Data regularization for a backward time-fractional diffusion problem

Liyan Wang, Jijun Liu*

Department of Mathematics, Southeast University, Nanjing, 210096, PR China

ARTICLE INFO

Article history:
Received 10 April 2012
Received in revised form 29 July 2012
Accepted 3 October 2012

Keywords:
Backward problem
Diffusion process
Fractional derivative
Regularization
Data fitting
Numerics

ABSTRACT

We investigate a backward problem for a time-fractional diffusion process in inhomogeneous media, which aims to determine the initial status of some physical field such as temperature for slow diffusion from its present measurement data. This problem is well-known to be ill-posed due to the rapid decay of the forward process. By using the eigenfunction expansion, we construct a new regularizing scheme with an explicit solution for the noisy input data with the number of truncation terms as a regularizing parameter. The convergence rate depending on the choice of strategy of the regularizing parameter is given based on the asymptotic behavior of the Mittag-Leffler function. Numerical implementations are presented to show the validity of the proposed scheme for several models.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Research on the backward problems for the diffusion process has been a particularly active field in the past thirty years, which aims to obtain the initial status of a physical field from its measurement data at the present time. The classical convection diffusion model is the parabolic equation

\[ \frac{\partial u}{\partial t} + v \cdot \nabla u = D \Delta u, \quad x \in \Omega \subset \mathbb{R}^m, \quad t > 0. \tag{1.1} \]

The backward problems for this equation have been studied extensively, see [1–6], while the use of the Carelman estimate to get the conditional stability for such a kind of problem can be found in [7].

On the other hand, there are also some slow diffusion processes in applied areas such as mass transfer with a long tail distribution, which cannot be predicted from the classical convection diffusion model \((1.1)\), see [8–10]. It was found that the continuous time random walking (CTRW) model for the particles can describe this long tail phenomenon, see [11–15]. In this case, the normal time derivative in \((1.1)\) should be replaced by the derivative of fractional order \(\gamma \in (0, 1)\):

\[ \frac{\partial^\gamma u}{\partial t^\gamma} + v \cdot \nabla u = D \Delta u, \quad x \in \Omega \subset \mathbb{R}^m, \quad t > 0. \tag{1.2} \]

Since the fractional derivative is nonlocal, this new model is suitable for describing these physical phenomena with memory effect. The fractional derivative of order \(\gamma\) in the sense of Caputo is defined by

\[ \frac{d^\gamma f}{dt^\gamma} := \frac{1}{\Gamma(1 - \gamma)} \int_0^t \frac{f'(s)}{(t - s)^\gamma} ds, \tag{1.3} \]

where \(\Gamma(\cdot)\) is the standard \(\Gamma\)-function. Note that the Caputo fractional derivative at time \(t\) is related to the information of \(f(s)\) for all \(s \in (0, t)\), which is called the “memory effect” of the fractional derivative.

* Corresponding author.
E-mail address: jjliu@seu.edu.cn (J. Liu).

0898-1221/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2012.10.001
For this so-called slow diffusion model, one of the most important research areas in recent years is the topic related to the backward problem. This paper focuses on the backward problem related to the following time-fractional diffusion model:

\[
\begin{align*}
\frac{\partial^\gamma u}{\partial t^\gamma} &= \nabla \cdot (a(x) \nabla u), \quad (x, t) \in \Omega \times (0, T], \\
u(x, t) &= 0, \quad x \in \partial \Omega, \ t \in [0, T], \\
u(x, 0) &= u_0(x), \quad x \in \Omega,
\end{align*}
\]

where \( \gamma \in (0, 1) \) and \( 0 < a_0 \leq a(x) \in C^1(\Omega) \), \( u_0(x) \) is the initial data and \( \Omega \subset \mathbb{R}^2 \) is a bounded domain. (1.4) describes the diffusion process in porous media, while the variable coefficient \( a(x) \) represents the inhomogeneous diffusion at the media.

Our backward problem is to approximate \( u(x, t) \) for \( t \in [0, T) \) from the measurement data \( g^\delta(x) \), the noise-contaminated data for the exact temperature \( g(x) := u(x, T) \) satisfying

\[
\|g^\delta(\cdot) - g(\cdot)\|_{L^2(\Omega)} \leq \delta
\]

for known error level \( \delta > 0 \). In [16], Liu and Yamamoto considered a 1-dimensional backward problem for the equation \( \frac{\partial^\gamma u}{\partial t^\gamma} = u_{xx} \). By modifying the 1-dimensional equation with a regularizing term \( u_{xxx} \), they constructed a regularizing strategy to recover the initial state \( u(x, 0) \), and find that the initial data for this fractional diffusion equation can be recovered in a more efficient way, comparing with the standard backward problem for the heat conduction equation [3,4] in the 1-dimensional case. The advantage of the regularizing scheme constructed there is that the regularizing solution can be expressed explicitly by eigenfunction expansion. However, for our higher spatial dimensional model (1.4), the introduction of higher spatial partial derivatives into the equation as regularizing terms becomes much more difficult, due to the nonsymmetric divergence operator \( \nabla \cdot (a(x) \nabla u) \).

