Journal of Computational Mathematics
Vol.xx, No.x, 2026, 1–22.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2512-m2025-0207

RECOVERING SPARSE INITIAL CONDITIONS IN NONLOCAL DIFFUSION*

Jinyan Tian and Fangfang Dou[1)]
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China
Emails: jytian2022@163.com, fangfdou@uestc.edu.cn

Abstract

This work addresses the inverse problem of recovering sparse initial data in spatial fractional parabolic equations. The associated solution operator for initial-boundary value problem is continuous and compact, implying severe ill-posedness. To overcome this, an ℓ_1-regularized formulation is considered. The misfit functional is shown to be differentiable and strictly convex, and the well-posedness of the regularized problem is established, including existence, uniqueness, stability, and convergence. A numerical algorithm is proposed and implemented, incorporating Nesterov's accelerated algorithm to enhance efficiency. Numerical experiments in both one and two spatial dimensions confirm the feasibility and accuracy of the proposed approach.

Mathematics subject classification: 65N21, 60H15, 65D12.
Key words: Sparse initial data, Spatial fractional diffusion, Nonlocal diffusion, ℓ_1-regularization, Nesterov's accelerated algorithm.

1. Introduction

Fractional partial differential equations (FPDEs), involving derivatives or integrals of non-integer order, arise in a wide range of scientific and engineering applications. Notable examples include biological systems [17], quantum mechanics [23], porous media flow [21], viscoelasticity [22], contaminant transport [24], polymer unzipping [8], probability theory, and mathematical finance [3]. Among these, FPDEs with fractional Laplacian operators form an important class of nonlocal models. The fractional Laplacian, as the infinitesimal generator of Lévy-stable diffusion processes [3], has been particularly useful in modeling anomalous diffusion and stochastic dynamics across disciplines.

In this paper, we investigate a parabolic equation involving a fractional Laplacian operator. To facilitate a better understanding of our model, we first introduce the fractional Laplacian operator [2, 7, 15]. Let $\Omega \subset \mathbb{R}^N, N = 1, 2$ is an open and bounded domain with Lipschitz boundary $\partial\Omega$ and $0 < s < 1$. We denote a space

$$\mathrm{L}^1_s(\mathbb{R}^N) := \left\{ u : \mathbb{R}^N \to \mathbb{R} \text{ is measurable and } \int_{\mathbb{R}^N} \frac{|u(x)|}{(1+|x|)^{N+2s}} \mathrm{d}x < \infty \right\}.$$

For $u \in \mathrm{L}^1_s(\mathbb{R}^N)$ and $\varepsilon > 0$, let

$$(-\Delta)^s_\varepsilon u(x) := C_{N,s} \int_{\{y\in\mathbb{R}^N, |y-x|>\varepsilon\}} \frac{u(x)-u(y)}{|x-y|^{N+2s}} \mathrm{d}y, \quad x \in \mathbb{R}^N,$$

* Received August 12, 2025 / Revised version received November 18, 2025 / Accepted December 19, 2025 / Published online April 13, 2026 /
[1)] Corresponding author

here the normalization constant

$$C_{N,s} = \frac{s2^{2s}\Gamma(2s+N)/2}{\pi^{N/2}\Gamma(1-s)},$$

and Γ is the usual Euler gamma function. For $u \in \mathrm{L}^1_s(\mathbb{R}^N)$, fractional Laplace operator $(-\Delta)^s$ is denoted by the following equation:

$$(-\Delta)^s u(x) = C_{N,s}\mathrm{P.V.}\int_{\mathbb{R}^N} \frac{u(x)-u(y)}{|x-y|^{N+2s}}\mathrm{d}y = \lim_{\varepsilon\to 0}(-\Delta)^s_\varepsilon u(x), \quad x \in \mathbb{R}^N,$$

where P.V. denotes the principal value of the integral.

Consider the spatial fractional-order parabolic equation

$$\begin{cases} \partial_t u + (-\Delta)^s u = 0 & \text{in } \Omega \times (0,T), \\ u = 0 & \text{in } \Omega^c \times (0,T), \\ u|_{t=0} = v & \text{in } \Omega, \end{cases} \tag{1.1}$$

where $T > 0$ is a fixed final time, $\Omega^c \subset \mathbb{R}^N$ denotes the complement of the bounded domain $\Omega \subset \mathbb{R}^N$, and the initial condition satisfies $v \in L^2(\Omega)$.

This model has been used to describe anomalous diffusion phenomena in various fields, including porous media, biology, and quantum mechanics. From a probabilistic perspective, the solution u represents the probability density of a particle undergoing a Lévy-type random walk with arbitrarily long jumps within Ω. This definition is intrinsic to the full space and does not rely on domain geometry, boundary conditions, or spectral data, unlike the spectral definition on bounded domains; and the nonlocality is completely transparent in the formulation, whereas the Fourier and spectral definitions hide this fact. Furthermore, it exhibits directly the singular kernel structure and the need for the principal value, which is useful for establishing regularity estimates and nonlocal elliptic properties. The homogeneous Dirichlet condition $u = 0$ in Ω^c indicates absorption at the boundary $\partial\Omega$, and the initial distribution v specifies the probability density of the particle's initial position. For further applications of this system, we refer the reader to [11] and the references therein.

Due to the extensive applications of the fractional partial differential equations in various fields, a substantial amount of research on their inverse problems has been conducted in recent years. In particular, taking the later-time observations of the solution $u(x,T)$ at the final time T as the measurement data is a popular choice in practice. It is widely used in inverse source terms, inverse potentials, and inverse initial value problems. Various inverse problems related to fractional diffusion equations have been widely studied. Jin *et al.* [13] investigated the simultaneous inversion of initial data, source terms, and potentials for subdiffusion models with unknown terminal times. Zhang and Zhou [25] applied the quasi-boundary value method to regularize mildly ill-posed backward problems in time-fractional diffusion equations. Jiang *et al.* [12] reconstructed the spatial component of a source term using internal observations. Kian *et al.* [14] studied the identification of source terms and initial conditions from boundary measurements in time-fractional diffusion models. Zheng and Zhang [26] proposed a fractional Tikhonov regularization method to address the backward problem for fractional parabolic equations.

For spatial-fractional models, Hrizi *et al.* [11] considered the reconstruction of the initial condition from final-time data, and proposed a least-squares optimization framework combined

with conjugate gradient methods. Ben Salah and Tatar [4] addressed the identification of initial values in space-time fractional diffusion equations. While these works provide important insights, most of them focus on smooth or non-sparse initial data, and often assume ideal observation conditions.

However, in many practical applications, such as localized laser heating, micro-scale thermal devices, or pollutant dispersion from point sources, the initial temperature distribution is spatially sparse – nonzero only in a small subset of the domain. This sparsity of the initial data reflects physical constraints or design conditions, and it often leads to solutions with distinct features in both space and time. The case of sparse initial data in spatial fractional parabolic equations remains relatively less explored, both in theory and computation. This motivates the present work, which aims to fill this gap through a rigorous analysis of the ℓ_1-regularized model and the development of an efficient numerical scheme tailored for the fractional setting.

Inverse problem (IP): Can we recover the initial sparse probability distribution v of the particle positions from the noisy observation data h^δ of the solution $u(\cdot, T)$?

The inverse problem of recovering sparse initial conditions in classical parabolic equations has been extensively studied. For example, Li *et al.* [16] considered the reconstruction of sparse initial data from sparsely sampled observations of the heat equation. Muoi [18] investigated the recovery of sparse initial conditions from final-time measurements, and employed Nesterov's accelerated algorithm to enhance computational performance. In contrast, the recovery of sparse initial data in fractional parabolic equations remains largely unexplored. To the best of our knowledge, there is a conspicuous lack of studies addressing this problem, both in theory and computation. For completeness, we mention two related works concerning sparse optimal control of fractional equations [5, 19], which share similar structural features but focus on control design rather than initial data reconstruction.

In this work, we prove that the solution operator associated with the spatial fractional parabolic equation is compact, which leads to the ill-posedness of the inverse problem. To recover sparse initial conditions, an ℓ_1-regularization framework is introduced based on the sparsity prior. The misfit functional is shown to possess key properties such as differentiability and strict convexity, and we establish the well-posedness of the regularized problem. For numerical implementation, we employ Nesterov's accelerated algorithm, which achieves a convergence rate of $\mathcal{O}(1/n^2)$ [20]. In contrast to prior work such as [11], we conduct numerical experiments in both one and two spatial dimensions to verify the accuracy and robustness of the proposed method. This study extends existing theoretical and computational results on sparse initial recovery from classical parabolic equations to the fractional setting. The proposed framework offers a unified and more general perspective for analyzing sparse inverse problems governed by nonlocal diffusion models.

The remainder of the paper is organized as follows. In Section 2, we analyze the continuity and compactness of the solution operator, which characterizes the degree of ill-posedness of the inverse problem. Section 3 introduces an ℓ_1-regularization approach to reformulate the inverse problem and establishes the well-posedness of the regularized model, including uniqueness and convergence rate results. Section 4 presents a series of numerical experiments in both one and two spatial dimensions to demonstrate the effectiveness and robustness of the proposed method with Nesterov's acceleration. Finally, Section 5 concludes the paper with a brief summary and possible directions for future work.

2. Ill-posedness for the IP

For convenience of readers, we first review some properties of fractional Laplace operator and spaces, which can be found in the literature [10, 11, 15] and the references therein.

Let $\{(\Phi_k, \lambda_k)\}_{k=1}^{\infty}$ denote the eigenpairs of the fractional Laplace operator $(-\Delta)^s$ with Dirichlet conditions. That is for all $k \in \mathbb{N}^*, (\Phi_k, \lambda_k)$ satisfies

$$\begin{cases} (-\Delta)^s \Phi_k = \lambda_k \Phi_k & \text{in } \Omega, \\ \Phi_k = 0 & \text{in } \mathbb{R}^N \setminus \Omega. \end{cases}$$

Here

$$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_k \le \cdots \quad \text{and} \quad \lim_{k\to\infty} \lambda_k = +\infty,$$

and $\{\Phi_k\}_{k=1}^{\infty}$ forms an orthonormal basis for $L^2(\Omega)$.

