Numerical Reconstruction of the Spatial Component in the Source Term of a Time-Fractional Diffusion Equation

Daijun Jiang* Yikan Liu† Dongling Wang‡

ABSTRACT

In this article, we are concerned with the analysis on the numerical reconstruction of the spatial component in the source term of a time-fractional diffusion equation. This ill-posed problem is solved through a stabilized nonlinear minimization system by an appropriately selected Tikhonov regularization. The existence and the stability of the optimization system are demonstrated. The nonlinear optimization problem is approximated by a fully discrete scheme, whose convergence is established under a novel result verified in this study that the $\mathcal{H}^1$-norm of the solution to the discrete forward system is uniformly bounded. The iterative thresholding algorithm is proposed to solve the discrete minimization, and several numerical experiments are presented to show the efficiency and the accuracy of the algorithm.

Keywords Time-fractional diffusion equation, Inverse source problem, Finite element method, Iterative thresholding algorithm

Mathematics subject classification (2010) 35R11, 65M32, 41A35

1 Introduction

Let $T > 0$ and $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) be a bounded convex polygonal domain. For $0 < \alpha < 1$, by $\partial^\alpha_t$ we denote the Caputo derivative defined as (see, e.g., [10, p. 91])

$$\partial^\alpha_t g(t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{g'(s)}{(t-s)^\alpha} \, ds,$$

where $\Gamma(\cdot)$ is the usual Gamma function. In this paper, we consider the following initial-boundary value problem for a time-fractional diffusion equation (TFDE) with the homogeneous Neumann boundary condition

$$\begin{cases}
(\partial^\alpha_t u - \triangle + 1)u(x,t) = f(x)\mu(t), & (x,t) \in \Omega \times (0,T), \\
u u(x,t) = 0, & (x,t) \in \partial \Omega \times (0,T),
\end{cases}$$

(1.1)

where $\partial_x u := \nabla u \cdot \nu$ and $\nu = (\nu_1, \ldots, \nu_d)$ denotes the unit outward normal derivative on the boundary $\partial \Omega$. The source term in (1.1) takes the form of separated variables, in which the spatial component $f(x)$ models the spatial distribution e.g. of the contaminant source, and the temporal component $\mu(t)$ describes the time evolution pattern.

The governing equation in (1.1) is a typical representative of a wide range of TFDEs, which were proposed as a powerful candidate for describing anomalous diffusion phenomena in heterogeneous media. In the past decades, TFDEs have been investigated thoroughly from both theoretical and numerical aspects. It reveals from [5,10,19] and other literature that TFDEs resemble their integer counterpart (i.e., $\alpha = 1$) in such qualitative aspects like analyticity and maximum principle, but show certain difference in the senses of limited smoothing effect, short-/long-time asymptotic behavior and weak unique continuation. For the numerical simulation of TFDEs, we refer e.g. to [6,9,11,14,21].

*School of Mathematics and Statistics & Hubei Key Laboratory of Mathematical Sciences, Central China Normal University, Wuhan 430079, China. E-mail: jiangdaijun@mail.ccnu.edu.cn
†Corresponding author. Research Center of Mathematics for Social Creativity, Research Institute for Electronic Science, Hokkaido University, N12W7, Kita-Ward, Sapporo 060-0812, Japan. E-mail: ykliu@es.hokudai.ac.jp
‡Department of Mathematics and Center for Nonlinear Studies, Northwest University, Xi’an, Shaanxi 710075, China. E-mail: wdymath@nwu.edu.cn
Simultaneously, various kinds of inverse problems for TFDEs have also gathered consistent popularity within the last decade. Owing to the limited smoothing property of the forward problem, usually the ill-posedness of inverse problems for TFDEs is less severe than that for the case of \( \alpha = 1 \), which can be witnessed from the backward problems. Nevertheless, due to the lack of techniques for analysis, for some inverse problems, one should relate TFDEs with other types of equations, so that one can indirectly obtain uniqueness and sometimes stability. For the inverse source problems on determining the spatial component, it turns out that the majority of existing literature dealt with the final observation data \( u(\cdot, T) \) except for \([29]\). We refer to Sakamoto and Yamamoto \([20]\) for the generic well-posedness, and \([21-26]\) etc. for the numerical reconstruction by various regularization methods. Recently, by a newly established unique continuation property for TFDEs, Jiang et al. \([5]\) proved the uniqueness for the same problem by the partial interior observation and developed an iterative thresholding algorithm. On reconstructing the temporal component in the source term, we refer e.g. to \([17, 18, 25]\). For a complete bibliography on inverse problems for TFDEs, see the topical review articles \([5, 12, 13, 15]\).

In the same direction of \([5]\), in this paper we mainly focus on the numerical aspect of the following inverse source problem.

**Problem 1.1.** Let \( \omega \subset \Omega \) be a nonempty open subdomain and \( u(f) \) be the solution to \([1.1]\). Provided that \( \mu \) is known on \([0, T]\), determine the spatial component \( f \) in \( \Omega \) by the partial interior observation of \( u(f) \) in \( \omega \times (0, T) \).

In fact, here we can also consider a more general formulation such as

\[
\partial_t^\alpha u(x, t) - \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u(x, t)}{\partial x_j} \right) + c(x)u(x, t) = f(x)\mu(t),
\]

where we assume certain regularity of \( a_{ij} \) and \( c \), and \((a_{ij})_{d \times d}\) is a strictly positive-definite matrix on \( \Omega \). However, we restrict ourselves to the formulation \([1.1]\) not only for its simplicity in the numerical simulation, but also due to the belief that the underlying ill-posedness are essentially the same.

On the other hand, we notice that the iteration method proposed in \([5]\) lacks detailed analysis especially on the convergence issue, although the numerical performance was discussed. From the practical viewpoint, it is obligatory to reformulate Problem \([1.1]\) in a discrete setting and investigate whether the resulting system inherits the corresponding properties of the continuous one. Moreover, we should verify the convergence of the discretized solution to the continuous one in some sense. This motivates the central topic of the present article, which turns out to be a very important and necessary supplementation to the investigation of Problem \([1.1]\).

To fully discretize system \([1.1]\), we shall employ the standard Galarkin method with piecewise linear finite element in space and the L1 scheme in time. This method is one of the most popular and successful numerical scheme for discretizing the subdiffusion problem, and it has been analyzed from various aspects. The optimal error estimate in \( L^2(\Omega) \) norm with respect to the regularity of the problem solution is established in \([6]\) for uniform time step size. In order to deal with the weak singularity of time-fractional diffusion equations, the L1-type scheme on graded time meshes is employed in the discrete scheme, and the optimal error estimate in \( L^\infty(\Omega) \) norm is established for \( d = 1 \) in \([21]\) and \( d = 2, 3 \) in \([9]\). For the convergence analysis, in this paper we further establish the uniform \( H^1(\Omega) \) error estimate on \([0, T]\) by using the Poincaré-Friedrichs inequality and the discrete fractional Grönwall’s inequality for the L1 method, which turns out to be novel to the best of our knowledge.

The remainder of this paper is organized as follows. Recalling key ingredients in theory, in Section \( 2 \) we interpret Problem \([1.1]\) as an optimization problem with Tikhonov regularization, and demonstrate its regularization effects. In Section \( 3 \) we discretize the optimization problem by a fully discrete finite element approximation and study its basic properties. Next, Section \( 4 \) is devoted to the convergence analysis of the solution to the discrete optimization problem. Finally, we propose the iterative thresholding algorithm to solve the discrete problem in Section \( 5 \) and implement several numerical examples to show the efficiency and accuracy of the algorithm.

### 2 Preliminary and Tikhonov regularization

In this section, we shall make general preparations and formulate Problem \([1.1]\) stated in Section \( 1 \) as a stabilized minimization system, and establish the unique existence of the solution as well as the stability of the minimization formulation.

Let \( L^2(\Omega) \) be a usual \( L^2 \)-space equipped with the inner product \((\cdot, \cdot)\), and \( H^\alpha(0, T), H^\gamma(\Omega) \) (\( \gamma \in \mathbb{R} \)) etc. denote Sobolev spaces (e.g., Adams \([1]\)). Throughout this paper, \( C > 0 \) stands for generic constants.
which may change line by line. In a Banach space $X$, we denote the weak convergence of a sequence $\{z_n\}$ to $z$ by

\[ z_n \rightharpoonup z \quad \text{in} \quad X \quad \text{as} \quad n \to \infty. \]

We first revisit some basic facts concerning the initial-boundary value problem (1.1). For the solution regularity, we define the fractional Sobolev spaces $\mathcal{H}^\alpha(0, T)$ as

\[ \mathcal{H}^\alpha(0, T) := \begin{cases} H^\alpha(0, T), & 0 < \alpha < \frac{1}{2}, \\ \left\{ g \in H^\frac{1}{2}(0, T); \int_0^T \frac{|g(t)|^2}{t} \, dt < \infty \right\}, & \alpha = \frac{1}{2}, \\ \left\{ g \in H^\alpha(0, T); g(0) = 0 \right\}, & \frac{1}{2} < \alpha < 1. \end{cases} \]

We also define the Riemann-Liouville integral operator $J_0^\alpha$ and the backward Riemann-Liouville integral operator $J_1^\alpha$ respectively as

\[ J_0^\alpha g(t) := \frac{1}{\Gamma(\alpha)} \int_0^t \frac{g(\tau)}{(t-\tau)^{1-\alpha}} \, d\tau, \quad J_1^\alpha g(t) := \frac{1}{\Gamma(\alpha)} \int_t^1 \frac{g(\tau)}{(\tau-t)^{1-\alpha}} \, d\tau \]

for $g \in L^2(0, T)$. We know that the Caputo derivative can be rewritten as $\partial_t^{\alpha} g = J_0^{1-\alpha} \frac{d}{dt} g$, and the following relation holds (see [5, Lemma 4.1])

\[ \int_0^T (J_0^\alpha g_1) g_2 \, dt = \int_0^T g_1 J_1^\alpha g_2 \, dt, \quad \forall g_1, g_2 \in L^2(0, T). \quad (2.1) \]