In recent years, the attempts to construct a regularizing scheme such that the regularizing solution can be expressed explicitly have received much attention. The advantage of this new idea is that the well-posedness of the regularizing problem is guaranteed automatically, the only remaining issue is the convergence analysis on the regularizing solution. Then the numerical computation of the regularizing solution for all \( t \in [0, T) \) is much easier, for example, see [1,17–19] for the mollification method, where the final measurement data are regularized by using the Dirichlet kernel or the \( \delta \)-mollification of the measurement data. We call such a kind of scheme as a data regularization technique.

In this paper, we use the data regularization technique to deal with the backward problem related to (1.4), that is, for given noisy data \( g^\delta(x) \) of \( u(x, T) \), we try to determine \( u(x, t) \) for \( t \in [0, T) \) approximately. By using the eigenfunction expansion, we solve this ill-posed problem by an optimization problem, which is essentially a regularizing scheme for the noisy input data with the number of truncation terms as the regularizing parameter. The Hölder convergence \( O(\delta^{\beta/2}) \) is established uniformly for all \( t \in [0, T) \), under the a priori information about the bound on \( \|u_0\|_{H^\beta(\Omega)} \). The implementation of such a kind of regularizing strategy has been applied for the standard heat equation in [20].

The organization of our paper is as follows. In Section 2, we propose the regularizing scheme for the backward problem of (1.4). Then we consider the choice of strategy for the regularizing parameter and establish the convergence rate in Section 3. Finally, the numerical implementations are given in Section 4.

2. Regularization for the backward problem

To construct our regularizing solution, we need the Mittag-Leffler function and its asymptotic behavior. The double-parameter Mittag-Leffler function is defined by

\[
E_{\gamma, \beta}(z) := \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\gamma k + \beta)}, \quad z \in \mathbb{C}, \gamma > 0, \beta \geq 0.
\]

As a direct result of Theorems 1.3 and 1.4 in [21], we have the following estimates for \( E_{\gamma, 1}(x) \) which will be used in our paper.

**Lemma 2.1.** (i) Assume that \( \gamma \in (0, 1) \), then

\[
E_{\gamma, 1}(x) = \frac{1}{\gamma} e^{x^{1/\gamma}} - \frac{1}{x^{\gamma}(1 - \gamma)} + O \left( \frac{1}{x^2} \right), \quad \text{as } x \to +\infty,
\]

\[
E_{\gamma, 1}(x) = -\frac{1}{x^{\gamma}(1 - \gamma)} + O \left( \frac{1}{x^2} \right), \quad \text{as } x \to -\infty.
\]

(ii) Assume that \( 0 < \gamma_0 < \gamma_1 < 1 \). Then there exist constants \( C_1, C_2 > 0 \) depending only on \( \gamma_0, \gamma_1 \) such that

\[
\frac{C_1}{\Gamma(1 - \gamma)} \frac{1}{1 - x} \leq E_{\gamma, 1}(x) \leq \frac{C_2}{\Gamma(1 - \gamma)} \frac{1}{1 - x}, \quad \text{for all } x \leq 0.
\]

These estimates are uniform for all \( \gamma \in [\gamma_0, \gamma_1] \).
The upper bound for $E_{γ,1}(x)$ in (ii) is obvious. As for the lower bound, noticing
\[-\frac{1}{x} = \frac{1}{1-x} + o \left( \frac{1}{1-x} \right), \quad x \to -\infty,\]
so it follows from the second asymptotic in (i) that
\[E_{γ,1}(x) = \frac{1}{\Gamma(1-γ)} \frac{1}{1-x} + o \left( \frac{1}{1-x} \right) \geq \frac{1}{2\Gamma(1-γ)} \frac{1}{1-x}, \quad x \in (-\infty, -L)\]
for some large $L > 0$. For all $x \in [-L, 0]$, it follows that
\[E_{γ,1}(x) \geq \min_{x \in [-L, 0]} E_{γ,1}(x) \geq \frac{C_1}{\Gamma(1-γ)} \frac{1}{1-x} \geq \frac{C_1}{\Gamma(1-γ)} \frac{1}{1-x}, \quad x \in (-\infty, -L)\]
for some $C_1 > 0$ small enough, noticing $E_{γ,1}(x) > 0$ in $[-L, 0]$. The lower bound in (ii) follows.

