leads to

$$\|P_h F_\tau^n(q^\dagger) - F_{h,\tau}^n(q^\dagger)P_h\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \le ch^2\ell_h^2 t_n^{-\alpha},$$
$$\|F_{h,\tau}^n(q^\dagger)A_h(q^\dagger)P_h - P_h F_\tau^n(q^\dagger)A(q^\dagger)\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \le ch^2\ell_h^2 t_n^{-2\alpha}.$$

These two estimates imply

$$\|\mathrm{I}_4\|_{L^\infty(\Omega)} \le ch^2\ell_h^2 t_n^{-\alpha}\|f\|_{L^\infty(\Omega)} + ch^2\ell_h^2 t_n^{-2\alpha}\|u_0\|_{L^\infty(\Omega)} \le ch^2\ell_h^2 \max(t_n^{-\alpha}, t_n^{-2\alpha}).$$

Combining the preceding estimates completes the proof of the lemma. □

The next result gives a discrete analogue of Lemma 3.

Lemma 10 *Let Ω be a convex C^2 domain, and Assumption 2 hold. Let $\{U_h^n(q^\dagger)\}_{n=0}^N$ be the solution of problem (3.7) with $q^\dagger$. Then for small $h > 0$, there exists $c > 0$, independent of τ, h, t_n and $q^\dagger$, such that*

$$\|\bar\partial_\tau^\alpha U_h^n(q^\dagger)\|_{W^{1,\infty}(\Omega)} \le c\max(t_n^{-\frac{\alpha}{2}}, t_n^{-\alpha}, t_n^{-2\alpha}), \quad n = 1, 2, \ldots, N.$$

Proof It follows from the inverse estimate on the FEM space X_h [38, equation (1.12)], Lemma 9 and the approximation property and $W^{1,\infty}(\Omega)$ stability of P_h in (3.5) that

$$\begin{aligned}\|\bar\partial_\tau^\alpha U_h^n(q^\dagger)\|_{W^{1,\infty}(\Omega)} &\le \|\bar\partial_\tau^\alpha U_h^n(q^\dagger) - P_h\bar\partial_\tau^\alpha U^n(q^\dagger)\|_{W^{1,\infty}(\Omega)} + \|P_h\bar\partial_\tau^\alpha U^n(q^\dagger)\|_{W^{1,\infty}(\Omega)}\\ &\le ch^{-1}\|\bar\partial_\tau^\alpha U_h^n(q^\dagger) - \bar\partial_\tau^\alpha U^n(q^\dagger)\|_{L^\infty(\Omega)} + c\|\bar\partial_\tau^\alpha U^n(q^\dagger)\|_{W^{1,\infty}(\Omega)}\\ &\le ct_n^{-\frac{\alpha}{2}} + ch\ell_h^2\max(t_n^{-\alpha}, t_n^{-2\alpha}) \le c\max(t_n^{-\frac{\alpha}{2}}, t_n^{-\alpha}, t_n^{-2\alpha}).\end{aligned}$$

This completes the proof of the lemma. □

We also have a discrete version of Lemma 4. This estimate is crucial to the error analysis.

Proposition 2 *Let Ω be a convex C^2 domain, and Assumption 2 hold. Let $\{U_h^n(q^\dagger)\}_{n=0}^N$ and $\{U_h^n(q_h^*)\}_{n=0}^N$ be the solutions of problem (3.7) with $q^\dagger$ and q_h^*, respectively. Then for small $\epsilon, h > 0$, there exists $c > 0$, independent of τ, h, and t_n, such that for all $n = 1, \ldots, N$*

$$\|\bar\partial_\tau^\alpha\big(U_h^n(q^\dagger) - U_h^n(q_h^*)\big)\|_{L^2(\Omega)} \le c\max(t_n^{-\alpha}, t_n^{-\frac{\alpha}{2}(1-2\epsilon)}, t_n^{-\alpha(1-\epsilon)}, t_n^{-\alpha(2-\epsilon)})$$
$$\|q^\dagger - q_h^*\|_{L^2(\Omega)}.$$

Proof Let $W_h^n := \bar\partial_\tau^\alpha\big(U_h^n(q_h^*) - U_h^n(q^\dagger)\big)$. Then $W_h^0 = \big(A_h(q^\dagger) - A_h(q_h^*)\big)U_h^0$ and

$$\bar\partial_\tau^\alpha W_h^n + A_h(q_h^*)W_h^n = \big(A_h(q^\dagger) - A_h(q_h^*)\big)\bar\partial_\tau^\alpha U_h^n(q^\dagger), \quad n = 1, 2, \ldots, N.$$