For later use, we show a technical lemma on the basis of (2.1)

**Lemma 2.1.** Let $g \in \mathcal{H}^\alpha(0, T) \cap C[0, T]$ and $h \in C^1[0, T]$ with $J_1^{1-\alpha} h(T) = 0$. Then

\[ \int_0^T (\partial_t^{\alpha} g) h \, dt = -g(0) J_1^{1-\alpha} h(0) - \int_0^T g (J_1^{1-\alpha} h)' \, dt. \]

**Proof.** Pick a sequence $\{g_n\} \subset C^\infty[0, T]$ such that $g_n \rightharpoonup g$ in $\mathcal{H}^\alpha(0, T) \cap C[0, T]$. Then for each $n = 1, 2, \ldots$, we utilize (2.1) and calculate

\[
\int_0^T (\partial_t^{\alpha} g_n) h \, dt = \int_0^T (J_0^{1-\alpha} g_n)' h \, dt = \int_0^T g_n J_1^{1-\alpha} h \, dt = g_n J_1^{1-\alpha} h \big|_{t=0}^{t=T} - \int_0^T g_n (J_1^{1-\alpha} h)' \, dt \\
= -g_n(0) J_1^{1-\alpha} h(0) - \int_0^T g_n (J_1^{1-\alpha} h)' \, dt,
\]

where the last equality follows from the assumption $J_1^{1-\alpha} h(T) = 0$. Then the proof is completed by passing $n \to \infty$ on both sides of the above identity.

On the basis of [5, Lemma 2.4] and [19, Theorem 2.2(i)], we summarize the well-posedness of the forward problem (1.1) as follows.

**Lemma 2.2.** Let $f \in L^2(\Omega)$ and $\mu \in L^\infty(0, T)$. Then the initial-boundary value problem (1.1) admits a unique solution $u(f) \in \mathcal{H}^\alpha(0, T; L^2(\Omega)) \cap C([0, T]; H^\alpha(\Omega)) \cap L^2(0, T, H^\alpha(\Omega))$. Moreover, there exists a constant $C > 0$ depending on $\Omega, T, \alpha$ and $\mu$ such that

\[ \|u(f)\|_{\mathcal{H}^\alpha(0, T; L^2(\Omega))} + \|u(f)\|_{C([0, T]; H^\alpha(\Omega))} + \|u(f)\|_{L^2(0, T, H^\alpha(\Omega))} \leq C\|f\|_{L^2(\Omega)}. \]

In [19, Theorem 2.2(i)], it was proved that the solution $u \in C([0, T]; H^{-2\gamma}(\Omega))$ for any $\gamma > \frac{d}{4} - 1$ provided that the source term belongs to $L^\infty(0, T; L^2(\Omega))$. Since we restrict the spatial dimensions to $d = 1, 2, 3$ throughout this paper, one can choose negative $\gamma$ and thus $u(f)$ makes sense in $C([0, T]; L^2(\Omega))$ under the assumptions of Lemma 2.2.

For the theoretical aspect of Problem (1.1), we recall the uniqueness result stated in [3, Theorem 2.6].

**Lemma 2.3.** Let $\omega \subset \Omega$ be an arbitrarily chosen open subdomain and $u(f)$ be the solution to (1.1). Assume that $f \in L^2(\Omega)$ and $\mu \in C^1[0, T]$ with $\mu(0) \neq 0$. Then $u = 0$ in $\omega \times (0, T)$ implies $f = 0$ in $\Omega$. 

3
Suppose that we are given the noisy observation data \( u^\delta \in L^2(\omega \times (0,T)) \) in practice. Usually, \( u^\delta \) satisfies
\[
\|u^\delta - u(f^*)\|_{L^2(\omega \times (0,T))} \leq \delta,
\]
where \( f^* \in L^2(\Omega) \) and \( \delta > 0 \) stand for the true source term and the noise level, respectively. To deal with the ill-posedness of Problem [1.1] we still adopt a classical Tikhonov regularization methodology as that in [5] to consider the following optimization problem with Tikhonov regularization
\[
\min_{f \in L^2(\Omega)} J(f), \quad J(f) := \|u(f) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \|f\|^2_{L^2(\Omega)},
\]
where \( \beta > 0 \) is the regularization parameter, and \( u = u(f) \) satisfies the initial condition \( u(x,0) = 0 \) and the variational system associated with the equation (1.1), for any \( f \in H^\alpha(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega)) \),
\[
\int_0^T \int_\Omega (\partial_t^\alpha u(f)) \varphi + \nabla u(f) \cdot \nabla \varphi + u(f) \varphi \, dx \, dt = \int_0^T \int_\Omega f \mu \varphi \, dx \, dt.
\]

Regarding the minimizer of (2.2), we first show the following unique existence result.

**Theorem 2.1.** For any \( u^\delta \in L^2(\omega \times (0,T)) \), there exists a unique minimizer \( f^* \in L^2(\Omega) \) to the optimization problem (2.2).

**Proof.** Since \( J(f) \) is nonnegative, we know that \( \inf_{f \in L^2(\Omega)} J(f) \) is finite. Thus there exists a minimizing sequence \( \{f^n\} \subset L^2(\Omega) \) such that
\[
\lim_{n \to \infty} J(f^n) = \inf_{f \in L^2(\Omega)} J(f).
\]
Then by the definition of \( J(f^n) \), it is obvious that \( \{f^n\} \) is uniformly bounded in \( L^2(\Omega) \). Therefore, there exist \( f^* \in L^2(\Omega) \) and a subsequence of \( \{f^n\} \), still denoted by \( \{f^n\} \), such that
\[
f^n \rightharpoonup f^* \quad \text{in} \quad L^2(\Omega) \quad \text{as} \quad n \to \infty.
\]
We shall prove that \( f^* \) is indeed the unique minimizer of (2.2).

Since each \( f^n \) corresponds with a solution \( u(f^n) \) to (1.1) with \( f = f^n \), it follows immediately from Lemma 2.2 that the sequence \( \{u(f^n)\} \) is also uniformly bounded in \( \partial_t^\alpha(u^n) \cap L^2(0,T;H^1(\Omega)) \). This, together with the compact embedding \( H^2(\Omega) \subset \subset L^2(\Omega) \), indicates the existence of some \( u^* \in \partial_t^\alpha(u^n) \cap L^2(0,T;H^2(\Omega)) \) and a subsequence of \( \{u(f^n)\} \), again still denoted by \( \{u(f^n)\} \), such that
\[
u(f^n) \rightharpoonup u^* \quad \text{in} \quad \partial_t^\alpha(u^n) \cap L^2(0,T;H^2(\Omega)).
\]

We claim \( u^* = u(f^*) \). Actually, we utilize the fact that
\[
\int_0^T \int_\Omega (\partial_t^\alpha u(f^n)) \varphi + \nabla u(f^n) \cdot \nabla \varphi + u(f^n) \varphi \, dx \, dt = \int_0^T \int_\Omega f^n \mu \varphi \, dx \, dt
\]
for all \( \varphi \in L^2(0,T;H^1(\Omega)) \). Since (2.4) implies
\[
\partial_t^\alpha u(f^n) \rightharpoonup \partial_t^\alpha u^*, \quad \nabla u(f^n) \rightharpoonup \nabla u^* \quad \text{in} \quad L^2(Q),
\]
we pass \( n \to \infty \) in (2.5) to obtain
\[
\int_0^T \int_\Omega (\partial_t^\alpha u^* \varphi + \nabla u^* \cdot \nabla \varphi + u^* \varphi) \, dx \, dt = \int_0^T \int_\Omega f^* \mu \varphi \, dx \, dt
\]
for all \( \varphi \in L^2(0,T;H^1(\Omega)) \). Next, we shall prove \( u^*(\cdot,0) = 0 \), which with (2.6) and the definition (2.3) of weak solutions implies \( u^* = u(f^*) \).

Indeed, it follows from Lemma 2.2 that \( u(f^n) \in \partial_t^\alpha(u^n) \cap L^2(\Omega) \cap C([0,T];L^2(\Omega)) \), which satisfies the assumption of Lemma 2.4. Then by taking \( \psi \in C^1(0,T) \) with \( J_\psi^\delta \to \psi(T) = 0 \) and \( v \in L^2(\Omega) \) arbitrarily, we employ Lemma 2.1 to deduce
\[
\int_\Omega \int_0^T \partial_t^\alpha u(f^n) \psi \, dx \, dt = - \int_\Omega u(f^n)(\cdot,0) J_\psi^\delta \psi(0) \, v \, dx - \int_\Omega \int_0^T u(f^n) \frac{d}{dt} J_\psi^\delta \psi \, v \, dx \, dt.
\]
Noting that \( u(f^n)(x,0) = 0 \) and letting \( n \to \infty \) in the above equation, we obtain
\[
\int_\Omega \int_0^T \partial_t^\alpha u^* \psi \, dx \, dt = - \int_\Omega \int_0^T u^* \frac{d}{dt} J_\psi^\delta \psi \, v \, dx \, dt.
\]
In a parallel manner, we also have

\[ \int_0^T \int_{\Omega} \partial_t^\alpha u \cdot \psi \, dtdx = - \int_0^T \int_{\Omega} u^*(\cdot,0) J_{L^2}^{1-\alpha}(\psi(0)) v \, dtdx - \int_0^T \int_{\Omega} u^* \frac{d}{dt} J_{L^2}^{1-\alpha}(\psi) v \, dtdx \]

for any \( \psi \in C^1[0,T] \) with \( J_{L^2}^{1-\alpha}(\psi(T)) = 0 \) and \( v \in L^2(\Omega) \), which, together with \( (2.8) \), implies that \( u^*(\cdot,0) = 0 \).