The fractional Sobolev space $H^s(\mathbb{R}^N)$ is defined as

$$H^s(\mathbb{R}^N) := \left\{ z \in L^2(\mathbb{R}^N) : \ |\xi|^s \mathcal{F}(u)(\xi) \in L^2(\mathbb{R}^N) \right\},$$

where $\mathcal{F}$ denotes the Fourier transform. The fractional Sobolev space $H^s(\Omega)$ is defined as

$$H^s(\Omega) = \left\{ u \in L^2(\Omega) : \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^2}{|x-y|^{N+2s}} \mathrm{d}x\mathrm{d}y < \infty \right\},$$

which is endowed with the norm

$$\|u\|_{H^s(\Omega)} = \left(\int_\Omega |u|^2 \mathrm{d}x + \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^2}{|x-y|^{N+2s}} \mathrm{d}x\mathrm{d}y \right)^{\frac{1}{2}}.$$

Additionally, $H_0^s(\Omega)$ is the Hilbert space defined by

$$H_0^s(\Omega) = \left\{ u \in H^s(\mathbb{R}^N) : u = 0 \ \text{ in } \ \mathbb{R}^N \setminus \Omega \right\} = \left\{ u \in H^s(\mathbb{R}^N) : \mathrm{supp}[u] \subset \overline{\Omega} \right\}.$$

It is endowed with the following norm:

$$\|u\|_{H_0^s(\Omega)} = \left(\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \frac{|u(x) - u(y)|^2}{|x-y|^{N+2s}} \mathrm{d}x\mathrm{d}y \right)^{\frac{1}{2}}.$$

$H^{-s}(\Omega) := (H_0^s(\Omega))^*$ is the dual space of $H_0^s(\Omega)$ with respect to the pivot space $L^2(\Omega)$, and $\langle \cdot, \cdot \rangle_{-s,s}$ is the dual pairing between $H^{-s}(\Omega)$ and $H_0^s(\Omega)$. Define the bilinear form

$$\mathfrak{a}^s(u,u) = \frac{C_{N,s}}{2} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \frac{|u(t,x) - u(t,y)|^2}{|x-y|^{N+2s}} dxdy. \tag{2.1}$$

Note that the norm $\|\cdot\|_{H_0^s(\Omega)}$ is equivalent to the norm induced by $\mathfrak{a}^s(\cdot,\cdot)$ on $H_0^s(\Omega)$.

We use standard Sobolev space $L^2(\Omega), H_0^1(\Omega)$ and $H^1(\Omega)$, see [6].

Definition 2.1 ([15]). *Let $v \in L^2(\Omega), f \in L^2(0,T; H^{-s}(\Omega))$. We say that*

$$u[v] \in C\big([0,T]; L^2(\Omega)\big) \cap L^2\big(0,T; H_0^s(\Omega)\big) \cap H^1\big(0,T; H^{-s}(\Omega)\big)$$

is a finite energy solution to the fractional parabolic system

$$\begin{cases} \partial_t u[v] + (-\Delta)^s u[v] = f & \text{in } \Omega \times (0,T), \\ u[v] = 0 & \text{in } \Omega^c \times (0,T), \\ u[v](\cdot,0) = v & \text{in } \Omega, \end{cases} \tag{2.2}$$

if $u[v](\cdot,0)=v$ *a.e in* Ω *and*

$$\int_0^T \langle f,\phi\rangle_{-s,s}dt=\int_0^T \langle \partial_t u[v],\phi\rangle_{-s,s}dt + \frac{C_{N,s}}{2}\int_0^T\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\frac{\big(u[v](x,t)-u[v](y,t)\big)\big(\phi(x,t)-\phi(y,t)\big)}{|x-y|^{N+2s}}dxdydt$$

holds for all $\phi\in L^2(0,T;H_0^s(\Omega))$.

Lemma 2.1 ([10]). *For a given* $v\in L^2(\Omega)$*, there exists a unique finite energy solution to problem* (2.2) *in the sense of Definition* 2.1*, which is given by*

$$u[v](x,t)=\sum_{k=1}^{\infty}u_k(t)\Phi_k(x),$$

where

$$u_k(t)=u_{0,k}e^{-\lambda_k t}+\int_0^t e^{-\lambda_k(t-r)}f_k(r)\mathrm{d}r,$$
$$u_{0,k}=\langle v,\Phi_k\rangle_{L^2(\Omega)},\quad f_k(t)=\langle f(\cdot,t),\Phi_k\rangle_{L^2(\Omega)}.$$

Especially, the solution of problem (2.2) *with homogeneous source, i.e.,*

$$\begin{cases}\partial_t u[v]+(-\Delta)^s u[v]=0 & in\ \Omega\times(0,T),\\ u[v]=0 & in\ \Omega^c\times(0,T),\\ u[v](\cdot,0)=v & in\ \Omega\end{cases}\tag{2.3}$$

has the form

$$u[v](x,t)=\sum_{k=1}^{\infty}e^{-\lambda_k t}\langle v,\Phi_k\rangle_{L^2(\Omega)}\Phi_k(x).\tag{2.4}$$

Lemma 2.2 ([11]). *The solution of* (2.3) *satisfies the following energy estimate:*

$$\|u[v]\|_{H^1(0,T;H^{-s}(\Omega))}+\|u[v]\|_{L^2(0,T;H_0^s(\Omega))}+\|u[v]\|_{C([0,T];L^2(\Omega))}\le C\|v\|_{L^2(\Omega)},\tag{2.5}$$

where C *is a positive constant, depending only on* Ω,N,s,T.

Lemma 2.3 (Uniqueness of IP, [11]). *Let* $u[v_1]$ *and* $u[v_2]$ *be two solutions to system* (2.3) *with initial values* $v_\ell\in L^2(\Omega),\ell=1,2$*. Then, if* $u[v_1](\cdot,T)=u[v_2](\cdot,T)$*, it follows that* $v_1=v_2$ *in* $L^2(\Omega)$.

Lemma 2.4. *Let* $h=u[v](\cdot,T)\in L^2(\Omega)$ *is given and* v *is unknown. Then, the solution to problem* (2.3) *can be expressed in the following form:*

$$u[v](x,t)=\sum_{k=1}^{\infty}e^{\lambda_k(T-t)}\langle h,\Phi_k\rangle_{L^2(\Omega)}\Phi_k(x).\tag{2.6}$$

Particularly,

$$v(x)=\sum_{k=1}^{\infty}e^{\lambda_k T}\langle h,\Phi_k\rangle_{L^2(\Omega)}\Phi_k(x).\tag{2.7}$$

Proof. From Eq. (2.4), the solution of problem (2.3)

$$\begin{cases} \partial_t u[v] + (-\Delta)^s u[v] = 0 & \text{in } \Omega \times (0,T), \\ u[v] = 0 & \text{in } \Omega^c \times (0,T), \\ u[v](\cdot,0) = v & \text{in } \Omega \end{cases}$$

can be written as

$$u[v](x,t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} \langle v, \Phi_k \rangle_{L^2(\Omega)} \Phi_k(x). \tag{2.8}$$

By setting $t = T$ and using the fact that h is the observation data at time T, we obtain

$$h = \sum_{k=1}^{\infty} e^{-\lambda_k T} \langle v, \Phi_k \rangle_{L^2(\Omega)} \Phi_k(x).$$

Multiplying both sides of the above equation by Φ_k and integrating with respect to x, we obtain

$$\langle h, \Phi_k \rangle_{L^2(\Omega)} = \langle v, \Phi_k \rangle_{L^2(\Omega)} e^{-\lambda_k T}.$$

Since the set of eigenfunctions $\{\Phi_k\}_{k=1}^{\infty}$ forms an orthonormal basis for $L^2(\Omega)$. From the fact that $e^{-\lambda_k T} > 0$, one can deduce that

$$\langle v, \Phi_k \rangle_{L^2(\Omega)} = \langle h, \Phi_k \rangle_{L^2(\Omega)} e^{\lambda_k T}. \tag{2.9}$$

Thus, $v(x)$ can be written as

$$v(x) = \sum_{k=1}^{\infty} e^{\lambda_k T} \langle h, \Phi_k \rangle_{L^2(\Omega)} \Phi_k(x).$$

Substituting (2.9) into (2.8), we have that

$$u[v](x,t) = \sum_{k=1}^{\infty} e^{\lambda_k (T-t)} \langle h, \Phi_k \rangle_{L^2(\Omega)} \Phi_k(x).$$

The proof is complete. □

By Lemma 2.4, the initial value v can be uniquely determined by the solution $u[v](\cdot, T)$. To reconstruct v from the noisy observation data h^δ of the solution at time T, the most common approach is to minimize the following functional:

$$\|\mathcal{K}(v) - h\|^2_{L^2(\Omega)},$$

where the operator $\mathcal{K}$ is defined as

$$\mathcal{K} : L^2(\Omega) \longrightarrow L^2(\Omega), \quad v \longmapsto u[v](\cdot, T).$$

It is easy to see that the linear operator $\mathcal{K}$ is a self-adjoint operator, and we prove that the operator $\mathcal{K}$ is continuous and compact.