Denote by $\{(λ_n, ϕ_n(x)) : n \in \mathbb{N}\}$ the eigen-system of the operator $-\nabla \cdot (a(x)\nabla ϕ_n)$, acting on $H^2(Ω) \cap H^1_0(Ω)$. That is, $ϕ_n(x)$ satisfies
\[
\begin{cases}
\nabla \cdot (a(x)\nabla ϕ_n(x)) + λ_n ϕ_n(x) = 0, & x \in Ω, \\
ϕ_n(x) = 0, & x \in \partial Ω.
\end{cases}
\]
It is well known that $\{ϕ_n(x)\}_{n=1}^∞$ constitutes the base of $L^2(Ω)$ and
\[0 < λ_1 \leq λ_2 \leq \cdots \leq λ_n \leq \cdots, \quad λ_n \to +\infty.\]
Assume that $\{ϕ_n(x)\}_{n=1}^∞$ is normal orthogonal, hence we have
\[u_0(x) = ∑_{n=1}^{∞} c_n ϕ_n(x), \quad c_n = ∫_Ω u_0(x)ϕ_n(x)dx, \quad n \in \mathbb{N}. \quad (2.1)\]
It is easy to show that $λ_n$ has the following representations in terms of the eigenfunction:
\[λ_n = ∥a^{1/2}\nabla ϕ_n∥_{L^2(Ω)}^2, \quad λ_n^2 = ∥\nabla \cdot (a^{1/2}\nabla ϕ_n)∥_{L^2(Ω)}^2\]
and
\[∫_Ω a^{1/2}\nabla ϕ_n \cdot \nabla ϕ_m dx = 0, \quad ∫_Ω \nabla \cdot (a^{1/2}\nabla ϕ_n) \nabla \cdot (a^{1/2}\nabla ϕ_m) dx = 0, \quad n \neq m.\]
Define the functions
\[ϕ_n(x, t) := E_{γ,1}(-λ_n t^γ)ϕ_n(x), \quad (2.2)\]
which satisfy
\[
\begin{cases}
\frac{∂^γ ϕ_n(x, t)}{∂t^γ} = \nabla \cdot (a(x)\nabla ϕ_n(x, t)), & (x, t) \in Ω × (0, T], \\
ϕ_n(x, t) = 0, & x \in \partial Ω, t \in [0, T], \\
ϕ_n(x, 0) = ϕ_n(x), & x \in Ω.
\end{cases} \quad (2.3)
\]
So the exact solution of (1.4) can be expressed as
\[u(x, t) = ∑_{n=1}^{∞} c_n ϕ_n(x, t). \quad (2.4)\]
Especially, the final value has the expansion
\[g(x) = u(x, T) = ∑_{n=1}^{∞} c_n ϕ_n(x, T). \quad (2.5)\]
In practice, only the noisy data $g^δ(x)$ of $g(x)$ is given and we can only compute finite terms of the series in (2.5). Therefore, $c_n$ are determined from the approximate equation to (2.5), i.e.,
\[∑_{n=1}^{M} c_n^δ ϕ_n(x, T) = g^δ(x), \quad (2.6)\]
where the truncation term number $M$ is unknown in advance which will affect the inversion results. So the solution to (2.6) for determining both $M$ and $\{c_n^δ\}$ is ill-posed, the regularization method should be applied to solve (2.6).
To this end, let
\[ C_{M, \delta}^{\epsilon, \delta} := (c_1^{\epsilon, \delta}, c_2^{\epsilon, \delta}, \ldots, c_M^{\epsilon, \delta}) \in \mathbb{R}^M \]
be the minimum norm solution of Equation (2.6) with discrepancy \( \epsilon \), i.e.,
\[ \left| \left| \sum_{n=1}^{M} c_n^{\epsilon, \delta} \Phi_n(\cdot, T) - g^\delta(\cdot) \right| \right| \leq \epsilon, \tag{2.7} \]
\[ \left| \left| C_{M, \delta}^{\epsilon, \delta} \right| \right|_{\mathbb{R}^M} = \inf \left\{ \left| \left| C_{n, \delta}^{\epsilon, \delta} \right| \right|_{\mathbb{R}^M} : \left| \left| \sum_{n=1}^{M} c_n^{\epsilon, \delta} \Phi_n(\cdot, T) - g^\delta(\cdot) \right| \right| \leq \epsilon \right\}. \tag{2.8} \]
where \( c_n^{\epsilon, \delta} := (c_1^{\epsilon, \delta}, c_2^{\epsilon, \delta}, \ldots, c_M^{\epsilon, \delta}) \). In the following we choose \( \epsilon = \delta \). The positive integer \( M := M(\delta) \) as the regularizing parameter will be specified later.

Define the operator \( K : \mathbb{R}^M \to L^2(\Omega) \) by
\[ (KC_{M, \delta}^{\epsilon, \delta})(x) := \sum_{n=1}^{M} c_n^{\epsilon, \delta} \Phi_n(x, T) \tag{2.9} \]
with its adjoint operator under the dual system \( \langle \mathbb{R}^M, \mathbb{R}^M \rangle \) and \( \langle L^2(\Omega), L^2(\Omega) \rangle \):
\[ K^*f = \left( \int_{\Omega} f(x)\Phi_1(x, T)dx, \int_{\Omega} f(x)\Phi_2(x, T)dx, \ldots, \int_{\Omega} f(x)\Phi_M(x, T)dx \right)^T. \tag{2.10} \]
The minimum norm solution \( C_{M, \delta}^{\epsilon, \delta} \) can be solved by Tikhonov regularization with the Morozov principle [22], i.e., \( C_{M, \delta}^{\epsilon, \delta} \) is the solution of the following equation
\[ (\alpha(\delta)I + K^*K)C_{M, \delta}^{\epsilon, \delta} = K^*g^\delta(x), \tag{2.11} \]
where the regularizing parameter \( \alpha = \alpha(\delta) \) is determined from the implicit system
\[ (\alpha I + K^*K)u_{M, \delta}^{\epsilon, \delta} = K^*g^\delta(x), \quad \|Ku_{M, \delta}^{\epsilon, \delta} - g^\delta\| = \delta, \]
which can be solved numerically by the classical Newton method [22] or the recently developed model function method [23,24].