By \( f^n \to f^* \) in \( L^2(\Omega) \) and \( (2.4) \), we employ the lower semi-continuity of the \( L^2 \)-norm to conclude

\[
J(f^*) = \|u(f^*) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \|f^*\|^2_{L^2(\Omega)}
\leq \liminf_{n \to \infty} \|u(f^n) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \liminf_{n \to \infty} \|f^n\|^2_{L^2(\Omega)}
\leq \liminf_{n \to \infty} J(f^n) = \inf_{f \in L^2(\Omega)} J(f),
\]

indicating that \( f^* \) is indeed a minimizer to the optimization problem \( (2.2) \). Furthermore, the uniqueness of \( f^* \) is readily seen from the convexity of \( J(f) \).

Next, we justify the stability of \( (2.2) \); namely, the minimization system \( (2.2) \) is indeed a stabilization for Problem \( (1.1) \) with respect to the perturbation in observation data.

**Theorem 2.2.** Let \( \{u^n_\ell\} \subset L^2(\omega \times (0,T)) \) be a sequence such that

\[ u^n_\ell \to u^\delta \quad \text{in} \quad L^2(0,T;L^2(\omega)) \quad \text{as} \quad \ell \to \infty, \quad (2.8) \]

and \( \{f^n_\ell\} \) be a sequence of minimizers of problems

\[ \min_{f \in L^2(\Omega)} J_\ell(f), \quad J_\ell(f) := \|u(f) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \|f\|^2_{L^2(\Omega)}, \quad \ell = 1, 2, \ldots. \]

Then \( \{f^n_\ell\} \) converges strongly in \( L^2(\Omega) \) to the minimizer of \( (2.2) \).

**Proof.** The unique existence of each \( f^n_\ell \) is guaranteed by Theorem \( 2.1 \). By definition, we have

\[ J_\ell(f^n_\ell) \leq J_\ell(f), \quad \forall f \in L^2(\Omega), \]

which implies the uniform boundedness of \( f^n_\ell \) in \( L^2(\Omega) \). Hence, there exist \( f^* \in L^2(\Omega) \) and a subsequence of \( \{f^n_\ell\} \), still denoted by \( \{f^n_\ell\} \), such that

\[ f^n_\ell \to f^* \quad \text{in} \quad L^2(\Omega) \quad \text{as} \quad \ell \to \infty. \]

Now it suffices to show that \( f^* \) is indeed the unique minimizer of \( (2.2) \). Actually, repeating the same argument as that in the proof of Theorem \( 2.1 \), we can derive

\[ u(f^n_\ell) \to u(f^*) \quad \text{in} \quad H^\alpha(0,T;L^2(\Omega)) \cap L^2(0,T;H^2(\Omega)) \quad \text{as} \quad \ell \to \infty, \]

up to taking a further subsequence. Combining the above convergence with \( (2.8) \), we obtain

\[ u(f^n_\ell) - u^\delta \to u(f^*) - u^\delta \quad \text{in} \quad L^2(0,T;L^2(\omega)) \quad \text{as} \quad \ell \to \infty. \]

Therefore, we get

\[ \|u(f^*) - u^\delta\|^2_{L^2(\omega \times (0,T))} \leq \liminf_{\ell \to \infty} \left( \|u(f^n_\ell) - u^\delta\|^2_{L^2(\omega \times (0,T))} \right). \quad (2.9) \]

For any \( f \in L^2(\Omega) \), again we take advantage of the lower semi-continuity of the \( L^2 \)-norm to deduce

\[
J(f^*) = \|u(f^*) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \|f^*\|^2_{L^2(\Omega)}
\leq \liminf_{\ell \to \infty} \|u(f^n_\ell) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \liminf_{\ell \to \infty} \|f^n_\ell\|^2_{L^2(\Omega)}
\leq \liminf_{\ell \to \infty} (\|u(f^n_\ell) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \|f^n_\ell\|^2_{L^2(\Omega)})
\leq \lim_{\ell \to \infty} (\|u(f) - u^\delta\|^2_{L^2(\omega \times (0,T))} + \beta \|f\|^2_{L^2(\Omega)})
\]

5
which verifies that \( f^* \) is the minimizer of (2.2).

Next, we shall prove \( \{ f^\ell \} \) converges to \( f^* \) strongly in \( L^2(\Omega) \) by contradiction. Assuming that it is not true, then we know that \( \{ \| f^\ell \|_{L^2(\Omega)} \} \) does not converge to \( \| f^* \|_{L^2(\Omega)} \). As \( f^\ell \to f^* \) in \( L^2(\Omega) \), by the weak lower semi-continuity of the norm, we have

\[
\| f^* \|_{L^2(\Omega)} \leq \liminf_{\ell \to \infty} \| f^\ell \|_{L^2(\Omega)}.
\]

Hence, setting \( A := \limsup_{\ell \to \infty} \| f^\ell \|_{L^2(\Omega)} \), we get

\[
A = \limsup_{\ell \to \infty} \| f^\ell \|_{L^2(\Omega)} > \liminf_{\ell \to \infty} \| f^\ell \|_{L^2(\Omega)} \geq \| f^* \|_{L^2(\Omega)},
\]

and there exists a subsequence \( \{ f^m \} \) of \( \{ f^\ell \} \) such that

\[
A = \lim_{m \to \infty} \| f^m \|_{L^2(\Omega)}.
\]

Now taking \( f = f^* \) in (2.10), we find that

\[
\| u(f^*) - u^\delta \|_{L^2(\omega \times (0,T))}^2 + \beta \| f^* \|_{L^2(\Omega)}^2 = \liminf_{m \to \infty} \left( \| u(f^m) - u^\delta \|_{L^2(\omega \times (0,T))}^2 + \beta \| f^m \|_{L^2(\Omega)}^2 \right) \\
= \liminf_{m \to \infty} \| u(f^m) - u^\delta \|_{L^2(\omega \times (0,T))}^2 + \beta \| f^* \|_{L^2(\Omega)}^2,
\]

This, together with (2.11), implies that

\[
\| u(f^*) - u^\delta \|_{L^2(\omega \times (0,T))}^2 = \liminf_{m \to \infty} \| u(f^m) - u^\delta \|_{L^2(\omega \times (0,T))}^2 + \beta \left( A^2 - \| f^* \|_{L^2(\Omega)}^2 \right) \\
> \liminf_{m \to \infty} \| u(f^m) - u^\delta \|_{L^2(\omega \times (0,T))}^2,
\]

which contradicts with (2.10). The proof of Theorem 2.2 is completed.

\[ \square \]

3 Fully Discrete Finite Element Approximation

In this section we propose a fully discrete finite element method for approximating the nonlinear optimization problem (2.2). We first introduce some appropriate time and space discretization. For the space discretization, we consider a shape regular triangulation \( \mathcal{T}_h \) of \( \Omega \) with a mesh size \( h \), consisting of tetrahedral elements (see [2]). Then we introduce the standard nodal finite element spaces of piecewise linear functions

\[
V_h = \{ v_h \in H^1(\Omega); v_h|_A \in P_1(A), \ \forall A \in \mathcal{T}_h \}, \\
S_h = \{ f_h \in L^2(\Omega); f_h|_A \in P_1(A), \ \forall A \in \mathcal{T}_h \},
\]

where \( P_1(A) \) is the space of linear polynomials on \( A \). It is obvious that \( V_h \subset S_h \).

To fully discretize the minimization problem (2.2), we also need the time discretization. To this end, we divide the time interval \([0, T]\) into \( M \) subintervals by the equidistant nodal points

\[
0 = t_0 < t_1 < \cdots < t_M = T
\]

with \( t_m = m \tau \), where \( \tau = T/M \) is the step length. Hereinafter, we additionally assume that the noisy observation data \( u^\delta \) satisfies \( u^\delta \in C([0, T]; L^2(\omega)) \), so that each \( u^\delta(\cdot, t_m) \) makes sense in \( L^2(\omega) \).

To discretize the Caputo derivative in time, we refer to [7] and use the following approximation:

\[
\partial_t^\alpha u(x, t_{m+1}) = \frac{1}{\Gamma(1 - \alpha)} \sum_{j=0}^{m} \int_{t_j}^{t_{j+1}} \frac{\partial_s u(x, s)}{(t_{m+1} - s)^\alpha} \, ds \\
= \frac{1}{\Gamma(1 - \alpha)} \sum_{j=0}^{m} \frac{u(x, t_{j+1}) - u(x, t_j)}{s} \int_{t_j}^{t_{j+1}} (t_{m+1} - s)^{-\alpha} \, ds + r_{\tau_{m+1}}^m \\
= \frac{1}{\Gamma(2 - \alpha)} \sum_{j=0}^{m} d_j \frac{u(x, t_{m+1} - j) - u(x, t_{m-j})}{\tau^\alpha} + r_{\tau_{m+1}}^m
\]
\begin{equation}
\gamma_j = \begin{cases} 
1, & j = 0, \\
(1 + \alpha) - 2j^{1 - \alpha} + (j - 1)^{1 - \alpha}, & j = 1, \ldots, m, \\
m^{1 - \alpha} - (m + 1)^{1 - \alpha}, & j = m + 1.
\end{cases}
\end{equation}

Following the same line as in [7], now we arrive at a fully discrete scheme for the forward problem 3.1: given \( u_h^0 = 0 \), find \( u_h^{m+1} \in V_h \) (\( m = 0, \ldots, M - 1 \)) such that

\begin{equation}
(1 + b_0) \left( u_h^{m+1}, \chi_h \right) + b_0 \left( \nabla u_h^{m+1}, \nabla \chi_h \right) = \sum_{j=0}^{m-1} \left( d_j - d_{j+1} \right) \left( u_h^{m-j}, \chi_h \right) + b_0 \left( f_h \mu^{m+1}, \chi_h \right)
\end{equation}

or equivalently, find \( u_h^m \in V_h \) (\( m = 1, \ldots, M \)) such that

\begin{equation}
\left( \partial_t^\alpha u_h^m, \chi_h \right) + \left( \nabla u_h^m, \nabla \chi_h \right) + \left( u_h^m, \chi_h \right) = \left( f_h \mu^m, \chi_h \right)
\end{equation}

holds for all \( \chi_h \in V_h \), where \( b_0 := \Gamma(2 - \alpha) \). Here and hereinafter, we understand \( f_h \in S_h \) as some approximation of \( f \) and \( \mu^m = \mu(t_m) \). To emphasize the dependency, we also denote the solution to 3.2 or 3.3 as \( u_h^m(f_h) \).