By the discrete Laplace transform, we obtain

$$W_h^n = F_{h,\tau}^n(q_h^*)\big(A_h(q^\dagger) - A_h(q_h^*)\big)U_h^0 + \tau\sum_{j=1}^n E_{h,\tau}^{n-j}(q_h^*)\big(A_h(q^\dagger) - A_h(q_h^*)\big)\bar\partial_\tau^\alpha U_h^j(q^\dagger).$$

Next we bound the two terms separately. Note that the following inequality holds

$$c_1\|A_h(q_h^*)^{\frac12}v_h\|_{L^2(\Omega)} \le \|\nabla v_h\|_{L^2(\Omega)} \le c_2\|A_h(q_h^*)^{\frac12}v_h\|_{L^2(\Omega)}, \quad \forall v_h \in X_h.$$

This ienquality, Lemma 5, the $W^{1,\infty}(\Omega)$ stability of P_h in (3.5), and the boundedness of the operator $A_h(q_h^*)^{-\frac12}$ in $L^2(\Omega)$ imply

$$\begin{aligned}
&\|F_{h,\tau}^n(q_h^*)\big(A_h(q^\dagger) - A_h(q_h^*)\big)U_h^0\|_{L^2(\Omega)}\\
&= \sup_{\|\varphi_h\|_{L^2(\Omega)}=1}\Big((q^\dagger - q_h^*)\nabla P_h u_0, \nabla A_h(q_h^*)^{\frac12}F_{h,\tau}^n(q_h^*)A_h(q_h^*)^{-\frac12}\varphi_h\Big)\\
&\le c\|q^\dagger - q_h^*\|_{L^2(\Omega)}\|\nabla P_h u_0\|_{L^\infty(\Omega)}\|A_h(q_h^*)F_{h,\tau}^n(q_h^*)\|_{L^2(\Omega)\to L^2(\Omega)}\\
&\quad\times\sup_{\|\varphi_h\|_{L^2(\Omega)}=1}\|A_h(q_h^*)^{-\frac12}\varphi_h\|_{L^2(\Omega)}\\
&\le ct_n^{-\alpha}\|q^\dagger - q_h^*\|_{L^2(\Omega)}.
\end{aligned}$$

Similarly, the boundedness of the operator $A(q_h^*)^{-\frac{1}{2}+\epsilon}$ in $L^2(\Omega)$ and Lemmas 5–10 lead to

$$\begin{aligned}
&\|E_{h,\tau}^{n-j}(q_h^*)\big(A_h(q^\dagger)-A_h(q_h^*)\big)\bar\partial_\tau^\alpha U_h^j(q^\dagger)\|_{L^2(\Omega)}\\
&\quad=\sup_{\|\varphi_h\|_{L^2(\Omega)}=1}\Big((q^\dagger-q_h^*)\nabla\bar\partial_\tau^\alpha U_h^j(q^\dagger),\nabla A_h(q_h^*)^{\frac{1}{2}-\epsilon}E_{h,\tau}^{n-j}(q_h^*)A_h(q_h^*)^{-\frac{1}{2}+\epsilon}\varphi_h\Big)\\
&\quad\le c\|q^\dagger-q_h^*\|_{L^2(\Omega)}\|\nabla\bar\partial_\tau^\alpha U_h^j(q^\dagger)\|_{L^\infty(\Omega)}\|A_h(q_h^*)^{1-\epsilon}E_{h,\tau}^{n-j}(q_h^*)\|_{L^2(\Omega)\to L^2(\Omega)}\\
&\qquad\times\sup_{\|\varphi_h\|_{L^2(\Omega)}=1}\|A_h(q_h^*)^{-\frac{1}{2}+\epsilon}\varphi_h\|_{L^2(\Omega)}\\
&\quad\le ct_{n-j}^{-1+\epsilon\alpha}\max(t_j^{-\frac{\alpha}{2}},t_j^{-\alpha},t_j^{-2\alpha})\|q^\dagger-q_h^*\|_{L^2(\Omega)}.
\end{aligned}$$

Finally, the preceding two estimates and the triangle inequality imply

$$\begin{aligned}
\|W_h^n\|_{L^2(\Omega)}&\le ct_n^{-\alpha}\|q^\dagger-q_h^*\|_{L^2(\Omega)}+c\|q^\dagger-q_h^*\|_{L^2(\Omega)}\tau\sum_{j=1}^n t_{n-j}^{-1+\epsilon\alpha}\max(t_j^{-\frac{\alpha}{2}},t_j^{-\alpha},t_j^{-2\alpha})\\
&\le c\max(t_n^{-\alpha},t_n^{-\frac{\alpha}{2}(1-2\epsilon)},t_n^{-\alpha(1-\epsilon)},t_n^{-\alpha(2-\epsilon)})\|q^\dagger-q_h^*\|_{L^2(\Omega)}.
\end{aligned}$$

This completes the proof of the lemma. □

Remark 4 Proposition 2 and Lemma 10 imply a uniform boundedness of the discrete fractional derivative $\bar\partial_\tau^\alpha U_h^N(q_h^*)$ in the $L^2(\Omega)$ norm, which directly follows from the box constraint of the admissible set $\mathcal{A}$ and the triangle inequality.