Proposition 2.1. *The linear operator $\mathcal{K} : L^2(\Omega) \longrightarrow L^2(\Omega)$ is continuous and compact.*

Proof. Firstly, we prove that the linear operator $\mathcal{K}$ is continuous. For any $v_1, v_2 \in L^2(\Omega)$, according to (2.5), we can obtain that

$$\begin{aligned} &\|\mathcal{K}(v_2) - \mathcal{K}(v_1)\|_{L^2(\Omega)} \\ &= \|\mathcal{K}(v_2 - v_1)\|_{L^2(\Omega)} \\ &\leq \|u[v_2 - v_1](x,t)\|_{C([0,T];L^2(\Omega))} \\ &\leq C\|v_2 - v_1\|_{L^2(\Omega)}. \end{aligned}$$

This means that the operator $\mathcal{K}$ is continuous.

Next, we show that operator $\mathcal{K}$ is compact. By (2.4) and the definition of the operator $\mathcal{K}$, we can get

$$\mathcal{K}(v) = \sum_{k=1}^{\infty} e^{-\lambda_k T}\langle v, \Phi_k\rangle_{L^2(\Omega)}\Phi_k(x), \quad \forall\, v \in L^2(\Omega). \tag{2.10}$$

We define the finite rank operator $\mathcal{K}_n$ as follows:

$$\mathcal{K}_n(v) := \sum_{k=1}^{n} e^{-\lambda_k T}\langle v, \Phi_k\rangle_{L^2(\Omega)}\Phi_k(x), \quad \forall\, v \in L^2(\Omega). \tag{2.11}$$

Through (2.10) and (2.11), we obtain that

$$\|\mathcal{K}(v) - \mathcal{K}_n(v)\|^2_{L^2(\Omega)} = \sum_{k=n+1}^{\infty} \left|\langle v, \Phi_k\rangle_{L^2(\Omega)}\right|^2 e^{-2\lambda_k T}.$$

From the properties of eigenvalues for the fractional Laplace operator, we have that

$$\|\mathcal{K}(v) - \mathcal{K}_n(v)\|^2_{L^2(\Omega)} \leq \frac{C_0}{e^{2\lambda_n T}}\|v\|^2_{L^2(\Omega)}.$$

Hence, $\|\mathcal{K}(v) - \mathcal{K}_n(v)\|_{L^2(\Omega)} \to 0$, as $n \to \infty$, in the sense of space $\mathcal{L}(L^2(\Omega); L^2(\Omega))$. □

Remark 2.1. The fractional Laplace operator admits several equivalent definitions. Our method applies to these formulations as well. In particular, the definition based on the Fourier transform is fully equivalent to the one adopted in this paper; detailed proofs can be found in [11] and the references therein. Moreover, for the spectral definition of the fractional Laplacian [9], an analogue of Lemma 2.4 can be derived: once the eigenvalues and eigenfunctions are known, the initial data can be recovered directly by inverting the solution representation.

3. l_1-regularization Method for the IP

From the above analysis we know that the considered IP is ill-posed, since $\mathcal{K}$ does not have a bounded inverse. Based on the sparsity of initial condition, we introduce l_1-regularization to deal with the ill-posedness. We reformulate the inverse problem as the following regularized optimization problem:

$$\min_{v \in L^2(\Omega)} J_{\alpha,h^\delta}(v) = F_{h^\delta}(v) + \alpha\Phi(v), \tag{3.1}$$

where

$$F_{h^\delta}(v) := \frac{1}{2}\|\mathcal{K}(v) - h^\delta\|^2_{L^2(\Omega)}, \quad \Phi(v) = \sum_{i\in\Lambda} w_i|\langle v, \psi_i\rangle_{L^2(\Omega)}|, \tag{3.2}$$

$\{\psi_i\}_{i\in\Lambda}$ is a basis of $L^2(\Omega)$ and $w_i \geqslant w_0 > 0$, for all $i \in \Lambda$ are weighted numbers.

In order to obtain the well-posedness of the regularized optimization problem, we first analyze some properties of the misfit functional.

Proposition 3.1. *Assume that $u[v]$ and p are the finite energy solutions to*

$$\begin{cases} \partial_t u[v] + (-\Delta)^s u[v] = f & \text{in } \Omega \times (0,T), \\ u[v] = 0 & \text{in } \Omega^c \times (0,T), \\ u[v](\cdot,0) = v & \text{in } \Omega, \end{cases} \tag{3.3}$$

$$\begin{cases} -\partial_t p + (-\Delta)^s p = g & \text{in } \Omega \times (0,T), \\ p = 0 & \text{in } \Omega^c \times (0,T), \\ p(\cdot,T) = z & \text{in } \Omega, \end{cases} \tag{3.4}$$

respectively, where $f, g \in L^2(0,T;H^{-s}(\Omega))$, and $v, z \in L^2(\Omega)$. Then we have

$$\int_0^T \langle f, p\rangle_{-s,s} dt + \int_\Omega v p(\cdot,0) dx = \int_0^T \langle g, u[v]\rangle_{-s,s} dt + \int_\Omega z u[v](\cdot,T) dx. \tag{3.5}$$

Proof. Let $t' = T - t$, we can check that there exists a unique finite energy solution to the problem (3.4). Combining Definition 2.1 and bilinear form (2.1), it is easy to obtain its weak solution with the form

$$\int_0^T \langle g, \phi\rangle_{-s,s} dt = \int_0^T \langle -\partial_t p, \phi\rangle_{-s,s} dt + \int_0^T \mathfrak{a}^s(p,\phi) dt, \quad \forall\, \phi \in L^2\big(0,T;H_0^s(\Omega)\big). \tag{3.6}$$

Set $\phi = u[v]$ in (3.6), we have

$$\int_0^T \langle g, u[v]\rangle_{-s,s} dt = \int_0^T \langle -\partial_t p, u[v]\rangle_{-s,s} dt + \int_0^T \mathfrak{a}^s(p, u[v]) dt.$$

Using integration by parts respect to t, we get

$$\begin{aligned} &\int_0^T \langle g, u[v]\rangle_{-s,s} dt + \int_\Omega z u[v](\cdot,T) dx - \int_\Omega v p(\cdot,0) dx \\ &= \int_0^T \langle \partial_t u[v], p\rangle_{-s,s} dt + \int_0^T \mathfrak{a}^s(p, u[v]) dt. \end{aligned} \tag{3.7}$$

Similarly, we obtain

$$\int_0^T \langle f, \phi\rangle_{-s,s} dt = \int_0^T \langle \partial_t u[v], \phi\rangle_{-s,s} dt + \int_0^T \mathfrak{a}^s(u[v], \phi) dt, \quad \forall\, \phi \in L^2\big(0,T;H_0^s(\Omega)\big) \tag{3.8}$$

for problem (3.3). Set $\phi = p$ in (3.8), it is observed that

$$\int_0^T \langle f, p\rangle_{-s,s} dt = \int_0^T \langle \partial_t u[v], p\rangle_{-s,s} dt + \int_0^T \mathfrak{a}^s(u[v], p) dt. \tag{3.9}$$

Combining (3.7) and (3.9) yields (3.5). □

Lemma 3.1. *For any $v \in L^2(\Omega)$, the misfit functional $F_h : L^2(\Omega) \to R$ given by (3.2) has the following properties:*

(i) F_h *is nonnegative, continuous, and compact.*

(ii) F_h *is Lipschitz continuous and Fréchet differentiable. Moreover, for any* $v \in L^2(\Omega)$, $\nabla F_h(v) = p[v](x,0), p[v](x,t)$ *is a finite energy solution to the following problem:*

$$\begin{cases} -\partial_t p[v] + (-\Delta)^s p[v] = 0 & \text{in } \Omega \times (0,T), \\ p[v] = 0 & \text{in } \Omega^c \times (0,T), \\ p[v](\cdot,T) = \mathcal{K}(v) - h & \text{in } \Omega. \end{cases} \tag{3.10}$$

(iii) F_h *is quadratically differentiable and strictly convex.*

Proof. (i) F_h is nonnegative. It follows from the definition of F_h and the fact that the operator $\mathcal{K}$ is continuous that F_h is also continuous. To prove that F_h is compact, i.e., to show that F_h maps a bounded sequence $\{v_n\} \subset L^2(\Omega)$ into a sequence $\{F_h(v_n)\}$ with strongly convergent subsequences. Since the sequence $\{v_n\} \subset L^2(\Omega)$ is bounded, there exists a weakly convergent subsequence $\{v_{n_k}\}$ of the sequence $\{v_n\}$. Without loss of generality, we set $v_{n_k} \rightharpoonup v_0$. Owing to the fact that operator $\mathcal{K}$ is compact, we know $\mathcal{K}(v_{n_k}) \to \mathcal{K}(v_0)$. Combined with the definition of F_h it follows that $F_h(v_{n_k}) \to F_h(v_0)$. This implies that F_h is compact.