On the basis of 3.2, now we are well prepared to propose the fully discrete finite element approximation of the nonlinear optimization 2.2:

\begin{equation}
\min_{h \in S_h} J_h,\tau(f_h), \quad J_h,\tau(f_h) := \sum_{m=0}^{M} c_m \left\| u_h^m(f_h) - u^\delta(\cdot, t_m) \right\|^2_{L^2(\Omega)} + \beta \| f_h \|^2_{L^2(\Omega)}.
\end{equation}

Here \( c_m \) (\( m = 0, 1, \ldots, M \)) are the coefficients of the composite trapezoidal rule for the time integration over \([0, T]\), i.e., \( c_0 = c_M = \frac{1}{2} \) and \( c_m = 1 \) for \( m = 1, \ldots, M - 1 \).

Before verifying the existence of a minimizer to the discrete minimization problem 3.4, we shall derive some useful a priori estimates for the discrete solution \( u_h^m(f_h) \) to 3.2 or 3.3. To do so, we need the following lemmas which allows us to use the energy-like methods in the analysis.

**Lemma 3.1** (see [23]). For the L1 numerical approximation 3.3 of the Caputo derivative, there holds

\begin{equation}
\left\| \partial_t^\alpha u_h^m \right\|^2_{L^2(\Omega)} \leq \left( 2u_h^m, \partial_t^\alpha u_h^m \right), \quad m \geq 1.
\end{equation}

**Lemma 3.2** (see [11]). Suppose that the nonnegative sequences \( \{v^n\} \) and \( \{g^n\} \) satisfy

\begin{equation}
\partial_t^\alpha v^n \leq \lambda_1 v^n + \lambda_2 v^{n-1} + g^n, \quad n \geq 1,
\end{equation}

where \( \lambda_1, \lambda_2 \) are given positive constants independent of the time step \( \tau \). Then there exists a positive constant \( \tau_\ast > 0 \) such that

\begin{equation}
v^n \leq 2 \left( e^{\theta} + \frac{\theta^n}{\Gamma(1 + \alpha)} \max_{0 \leq j \leq n} g^j \right) E_\alpha(2\lambda\tau_\ast^n)
\end{equation}

for \( 0 < \tau < \tau_\ast \), where \( E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k + 1 + \alpha)} \) is the Mittag-Leffler function and \( \lambda = \lambda_1 + \lambda_2/\tau_\ast^{1+\alpha} \).

Just as the exponential function plays a fundamental role in the integral order equation, the Mittag-Leffler function naturally appears in the fractional order calculus, and it has many nice properties similar to the exponential function. In fact, we have \( E_\alpha(z) = e^z \) for all \( z \in \mathbb{C} \). Let us mention a couple of properties that we will use later, and refer to the recent monograph [11] for further details.

\footnote{For historical reasons, the numerical scheme in 3.4 is often referred to as L1 numerical approximation for Caputo fractional derivative. One possible explanation is that in this scheme we replace the first derivative with a first order linear interpolation in each cell \([t_j, t_{j+1}]\), and is called L1 scheme. This name has nothing to do with the usual notation \( L^1(\Omega) \) for integrable function in \( \Omega \).}
For $\alpha > 0$, the Mittag-Leffler function $E_\alpha(z)$ is an entire function. In particular, for real variable $\xi \in \mathbb{R}$ and $\alpha \in (0, 1)$, $E_\alpha(\xi)$ is strictly positive and monotone increasing. Moreover, we have $E_\alpha(0) > 0$, $E_\alpha(0) = 1$,

$$\lim_{\xi \to -\infty} E_\alpha(\xi) = 0, \quad \lim_{\xi \to +\infty} E_\alpha(\xi) = +\infty,$$  
$$\frac{d}{d\xi} E_\alpha(\xi) = \frac{1}{\alpha} E_{\alpha-1}(\xi),$$

where $E_{\alpha,\beta}(\xi) := \sum_{k=0}^{\infty} \frac{\xi^k}{\Gamma(\alpha+k)}$ is the two parameter Mittag-Leffler function. These properties can be derived from the definition of $E_\alpha(\xi)$ and the basic properties of gamma function $\Gamma(\xi)$ for $\xi > 0$. In view of the above properties for Mittag-Leffler functions, we know that the function $E_\alpha(2\lambda t^\alpha_m)$ involved in Lemma 3.2 is greater than one.

Now we proceed to the proof of the uniform $H^1$-norm estimate for the solution $u^m_h(f_h)$ to the fully discrete scheme (3.2) or (3.3).

Lemma 3.3. The fully discrete scheme (3.2) or (3.3) is unconditional stable. Moreover, the following estimates hold for $m = 1, \ldots, M$:

(i) $\|u^m_h(f_h)\|_{L^2(\Omega)} \leq C_0 \|f_h\|_{L^2(\Omega)}$,  
(ii) $\|\nabla u^m_h(f_h)\|_{L^2(\Omega)} \leq C_1 \|f_h\|_{L^2(\Omega)}$,

where $C_0, C_1 > 0$ are constants independent of $h$ and $\tau$.

Proof. We adopt an inductive argument to prove (i) in (3.6). First, for $m = 1$, it follows immediately from (3.2) and $u^0_h = 0$ that

$$(1 + b_0) \|u^1_h(f_h)\|_{H^1(\Omega)}^2 \leq (1 + b_0) \|u^1_h(f_h)\|_{L^2(\Omega)}^2 + b_0 \|\nabla u^1_h(f_h)\|_{L^2(\Omega)}^2 = b_0 (f_h \mu_1, u^1_h(f_h)) \leq C_0 \|f_h\|_{L^2(\Omega)} \|u^1_h(f_h)\|_{H^1(\Omega)},$$

which implies

$$\|u^1_h(f_h)\|_{H^1(\Omega)} \leq C_0 \|f_h\|_{L^2(\Omega)}.\quad (3.7)$$

Next, assuming that (i) in (3.6) holds for some $m \geq 1$, we shall prove that it also holds for $m + 1$. Choosing $\chi_h = u^{m+1}_h(f_h)$ in (3.2) and employing the monotone decreasing property of the sequence $\{d_j\}$, we obtain

$$(1 + b_0) \|u^{m+1}_h(f_h)\|_{L^2(\Omega)}^2 + b_0 \|\nabla u^{m+1}_h(f_h)\|_{L^2(\Omega)}^2 \leq \left( \sum_{j=0}^{m-1} (d_j - d_{j+1}) \right) \|u^{m-j}_h(f_h)\|_{L^2(\Omega)} + C_0 \|f_h\|_{L^2(\Omega)} \|u^{m+1}_h(f_h)\|_{L^2(\Omega)}$$

$$\leq (C_0 (1 - d_m) + C_0 b_0) \|f_h\|_{L^2(\Omega)} \|u^{m+1}_h(f_h)\|_{L^2(\Omega)}.$$

By the same argument as that in the proof of Lemma 4.1, we can suitably choose $C_0 > 0$ such that

$$C_0 (1 - d_m) + C_0 b_0 \leq C_0$$

holds uniformly for all $m = 1, 2, \ldots$. Then we immediately obtain $\|u^{m+1}_h(f_h)\|_{L^2(\Omega)} \leq C_0 \|f_h\|_{L^2(\Omega)}$.

We now prove the assertion (ii) in (3.6). Taking $\chi_h = 2\partial_\nu u^m_h$ in (3.3) yields

$$\big(\partial_\nu u^m_h, 2\partial_\nu u^m_h \big) + \big(\nabla u^m_h, \nabla (2\partial_\nu u^m_h) \big) + \big(u^m_h, 2\partial_\nu u^m_h \big) = (f_h \mu^m, 2\partial_\nu u^m_h), \quad (3.8)$$

where

$$\big(\nabla u^m_h, \nabla (2\partial_\nu u^m_h) \big) = \big(\nabla u^m_h, 2\partial_\nu \nabla u^m_h \big) \geq \partial_\nu \|\nabla u^m_h\|_{L^2(\Omega)}$$

by (3.3). Then (3.8) leads to

$$2 \|\partial_\nu u^m_h\|_{L^2(\Omega)}^2 + \partial_\nu \|\nabla u^m_h\|_{L^2(\Omega)}^2 \leq 2 \|\partial_\nu u^m_h\|_{L^2(\Omega)}^2 + \|u^m_h\|_{L^2(\Omega)}^2 + \|f_h \mu^m\|_{L^2(\Omega)}^2.$$

Hence, we have that

$$\partial_\nu \|\nabla u^m_h\|_{L^2(\Omega)} \leq \|u^m_h\|_{L^2(\Omega)} + \|f_h \mu^m\|_{L^2(\Omega)} \leq C_2 \left( \|\nabla u^m_h\|_{L^2(\Omega)}^2 + \int_{\Omega} u^m_h \, dx \right)^2 + \|f_h \mu^m\|_{L^2(\Omega)}^2.$$
Consequently, \( \chi \) for all \( \chi \) we see that.