Now let
\[ F_{M, \delta}^{\epsilon, \delta}(x, T) := \sum_{n=1}^{M} c_n^{\epsilon, \delta} \Phi_n(x, T) \tag{2.12} \]
and consider the problem
\[ \begin{aligned}
&\frac{\partial^\gamma u_{M, \delta}^{\epsilon, \delta}(x, t)}{\partial t^\gamma} = \nabla \cdot (\alpha(x)\nabla u_{M, \delta}^{\epsilon, \delta}(x, t)), \quad (x, t) \in \Omega \times (0, T], \\
&u_{M, \delta}^{\epsilon, \delta}(x, t) = 0, \quad x \in \partial \Omega, t > 0, \\
&u_{M, \delta}^{\epsilon, \delta}(x, T) = F_{M, \delta}^{\epsilon, \delta}(x, T), \quad x \in \Omega.
\end{aligned} \tag{2.13} \]
The unique solution to this backward problem can be expressed explicitly as
\[ u_{M, \delta}^{\epsilon, \delta}(x, t) = \sum_{n=1}^{M} c_n^{\epsilon, \delta} \Phi_n(x, t), \quad t \in [0, T]. \tag{2.14} \]
The function \( u_{M, \delta}^{\epsilon, \delta}(x, t) \) will be taken as the regularizing solution for \( u(x, t) \) with the regularizing parameter \( M \) specified in terms of the noise level \( \delta \) in the following.

3. Convergence analysis on the regularizing solution

In this section, we will establish the uniform convergence rate of the regularizing solution \( u_{M, \delta}^{\epsilon, \delta}(x, t) \) for all \( t \in [0, T] \). This fact is quite different from the classical regularization where the convergence rate depends on \( t \). Moreover, to get the convergence rate at \( t = 0 \), some more stronger regularity of \( u_0(x) \) should be assumed, see [25].

Firstly, we consider the approximation error.
Theorem 3.1. Assume that $u_0 \in H^2_0$ with $\|u_0\|_{H^2_0} \leq U_p$ for $p = 1$ or $p = 2$. Then for arbitrary $\delta > 0$ and positive integer $M$, it holds

$$
\|u_M^{\delta, \cdot} (\cdot, t) - u(\cdot, t)\|_{L^2(\Omega)}^2 \leq \left( \frac{3 + 2C_2 + C_1C_2}{C_1(1 + \lambda_1t^\gamma)} \right)^2 \left[ (1 + \lambda_MT^\gamma)\delta + \frac{U_p}{\lambda_M^{p/2}} \right], \quad t \in [0, T].
$$

(3.1)

Proof. A direct calculation gives

$$
|u_M^{\delta, \cdot} (x, t) - u(x, t)|^2 = \left( \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n) \Phi_n(x, t) \right)^2 + \left( \sum_{n=M+1}^\infty c_n \Phi_n(x, t) \right)^2
+ 2 \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n) \Phi_n(x, t) \sum_{n=M+1}^\infty c_n \Phi_n(x, t).
$$

By the orthogonality of $\{\Phi_n(\cdot, t)\}_{n=1}^\infty$ and Lemma 2.1, we obtain

$$
\|u_M^{\delta, \cdot} (\cdot, t) - u(\cdot, t)\|_{L^2(\Omega)}^2 = \left( \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n) \Phi_n(x, t) \right)^2 + \left( \sum_{n=M+1}^\infty c_n \Phi_n(x, t) \right)^2
\leq \left( \frac{C_2}{1 + \lambda_1t^\gamma} \right)^2 \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n)^2 + \left( \frac{C_2}{1 + \lambda_MT^\gamma} \right)^2 \sum_{n=M+1}^\infty c_n^2. \tag{3.2}
$$

On the other hand, the following decomposition

$$
\sum_{n=1}^M (c_n^{\delta, \cdot} - c_n) \Phi_n(x, t) = \sum_{n=1}^M c_n^{\delta, \cdot} \Phi_n(x, t) - g^\delta(x) + g(x) + \sum_{n=M+1}^\infty c_n \Phi_n(x, t)
$$

yields that

$$
\frac{1}{3} \left( \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n)^2 \Phi_n(x, t) \right)^2 \leq \left( \sum_{n=1}^M c_n^{\delta, \cdot} \Phi_n(x, t) - g^\delta(x) \right)^2 + (g(x) - g(x))^2 + \left( \sum_{n=M+1}^\infty c_n \Phi_n(x, t) \right)^2.
$$

Using (2.7), we have

$$
\frac{1}{3} \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n)^2 \Phi_n(x, t) \leq 2\delta^2 + \sum_{n=M+1}^\infty c_n^2 \Phi_n(x, t).
$$

By Lemma 2.1, it follows that

$$
\frac{1}{3} \left( \frac{C_1}{1 + \lambda_MT^\gamma} \right)^2 \sum_{n=1}^M (c_n^{\delta, \cdot} - c_n)^2 \leq 2\delta^2 + \left( \frac{C_2}{1 + \lambda_MT^\gamma} \right)^2 \sum_{n=M+1}^\infty c_n^2,
$$

that is,

$$
\sum_{n=1}^M (c_n^{\delta, \cdot} - c_n)^2 \leq \frac{6}{C_1^2} (1 + \lambda_MT^\gamma)^2 \delta^2 + \frac{3C_2^2}{C_1^2} \sum_{n=M+1}^\infty c_n^2. \tag{3.3}
$$