(ii) We first prove that F_h is Fréchet differentiable and $\nabla F_h(v) = p[v](x,0)$. Let $u[\delta v]$ denote the solution to the following problem:

$$\begin{cases} \partial_t u[\delta v] + (-\Delta)^s u[\delta v] = 0 & \text{in } \Omega \times (0,T), \\ u[\delta v] = 0 & \text{in } \Omega^c \times (0,T), \\ u[\delta v](\cdot,0) = \delta v & \text{in } \Omega. \end{cases} \tag{3.11}$$

Then for a small variational δv, we can get

$$\begin{aligned} & F_h(v+\delta v) - F_h(v) \\ &= \frac{1}{2}\|\mathcal{K}(v+\delta v) - h\|^2_{L^2(\Omega)} - \frac{1}{2}\|\mathcal{K}(v) - h\|^2_{L^2(\Omega)} \\ &= \int_\Omega \big(u[v](\cdot,T) - h\big) u[\delta v](\cdot,T) dx + o\big(\|\delta v\|_{L^2(\Omega)}\big). \end{aligned}$$

By applying Proposition 3.1 to (3.10) and (3.11), we obtain

$$\int_\Omega \big(u[v](\cdot,T) - h\big) u[\delta v](\cdot,T) dx = \int_\Omega p[v](\cdot,0)\delta v dx. \tag{3.12}$$

This implies

$$F_h(v+\delta v) - F_h(v) = \langle p[v](\cdot,0), \delta v\rangle_{L^2(\Omega)} + o\big(\|\delta v\|_{L^2(\Omega)}\big).$$

Thus, we achieve that

$$\nabla F_h(v) = p[v](x,0).$$

Next, we prove that ∇F_h is Lipschitz continuous. For any $v_1, v_2 \in L^2(\Omega)$, we have $\nabla F_h(v_1) = p[v_1](x,0)$ and $\nabla F_h(v_2) = p[v_2](x,0)$, where $p[v_j], j = 1,2$ are the solutions to system

$$\begin{cases} -\partial_t p[v_j] + (-\Delta)^s p[v_j] = 0 & \text{in } \Omega \times (0,T), \\ p[v_j] = 0 & \text{in } \Omega^c \times (0,T), \\ p[v_j](\cdot,T) = \mathcal{K}(v_j) - h & \text{in } \Omega \end{cases} \tag{3.13}$$

with $j = 1, 2$, respectively. Then $\delta p := p[v_2] - p[v_1]$ satisfies

$$\begin{cases} -\partial_t \delta p + (-\Delta)^s \delta p = 0 & \text{in } \Omega \times (0,T), \\ \delta p = 0 & \text{in } \Omega^c \times (0,T), \\ \delta p(\cdot, T) = \mathcal{K}(v_2) - \mathcal{K}(v_1) & \text{in } \Omega. \end{cases} \tag{3.14}$$

From the estimate (2.5) and the continuity of the operator $\mathcal{K}$, one can get that

$$\begin{aligned} &\|p[v_2](x,0) - p[v_1](x,0)\|_{L^2(\Omega)} \\ &= \|\delta p(x,0)\|_{L^2(\Omega)} \le \|\delta p(x,t)\|_{C([0,T];L^2(\Omega))} \\ &\le C\|\mathcal{K}(v_2) - \mathcal{K}(v_1)\|_{L^2(\Omega)} \le C^{'}\|v_2 - v_1\|_{L^2(\Omega)}. \end{aligned}$$

Thus, ∇F_h is Lipschitz continuous.

(iii) According to (ii), for any v and $v+\delta v, \nabla F_h(v+\delta v) = p[v+\delta v](x,0), \nabla F_h(v) = p[v](x,0)$, $\delta p = p[v+\delta v] - p[v]$ is the solution to the following problems:

$$\begin{cases} -\partial_t \delta p + (-\Delta)^s \delta p = 0 & \text{in } \Omega \times (0,T), \\ \delta p = 0 & \text{in } \Omega^c \times (0,T), \\ \delta p(\cdot, T) = \mathcal{K}(v+\delta v) - \mathcal{K}(v) & \text{in } \Omega. \end{cases} \tag{3.15}$$

Since the operator $\mathcal{K}$ is linear, problem (3.15) is equivalent to the problem

$$\begin{cases} -\partial_t \delta p + (-\Delta)^s \delta p = 0 & \text{in } \Omega \times (0,T), \\ \delta p = 0 & \text{in } \Omega^c \times (0,T), \\ \delta p(\cdot, T) = u[\delta v](\cdot, T) & \text{in } \Omega. \end{cases} \tag{3.16}$$

$u[\delta v](x,t)$ is the solution to problem (3.11). Denote that $\mathcal{K}_0 : L^2(\Omega) \to L^2(\Omega), \delta v \mapsto \mathcal{K}_0 \delta v := \delta p(x, 0; \delta v)$, then the operator $\mathcal{K}_0$ is a bounded linear operator,

$$\nabla F_h(v + \delta v) - \nabla F_h(v) = \mathcal{K}_0 \delta v.$$

Thus, for any $v \in L^2(\Omega)$, $\nabla F_h(v)$ is Fréchet differentiable, and $\nabla^2 F_h(v) = \mathcal{K}_0$. By applying Proposition 3.1 to (3.11) and (3.16), we obtain that for any $\delta v \in L^2(\Omega)$, we obtain that for any $\delta v \in L^2(\Omega)$,

$$\int_\Omega \delta p(x,T) u[\delta v](x,T) dx = \int_\Omega \delta v \delta p(x,0) dx \;\; \Leftrightarrow \;\; \|u[\delta v](x,T)\|^2_{L^2(\Omega)} = \langle \mathcal{K}_0 \delta v, \delta v \rangle_{L^2(\Omega)}.$$

Thus, for any $v \in L^2(\Omega), \nabla^2 F_h(v) = \mathcal{K}_0$ is positive definite, and F_h is strictly convex. □

Next, we address the existence and uniqueness, stability, and convergence of the regularized optimization problem.

Lemma 3.2. *The functional $\Phi(v)$ given by (3.2) has the following properties:*

(i) $\Phi(v)$ is nonnegative, proper, convex and weakly lower semi-continuous.

(ii) $\Phi(v)$ is coercive. Furthermore, there exists a constant c_1, such that for any $v \in L^2(\Omega)$,

$$\Phi(v) \ge c_1 \|v\|_{L^2(\Omega)}.$$

(iii) If the sequence $\{v_n\}_{n\in\mathbb{N}} \subset L^2(\Omega)$ *converges weakly to* $v \in L^2(\Omega)$ *and* $\Phi(v_n)$ *converges strongly to* $\Phi(v)$*, then* $\Phi(v_n - v)$ *converges strongly to zero.*

Proof. The proof can be found in [18] and the references therein. □

Theorem 3.1. *For each regularization parameter* $\alpha > 0$*, the regularization problem* (3.1) *has a unique minimizer.*

Proof. Noting that the objective functional $J_{\alpha,h^\delta}(v) \geq 0$ of regularization problem (3.1) is bounded below. By Lemma 3.1, F_{h^δ} is continuous and convex, so it is also weakly lower semi-continuous. Combined with the weakly lower semi-continuity of Φ, the objective functional is also weakly lower semi-continuous. Moreover, the objective functional is coercive, and thus there exists a minimizer. Finally, F_{h^δ} is strictly convex and $\Phi(v)$ is convex, and so the objective functional is strictly convex, and hence there exists a unique minimizer for the regularization problem. □

Theorem 3.2 (Stability). *Fixed* $\alpha > 0$*. Suppose that sequence* $\{h_n\}_{n=1}^{\infty}$ *converges to* h^δ *in* $L^2(\Omega)$ *and* $v_n = \operatorname{argmin}_{v\in L^2(\Omega)} J_{\alpha,h_n(v)}$*, then* $\{v_n\}_{n=1}^{\infty}$ *converges to the minimizer of* $J_{\alpha,h^\delta(v)}$*.*

Proof. Owing to $v_n = \operatorname{argmin}_{v\in L^2(\Omega)} J_{\alpha,h_n(v)}$, we have that for any $v \in L^2(\Omega)$,

$$\frac{1}{2}\|\mathcal{K}(v_n) - h_n\|^2 + \alpha\Phi(v_n) \leq \frac{1}{2}\|\mathcal{K}(v) - h_n\|^2 + \alpha\Phi(v).$$

Therefore, the sequence $\{\Phi(v_n)\}$ is bounded. By Lemma 3.2, it follows that $\{v_n\}$ is also bounded. Consequently, there exist a weakly convergent subsequence $\{v_{n_k}\}$ of $\{v_n\}$ and $v^* \in L^2(\Omega)$, which satisfies

$$v_{n_k} \rightharpoonup v^*.$$

Combined with the compactness of the operator $\mathcal{K}$, we have

$$\mathcal{K}(v_{n_k}) \to \mathcal{K}(v^*).$$

From the fact that $h_n \to h^\delta$, one can check that

$$\mathcal{K}(v_{n_k}) - h_{n_k} \to \mathcal{K}(v^*) - h^\delta. \tag{3.17}$$

According to the continuity of $\|\cdot\|_{L^2(\Omega)}$ and the weakly lower semi-continuity of $\Phi(\cdot)$, it is shown that

$$\begin{aligned} J_{\alpha,h^\delta}(v^*) &= \frac{1}{2}\|\mathcal{K}(v^*) - h^\delta\|^2 + \alpha\Phi(v^*) \\ &\leq \frac{1}{2}\liminf_{n_k}\|\mathcal{K}(v_{n_k}) - h_{n_k}\|^2 + \alpha\liminf_{n_k}\Phi(v_{n_k}) \\ &\leq \liminf_{n_k}\left(\frac{1}{2}\|\mathcal{K}(v_{n_k}) - h_{n_k}\|^2 + \alpha\Phi(v_{n_k})\right) \\ &\leq \lim_{n_k}\left(\frac{1}{2}\|\mathcal{K}(v) - h_{n_k}\|^2 + \alpha\Phi(v)\right) = J_{\alpha,h^\delta}(v) \end{aligned} \tag{3.18}$$

for any $v \in L^2(\Omega)$. This means that v^* is a minimizer of J_{α,h^δ}.