Now it suffices to show actually strong, i.e.,

\[
\text{In view of the norm equivalence in finite-dimensional spaces, the above two weak convergence results are}
\]

\[
\text{Now applying Lemma 3.2 and (3.7), we obtain that}
\]

\[
\text{This completes the proof.}
\]

The next theorem provides the existence of a solution to the discrete system (3.4).

**Theorem 3.1.** For each fixed \( \tau > 0 \) and \( h > 0 \), there exists at least a minimizer to the discrete system (3.4).\n
**Proof.** It is readily seen from the non-negativity of \( J_{h, \tau}(f_h) \) that \( \inf J_{h, \tau}(f_h) \) is finite. Then there exists a minimizing sequence \( \{f_h^n\} \subset S_h \) such that

\[
\lim_{n \to \infty} J_{h, \tau}(f_h^n) = \inf_{f_h \in S_h} J_{h, \tau}(f_h).
\]

Then \( \{f_h^n\} \) is uniformly bounded in \( S_h \subset L^2(\Omega) \). Therefore, there exist \( f_h^* \in S_h \) and a subsequence of \( \{f_h^n\} \), still denoted by \( \{f_h^n\} \), such that

\[
f_h^n \rightharpoonup f_h^* \quad \text{in } L^2(\Omega) \quad \text{as } n \to \infty.
\]

Similarly to the proof of Theorem 2.1 we shall prove that \( f_h^* \) is a minimizer of (3.3).

Since \( \tau > 0 \) is now a fixed constant, if follows from Lemma 3.3 that for each \( m = 1, \ldots, M \), the sequence \( \{u_h^m(f_h^n)\} \) is uniformly bounded in \( V_h \subset L^2(\Omega) \) with respect to \( n \). Again, for each \( m = 1, \ldots, M \), this indicates the existence of \( u_h^{m, *}_n \in V_h \) and a subsequence of \( \{u_h^m(f_h^n)\} \), still denoted by \( \{u_h^m(f_h^n)\} \), such that

\[
u_h^m(f_h^n) \rightharpoonup u_h^{m, *}_n \quad \text{in } L^2(\Omega) \quad \text{as } n \to \infty, \quad m = 1, \ldots, M.
\]

In view of the norm equivalence in finite-dimensional spaces, the above two weak convergence results are actually strong, i.e.,

\[
u_h^m(f_h^n) \to u_h^{m, *}_n \quad \text{in } H^1(\Omega), \quad m = 1, \ldots, M,
\]

\[
f_h^n \to f_h^* \quad \text{in } H^1(\Omega)
\]

as \( n \to \infty \). (3.9)

Now it suffices to show \( \nu_h^m = u_h^m(f_h^*) \). In fact, by definition 3.2 and up to taking a further subsequence, we see that \( \{u_h^m(f_h^n)\} \) satisfies

\[
(1 + b_0) \left( u_h^{m,+}(f_h^n), \chi_h \right) + b_0 \left( \nabla u_h^{m,+}(f_h^n), \nabla \chi_h \right) = \sum_{j=0}^{m-1} \left( d_j - d_{j+1} \right) \left( u_h^{m-j}(f_h^n), \chi_h \right) + b_0 \left( f_h^n, \mu_h^{m+1} \right)
\]

for all \( \chi_h \in V_h \). Passing \( n \to \infty \) in the above equation and using (3.3), we obtain

\[
(1 + b_0) \left( u_h^{m,+}, \chi_h \right) + b_0 \left( \nabla u_h^{m,+}, \nabla \chi_h \right) = \sum_{j=0}^{m-1} \left( d_j - d_{j+1} \right) \left( u_h^{m-j}, \chi_h \right) + b_0 \left( f_h^*, \mu_h^{m+1} \right)
\]

for all \( \chi_h \in V_h \). By definition 3.2, again, we conclude \( u_h^{m, *}_n = u_h^m(f_h^*) \) by setting \( u_h^{m, *}_n = 0 \) artificially.

Finally, collecting the above results, we again employ the lower semi-continuity to deduce

\[
J_{h, \tau}(f_h^n) = \tau \sum_{m=0}^{M} c_m \left\| u_h^m(f_h^n) - u_h^m(\cdot, t_m) \right\|_{L^2(\omega)}^2 + \beta \| f_h^n \|_{L^2(\Omega)}^2
\]

\[
= \lim_{n \to \infty} \tau \sum_{m=0}^{M} c_m \left\| u_h^m(f_h^n) - u_h^m(\cdot, t_m) \right\|_{L^2(\omega)}^2 + \beta \lim_{n \to \infty} \| f_h^n \|_{L^2(\Omega)}^2
\]

\[
\leq \lim_{n \to \infty} \inf_{f_h \in S_h} J_{h, \tau}(f_h^n) = \inf_{f_h \in S_h} J_{h, \tau}(f_h).
\]

Consequently, \( f_h^* \) is indeed a minimizer of \( J_{h, \tau}(f_h) \) over \( S_h \).
4 Convergence of the Fully Discrete Approximation

This section is devoted to the convergence analysis of the fully discrete finite element approximation \[3.4\]. In order to relate \[3.4\] with the continuous minimization problem \[2.2\], we start with the introduction of a few auxiliary tools and useful results.

First, we recall the standard \(L^2\) projection \(P_h\): \(L^2(\Omega) \to V_h\) and Ritz projection \(R_h\): \(H^1(\Omega) \to V_h\), defined respectively by

\[
(P_h \varphi, \chi_h) = (\varphi, \chi_h), \quad \forall \chi_h \in V_h,
\]

\[
(\nabla R_h \varphi, \chi_h) = (\nabla \varphi, \chi_h), \quad \forall \chi_h \in V_h.
\]

The operators \(P_h\) and \(R_h\) satisfy the following approximation properties (see \[2.22\]):

\[
\|(I - P_h) \varphi\|_{H^r(\Omega)} \leq C h^{1-\gamma} \|\varphi\|_{H^1(\Omega)}, \quad \forall \varphi \in H^1(\Omega), \quad \gamma \in [0, 1],
\]

\[
\|(I - R_h) \varphi\|_{L^2(\Omega)} + h \|\nabla (\varphi - R_h \varphi)\|_{L^2(\Omega)} \leq C h^{2} \|\varphi\|_{H^r(\Omega)}, \quad \forall \varphi \in H^r(\Omega), \quad \gamma \in [1, 2].
\]

The next lemma is a useful classical approximation result (see \[2.28\]).

**Lemma 4.1.** For a Banach space \(X\) and a function \(g \in C([0,T]; X)\), define a step function approximation by

\[
S_\Delta g(x,t) = \sum_{m=1}^{M} 1_{(t_{m-1}, t_m]}(l) g(x,t_m),
\]

where \(1_{(t_{m-1}, t_m]}\) is the characteristic function of \((t_{m-1}, t_m]\). Then we have the convergence

\[
\lim_{\tau \to 0} \int_{0}^{T} \|S_\Delta g(\cdot, t) - g(\cdot, t)\|_X^2 \, dt = 0.
\]

In order to demonstrate the convergence of the finite element approximation \[3.4\] to the continuous minimization problem \[2.2\], we shall turn to the following key error estimate, which is a straightforward consequence of \[1.3\] Theorem 3.6.

**Lemma 4.2.** Let \(u(f)\) be the solution to system \[1.3\], where \(f \in L^2(\Omega)\) and \(\mu \in C^1[0,T]\). Then for the solution \(\{u_h^m(P_h f)\}\) to the fully discrete scheme \[3.2\] with \(f_h = P_h f\), we have the following error estimate: for any \(m = 1, \ldots, M\), there holds

\[
\|u(f)(\cdot, t_m) - u_h^m(P_h f)\|_{L^2(\Omega)} \leq C \|f\|_{L^2(\Omega)} \|\mu\|_{C^1[0,T]} \left( h^2 \ln h^2 + t_m^{a-1} \right).
\]

**Remark 4.1.** According to \[6\] Theorem 3.6, the source term \(f \mu\) should satisfy

\[
\int_{0}^{t} (t - s)^{a-1} \|f \mu(s)\|_{L^2(\Omega)} \, ds < \infty
\]

for \(0 < t \leq T\). This is automatically guaranteed by the assumption \(f \in L^2(\Omega)\) and \(\mu \in C^1[0,T]\). Therefore, the term including \(\int_{0}^{t} (t - s)^{a-1} \|f \mu(s)\|_{L^2(\Omega)} \, ds\) \(\leq \alpha\) \(\|f \mu\|_{L^2(\Omega)} \|\mu\|_{C^1[0,T]} \left( h^2 \ln h^2 + t_m^{a-1} \right)\) \(\leq \alpha\) \(\|f \mu\|_{L^2(\Omega)} \|\mu\|_{C^1[0,T]} \left( h^2 \ln h^2 + t_m^{a-1} \right)\).

The following lemma provides a crucial strong convergence.