4 Numerical experiments and discussions

Now we present numerical results for the time fractional diffusion model. We employ the standard conjugate gradient method [5] to solve the discrete optimization problem. The gradient J'_γ is computed using an alternative adjoint technique (following [11, Section 5]). The details of the algorithm are described in Algorithm 1 in the appendix for completeness. The noise data z^δ is generated by

$$z^\delta(x)=u(q^\dagger)(x,T)+\epsilon\|u(q^\dagger)(T)\|_{L^\infty(\Omega)}\xi(x),\quad x\in\Omega,$$

where the noise $\xi(x)$ follows the standard Gaussian distribution and $\epsilon>0$ is the relative noise level. To measure the convergence of the approximation q_h^*, we employ two metrics, i.e., $e_q=\|q^\dagger-q_h^*\|_{L^2(\Omega)}$ and $e_u=\|u(q^\dagger)(T)-u_h^N(q_h^*)\|_{L^2(\Omega)}$. Throughout, we choose the algorithmic parameters h, τ and γ as $h=O(\sqrt{\delta})$, $\tau=O(\delta)$ and $\gamma=O(\delta^2)$, following Theorem 2.

First, we consider the one-dimension case with different combinations of the terminal time T and the fractional order α.

Table 1 Numerical results for Example 1

(a) Results for case (a), with T fixed at $T = 1.00$, and varying α. The computation is initialized with the mesh size $h = 1/150$ and the time step size $\tau = 1/180$.

α	ϵ / γ	5.00e−2 / 7.50e−8	3.00e−2 / 2.70e−8	1.00e−2 / 3.00e−9	5.00e−3 / 7.50e−10	1.00e−3 / 3.00e−11	Rate
0.25	e_q	7.50e−2	4.99e−2	3.02e−2	2.41e−2	5.88e−3	0.61
	e_u	3.70e−3	2.68e−3	1.19e−3	5.85e−4	1.42e−4	0.84
0.50	e_q	5.77e−2	4.08e−2	2.61e−2	1.84e−2	5.68e−3	0.57
	e_u	6.73e−3	3.35e−3	1.55e−3	7.59e−4	1.69e−4	0.92
0.75	e_q	5.73e−2	4.13e−2	2.60e−2	1.89e−2	5.91e−3	0.56
	e_u	6.77e−3	3.38e−3	1.55e−3	7.80e−4	1.91e−4	0.89

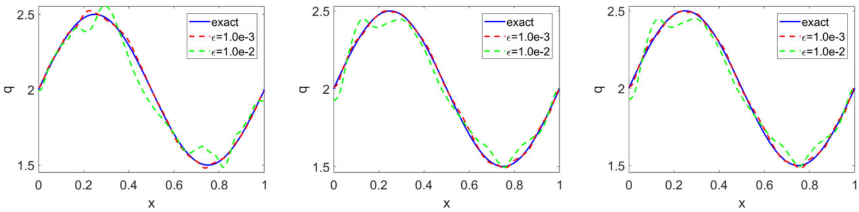
(b) Results for case (b) with α fixed at 0.5, and varying T. The computation is initialized with the mesh size $h = 1/150$ and the time step size $\tau = 1/180$

T	ϵ / γ	5.00e−2 / 7.50e−8	3.00e−2 / 2.70e−8	1.00e−2 / 3.00e−9	5.00e−3 / 7.50e−10	1.00e−3 / 3.00e−11	Trend
10^{-5}	e_q	9.05e−2	9.86e−2	1.32e−1	1.06e−1	1.30e−1	–
	e_u	2.12e−3	1.84e−3	1.95e−3	1.68e−3	1.49e−3	–
3.00	e_q	1.08e−1	8.08e−2	6.25e−2	2.31e−2	8.34e−3	↘
	e_u	5.43e−3	4.00e−3	2.33e−3	7.84e−4	1.82e−4	↘
5.00	e_q	9.90e−2	7.94e−2	5.32e−2	2.66e−2	7.74e−3	↘
	e_u	5.73e−3	3.80e−3	1.63e−3	8.20e−4	1.81e−4	↘

Example 1 Let $\Omega = (0, 1)$, $q^\dagger(x) = 2 + \frac{1}{2}\sin(2\pi x)$, $u_0(x) = x(1-x)$ and $f \equiv 10$. Consider the following two cases: (a) $T = 1.00$ and $\alpha = 0.25$, 0.50 and 0.75; (b) $\alpha = 0.50$ and $T = 10^{-5}$, 3.00 and 5.00.