Next, we prove that $\Phi(v_{n_k}) \to \Phi(v^*)$. By contradiction. Assume that $\Phi(v_{n_k}) \nrightarrow \Phi(v^*)$, it must be

$$\Phi(v^*) \leq \liminf_{n_k} \Phi(v_{n_k}) < \limsup_{n_k} \Phi(v_{n_k}), \tag{3.19}$$

we set $A := \limsup_{n_k} \Phi(v_{n_k})$. There exist a subsequence $\{v_{n_{k,l}}\}$ of sequence $\{v_{n_k}\}$ and

$$A = \lim_{n_{k,l}} \Phi(v_{n_{k,l}}).$$

Let $v = v^*$ in (3.18), we can get that

$$\begin{aligned}
&\frac{1}{2}\|\mathcal{K}(v^*) - h^\delta\|^2 + \alpha\Phi(v^*) \\
&= \liminf_{n_{k,l}} \left(\frac{1}{2}\|\mathcal{K}(v_{n_{k,l}}) - h_{n_{k,l}}\|^2 + \alpha\Phi(v_{n_{k,l}})\right) \\
&= \frac{1}{2}\liminf_{n_{k,l}} \|\mathcal{K}(v_{n_{k,l}}) - h_{n_{k,l}}\|^2 + \alpha A.
\end{aligned}$$

This together with the inequality (3.19) implies that

$$\begin{aligned}
\frac{1}{2}\|\mathcal{K}(v^*) - h^\delta\|^2 &= \liminf_{n_{k,l}} \frac{1}{2}\|\mathcal{K}(v_{n_{k,l}}) - h_{n_{k,l}}\|^2 + \alpha\big(A - \Phi(v^*)\big) \\
&> \liminf_{n_{k,l}} \frac{1}{2}\|\mathcal{K}(v_{n_{k,l}}) - h_{n_{k,l}}\|^2.
\end{aligned}$$

This contradicts the convergence of (3.17). It follows from Lemma 3.2 that $v_{n_k} - v^*$ converges to 0 with respect to Φ, and thus we have $v_{n_k} \to v^*$.

Finally, we show that v_n converges to v^*. If v_n does not converge to v^*, there exists a subsequence $\{v_{m_k}\}$ of sequence $\{v_n\}$ satisfying

$$\|v_{m_k} - v^*\| \geq \varepsilon, \tag{3.20}$$

where ε is an arbitrarily small positive number. Also since $\{v_{m_k}\}$ is a bounded sequence, there exists a weakly convergent subsequence, which by the previous proof converges to the minimizer of J_{α,h^δ}. we can know that the minimizer of J_{α,h^δ} exists and is unique from Theorem 3.1, so the subsequence $\{v_{m_k}\}$ converges to v^*, which contradicts (3.20). □

Definition 3.1. *Assume that $v^* \in L^2(\Omega)$ satisfies: $\Phi(v^*) = \inf\{\Phi(v) : \mathcal{K}(v) = h\}$, We call v^* a Φ-minimizing solution of $\mathcal{K}(v) = h$.*

Theorem 3.3 (Convergence). *Let $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ satisfy*

$$\alpha(\delta) \;\to\; 0, \quad \frac{\delta}{\alpha(\delta)} \;\to\; 0 \quad as \quad \delta \;\to\; 0.$$

Suppose that

$$h^{\delta_n} \in L^2(\Omega) : \|h^\delta - h^{\delta_n}\| \leq \delta_n, \quad \delta_n \;\to\; 0. \tag{3.21}$$

Set $\alpha_n = \alpha(\delta_n), v_n = \operatorname{argmin}_{v\in L^2(\Omega)} J_{\alpha_n,h^{\delta_n}}(v)$. Then there exist a Φ-minimizing solution v^* and a subsequence of $\{v_n\}$ that converges to v^* in $L^2(\Omega)$. Furthermore, $\{v_n\}$ converges to v^* in $L^2(\Omega)$ if the Φ-minimizing solution is unique.

Proof. Let $\tilde{v}$ be any solution satisfying $\mathcal{K}(v) = h^\delta$. By the definition of $\{v_n\}$, it follows that

$$\begin{aligned}
&\frac{1}{2}\left\|\mathcal{K}(v_n) - h^{\delta_n}\right\|^2 + \alpha_n \Phi(v_n) \\
\leq\ &\frac{1}{2}\|\mathcal{K}(\tilde{v}) - h^{\delta_n}\|^2 + \alpha_n \Phi(\tilde{v}) \\
\leq\ &\frac{1}{2}\delta_n^2 + \alpha_n \Phi(\tilde{v}).
\end{aligned}$$

We deduce that

$$\left\|\mathcal{K}(v_n) - h^{\delta_n}\right\| \ \to\ 0, \quad \Phi(v_n) \leq \frac{\delta_n^2}{2\alpha_n} + \Phi(\tilde{v}). \tag{3.22}$$

Taking the supremum on both sides of the last inequality yields

$$\limsup_n \Phi(v_n) \leq \Phi(\tilde{v}). \tag{3.23}$$

This means that $\{\Phi(v_n)\}$ is a bound sequence. Thus, according to Lemma 3.2, $\{v_n\}$ is also bound. There exist a $v^* \in L^2(\Omega)$ and $\{v_{n_k}\} \subset \{v_n\}$ satisfying

$$v_{n_k} \ \rightharpoonup\ v^*.$$

Since operator $\mathcal{K}$ is compact, we have that

$$\mathcal{K}(v_{n_k}) \ \to\ \mathcal{K}(v^*).$$

Owing to

$$\left\|\mathcal{K}(v_{n_k}) - h^\delta\right\| \leq \left\|\mathcal{K}(v_{n_k}) - h^{\delta_{n_k}}\right\| + \left\|h^{\delta_{n_k}} - h^\delta\right\| \ \to\ 0,$$

which implies that $\mathcal{K}(v^*) = h^\delta$. Next, we prove that v^* is a Φ-minimizing solution of $\mathcal{K}(v) = h^\delta$. By the weakly lower semi-continuity of Φ, it follows that

$$\Phi(v^*) \leq \liminf_{n_k} \Phi(v_{n_k}).$$

Combining with inequality (3.23), we can find that v^* is a Φ-minimizing solution of $\mathcal{K}(v) = h^\delta$ and $\lim_{n_k} \Phi(v_{n_k}) = \Phi(v^*)$. Hence, one can get that $v_{n_k} \to v^*$. At last, $v_n \to v^*$ can be proven by method of contradiction which is similar to Theorem 3.2. □

4. Numerical Approximations

Noting that the objective functional of regularized optimization problem (3.1) is strictly convex, we adopt Nesterov's accelerated algorithm as complement of numerical experiments. To solve the optimization problem (3.1) via Nesterov's accelerated algorithm, both the forward problem (2.3) and the Fréchet derivative of the functional $F_{h^\delta}(v)$ must be resolved. From Theorem 3.1 we know $F_{h^\delta}(v) = p[v](x, 0)$, where $p[v](x, t)$ satisfies

$$\begin{cases}
-\partial_t p[v] + (-\Delta)^s p[v] = 0 & \text{in } \Omega \times (0, T), \\
p[v] = 0 & \text{in } \Omega^c \times (0, T), \\
p[v](\cdot, T) = \mathcal{K}(v) - h^\delta & \text{in } \Omega.
\end{cases} \tag{4.1}$$

Assume that the exact initial condition has a sparse representation on the basis composed of piecewise polynomial ansatz functions. We employ the finite element method to numerically solve the forward problem (2.3) and the problem (4.1).

First, we discretize the temporal and spatial domains. Let $k := \Delta t = T/M$, which partitions the time interval $[0, T]$ into M equal subintervals. Define $u^m = u(\cdot, t_m), t_m = m\Delta t$, $m = 0, \dots, M$. The time derivative is discretized using the backward Euler method $\partial_t u \approx (u^m - u^{m-1})/\Delta t, m = 1, \dots, M$. For spatial discretization, consider the spatial domain $\Omega \subset \mathbb{R}^N$ with $N = 1, 2$. For the one-dimensional case ($N = 1$), Ω is partitioned into equal intervals, whereas for the two-dimensional case ($N = 2$), we consider a uniform triangular mesh $\mathcal{T}_{\tilde{h}}$ on Ω, where $\tilde{h}$ denotes the maximum diameter of elements in $\mathcal{T}_{\tilde{h}}$.

We define the discrete space $\mathbb{V}_{\tilde{h}}$ using standard continuous piecewise linear functions. We introduce the basis functions $\{\psi_1, \cdots, \psi_{\tilde{N}}\} \subset \mathbb{V}_{\tilde{h}}$ corresponding to the nodes $\{x_1, \cdots, x_{\tilde{N}}\}$. Thus, the solution $u[v](x, t)$ can be expressed as

$$u_{\tilde{h}}[v](x, t_m) = \sum_{j=1}^{\tilde{N}} u(x_j, t_m)\psi_j(x) = \sum_{j=1}^{\tilde{N}} u_j^m \psi_j(x).$$

The weak form of the forward problem is

$$\begin{cases} \displaystyle\int_\Omega \frac{u^m - u^{m-1}}{\Delta t}\phi_{\tilde{h}} dx + \mathfrak{a}^s\big(u^m, \phi_{\tilde{h}}\big) = 0, \quad \forall\, \phi_{\tilde{h}} \in \mathbb{V}_{\tilde{h}}, \\ \displaystyle u^0 = \sum_{j=1}^{\tilde{N}} v(x_j)\psi_j(x) = \sum_{j=1}^{\tilde{N}} v_j \psi_j(x), \end{cases} \qquad m = 1, \dots, M, \tag{4.2}$$

$\mathfrak{a}^s$ is the bilinear form defined by (2.1). The weak form of problem (4.1) is

$$\begin{cases} \displaystyle\int_\Omega \frac{p^m - p^{m+1}}{\Delta t}\phi_{\tilde{h}} dx + \mathfrak{a}^s\big(p^m, \phi_{\tilde{h}}\big) = 0, \quad \forall\, \phi_{\tilde{h}} \in \mathbb{V}_{\tilde{h}}, \\ \displaystyle p^M = u^M - \sum_{j=1}^{\tilde{N}} h^\delta(x_j)\psi_j(x) = u^M - \sum_{j=1}^{\tilde{N}} h_j^\delta \psi_j(x), \end{cases} \qquad m = 0, \dots, M-1. \tag{4.3}$$

Let $\phi_{\tilde{h}} = \psi_i, i = 1, \dots, \tilde{N}$ in (4.2) and (4.3). Then solving the forward problem (2.3) and problem (4.1) is equivalent to solving the following linear systems:

$$\big(kA_{\tilde{h}}^s + M_{\tilde{h}}\big)U^m = M_{\tilde{h}} U^{m-1}, \tag{4.4}$$

$$\big(kA_{\tilde{h}}^s + M_{\tilde{h}}\big)p^m = M_{\tilde{h}} p^{m+1}, \tag{4.5}$$

where U^m denotes the approximate solution to the forward problem at time t_m, the stiffness matrix

$$A_{\tilde{h}}^s = \big(a_{ij}^s\big)_{\tilde{N}\times\tilde{N}}, \quad a_{ij}^s = \mathfrak{a}^s(\psi_i, \psi_j) = \frac{C_{N,s}}{2}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\big(\psi_i(x) - \psi_i(y)\big)\big(\psi_j(x) - \psi_j(y)\big)}{|x - y|^{N+2s}} dxdy,$$

and the mass matrix

$$M_{\tilde{h}} = (m_{ij})_{\tilde{N}\times\tilde{N}}, \quad m_{ij} = \int_\Omega \psi_j(x)\psi_j(x) dx.$$

The main challenge in solving the discrete problem via the finite element method lies in computing the stiffness matrix $A_{\tilde{h}}^s$, as the fractional Laplace operator is a non-local operator requiring the evaluation of singular integrals. one can see [1].