Noticing $\lambda_n^2 = \|\nabla \cdot (a\nabla \varphi_n)\|_{L^2(\Omega)}^2$ and the orthogonality of $\nabla \cdot (a\nabla \varphi_n)$, it follows that

$$
\lambda_n^2 \sum_{n=M+1}^\infty c_n^2 \leq \sum_{n=M+1}^\infty c_n^2 \lambda_n^2 \leq \sum_{n=1}^\infty c_n^2 \|\nabla \cdot (a\nabla \varphi_n)\|_{L^2(\Omega)}^2 = \left( \sum_{n=1}^\infty c_n \nabla \cdot (a\nabla \varphi_n) \right)^2_{L^2(\Omega)}^2,
$$

i.e.,

$$
\lambda_n^2 \sum_{n=M+1}^\infty c_n^2 \leq \left( \nabla \cdot \left( a\nabla \sum_{n=1}^\infty c_n \varphi_n \right) \right)^2_{L^2(\Omega)} \leq \|\nabla \cdot (a\nabla u_0)\|_{L^2(\Omega)}^2 \leq C U_2^2. \tag{3.4}
$$
Similarly, we have
\[
\lambda_M \sum_{n=M+1}^{\infty} c_n^2 \leq \sum_{n=1}^{\infty} c_n^2 \lambda_n = \sum_{n=1}^{\infty} c_n^2 \| \sqrt{a} \nabla \varphi_n \|_{L^2(\Omega)}^2 = \| \sqrt{a} \nabla u_0 \|_{L^2(\Omega)}^2 \leq C U_1^2
\] (3.5)
due to the orthogonality of \( \sqrt{a} \nabla \varphi_n \). Combining (3.2)–(3.5), we obtain
\[
\| u_M^{\delta}(\cdot, t) - u(\cdot, t) \|_{L^2(\Omega)}^2 \leq \frac{C_2^2}{C_1^2 (1 + \lambda_1 T')^2} \left[ 6(1 + \lambda_M T')^2 \delta^2 + \left( 3C_2^2 + C_1^2 \left( \frac{1 + \lambda_1 T'}{1 + \lambda_M T'} \right)^2 \right) \frac{U_p^2}{\lambda_M^2} \right].
\]
which yields
\[
\| u_M^\delta(\cdot, t) - u(\cdot, t) \|_{L^2(\Omega)}^2 \leq \frac{(3 + 2C_2 + C_1)C_2}{C_1(1 + \lambda_1 T')} \left[ (1 + \lambda_M T') \delta + \frac{U_p}{\lambda_M^p} \right], \quad t \in [0, T].
\] (3.6)
The proof of Theorem 3.1 is completed. \( \Box \)

Using the error estimate (3.1), we can establish the convergence rate. We take \( \lambda_M(\delta) \approx \frac{1}{T \gamma h(\delta)} \) in (3.1), where \( h(\delta) \) is a positive function satisfying \( h(\delta) \to 0 \) as \( \delta \to 0 \). Then it follows that
\[
\| u_M^\delta(\cdot, t) - u(\cdot, t) \|_{L^2(\Omega)} \leq \frac{C_0(\gamma, p, T)}{1 + \lambda_1 T'} \left[ \delta + \frac{\delta}{h(\delta)} + h(\delta)^{p/2} \right]
\]
with the constant \( C_0(\gamma, p, T) \). By choosing \( h(\delta) := \delta^{p/2} \), we have the following result.

**Theorem 3.2.** Assume that \( u_0 \in H^p_0 \) with \( \| u_0 \|_{H^p_0} \leq U_p \) for \( p = 1 \) or \( p = 2 \). If the truncation term number \( M = M(\delta) \) is chosen such that the eigenvalue satisfies
\[
\lambda_M(\delta) \approx \frac{1}{T \gamma \delta^{p/2}},
\] (3.7)
then it holds the convergence rate
\[
\| u_M^\delta(\cdot, t) - u(\cdot, t) \|_{L^2(\Omega)} \leq \frac{3C_0(\gamma, p, T)}{1 + \lambda_1 T'} \delta^{p/2}, \quad t \in [0, T].
\] (3.8)

**Remark 3.1.** Since the relation between \( \lambda_M \) and \( M \) is theoretically known noticing that \( a(x) \) is known, this theorem gives a strategy for the choice of \( M \) in terms of \( \delta \) theoretically. When we make numerical implementations, the determination of \( M \) depends on the numerical computation of the eigen-system of the known operator \( -\nabla \cdot (a(x) \nabla \phi) \). Although the eigenfunctions are computationally very expensive to obtain for general domain \( \Omega \) and variable coefficient \( a(x) \), the regularizing solution for all \( t > 0 \) can be expressed explicitly based on the eigen-system. In our numerical implementations, the eigen-system is computed using the standard program in Matlab. Due to the decay property of the diffusion process, the value of \( M \) need not be quite large in numerical implementations.