In this example, the exact data $z^\dagger = u(q^\dagger)(T)$ is obtained using a fine grid with a spatial mesh size $h = 1/1600$ and the number $N = 1280$ of time steps. The lower and upper bounds of the admissible set $\mathcal{A}$ are taken to be 1.5 and 3.0, respectively. The convergence of the conjugate gradient method is achieved within tens of iterations. The numerical results for Example 1 are presented in Table 1, with the rate estimated by applying the linear least-square approach of the error and noise level in a double logarithmic scale. Note that when the terminal time T is sufficiently large, Theorem 2 predicts a convergence rate $O(\delta^{\frac{1}{4}})$ at best for e_q, and $O(\delta)$ for e_u of the state approximation (if the parameters are chosen properly). For case (a), both e_q and e_u exhibit a clear decay property as the noise level δ tends to zero but the empirical rate of e_q is faster than the theoretical one, indicating a room for further improvement in the convergence analysis. In Fig. 1, we present the numerical reconstructions of case (a) at different noise levels. The results for case (b) in Table 1 show that the convergence behaviors of both e_q and e_u fail to hold due to the loss of the conditional stability estimate in Theorems 1 and 2 when the terminal time T is small.

Next we give a two-dimensional example.



(a) $\alpha = 0.25$ (b) $\alpha = 0.50$ (c) $\alpha = 0.75$

Fig. 1 The numerical reconstructions for Example 1(a): exact (real line); reconstruction (dot line)

Table 2 Numerical results for Example 2, initialized the mesh size $h = 1/10$ and the time step size $\tau = 1/50$

α	ϵ γ	5.00e−2 1.00e−9	3.00e−2 3.60e−10	1.00e−2 4.00e−11	5.00e−3 1.00e−11	Rate
0.25	e_q	2.12e−2	1.32e−2	7.12e−3	6.27e−3	0.53
	e_u	3.43e−5	1.35e−5	6.16e−6	5.02e−6	0.80
0.50	e_q	1.76e−2	1.03e−2	6.65e−3	6.11e−3	0.44
	e_u	3.71e−5	1.47e−5	6.46e−5	5.20e−6	0.83
0.75	e_q	2.32e−2	1.23e−2	7.02e−3	6.11e−3	0.56
	e_u	3.35e−5	1.45e−5	6.26e−6	4.86e−6	0.81

Example 2 Let $\Omega = (0,1)^2$, $q^\dagger(x,y) = 1 + \frac{1}{2}y(1-y)\sin(\pi x)$, $u_0(x,y) = x(1-x)y(1-y)$ and $f \equiv \frac{1}{10}\sin(\pi x)\sin(\pi y)$. Fix $T = 0.2$, and $\alpha = 0.25$, 0.50 and 0.75.

In this example, we divide the domain Ω into $M \times M$ (with $h = 1/M$) equal small squares and connect the bottom-left and top-right corners to form a triangulation of the domain Ω. Like before, the exact data $z^\dagger = u(q^\dagger)(T)$ is obtained by solving the direct problem on a fine grid with $h = 1/100$ and $\tau = 1/2000$. The inversion is carried out over a sequence of coarse space-time grids with the initial mesh size $h = 1/10$ and time step size $\tau = 1/50$. The lower and upper bounds of the admissible set $\mathcal{A}$ are taken to be 0.9 and 1.5, respectively. The numerical results are shown in Table 2. Numerically we again observe a steady convergence for both e_q and e_u: the convergence rate of e_u is slightly slower than the first order; but the convergence rate of e_q is again much higher than the theoretical one $O(\delta^{1/4})$. Figure 2 shows exemplary reconstructions and the pointwise error $e := q^\dagger - q_h^*$ at two noise levels $\epsilon = 1.00\text{e}{-2}$ and 5.00e−3. Numerically, the error mostly concentrate near the domain boundary $\partial\Omega$.

A Approximation property of Lagrange interpolation

In this appendix, we prove the approximation property (3.3) and (3.4) of the Lagrange nodal interpolation operator Π_h in a convex smooth domain.