The Nesterov's accelerated algorithm used in this paper (Algorithm 4.1) was proposed in [18, 20]. We make the weighting function

$$w_i = \left(\int_\Omega \psi_i^2(x) dx \right)^{-1}$$

in optimization problem (3.2), and thus functional $\Phi(v)$ can be numerically discretized as $\sum_{j=1}^{\tilde{N}} |v(x_j)|$ in the finite method. The soft-thresholding operator is defined as

$$(\mathrm{S}_{\lambda\mu}(v))_i = \mathrm{sgn}(v_i) \max(|v_i| - \lambda\mu, 0) = \mathrm{sgn}(v_i)(|v_i| - \lambda\mu)_+.$$

Algorithm 4.1: Nesterov's Accelerated Algorithm.

Input: $v^0, A_0, z^0, w^0, \mu, \alpha, s^0, \underline{s}, \overline{s}, \eta, Tol, Maxiter$.
for $n = 0, 1, 2, \ldots, Maxiter$ **do**
  $a_{n+1} \leftarrow \dfrac{1 + \sqrt{1 + 2A_n s^n}}{s^n}$.
  $y^n \leftarrow \dfrac{A_n v^n + a_{n+1} w^n}{A_n + a_{n+1}}$.
  repeat
    $v^{n+1} \leftarrow \mathbb{S}_{\alpha/s^n} \left(y^n - \dfrac{1}{s^n} F'_{h^\delta}(y^n) \right)$.
    if $\left\| F'_{h^\delta}(v^{n+1}) - F'_{h^\delta}(y^n) \right\|^2 > s^n \left\langle F'_{h^\delta}(y^n) - F'_{h^\delta}(v^{n+1}), y^n - v^{n+1} \right\rangle$ **then**
      $s^n \leftarrow s^n \cdot \mu$.
    end
  until $\left\| F'_{h^\delta}(v^{n+1}) - F'_{h^\delta}(y^n) \right\|^2 \le s^n \left\langle F'_{h^\delta}(y^n) - F'_{h^\delta}(v^{n+1}), y^n - v^{n+1} \right\rangle$ *or* $s^n \notin [\underline{s}, \overline{s}]$
  $A_{n+1} \leftarrow A_n + a_{n+1}$.
  $w^{n+1} \leftarrow \mathbb{S}_{\alpha A_{n+1}} \left(z^n - a_{n+1} F'_{h^\delta}(v^{n+1}) \right)$.
  $s^{n+1} \leftarrow s^n / \mu$.
  if $\left\| v^{n+1} - \mathbb{S}_{\eta\alpha} \left(v^{n+1} - \eta F'_{h^\delta}(v^{n+1}) \right) \right\| \le Tol$ **then**
    Break
  end
end
Output: $v = v^{end}$.

Next, several numerical experiments will be conducted to demonstrate the effectiveness of the proposed algorithm in 1D and 2D spatial domains. The solution $u[v_{true}](x, T)$ of the forward problem with the initial value v_{true} is used as the measurement data h. To test the stability of the algorithm, a random perturbation is added to the measurement data h to obtain the noisy data

$$h^\delta(x) = h(x) + \delta h(x) * 2\mathrm{rand}\big(\mathrm{size}\big(h(x)\big) - 1\big),$$

where the parameter δ represents the noise level, and $2\mathrm{rand}(\mathrm{size}(h(x)) - 1)$ generates a random array in the interval $[-1, 1]$. The absolute error r_1 and relative error r_2 between the exact initial value and the reconstructed initial value are defined as

$$r_1 = \|v_{true} - v_c\|_{L^2(\Omega)} = \sqrt{\sum_{i=1}^{\tilde{N}} \big(v_{true}(x_i) - v_c(x_i)\big)^2},$$

$$r_2 = \frac{\|v_{true} - v_c\|_{L^2(\Omega)}}{\|v_{true}\|_{L^2(\Omega)}} = \frac{\sqrt{\sum_{i=1}^{\widetilde{N}} \left(v_{true}(x_i) - v_c(x_i)\right)^2}}{\sqrt{\sum_{i=1}^{\widetilde{N}} \left(v_{true}(x_i)\right)^2}}.$$

Here, v_{true} denotes the exact initial value, and v_c represents the reconstructed initial value.

In the Examples 4.1 and 4.2, we consider $\Omega = (-2, 2), \widetilde{N} = 400, M = 15$ with the initial guess $v^0 = 0, T = 0.3$.

Example 4.1. Assume that the exact initial value is

$$v^1_{true}(x) = \begin{cases} x + 0.5, & x \in (-0.5, 0], \\ 0.5 - x, & x \in (0, 0.5), \\ 0, & \text{otherwise.} \end{cases}$$

The numerical inversion results for different fractional orders s are presented in Fig. 4.1. As shown in this figure, all fractional orders s exhibit favorable inversion performance. Specifically, when s=0.3, the singularity of the image is well reconstructed under the noise level δ<0.005 with significantly better inversion results than those of $s = 0.5$ and 0.7. However, the inversion accuracy deteriorates notably at $\delta = 0.01$, indicating that a smaller fractional order s facilitates superior singularity reconstruction but renders the inversion more susceptible to noise interference. Additionally, we investigated the impact of the final time on the inversion performance. With the fractional order fixed at $s = 0.5$, we examined the absolute and relative errors under different noise levels and varying final time T. As shown in Table 4.1, the inversion performance is optimal when $T = 0.3$, but the other results remain acceptable.

Table 4.1: Errors of Example 4.1 with fixed $s = 0.5$ under different final times and noise levels. The first and seconde rows show absolute errors and relative errors, respectively.

T	$\delta = 1 \times 10^{-2}$	$\delta = 2 \times 10^{-2}$	$\delta = 4 \times 10^{-2}$
0.3	0.0876	0.1384	0.2221
	0.0300	0.0479	0.0769
0.5	0.1897	0.2041	0.2244
	0.0657	0.0707	0.0777
0.75	0.2483	0.2489	0.2680
	0.0860	0.0862	0.0926
1	0.4176	0.4496	0.4669
	0.1447	0.1557	0.1617

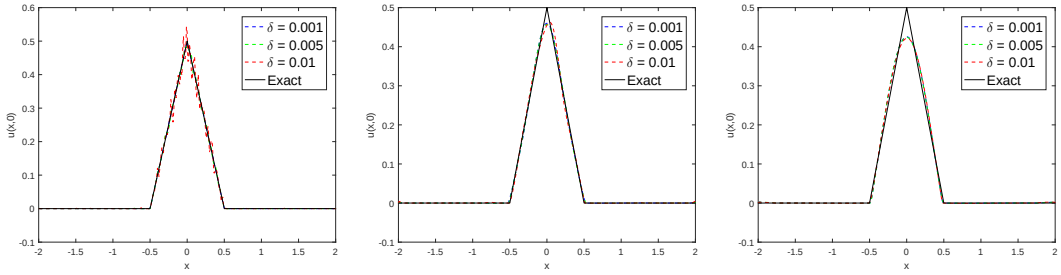


Fig. 4.1. Reconstructed initial values under different fractional orders s, from left to right, the fractional orders are $0.3, 0.5, 0.7$.

Example 4.2. Assume that the exact initial value is

$$v_{true}^2(x) = \begin{cases} 0.25 - x^2, & x \in (-0.5, 0.5), \\ 0, & \text{otherwise.} \end{cases}$$

We investigated the influence of different orders $s \in (0, 1)$ on the inversion performance. To avoid the impact of other parameters, we fixed the noise level at $\delta = 0$. In Fig. 4.2, we present the variation of relative errors for different orders s. It can be seen from Fig. 4.2 that smaller s generally yields better inversion results, and overall, the relative errors are all controlled within 10%, which is acceptable. Fig. 4.3 shows the inversion results for fractional orders $s = 0.3, 0.5, 0.7$, where all orders demonstrate highly satisfactory inversion performance, indicating that our inversion method is highly effective and stable and Table 4.2 presents the absolute errors and relative errors under different fractional orders and noise levels. From this table, we observe that the errors increase as s becomes larger; nevertheless, the overall accuracy remains satisfactory.

In the following examples, we let $\Omega = (-1, 1) \times (-1, 1) \subset \mathbb{R}^2$ with $T = 0.3$. The domain Ω is triangulated into 1368 elements with $\widetilde{N} = 645, M = 15$, and the initial guess is set as $v^0 = 0$.