4. **Numerical examples**

In this section, we make some numerical implementations on our inversion scheme. In the first three examples, we always take \( \gamma = 1/2 \) and \( \Omega = (0, \pi)^2 \subset \mathbb{R}^2 \), and consider the following model
\[
\begin{aligned}
\frac{\partial^\gamma u}{\partial t^\gamma} &= \nabla \cdot (a(x) \nabla u), \quad (x, t) \in \Omega \times (0, T], \\
u(x, t) &= 0, \quad x \in \partial \Omega, \quad t \in [0, T], \\
u(x, 0) &= u_0(x), \quad x \in \Omega
\end{aligned}
\] (4.1)
for different configurations of \( a(x) \) and \( u_0(x) \). In this case, it follows that
\[
E_{1/2,1}(t) = e^{\gamma^2} \text{erfc}(-t) = e^{\gamma^2} \frac{2}{\sqrt{\pi}} \int_{-t}^{\infty} e^{-s^2} ds.
\] (4.2)
Since our inversion scheme is based on the eigenfunction expansion, the rectangular spatial domain does not lose generality, noticing that the eigenfunction defined by
\[
\begin{aligned}
-\nabla \cdot (a(x) \nabla \varphi_n(x)) &= \lambda_n \varphi_n(x), \quad x \in \Omega, \\
\varphi_n(x) &= 0, \quad x \in \partial \Omega
\end{aligned}
\] (4.3)
can always be solved numerically. However, the approximation behavior of the eigen-system will have some influence on our reconstruction scheme, as seen in Example 3 in this section. The last example is devoted to a general domain \( \Omega \) with variable diffusion coefficient \( a(x) \) and the order of fractional derivative \( \gamma \neq 1/2 \).

In the following we will change the notation \( x = (x_1, x_2) \in \mathbb{R}^2 \) as \( (x, y) \in \mathbb{R}^2 \).

**Example 1.** Consider the case \( a(x, y) = 1 \) and \( u_0(x, y) = xy(\pi - x)(\pi - y) \in H^p_0(\Omega) \) with \( p = 2 \). The eigen-system of \( -\Delta \) is

\[
\lambda_{n,m} = n^2 + m^2, \quad \varphi_{n,m}(x, y) = \frac{2}{\pi} \sin nx \sin my, \quad n, m = 1, 2, \ldots .
\]

For \( \gamma = \frac{1}{2} \), the unique solution of (4.1) can be expressed as

\[
u(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} d_{n,m} E_{1/2,1}(- (n^2 + m^2)t^{1/2}) \sin nx \sin my \quad (4.4)
\]

with the coefficients

\[
d_{n,m} = \frac{8(1 - (-1)^n)(1 - (-1)^m)}{\pi n^m m^3}, \quad n, m = 1, 2, \ldots . \quad (4.5)
\]

The high-frequency amplitude is small in the above Fourier expansion, so we can apply the truncation with a finite number of terms

\[
u(x, y, t) \approx \sum_{n=1}^{10} \sum_{m=1}^{10} d_{n,m} E_{1/2,1}(- (n^2 + m^2)t^{1/2}) \sin nx \sin my \quad (4.6)
\]

to generate the final measurement data at \( T = 0.5 \) with noise by

\[
u^\delta(x, y, T) = \nu(x, y, T) + \frac{1}{\pi} \delta \cdot \text{rand}(x, y), \quad (4.7)
\]

where \( \text{rand}(x, y) \in [-1, 1] \) for \( (x_i, y_j) \in [0, \pi]^2 \) with \( i, j = 1, \ldots, 100 \) is the standard random number and \( \delta \) is the error level. Using the noisy data simulated from (4.7), we compute the regularizing solution

\[
u_M^\delta(x, y, t) \approx \sum_{n=1}^{M} \sum_{m=1}^{M} c_{n,m} E_{1/2,1}(- (n^2 + m^2)t^{1/2}) \sin nx \sin my \quad (4.8)
\]

from our regularizing scheme. For \( \delta = 0.01, T = 0.5 \), we can estimate \( M \) from

\[
\lambda_M = 2M^2 \approx \frac{C}{T^{1/2} \gamma^{1/2}} = \frac{C}{T^{1/2} \gamma^{1/2}}
\]

and \( \{c_{n,m}, n, m = 1, 2, \ldots , M\} \) can be computed from (2.11).

We present our reconstructions at \( t = 0 \) from the final measurement time \( T = 0.5 \) in Fig. 1 with fixed noise level \( \delta = 0.01 \). To explain our results, we give the exact initial distribution, noisy measurement data at \( T = 0.5 \) and our reconstruction for \( M = 5 \) in the top line of Fig. 1, respectively, while the reconstruction error distributions for different \( M \) are shown in the bottom line of Fig. 1.

It can be seen that the initial status almost disappears in the noisy measurement data at \( T = 0.5 \), which comes from the quick decay of direction of the diffusion process. Using our reconstruction scheme, the best reconstruction is obtained for \( M = 5 \).

In Fig. 2 (left), we give the truncation-error curves of \( \|\nu_M^\delta(\cdot, 0) - \nu(\cdot, 0)\|_{L^2(\Omega)} \) with different \( \delta \) and \( M \). From the numerical behavior we can find that \( M = 5 \) is the optimal parameter for \( \delta = 0.01 \), while \( M = 3 \) is optimal for noisy level \( \delta = 0.03, 0.05 \). These facts show that \( M \) is indeed the regularizing parameter.