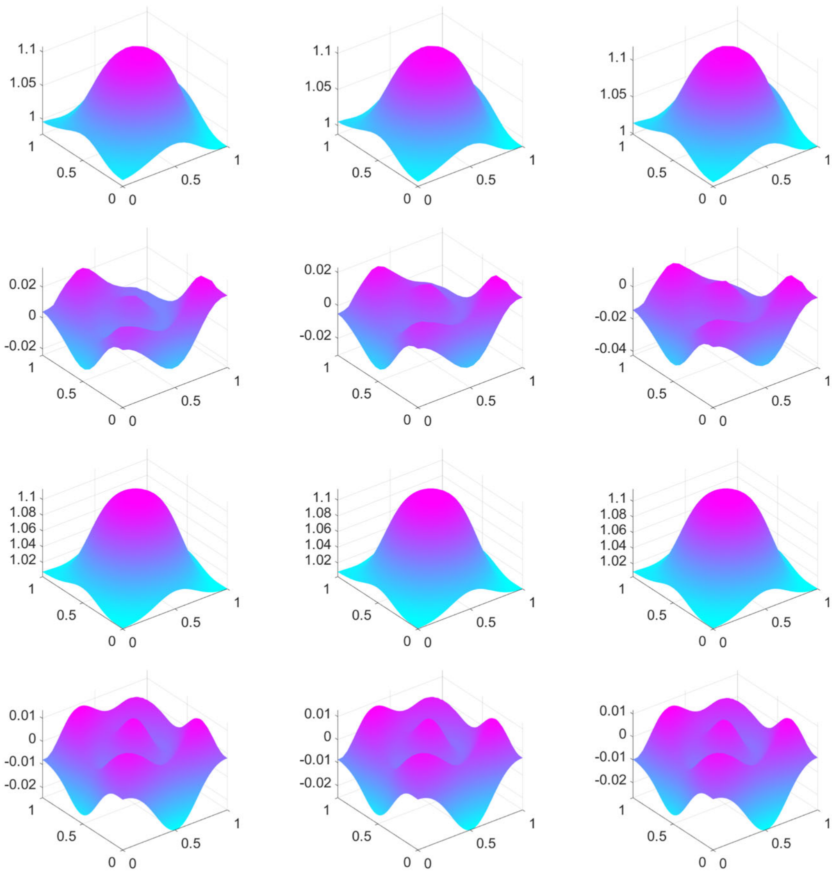


Fig. 2 The numerical reconstructions for Example 2 with $\epsilon = 3.00\text{e}{-2}$ (top two rows) and $5.00\text{e}{-3}$ (bottom two rows) and the pointwise error $e := q^\dagger - q_h^*$. From left to right, the results are for $\alpha = 0.25$, 0.50 and 0.75

Lemma 11 *Let Ω be a convex and smooth domain. Let the polygon Ω_h, the FEM space V_h, and the Langrange interpolation operator $\Pi_h : C(\overline{\Omega}) \to V_h$ be defined as in Sect. 3.1. Then the estimates (3.3) and (3.4) hold.*

Proof To show the estimate (3.4), it suffices to show

$$\|v - \Pi_h v\|_{L^\infty(\Omega\setminus\Omega_h)} + h\|\nabla(v - \Pi_h v)\|_{L^\infty(\Omega\setminus\Omega_h)} \le ch\|v\|_{W^{1,\infty}(\Omega)}.$$

By the construction of the space V_h, we observe

$$\|\nabla \Pi_h v\|_{L^\infty(\Omega\setminus\Omega_h)} \le \|\nabla \Pi_h v\|_{L^\infty(\Omega_h)} \le \|\nabla v\|_{L^\infty(\Omega_h)} \le \|\nabla v\|_{L^\infty(\Omega)}. \qquad \text{(A.1)}$$

Moreover, since $\mathrm{dist}(x, \Omega_h) \leq ch^2$ for all $x \in \partial\Omega$, from the estimate (A.1), we derive

$$\begin{aligned}\|v - \Pi_h v\|_{L^\infty(\Omega\backslash\Omega_h)} &\leq \|v - \Pi_h v\|_{L^\infty(\partial\Omega_h)} + ch^2\|\nabla(v - \Pi_h v)\|_{L^\infty(\Omega\backslash\Omega_h)} \\ &\leq ch\|v\|_{W^{1,\infty}(\Omega_h)} + ch^2\|\nabla v\|_{L^\infty(\Omega\backslash\Omega_h)} + ch^2\|\nabla \Pi_h v\|\|_{L^\infty(\Omega\backslash\Omega_h)} \\ &\leq ch\|v\|_{W^{1,\infty}(\Omega)}.\end{aligned}$$