Example 4.3. Assume that the exact initial value is

$$v_{true}^3(x) = \begin{cases} (0.6 - x^2 - y^2)^2, & x^2 + y^2 \le 0.6, \\ 0, & \text{otherwise.} \end{cases}$$

We also investigated the influence of different fractional orders s on the inversion performance. To avoid the impact of other parameters, the noise level was fixed at $\delta = 0$. In Fig. 4.4 the

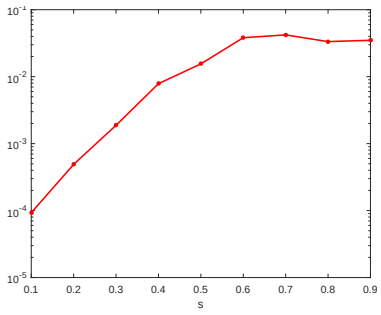
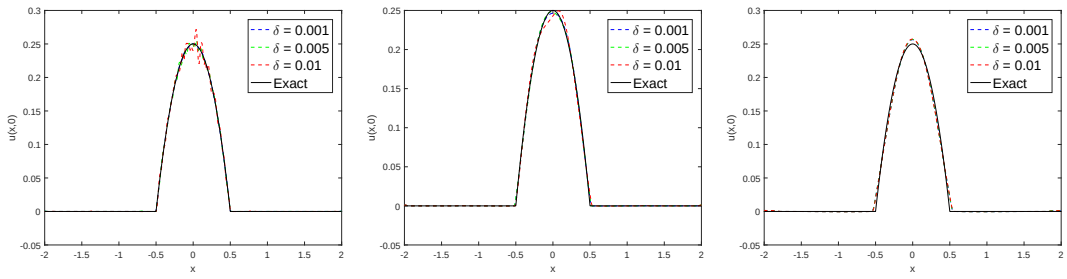


Fig. 4.2. Errors under different fractional orders s.



(a) (b) (c)

Fig. 4.3. Reconstructed initial values under different fractional orders s, from left to right, the fractional orders are $0.3, 0.5, 0.7$.

Table 4.2: Errors of Example 4.2 with fixed $T = 0.3$ under different fractional orders and noise levels. The first and second rows show absolute errors and relative errors, respectively.

s	$\delta = 1 \times 10^{-3}$	$\delta = 5 \times 10^{-3}$	$\delta = 1 \times 10^{-2}$
0.1	0.0016	0.0069	0.0157
	0.0009	0.0038	0.0086
0.3	0.0073	0.0320	0.0739
	0.0040	0.0175	0.0405
0.5	0.0288	0.0288	0.0791
	0.0158	0.0158	0.0355
0.7	0.0769	0.0767	0.0791
	0.0421	0.0420	0.0433
0.9	0.0635	0.0701	0.0834
	0.0348	0.0384	0.0457

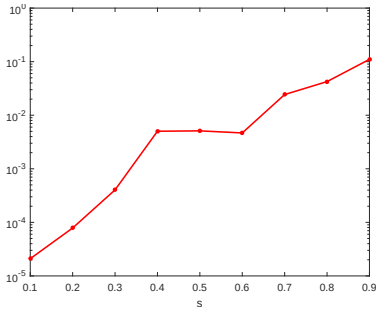


Fig. 4.4. The errors for different fractional orders s.

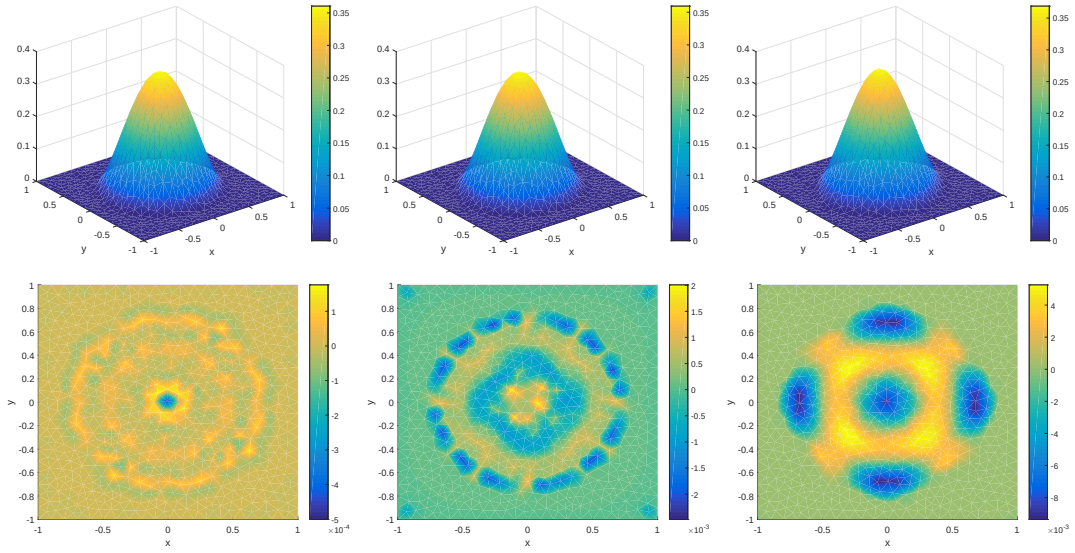
variation of relative errors for different orders $s \in (0, 1)$ is presented. It is evident that, consistent with the one-dimensional case, smaller s yields better inversion results. In general, the relative errors remain within 10%, which is acceptable. Fig. 4.5 shows the inversion results at noise level $\delta = 0$ for different fractional orders s, respectively, where pointwise errors calculation method presented below are denoted as r_1. Fig. 4.6 considers the initial value inversion for $s = 0.5$ under different noise levels, demonstrating that our proposed method remains highly effective and stable in 2D case.

Next, we consider several numerical examples with more general and sparser initial conditions.

Example 4.4. Assume that the exact initial value is

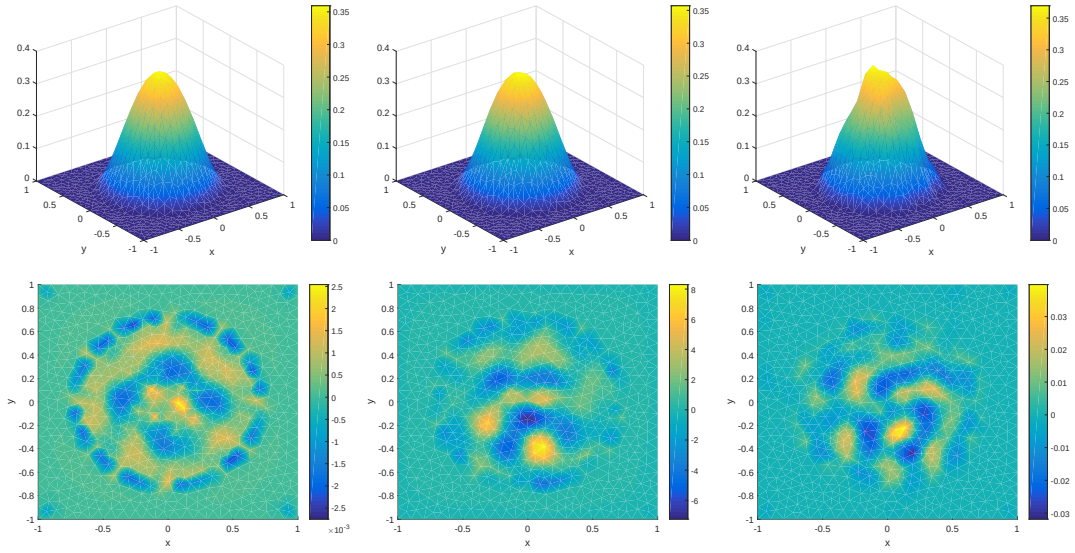
$$v^4_{true}(x) = \begin{cases} 10, & x^2 + y^2 \leq 0.09, \\ 0, & \text{otherwise.} \end{cases}$$

In this example, the support of the exact initial value is a disk centered at the origin, which has a sparse representation on the basis composed of piecewise polynomial ansatz functions. As shown in Fig. 4.7, our numerical method cannot only well reconstruct the support of the sparse initial value but also recover the function values.



(a) (b) (c)

Fig. 4.5. Reconstructed initial values (first row) and pointwise errors (second row) for different fractional orders s, with the fractional orders s from left to right as $0.3, 0.5, 0.7$.

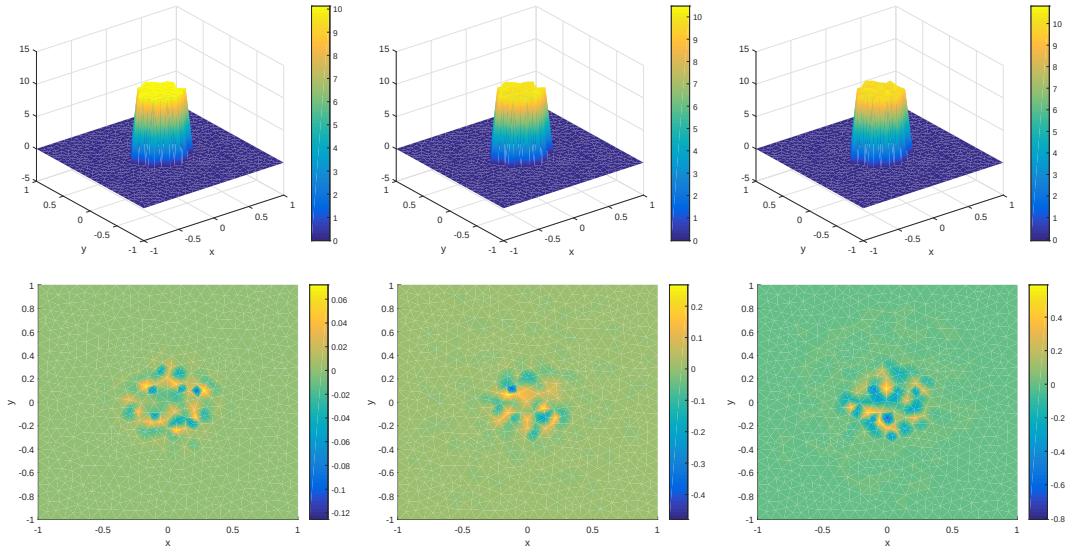


(a) (b) (c)

Fig. 4.6. Fixed $s = 0.5$. Reconstructed initial values (first row) and pointwise errors (second row) under different noise levels, with the noise levels arranged from left to right as $0.001, 0.005, 0.01$.