To support our convergence rate \( \delta^{p/2} \) for \( p = 2 \), we use the standard bisection scheme to analyze the convergence order, which needs to show

\[
\lim_{\delta \to 0} \frac{\|\nu_M^{1/2, \delta}(\cdot, 0) - \nu(\cdot, 0)\|_{L^2(\Omega)}}{\|\nu_M^{\delta}(\cdot, 0) - \nu(\cdot, 0)\|_{L^2(\Omega)}} = \sqrt{\frac{1}{2}}.
\]

The limitation behavior of the left-hand side for \( M = 5 \) is shown in Fig. 2 (right). It can be seen that the convergence order is indeed \( \delta^{1/2} \) for small \( \delta \).
Fig. 1. Reconstruction results at $t = 0$ from noisy measurement data at $T = 0.5$, for fixed level $\delta = 0.01$ with different regularizing parameters $M$ for Example 1.

Example 2. We consider a picture with jump discontinuity as our initial data $u(x, y, 0)$ with the constant diffusion $a(x, y) \equiv 1$. The exact model picture is shown in Fig. 3 (top-left). The grey level of this picture is taken as $u_0(x, y)$. Obviously $u_0(x, y) \not\in H^1_0(\Omega)$. The noisy measurement data at $T = 0.5$ are generated in the same way as that in Example 1. Due to the smooth effect of diffusion, the interfaces of the picture have been completely smoothed, see Fig. 3 (top-middle).

Although this initial configuration does not meet the regularity assumption for our theoretical estimate on the convergence rate, numerical implementations show that our scheme still works well. We give the reconstruction from these noisy data using $M = 13$ in Fig. 3 (top-right). Despite the smooth procedure, the interfaces and the homogeneity of the initial distribution are reconstructed very well.

In the bottom line of Fig. 3, we show the error distribution of our reconstruction for different $M$. Also it is found that $M = 13$ is the optimal value for our reconstruction. The other interesting observation is that the maximum error always appears in the interfaces of the picture for different $M$, which comes from the singularity of the initial distribution at these
Fig. 3. Reconstruction results for $t = 0$ from noisy measurement data at $T = 0.5$, for fixed $\delta = 0.01$ with different regularizing parameters $M$ for Example 2.

Fig. 4. Error curves with respect to truncation term $M$ (left) and convergence order behavior (right) for Example 2.

points, noticing that our convergence estimate requires the \textit{a priori} regularity assumptions on $u_0(x, y)$. These observations support our theoretical results.

The error distributions with respect to $M$ and the behavior of $\frac{\|u_{M}^{\delta/2}(\cdot, 0) - u(\cdot, 0)\|_{L^2(\Omega)}}{\|u_{M}^{\delta}(\cdot, 0) - u(\cdot, 0)\|_{L^2(\Omega)}}$ in terms of error level $\delta$ are given in Fig. 4, left and right, respectively. Notice, we do not have the limit value $\frac{\delta}{\sqrt{\xi}}$ as $\delta \to 0$ in the right figure of Fig. 4, since $u_0(x, y) \not\in H^p(\Omega)$. However, the phenomenon that the limit value is 1 as $\delta \to 0$ is natural, which means

$$\left\| u_{M}^{\delta/2}(\cdot, 0) - u(\cdot, 0) \right\|_{L^2(\Omega)} \approx \left\| u_{M}^{\delta}(\cdot, 0) - u(\cdot, 0) \right\|_{L^2(\Omega)}$$

for small $\delta$. Indeed, the truncation error has the main effect for very small $\delta$, and therefore the decay of $\delta$ does not affect the approximate error any more for small input error $\delta$. 
Example 3. In this example, we consider the inhomogeneous media, i.e., the diffusion coefficient depends on the point locations. We set
\[ a(x, y) = 2x + y + 1, \quad (x, y) \in \Omega, \]
while the initial status \( u_0(x, y) \) is the same as that in Example 1.

In this case, the eigenfunctions should be solved numerically. To avoid the error coming from both the approximate eigenfunction and finite terms truncation, we solve the forward diffusion system by the finite difference scheme directly, for details, see [26].

Under this configuration, the nonconstant diffusion coefficient in \( \Omega \) has some influence on the measurement data. It is obvious that \( a(x, y) \) takes its maximum value \( 2\pi + \pi + 1 \) in \( \Omega \), which means a strong diffusion near the corner \((\pi, \pi)\). Consequently, it can be found that the values \( u(x, y, 0.5) \) in the picture of Fig. 5 (top-middle) are almost zero in the bottom-right part due to the strong diffusion.

The reconstructions using noisy data \( u^\delta(x, y, 0.5) \) with \( \delta = 0.01 \) (top-right in Fig. 5) for different \( M \) are given in the bottom line of Fig. 5. Obviously, different from the cases in Examples 1 and 2 with constant diffusion coefficient, the symmetric property of \( u_0(x, y) \) cannot be recovered due to the inhomogeneous diffusion of the media. Also, we find that \( M = 5 \) is the optimal value of the regularizing parameter.