**Lemma 4.3.** Let \(\{f_h\}_{h > 0}\) be a sequence in \(S_h\) which converges weakly to some \(f \in L^2(\Omega)\) as \(h \to 0\). Let \(\{u_h^m(f_h)\}\) and \(u(f)\) be the solutions of \[3.2\] and \[1.3\], respectively. Then the following convergence holds:

\[
\lim_{h, \tau \to 0} \sum_{m=0}^{M} c_m \|u_h^m(f_h) - u^\delta(\cdot, t_m)\|_{L^2(\Omega)}^2 = \|u(f) - u^\delta\|_{L^2(\Omega)}^2.
\]

**Proof.** For simplicity, in this proof we abbreviate \(u_m := u(f)(\cdot, t_m)\).

Since we have \(u(f) \in C([0,T]; L^2(\Omega))\) by \[2.2\] first we can directly apply Lemma 4.1 to obtain

\[
\lim_{\tau \to 0} \sum_{m=0}^{M} c_m \|u - u^\delta(\cdot, t_m)\|_{L^2(\Omega)}^2 = \|u(f) - u^\delta\|_{L^2(\Omega)}^2.
\]
Using this convergence, (4.9) follows immediately if we can demonstrate
\[
\lim_{h,\tau \to 0} \sum_{m=0}^{M} c_m \| u_h^m(f_h) - u^h(\cdot, t_m) \|_{L^2(\omega)}^2 = \lim_{\tau \to 0} \sum_{m=0}^{M} c_m \| u^m - u^h(\cdot, t_m) \|_{L^2(\omega)}^2. \tag{4.6}
\]

In order to show (4.6), we utilize the a priori estimates of \( u_h^m(f_h) \) and \( u^m \) to deduce
\[
\left| \sum_{m=0}^{M} c_m \| u_h^m(f_h) - u^h(\cdot, t_m) \|_{L^2(\omega)}^2 - \sum_{m=0}^{M} c_m \| u^m - u^h(\cdot, t_m) \|_{L^2(\omega)}^2 \right| \\
\leq \left( \sum_{m=0}^{M} \| u_h^m(f_h) - u^m \|_{L^2(\omega)}^2 \right)^{\frac{1}{2}} \left( \sum_{m=0}^{M} \| u_h^m(f_h) + u^m - 2u^h(\cdot, t_m) \|_{L^2(\omega)}^2 \right)^{\frac{1}{2}} \\
\leq C \left( \sum_{m=0}^{M} \| u_h^m(f_h) - u^m \|_{L^2(\omega)}^2 \right)^{\frac{1}{2}}.
\]

Clearly, the convergence in (4.6) follows if there holds
\[
\lim_{h,\tau \to 0} \sum_{m=0}^{M} \| u_h^m(f_h) - u^m \|_{L^2(\omega)}^2 = 0. \tag{4.7}
\]

To this end, we split the error \( u_h^m(f_h) - u^m \) into two parts:
\[
u_h^m(f_h) - u^m = (u_h^m(f_h) - u_h^m(P_h f)) + (u_h^m(P_h f) - u^m) =: \eta_h^m + \theta_h^m.
\]

From the estimate (4.2) and Remark 4.1, we can easily see that \( \| \theta_h^m \|_{L^2(\Omega)} \to 0 \) as \( h, \tau \to 0 \).

Now we concentrate on the estimate of \( \eta_h^m \). By the definitions of \( u_h^m(f_h) \) and \( u_h^m(P_h f) \), it is readily seen that \( \eta_h^m \) satisfies
\[(1 + b_0) (\eta_h^{m+1}, \chi_h) + b_0 (\nabla\eta_h^{m+1}, \nabla\chi_h) = -\sum_{j=0}^{m-1} (d_j - d_{j+1}) (\eta_h^{m-j}, \chi_h) + b_0 ((f_h - P_h f)\mu_h^{m+1}, \chi_h), \tag{4.8}
\]
for all \( \chi_h \in V_h \). By the assumption of Lemma 4.3 and 4.1, we have
\[f_h \rightharpoonup f, \quad P_h f \rightharpoonup f \quad \text{in} \quad L^2(\Omega) \quad \text{as} \quad h \to 0,
\]
which indicates \( f_h - P_h f \to 0 \) in \( L^2(\Omega) \) as \( h \to 0 \). Meanwhile, by Lemma 4.3 and \( f_h, P_h f \in S_h \), we estimate
\[
\| \eta_h^m \|_{H^1(\Omega)} \leq C \| f_h - P_h f \|_{L^2(\Omega)} \leq C, \quad m = 1, \ldots, M.
\]
Then for each \( m = 1, \ldots, M \), there exist some element \( \eta^m \in H^1(\Omega) \) and a subsequence of \( \{ \eta_h^m \} \), still denoted by \( \{ \eta_h^m \} \), such that
\[\eta_h^m \rightharpoonup \eta^m \quad \text{in} \quad H^1(\Omega), \quad \eta_h^m \to \eta^m \quad \text{in} \quad L^2(\Omega) \quad \text{as} \quad h \to 0, \quad m = 1, \ldots, M. \tag{4.9}
\]

Picking any \( \chi \in H^1(\Omega) \), we take \( \chi_h = R_h \chi \in V_h \) as the test function in (4.8). Then it follows from (4.2) that \( \chi_h \rightharpoonup \chi \) in \( H^1(\Omega) \) as \( h \to 0 \). Together with (4.9), we derive
\[\left( \nabla\eta_h^{m+1}, \nabla\chi_h \right) = \left( \nabla\eta^{m+1}_h, \nabla\chi \right) + \left( \nabla\eta^{m+1}_h, \nabla(\chi_h - \chi) \right) \to \left( \nabla\eta^{m+1}_*, \nabla\chi \right) \quad \text{as} \quad h \to 0.
\]

Using (4.9) again, we pass \( h \to 0 \) in equation (4.8) to deduce
\[(1 + b_0) (\eta^m_*, \chi) + b_0 (\nabla\eta^m_*, \nabla\chi) = -\sum_{j=0}^{m-1} (d_j - d_{j+1}) (\eta^{m-j}_*, \chi),
\]
which turns out to be the semidiscrete scheme in time for the original problem (4.1) with \( f = 0 \). Then it follows immediately from (4.2) that \( \| \eta^m_* \|_{L^2(\Omega)} = 0 \) or equivalently \( \eta^m_0 = 0 \).

Consequently, we obtain \( \| \eta_h^m \|_{L^2(\Omega)} \to 0 \) as \( h \to 0 \), and finally conclude
\[
\| u_h^m(f_h) - u^m \|_{L^2(\Omega)} \leq \| \theta_h^m \|_{L^2(\Omega)} + \| \eta_h^m \|_{L^2(\Omega)} \to 0
\]
as \( h, \tau \to 0 \). This completes the proof of (4.7). \( \square \)
We conclude this section with the convergence of the fully discrete finite element approximation \( \text{(3.4)} \) to the continuous minimization problem \( \text{(2.2)} \).

**Theorem 4.1.** Let \( \{f_h^*\}_{h>0} \) be the minimizers to the discrete minimization problem \( \text{(3.4)} \). Then there exists a subsequence of \( \{f_h^*\}_{h>0} \) which converges weakly in \( L^2(\Omega) \) to the minimizer of the continuous problem \( \text{(2.2)} \), as \( h, \tau \to 0 \).

**Proof.** It is routine to get the boundedness of \( \{f_h^*\} \) in \( L^2(\Omega) \). Thus there exist \( f^* \in L^2(\Omega) \) and a subsequence of \( \{f_h^*\} \), still denoted by \( \{f^*_h\} \), such that \( f^*_h \to f^* \) in \( L^2(\Omega) \) as \( h \to 0 \).

Now it suffices to show that \( f^* \) is the minimizer of the continuous problem \( \text{(2.2)} \). For any \( f \in L^2(\Omega) \), we have from \( \text{(4.1)} \) that

\[
P_h f \in S_h, \quad \lim_{h \to 0} \|f - P_h f\|_{L^2(\Omega)} = 0. \tag{4.10}
\]

Since \( f^*_h \) is the minimizer of \( J_{h, \tau}(f_h) \) over \( S_h \), we obtain

\[
J_{h, \tau}(f^*_h) \leq J_{h, \tau}(P_h f). \tag{4.11}
\]

Then using Lemma \( \text{(4.3)} \), \( \text{(4.10)} \), \( \text{(4.11)} \) and the lower semi-continuity of the \( L^2 \)-norm, we deduce

\[
J(f^*) = \|u(f^*) - u^\delta\|_{L^2(\omega \times (0, T))}^2 + \beta \|f^*\|_{L^2(\Omega)}^2
\leq \lim_{h, \tau \to 0} \tau \sum_{m=0}^{M} c_m \|u^n_h(f^*_h) - u^\delta(\cdot, t_m)\|_{L^2(\omega)}^2 + \beta \liminf_{h \to 0} \|f^*_h\|_{L^2(\Omega)}^2
\leq \liminf_{h, \tau \to 0} J_{h, \tau}(f^*_h) \leq \liminf_{h, \tau \to 0} J_{h, \tau}(P_h f)
\leq \lim_{h, \tau \to 0} \left\{ \tau \sum_{m=0}^{M} c_m \|u^n_h(P_h f) - u^\delta(\cdot, t_m)\|_{L^2(\omega)}^2 + \beta \|P_h f\|_{L^2(\Omega)}^2 \right\}
= \|u(f) - u^\delta\|_{L^2(\omega \times (0, T))}^2 + \beta \|f\|_{L^2(\Omega)}^2 = J(f),
\]

which implies that \( f^* \) is indeed a minimizer of the continuous problem \( \text{(2.2)} \).

\[\square\]

## 5 Iterative Thresholding Algorithm and Numerical Examples

In this section, we develop an efficient iterative thresholding algorithm to solve the discrete minimization problem \( \text{(3.4)} \) and present several numerical experiments to show its efficiency and accuracy.