Next, we prove the estimate (3.3). The standard trace lemma [9, Theorem 1.6.6] implies

$$\|v - \Pi_h v\|_{L^2(\partial\Omega_h)} + h\|\nabla(v - \Pi_h v)\|_{L^2(\partial\Omega_h)} \leq ch^{\frac{3}{2}}\|v\|_{H^2(\Omega_h)} \leq ch^{\frac{3}{2}}\|v\|_{H^2(\Omega)}. \tag{A.2}$$

Since $\mathrm{dist}(x, \Omega_h) \leq ch^2$ for all $x \in \partial\Omega$ and $\Pi_h v$ is piecewise linear, we have

$$\begin{aligned}\|v - \Pi_h v\|^2_{L^2(\Omega\backslash\Omega_h)} &\leq ch\|v - \Pi_h v\|^2_{L^2(\partial\Omega_h)} + ch^2\|\nabla(v - \Pi_h v)\|^2_{L^2(\Omega\backslash\Omega_h)} \\ &\leq ch\|v - \Pi_h v\|^2_{L^2(\partial\Omega_h)} + ch^3\|\nabla(v - \Pi_h v)\|^2_{L^2(\partial\Omega_h)} \\ &\quad + ch^4\|v\|^2_{H^2(\Omega\backslash\Omega_h)}.\end{aligned}$$

Then applying (A.2) gives $\|v - \Pi_h v\|_{L^2(\Omega\backslash\Omega_h)} \leq ch^2\|v\|_{H^2(\Omega)}$. The bound on $\|\nabla(v - \Pi_h v)\|_{L^2(\Omega)}$ follows similarly. □

B Conjugate gradient method

Now we briefly describe the conjugate gradient algorithm [5] for minimizing the regularized problem (3.1)–(3.2). The main effort of the algorithm at each step is to compute the gradient $J'_\gamma(q)$ of the objective J_γ. This can be achieved using the adjoint technique. Specifically, let $v(q)$ solve the modified adjoint equation [11]

$$\begin{cases} {}^R\partial^\alpha_{T+\tau} v - \nabla\cdot(q\nabla v) = (u(q)(T) - z^\delta)\delta_T(t), & \text{in } \Omega\times(0, T+\tau), \\ v = 0, & \text{on } \partial\Omega\times(0, T+\tau), \\ v(T+\tau) = 0, & \text{in } \Omega. \end{cases} \tag{B.1}$$

Here $\delta_T(\cdot)$ is a Dirac function in t concentrated at $t = T$, and ${}^R_t\partial^\alpha_T v$ is defined by ${}^R_t\partial^\alpha_T v(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{\Gamma(1-\alpha)}\int_t^T (s-t)^{-\alpha}u(s)\mathrm{d}s$. Then the $L^2(\Omega)$ gradient $J'_\gamma(q)$ is given by

$$J'_\gamma(q) = \int_0^{T+\tau} \nabla u(q)\cdot\nabla v(q)\,\mathrm{d}t - \gamma\Delta q, \tag{B.2}$$

and the descent direction $g^k = -(-\Delta + I)^{-1}J'_\gamma(q^k)$ (equipped with a zero Neumann boundary condition). The conjugate gradient direction d_k is given by

$$d_k = \beta_k d_{k-1} + g^k, \quad \text{with } \beta_k = \|g^k\|^2_{L^2(\Omega)}/\|g^{k-1}\|^2_{L^2(\Omega)}, \tag{B.3}$$

with the convention $\beta_0 = 0$. To select the step size s at Step 6, we employ a simple strategy by linearizing the direct problem (1.1) along the direction d_k. The operator $P_{\mathcal{A}}$ at line 7 denotes the pointwise projection into the set $\mathcal{A}$.

Algorithm 1 Conjugate gradient method for problem (3.1)–(3.2).

1: Set the maximum iteration number K, and choose q^0.
2: **for** $k = 1, \ldots, K$ **do**
3: Solve for $u(q^k)$ the solution to problem (1.1) with $q = q^k$.
4: Solve for $v(q^k)$ the solution to the modified adjoint problem (B.1) with $q = q^k$;
5: Compute the gradient $J'_\gamma(q^k)$ via (B.2), and the descent direction d_k via (B.3);
6: Compute the step length s_k;
7: Update the diffusion coefficient by $q^{k+1} = P_{\mathcal{A}}(q^k + s_k d_k)$;
8: Check the stopping criterion.
9: **end for**