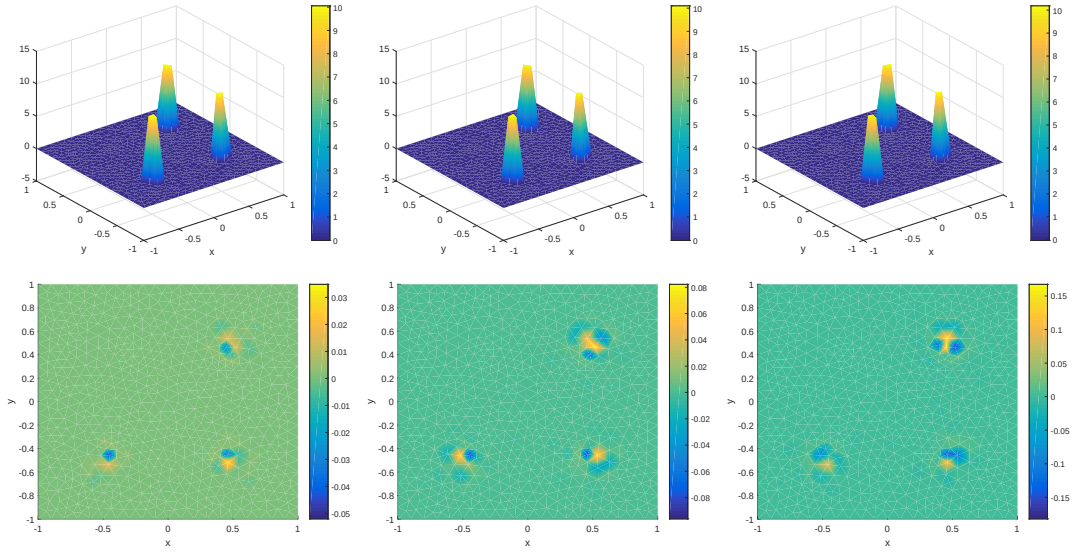
Example 4.5. Assume that the exact initial value is

$$v_{true}^5(x) = \begin{cases} 10, & |x-0.5| \le 0.1, \quad |y-0.5| \le 0.1, \\ 10, & |x+0.5| \le 0.1, \quad |y+0.5| \le 0.1, \\ 10, & |x-0.5| \le 0.1, \quad |y+0.5| \le 0.1, \\ 0, & \text{otherwise.} \end{cases}$$



(a) (b) (c)

Fig. 4.7. Fixed $s = 0.3$. Reconstructed initial values (first row) and pointwise errors (second row) under different noise levels, with the noise levels arranged from left to right as $0.001, 0.005, 0.01$.



(a) (b) (c)

Fig. 4.8. Fixed $s = 0.3$. Reconstructed initial values (first row) and pointwise errors (second row) under different noise levels, with the noise levels arranged from left to right as $0.001, 0.005, 0.01$.

In this example, the support of the exact initial value consists of three highly scattered small rectangles. As shown in Fig. 4.8, for initial values whose supports are composed of multiple scattered small regions, our numerical algorithm still exhibits high effectiveness and stability.

For the last two extremely non-smooth examples, we only show the case when $s = 0.3$, since the relative errors of the inverted initial values become larger as s increases. This may because the solution to the fractional parabolic equation becomes smoother as well as $u(\cdot, T)$

when s is larger. Therefore, it is difficult to obtain an accurate approximation of non-smooth functions from smooth data. The initial guesses are all set to 0, and despite the poor initial guesses, our algorithm still stably and effectively converges to satisfactory inversion numerical solutions. For sparse initial values with supports of different shapes, the inversion performance is consistently favorable. It is well-known that when the fractional order s approaches 1^-, the fractional Laplace operator is equivalent to the classical Laplace operator. For the classical parabolic equation, as shown in the literature [18], the relative error of reconstructing the initial value is large. Additionally, that study only considered the inversion performance under the noise level $\delta = 0.0001$ and final time $T = 0.3$.

5. Conclusion

In this paper, the l_1-regularization method is used to solve the problem of recovering sparse initial condition for the spatial fractional Laplace equation. It is proved that the misfit functional is differentiable and strictly convex, and the regularization problem is well-posed. Moreover, the convergence for the regularized solution is proven. Numerical examples are given to verify that the proposed algorithm is highly effective and stable both in 1D and 2D cases. The exact sparse initial conditions provided in this paper are all non-smooth, which leads to the inversion performance for larger s being slightly worse than that for smaller s. This may be attributed to the properties of the solutions to the spatial fractional Laplace equation: when s increases, the solutions to the equation become smoother, so the observation data $u(\cdot, T)$ will also be smoother. It is difficult to recover non-smooth sparse initial conditions from smooth measurement data.

Acknowledgements. This work is partially supported by the NSFC (Grant No. 12571449).

References

[1] G. Acosta, F.M. Bersetche, and J.P. Borthagaray, A short FE implementation for a 2D homogeneous Dirichlet problem of a fractional Laplacian, *Comput. Math. Appl.*, **74**:4 (2017), 784–816.

[2] H. Antil, D. Verma, and M. Warma, External optimal control of fractional parabolic PDEs, *ESAIM Control Optim. Calc. Var.*, **26** (2020), 20.

[3] D. Applebaum, Lévy processes – from probability to finance and quantum groups, *Notices Amer. Math. Soc.*, **51**:11 (2004), 1336–1347.

[4] M. Ben Salah and S. Tatar, Identification of the initial value for a space-time fractional diffusion equation, *arXiv:2412.05387*, 2024.

[5] F. Bersetche, F. Fuica, E. Otárola, and D. Quero, Fractional, semilinear, and sparse optimal control: A priori error bounds, *Appl. Math. Optim.*, **91**:1 (2025), 20.

[6] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2011.

[7] L. Caffarelli and L. Silvestre, An extension problem related to the fractional Laplacian, *Comm. Partial Differential Equations*, **32**:8 (2007), 1245–1260.

[8] A. Carpinteri and F. Mainardi, *Fractals and Fractional Calculus in Continuum Mechanics*, Springer, 1997.

[9] X. Cheng, Z. Li, and M. Yamamoto, Asymptotic behavior of solutions to space-time fractional diffusion-reaction equations, *Math. Methods Appl. Sci.*, **40** (2017), 1019–1031.

[10] C. Glusa and E. Otárola, Error estimates for the optimal control of a parabolic fractional PDE, *SIAM J. Numer. Anal.*, **59**:2 (2021), 1140–1165.

[11] M. Hrizi, M. Ben Salah, and M. Hassine, Determination of the initial density in nonlocal diffusion from final time measurements, *Discrete Contin. Dyn. Syst. Ser. S*, **15**:6 (2022), 1469–1498.

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

[13] B. Jin, Y. Kian, and Z. Zhou, Inverse problems for subdiffusion from observation at an unknown terminal time, *SIAM J. Appl. Math.*, **83**:4 (2023), 1496–1517.

[14] Y. Kian, É. Soccorsi, and F. Triki, Logarithmic stable recovery of the source and the initial state of time-fractional diffusion equations, *SIAM J. Math. Anal.*, **55**:4 (2023), 3888–3902.

[15] T. Leonori, I. Peral, A. Primo, and F. Soria, Basic estimates for solutions of a class of nonlocal elliptic and parabolic equations, *Discrete Contin. Dyn. Syst.*, **35**:12 (2015), 6031–6068.

[16] Y. Li, S. Osher, and R. Tsai, Heat source identification based on constrained minimization, *Inverse Probl. Imaging*, **8**:1 (2014), 199–221.

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

[18] P.Q. Muoi, ℓ^1-weighted regularization for the problem of recovering sparse initial conditions in parabolic equations from final measurements, *Appl. Numer. Math.*, **183** (2023), 301–316.

[19] E. Otárola, An adaptive finite element method for the sparse optimal control of fractional diffusion, *Numer. Methods Partial Differential Equations*, **36**:2 (2020), 302–328.

[20] Q.M. Pham, N.H. Dinh, P. Maass, and M. Pidcock, Descent gradient methods for nonsmooth minimization problems in ill-posed problems, *J. Comput. Appl. Math.*, **298** (2016), 105–122.

[21] L. Plociniczak, Analytical studies of a time-fractional porous medium equation: Derivation, approximation and applications, *Commun. Nonlinear Sci. Numer. Simul.*, **24** (2015), 169–183.

[22] Y.A. Rossikhin and M.V. Shitikova, Application of fractional calculus for dynamic problems of solid mechanics: Novel trends and recent results, *Appl. Mech. Rev.*, **63** (2010), 010801.

[23] Y. Zhang, X. Liu, M.R. Belić, W. Zhong, Y. Zhang, and M. Xiao, Propagation dynamics of a light beam in afractional Schrödinger equation, *Phys. Rev. Lett.*, **115** (2015), 180403.

[24] Y. Zhang, M.M. Meerschaert, and R.M. Neupauer, Backward fractional advection-dispersion model for contaminant source prediction, *Water Resour. Res.*, **52**:3 (2016), 2462–2473.

[25] Z. Zhang and Z. Zhou, Numerical analysis of backward subdiffusion problems, *Inverse Problems*, **36** (2020), 105006.

[26] G.H. Zheng and Q.G. Zhang, Solving the backward problem for space-fractional diffusion equation by a fractional Tikhonov regularization method, *Math. Comput. Simulation*, **148** (2018), 37–47.