Nearly all effective iterative methods for solving nonlinear optimization problems need the information of the derivatives of the concerned objective functional. We shall first derive the derivative of the discrete nonlinear functional \( J_{h, \tau}(f_h) \). Since \( u^n_h(f_h) \) is linear with respect to \( f_h \) in view of \( \text{(3.2)} \), we have \( u^n_h(f_h)p_h = u^n_h(p_h) \) for any \( p_h \in S_h \). Hence, it is straightforward to obtain

\[
J_{h, \tau}(f_h)p_h = 2\tau \sum_{m=0}^{M} c_m \int_\omega (u^n_h(f_h) - u^\delta) u^n_h(p_h) \; dx + 2\beta (f_h, p_h). \tag{5.1}
\]

Needless to say, it is extremely expensive to use the above formula directly to evaluate the derivatives, since computing the derivative at one fixed point \( f_h \) needs to solve equation \( \text{(3.2)} \) once for every direction \( p_h \in S_h \). In order to reduce the computational costs for computing the derivatives, we recall the backward Riemann-Liouville derivative

\[
D^\alpha_t g(t) := -\frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_t^T \frac{g(s)}{(s-t)^\alpha} \; ds = \frac{d}{dt} J^{1-\alpha}_t g(t)
\]
and introduce the following adjoint system

\[
\begin{cases}
(D^\alpha_t - \Delta + 1)v = 1_\omega (u(f) - u^\delta) \quad & \text{in } Q, \\
J^{1-\alpha}_t v = 0 \quad & \text{in } \Omega \times \{ T \}, \\
\partial_n v = 0 \quad & \text{on } \partial\Omega \times (0, T),
\end{cases}
\tag{5.2}
\]

where \( 1_\omega \) denotes the characterization function of \( \omega \). As before, we still denote the solution of \( \text{(5.2)} \) as \( v(f) \) to emphasize its dependency upon \( f \).
Lemma 5.1. Let \( f, p \in L^2(\Omega) \) and \( u, v \) be the solutions of systems (3.1) and (5.2) respectively. Then there holds
\[
\int_0^T \int_\Omega (u(f) - u^\delta) \, dx \, dt = \int_0^T \int_\Omega p \, v(f) \, dx \, dt.
\]

Proof. From the variational system (2.3), we have for any \( \varphi \in H^\alpha(0,T; L^2(\Omega)) \cap L^2(0,T; H^1(\Omega)) \) that
\[
\int_0^T \int_\Omega \left( \partial_t^\alpha u(p) \varphi + \nabla u(p) \cdot \nabla \varphi + u(p) \varphi \right) \, dx \, dt = \int_0^T \int_\Omega p \, \varphi \, dx \, dt. \tag{5.3}
\]

On the other hand, it is routine to get the variational system associated with (5.2) for any \( \psi \in H^\alpha(0,T; L^2(\Omega)) \cap L^2(0,T; H^1(\Omega)) \),
\[
\int_0^T \int_\Omega \left( D_t^\alpha v(f) \psi + \nabla v(f) \cdot \nabla \psi + v(f) \psi \right) \, dx \, dt = \int_0^T \int_\Omega (u(f) - u^\delta) \psi \, dx \, dt. \tag{5.4}
\]

Now letting \( \varphi = v(f) \) and \( \psi = u(p) \) in (5.3) and (5.4) respectively, we obtain
\[
\int_0^T \int_\Omega \partial_t^\alpha u(p) v(f) \, dx \, dt - \int_0^T \int_\Omega p \, v(f) \, dx \, dt = \int_0^T \int_\Omega D_t^\alpha v(f) u(p) \, dx \, dt - \int_0^T \int_\Omega (u(f) - u^\delta) u(p) \, dx \, dt. \tag{5.5}
\]

Further, by the integration by parts, the relation (2.1) and noting that
\[
u(p)(x,0) = 0, \quad (J_{T-}^{1-\alpha} v(f))(\cdot, T) = 0,
\]
we have
\[
\int_\Omega \int_0^T \partial_t^\alpha u(p) v(f) \, dx \, dt = \int_\Omega \int_0^T J_{T-}^{1-\alpha} \partial_t u(p) v(f) \, dx \, dt = \int_\Omega \int_0^T \partial_t u(p) J_{T-}^{1-\alpha} v(f) \, dx \, dt
\]
\[
= - \int_\Omega \int_0^T u(p) \, \frac{d}{dt} J_{T-}^{1-\alpha} v(f) \, dx \, dt = \int_\Omega \int_0^T u(p) D_t^\alpha v(f) \, dx \, dt.
\]

Substituting the above equality into (5.5), we arrive at the desired equality. \( \square \)

Writing the above equality into its discrete counterpart as that in previous sections, we have
\[
\tau \sum_{m=0}^M c_m \int_\Omega (u_h^m(f_h) - u^\delta_h) u_h^m(p_h) \, dx = \tau \sum_{m=0}^M c_m (\mu_h^m v_h^m(f_h), p_h).
\]

Therefore, we can rewrite (5.1) as
\[
J_{h,T}^T(f_h)p_h = 2\tau \sum_{m=0}^M c_m (\mu_h^m v_h^m(f_h), p_h) + 2\beta (f_h, p_h) = 2 \left( \tau \sum_{m=0}^M c_m \mu_h^m v_h^m(f_h) + \beta f_h, p_h \right).
\]

Since \( p_h \in S_h \) was chosen arbitrarily, the above equality gives the following necessary condition for some element \( f_h^* \in S_h \) to be a minimizer of the discrete optimization problem (3.4):
\[
\tau \sum_{m=0}^M c_m \mu_h^m v_h^m(f_h^*) + \beta f_h^* = 0.
\]

Following the same line as that in (3.4), this variational equation can be solved by the following iterative thresholding algorithm
\[
f_h^{k+1} = \frac{L}{L + \beta} f_h^k - \frac{\tau}{L + \beta} \sum_{m=0}^M c_m \mu_h^m v_h^m(f_h^k), \quad k = 0, 1, \ldots, \tag{5.6}
\]
where \( L > 0 \) is a tuning parameter for the convergence. The choice of \( L \) should be larger than the operator norm of the forward operator
\[
A : L^2(\Omega) \rightarrow L^2(\omega \times (0,T)), \quad f \mapsto u(f)_{\omega \times (0,T)},
\]

13
and we refer to [3] for details. It is obvious that at each step of \([5.6]\), we only need to solve equation \([3.2]\) for \(u_h^\alpha(f_h^\alpha)\) and then the discrete formulation of system \([5.2]\) for \(v_h^m(f_h^\alpha)\) subsequently, which does not involve heavy computational costs. By the way, although there appears the backward Riemann-Liouville derivative \(D_{-t}^{\alpha}\) in \([5.2]\), we know that the solution \(v(f)\) coincides with the following problem with a backward Caputo derivative

\[
\left\{ \begin{array}{ll}
(\partial_{t}^{\alpha} - \triangle + 1)v = \mathbf{1}_\omega (u(f) - u^\delta) & \text{in } Q, \\
v = 0 & \text{in } \Omega \times \{T\}, \\
\partial_{\nu}v = 0 & \text{on } \partial\Omega \times (0, T),
\end{array} \right.
\]  

(5.7)

where

\[\partial_{t}^{\alpha}g(t) := -\frac{1}{\Gamma(1 - \alpha)} \int_{t}^{T} g'(s) (s-t)^{\alpha} \text{ds}.\]

Therefore, instead of \([5.2]\) we just discretize system \([5.7]\) to compute \(v_h^m(f_h)\) by almost the same discrete method for \(u_h^\alpha(f_h)\).

We are now ready to propose the iterative thresholding algorithm for the reconstruction.

**Algorithm 5.1.** Choose a tolerance \(\epsilon > 0\), a regularization parameter \(\beta > 0\) and a tuning parameter \(L > 0\). Give an initial guess \(f_h^0\), and set \(k = 0\).

1. Compute \(f_h^{k+1}\) according to the iterative update \([5.6]\).

2. If \(\|f_h^{k+1} - f_h^k\|_{L^2(\Omega)} / \|f_h^k\|_{L^2(\Omega)} < \epsilon\), stop the iteration. Otherwise, update \(k \leftarrow k + 1\) and return to Step 1.

Next we shall present several numerical experiments by applying Algorithm \([5.1]\) to solve Problem \([1.1]\) for \(d = 1, 2\). In general, we set \(\Omega = (0, 1)^d, T = 1\) and specify various parameters involved in Algorithm \([5.1]\) as follows. Without loss of generality, we always choose the initial guess \(f_h^0\) as a constant, e.g., \(f_h^0 \equiv 2\). With the true source term \(f^*\) and thus the noiseless data \(u(f^*)\), the noisy data \(u^\delta\) is generated as

\[u^\delta(x, t) = (1 + \delta \text{rand}(-1, 1)) u(f^*)(x, t), \quad x \in \omega, \quad 0 < t \leq T,\]

where \text{rand}(-1, 1) denotes the random number uniformly distributed on \([-1, 1]\). For simplicity, in all examples we set the regularization parameter \(\beta = 10^{-4}\). Besides the illustrative figures, we mainly evaluate the numerical performance of Algorithm \([5.1]\) by the number \(K\) of iterations and the relative \(L^2\) error

\[\text{err} := \frac{\|f_h^K - f^*\|_{L^2(\Omega)}}{\|f^*\|_{L^2(\Omega)}},\]

where we understand \(f_h^K\) as the result of the numerical reconstruction.

In the one-dimensional case, we divide the space-time region \(\Omega \times [0, T] = [0, 1]^2\) into \(40 	imes 40\) equidistant meshes. We set the tuning parameter \(L\), the stopping criteria \(\epsilon\) and the known component \(\mu(t)\) as \(L = 1, \epsilon = 2 \times 10^{-3}\) and \(\mu(t) = 5 + 10 \in T\) respectively in Algorithm \([5.1]\). Except for the factors mentioned above, we shall test the numerical performance with different choices of exact source terms \(f^*\), fractional orders \(\alpha\), noise levels \(\delta\) and observation subdomains \(\omega\).

**Example 5.1.** In this example, we fix the observation subdomain \(\omega\) and the noise level \(\delta\) as \(\omega = \Omega \setminus [1/20, 19/20]\) and \(\delta = 1\%\), respectively. We test Algorithm \([5.1]\) with different fractional orders \(\alpha\) and exact source terms \(f^*\) as follows:

(a) \(\alpha = 0.3, \ f^*(x) = \sin(\pi x/2) + x^2 + 1\).

(b) \(\alpha = 0.5, \ f^*(x) = \sin(\pi x) - 2\).

Figure \([7]\) compares the true source terms \(f^*\) with the corresponding reconstructions \(f_h^K\), and also shows the iteration steps \(K\) and the relative errors in the caption.