Funding Open access funding provided by The Hong Kong Polytechnic University

References

1. Adams, E.E., Gelhar, L.W.: Field study of dispersion in a heterogeneous aquifer: 2. Spatial moments analysis. Water Resour. Res. **28**(12), 3293–3307 (1992). https://doi.org/10.1029/92WR01757
2. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces, 2nd edn. Elsevier/Academic Press, Amsterdam (2003)
3. Alessandrini, G.: A small collection of open problems. Rend. Istit. Mat. Univ. Trieste **52**, 591–600 (2020). https://doi.org/10.1007/s11118-018-9751-1
4. Alessandrini, G., Vessella, S.: Error estimates in an identification problem for a parabolic equation. Boll. Un. Mat. Ital. C (6) **4**(1), 183–203 (1985)
5. Alifanov, O.M., Artyukhin, E.A., Rumyantsev, S.V.: Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems. Begell House, New York (1995)
6. Bakaev, N.Y.: Maximum norm resolvent estimates for elliptic finite element operators. BIT **41**(2), 215–239 (2001). https://doi.org/10.1023/A:1021934205234
7. Bakaev, N.Y., Thomée, V., Wahlbin, L.B.: Maximum-norm estimates for resolvents of elliptic finite element operators. Math. Comput. **72**(244), 1597–1610 (2003). https://doi.org/10.1090/S0025-5718-02-01488-6
8. Bonito, A., Cohen, A., DeVore, R., Petrova, G., Welper, G.: Diffusion coefficients estimation for elliptic partial differential equations. SIAM J. Math. Anal. **49**(2), 1570–1592 (2017). https://doi.org/10.1137/16M1094476
9. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Springer, New York (2008). https://doi.org/10.1007/978-0-387-75934-0
10. Cheng, J., Nakagawa, J., Yamamoto, M., Yamazaki, T.: Uniqueness in an inverse problem for a one-dimensional fractional diffusion equation. Inverse Probl. **25**(11), 115,002 (2009). https://doi.org/10.1088/0266-5611/25/11/115002

11. Cheng, X., Yuan, L., Liang, K.: Inverse source problem for a distributed-order time fractional diffusion equation. J. Inverse Ill-Posed Probl. **28**(1), 17–32 (2020). https://doi.org/10.1515/jiip-2019-0006
12. Crouzeix, M., Larsson, S., Thomée, V.: Resolvent estimates for elliptic finite element operators in one dimension. Math. Comput. **63**(207), 121–140 (1994). https://doi.org/10.2307/2153565
13. Crouzeix, M., Thomée, V.: The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces. Math. Comput. **48**(178), 521–532 (1987). https://doi.org/10.2307/2007825. (https://doi-org.ezproxy.lb.polyu.edu.hk/10.2307/2007825)
14. Egger, H., Hofmann, B.: Tikhonov regularization in Hilbert scales under conditional stability assumptions. Inverse Probl. **34**(11), 115015, 17 (2018). https://doi.org/10.1088/1361-6420/aadef4
15. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of Inverse Problems. Kluwer Academic Publishers Group, Dordrecht (1996)
16. Fujita, H., Suzuki, T.: Evolution problems. In: Handbook of Numerical Analysis, Vol. II, Handb. Numer. Anal., vol. II, pp. 789–928. North-Holland, Amsterdam (1991)
17. Giona, M., Roman, H.E.: Fractional diffusion equation for transport phenomena in random media. Phys. A Stat. Mech. Appl. **185**(1–4), 87–97 (1992). https://doi.org/10.1016/0378-4371(92)90441-R
18. Hatano, Y., Hatano, N.: Dispersive transport of ions in column experiments: an explanation of long-tailed profiles. Water Resour. Res. **34**(5), 1027–1033 (1998). https://doi.org/10.1029/98WR00214
19. Isakov, V.: Inverse parabolic problems with the final overdetermination. Commun. Pure Appl. Math. **44**(2), 185–209 (1991). https://doi.org/10.1002/cpa.3160440203
20. Ito, K., Jin, B.: Inverse Problems: Tikhonov Theory and Algorithms. World Scientific Publishing Co. Pte. Ltd., Hackensack (2015)
21. Jin, B.: Fractional Differential Equations—An Approach via Fractional Derivatives, Applied Mathematical Sciences, vol. 206. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-76043-4
22. Jin, B., Zhou, Z.: Error analysis of finite element approximations of diffusion coefficient identification for elliptic and parabolic problems. SIAM J. Numer. Anal. **59**(1), 119–142 (2021). https://doi.org/10.1137/20M134383X
23. Jin, B., Zhou, Z.: Numerical estimation of a diffusion coefficient in subdiffusion. SIAM J. Control. Optim. **59**(2), 1466–1496 (2021). https://doi.org/10.1137/19M1295088
24. Jin, B., Zhou, Z.: Numerical Treatment and Analysis of Time-Fractional Evolution Equations. Springer, Cham (2023)
25. Jin, B., Zhou, Z.: Recovery of a space-time-dependent diffusion coefficient in subdiffusion: stability, approximation and error analysis. IMA J. Numer. Anal. **43**(4), 2496–2531 (2023). https://doi.org/10.1093/imanum/drac051
26. Keung, Y.L., Zou, J.: Numerical identifications of parameters in parabolic systems. Inverse Probl. **14**(1), 83–100 (1998). https://doi.org/10.1088/0266-5611/14/1/009
27. Kilbas, A.A., Srivastava, H.M., Trujillo, J.J.: Theory and Applications of Fractional Differential Equations. Elsevier Science B.V., Amsterdam (2006)
28. Larsson, S., Thomée, V.: Partial Differential Equations with Numerical Methods. Springer, Berlin (2003)
29. Li, B., Sun, W.: Maximal L^p analysis of finite element solutions for parabolic equations with nonsmooth coefficients in convex polyhedra. Math. Comput. **86**(305), 1071–1102 (2017). https://doi.org/10.1090/mcom/3133
30. Li, G., Gu, W., Jia, X.: Numerical inversions for space-dependent diffusion coefficient in the time fractional diffusion equation. J. Inverse Ill-Posed Probl. **20**(3), 339–366 (2012). https://doi.org/10.1515/jip-2011-0012
31. Li, G., Zhang, D., Jia, X., Yamamoto, M.: Simultaneous inversion for the space-dependent diffusion coefficient and the fractional order in the time-fractional diffusion equation. Inverse Probl. **29**(6), 065014, 36 (2013). https://doi.org/10.1088/0266-5611/29/6/065014
32. Lions, J.L., Magenes, E.: Non-Homogeneous Boundary Value Problems and Applications. Vol. I. Die Grundlehren der mathematischen Wissenschaften, Band 181. Springer, New York–Heidelberg (1972). (**Translated from the French by P. Kenneth**)
33. Lubich, C.: Discretized fractional calculus. SIAM J. Math. Anal. **17**(3), 704–719 (1986). https://doi.org/10.1137/0517050
34. Metzler, R., Klafter, J.: The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Rep. **339**(1), 1–77 (2000). https://doi.org/10.1016/S0370-1573(00)00070-3
35. Nigmatullin, R.R.: The realization of the generalized transfer equation in a medium with fractal geometry. Physica Status Solidi (B) **133**(1), 425–430 (1986). https://doi.org/10.1002/pssb.2221330150