**Example 5.2.** In this example, we fix \(\alpha = 0.8\) and the true source term \(f^*(x) = -\sin(\pi x) + x + 4\), and test different combinations of the observation subdomain \(\omega\) and the noise level \(\delta\) to observe their influence on the numerical performance. We first fix the noise level \(\delta = 1\%\) and shrink the size of \(\omega\) from \(\Omega \setminus [1/10, 9/10]\), \(\Omega \setminus [1/20, 19/20]\) to \(\Omega \setminus [1/40, 39/40]\). Next, we keep \(\omega = \Omega \setminus [1/20, 19/20]\) and enlarge the noise level \(\delta\) from 0.5\%, 1\%, 2\% to 4\%. The resulting numerical performance of the reconstructions is listed in Table \([7]\).
We can see from Figures 1 that with different fractional orders $\alpha$ and 1% noise in the observation data, the numerical reconstructions $f^K_h$ appear to be quite satisfactory in view of the ill-posedness of Problem 1.1, regardless of the casual choice of the initial guess $f^0_h$. In addition, we can observe from Table 1 that Algorithm 5.1 has two important advantages, namely, its strong robustness against the oscillating noise in observation data, and its insensitivity to the smallness of the observation subdomain $\omega$.

Now we consider the more challenging two-dimensional case, where we divide the space-time region $\Omega \times [0, T] = [0, 1]^3$ into $40 \times 40 \times 20$ equidistant meshes. Here we set the tuning parameter $L$ and the known component $\mu(t)$ in Algorithm 5.1 as $L = 2$ and $\mu(t) = 1 + 10\pi t^2$ respectively. Analogously to the one-dimensional counterpart, we evaluate the numerical performance of Algorithm 5.1 from various aspects, including different combinations of the true source terms, noise levels and observation subdomains.

Example 5.3. Parallel to Example 5.1, first we fix the observation subdomain $\omega$ and the noise level $\delta$ as $\omega = \Omega \setminus [1/10, 9/10]^2$ and $\delta = 1\%$, respectively. We take the stopping criteria $\epsilon = \delta/3$ and specify two pairs of fractional orders and true source terms as follows:

(a) $\alpha = 0.3$, $f^*(x) = \sin(x_1) + \sin(x_2) + 1$.

(b) $\alpha = 0.5$, $f^*(x) = \cos(\pi x_1) \cos(\pi x_2) + 2$.

In Figure 2, we illustrate the true source terms $f^*$ with the corresponding reconstructions $f^K_h$, and the iteration steps $K$ and the relative errors are shown in the caption.

Example 5.4. In this example we fix $\alpha = 0.8$, $f^*(x) = \exp((x_1 + x_2)/4) + 1$ and test Algorithm 5.1 with various choices of noise levels $\delta$ and observation subdomains $\omega$ in a similar manner as that in Example 5.2. Here we set $\epsilon = \delta/5$ as the stopping criteria. For the choice of $\omega$, we not only adjust its size, but also change its coverage of the boundary $\partial \Omega$. The resulting numerical performance of the reconstructions is listed in Table 2.

It is readily seen from Examples 5.3–5.4 that Algorithm 5.1 also works efficiently and accurately in the two-dimensional case. It inherits almost all advantages witnessed in the one-dimensional tests in view of its strong robustness against oscillating noise as well as its insensitivity to the smallness of the observation subdomain.
Figure 2: True source terms $f^*$ and their reconstructions $f^K_h$ obtained in Example 5.3. Left: Case (a), $K = 20$, err = 6.26%. Right: Case (b), $K = 36$, err = 7.17%.

Table 2: Numerical performance in Example 5.4 under various combinations of the observation subdomain $\omega$ and the noise level $\delta$.

| $\delta$ | $\omega$          | err  | $K$  |
|----------|-------------------|------|------|
| 1%       | $\Omega \setminus [1/10, 9/10]^2$ | 2.63% | 13   |
| 1%       | $\Omega \setminus [1/20, 19/20]^2$ | 4.94% | 27   |
| 1%       | $\Omega \setminus [0.9/10 \times 1/10, 9/10]$ | 4.36% | 9    |
| 0.5%     | $\Omega \setminus [1/10, 9/10]^2$ | 2.47% | 24   |
| 1%       | $\Omega \setminus [1/10, 9/10]^2$ | 3.61% | 15   |
| 2%       | $\Omega \setminus [1/10, 9/10]^2$ | 5.06% | 7    |
| 4%       | $\Omega \setminus [1/10, 9/10]^2$ | 6.58% | 5    |

Acknowledgement The authors thank Professor Bangti Jin (University College London) for his constructive discussions, and appreciate the valuable comments by the anonymous referees.

Funding information The first author is supported by National Natural Science Foundation of China (NSFC, Nos. 11871240, 11401241, 11571265) and NSFC-RGC (China-Hong Kong, No. 11661161017). The second author is supported by Japan Society for the Promotion of Science KAKENHI Grant Number JP15H05740. The last author is supported by NSFC (Nos. 11871057 and 11501447).

References

[1] Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)
[2] Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
[3] Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems. Comm. Pure Appl. Math. 57, 1413–1457 (2004)
[4] Gorenflo, R., Kilbas, A.A., Mainardi, F., Rogosin, S.V.: Mittag-Leffler Functions, Related Topics and Applications. Springer, Berlin (2014)
[5] Jiang, D., Li, Z., Liu, Y., Yamamoto, M.: Weak unique continuation property and a related inverse source problem for time-fractional diffusion-advection equations. Inverse Problems 33, 055013 (2017)
[6] Jin, B., Lazarov, R., Zhou, Z.: Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. SIAM J. Sci. Comput. 38, A146–A170 (2016)
[7] Jin, B., Lazarov, R., Liu, Y., Zhou, Z.: The Galerkin finite element method for a multi-term time-fractional diffusion equation. J. Comput. Phys. 281, 825–843 (2015)
[8] Jin, B., Rundell, W.: A tutorial on inverse problems for anomalous diffusion processes. Inverse Problems 31, 035003 (2015)
[9] Kopteva, N.: Error analysis of the L1 method on graded and uniform meshes for a fractional-derivative problem in two and three dimensions. Math. Comp. 88, 2135–2155 (2019)
[10] Kilbas, A., Srivastava, H., Trujillo, J.: Theory and Applications of Fractional Differential Equations. Elsevier, Amsterdam (2006)
[11] Li, D., Liao, H.-L., Sun, W., Wang, J., Zhang, J.: Analysis of L1-Galerkin FEMs for time-fractional nonlinear parabolic problems. Commun. Comput. Phys. 24, 86–103 (2018)
[12] Li, Z., Yamamoto, M.: Inverse problems of determining coefficients of the fractional partial differential equations. In: Kochubei, A., Luchko, Y. (eds.) Handbook of Fractional Calculus with Applications. Volume 2: Fractional Differential Equations, pp. 443–464. De Gruyter, Berlin (2019)
[13] Li, Z., Liu, Y., Yamamoto, M.: Inverse problems of determining parameters of the fractional partial differential equations. In: Kochubei, A., Luchko, Y. (eds.) Handbook of Fractional Calculus with Applications. Volume 2: Fractional Differential Equations, pp. 431–442. De Gruyter, Berlin (2019)
[14] Lin, Y., Xu, C.: Finite difference/spectral approximations for the time-fractional diffusion equation. J. Comput. Phys. 225, 1533–1552 (2007)
[15] Liu, Y., Li, Z., Yamamoto, M.: Inverse problems of determining sources of the fractional partial differential equations. In: Kochubei, A., Luchko, Y. (eds.) Handbook of Fractional Calculus with Applications. Volume 2: Fractional Differential Equations, pp. 411–430. De Gruyter, Berlin (2019)
[16] Liu, Y., Rundell, W., Yamamoto, M.: Strong maximum principle for fractional diffusion equations and an application to an inverse source problem. Fract. Calc. Appl. Anal. 19, 888–906 (2016)
[17] Liu, Y., Zhang, Z.: Reconstruction of the temporal component in the source term of a (time-fractional) diffusion equation. J. Phys. A 50, 305203 (2017)
[18] Ruan, Z., Wang, Z.: Identification of a time-dependent source term for a time fractional diffusion problem. Appl. Anal. 96, 1638–1655 (2017)
[19] Sakamoto, K., Yamamoto, M.: Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. J. Math. Anal. Appl. 382, 426–447 (2011)
[20] Sakamoto, K., Yamamoto, M.: Inverse source problem with a final overdetermination for a fractional diffusion equation. Math. Control Relat. Fields 1, 509–518 (2011)
[21] Stynes, M., O’Riordan, E., Gracia, J.L.: Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. SIAM J. Numer. Anal. 55, 1057–1079 (2017)
[22] Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, Springer Series in Computational Mathematics, vol. 25. Springer, Berlin (2006)
[23] Wang, D., Xiao, A., Zou, J.: Long time behavior of numerical solutions to nonlinear fractional ODEs. ESAIM Math. Model. Numer. Anal. 54, 335–358 (2020)
[24] Wang, W., Yamamoto, M., Han, B.: Numerical method in reproducing kernel space for an inverse source problem for the fractional diffusion equation. Inverse Problems 29, 095009 (2013)
[25] Wei, T., Li, X.L., Li, Y.S.: An inverse time-dependent source problem for a time-fractional diffusion equation. Inverse Problems 32, 085003 (2016)
[26] Wei, T., Wang, J.: A modified quasi-boundary value method for an inverse source problem of the time-fractional diffusion equation. Appl. Numer. Math. 78, 95–111 (2014)
[27] Xie, J., Zou, J.: Numerical reconstruction of heat fluxes. SIAM J. Numer. Anal. 43, 1504-C1535 (2005)
[28] Zeidler, E.: Nonlinear Functional Analysis and Its Applications, vol. 2. Springer, New York (1985)
[29] Zhang, Y., Xu, X.: Inverse source problem for a fractional diffusion equation. Inverse Problems 27, 035010 (2011)