36. Palencia, C.: Maximum norm analysis of completely discrete finite element methods for parabolic problems. SIAM J. Numer. Anal. **33**(4), 1654–1668 (1996). https://doi.org/10.1137/S0036142993259779
37. Stewart, H.B.: Generation of analytic semigroups by strongly elliptic operators. Trans. Am. Math. Soc. **199**, 141–162 (1974). https://doi.org/10.2307/1996879
38. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, 2nd edn. Springer, Berlin (2006)
39. Triki, F.: Coefficient identification in parabolic equations with final data. J. Math. Pures Appl. **9**(148), 342–359 (2021). https://doi.org/10.1016/j.matpur.2021.02.004
40. Wang, L., Zou, J.: 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). https://doi.org/10.3934/dcdsb.2010.14.1641
41. Yamamoto, M., Zou, J.: Simultaneous reconstruction of the initial temperature and heat radiative coefficient. Inverse Probl. **17**(4), 1181 (2001). https://doi.org/10.1088/0266-5611/17/4/340
42. Zhang, Z.: An undetermined coefficient problem for a fractional diffusion equation. Inverse Probl. **32**(1), 015011, 21 (2016). https://doi.org/10.1088/0266-5611/32/1/015011
43. Zhang, Z., Zhang, Z., Zhou, Z.: Identification of potential in diffusion equations from terminal observation: analysis and discrete approximation. SIAM J. Numer. Anal. **60**(5), 2834–2865 (2022). https://doi.org/10.1137/21M1446708

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Bangti Jin[1] · Xiliang Lu[2] · Qimeng Quan[3] · Zhi Zhou[4]

✉ Zhi Zhou
zhizhou@polyu.edu.hk

Bangti Jin
b.jin@cuhk.edu.hk

Xiliang Lu
xllv.math@whu.edu.cn

Qimeng Quan
qimengquan@um.edu.mo

[1] Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, People's Republic of China

[2] School of Mathematics and Statistics, National Center for Applied Mathematics in Hubei and Hubei Computational Science Key Laboratory, Wuhan University, Wuhan 430072, People's Republic of China

[3] State Key Laboratory of Internet of Things for Smart City, University of Macau, Avenida da Universidade, Taipa, Macau, People's Republic of China

[4] Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, People's Republic of China