It is reasonable to expect similar behavior between the reconstruction error and regularizing parameter \( M \). We present the error distribution in terms of \( M \) in the left of Fig. 6 and the convergence behavior in the right of Fig. 6. The theoretical convergence rate is also supported numerically for reasonably small \( \delta \) (around \( 10^{-2} \)) in this inhomogeneous diffusion case. However, for quite small \( \delta \), the numerical behavior for convergence order is not so satisfactory. The reasons come from our numerical approximation for the eigenvalue system and the simulated data for final measurement by the finite difference scheme. The errors arising in the numerical eigen-system cause new errors for our reconstruction and therefore change the behavior of convergence order.

Finally we give an example with general domain \( \Omega \) and also variable diffusion coefficient \( a(x, y) \). In this case, the eigen-system should be solved numerically.

Example 4. Consider the problem (4.1) with \( a(x, y) = x^2 + y^2 + 1, \gamma = \frac{4}{5} \) and
\[ \partial \Omega := \{(x, y) = (x(\theta), y(\theta)) := r(\theta)(\cos \theta, \sin \theta), \theta \in [0, 2\pi]\} \]
for radius \( r(\theta) = \left( (1 - 2 \sin \frac{\theta}{2})^2 + \sin^2 \theta \right)^{1/2} \), see Fig. 7 for the configuration of \( \Omega \). We choose the initial value \( u(x, y, 0) = u_0(x, y) \) for \((x, y) \in \tilde{\Omega} \) such that
\[
\quad u_0(x, y) = 0, \quad (x, y) \in \partial \Omega. \tag{4.9}
\]

To avoid the “inverse crime”, namely, the simulated inversion input data for the inverse problem should be generated from a completely different scheme from the inversion scheme applied. Here we solve the direct problem (4.1) by the difference scheme directly to get \( u(x, y, t) \). To do this, we extend the domain \( \tilde{\Omega} \) with general shape to a rectangular domain
\[
\tilde{\Omega} = \{(x, y) \in \mathbb{R}^2 : 0 \leq |x|, |y| \leq 1.2\}
\]
satisfying \( \tilde{\Omega} \subseteq \tilde{\Omega} \). Then we define
\[
\quad u(x, y, t) \equiv 0, \quad (x, y) \in \tilde{\Omega} \setminus \Omega, \ t > 0, \quad \tag{4.10}
\]
noticing the boundary condition \( u(x, y, t)|_{\partial \Omega} = 0 \) and the compatible condition (4.9) for initial value \( u_0(x, y) \). Then we divide \( \tilde{\Omega} \) as \( P1 \times P2 \) grids by \( x(i) := -1.2 + (i - 1)h_1, y(j) := -1.2 + (j - 1)h_2 \) with \( h_1 = 24/P1, h_2 = 24/P2 \) for \( i = 1, 2, \ldots, P1 + 1, j = 1, 2, \ldots, P2 + 1 \). In our implementations, we fix \( P1 = P2 = 30 \) and take the geometric center \( Q_{i,j}(x^i, y^j), i, j = 1, \ldots, 30 \) of each sub-rectangle as the points where we compute the function values. We use the scheme
\[
\quad \partial_t (a(x, y)\partial_x u)|_{Q_{i,j}} \approx \delta_t (a(x, y)\partial_x u)|_{Q_{i,j}} \nonumber \\
\quad := \frac{1}{h_1^2} (a_{i+1/2,j}u_{i+1,j} - (a_{i+1/2,j} + a_{i-1/2,j})u_{i,j} + a_{i-1/2,j}u_{i-1,j})
\]
to compute the spatial derivative along the \( x \) direction for each \( t > 0 \). To guarantee (4.10), we set \( a_{i\pm 1/2,j} = 0 \) if \((x^i, y^j) \in \tilde{\Omega} \setminus \Omega \). The spatial derivative along the \( y \) direction is treated analogously. Then the direct problem in \( \tilde{\Omega} \) is solved using the scheme for a rectangular domain in [26] from which we get \( u(x, y, t) \) for \((x, y) \in \tilde{\Omega} \). We will consider the following two cases:

Case 1: \( T = 1, \ u_0(x, y) = \sum_{j=1}^4 \psi_j(x, y), \ (x, y) \in \Omega, \) where \( \psi_j(x, y) \) is the \( j \)-th eigenfunction defined in Section 2. For this general domain, \( \psi_j(x, y) \) is solved numerically using Matlab.
Fig. 8. Reconstruction results at $t = 0$ from noisy measurement data for fixed $\delta = 0.05$ for case 1 with different regularizing parameters $M$ for Example 4.

Case 2: $T = 0.5$, $u_0(x, y) = (\rho - r(t))^2$, $(x, y) \in \Omega$, where $(\rho, \theta)$ is the polar coordinate of $\Omega$, i.e.,

$$\Omega = \{(\rho, \theta) : 0 \leq \rho \leq r(\theta) : \theta \in [0, 2\pi]\}$$

with $r(\theta)$ the polar radius of $\partial \Omega$.

For these two cases, (4.9) is satisfied. We make a partition for time domain $[0, T]$ as 1000 subintervals. Noticing that for $\gamma \neq 1/2$ in this example, $E_{\gamma,1}(-\lambda_{n,m}\gamma^2)$ cannot be computed from any integral expression, so we use the standard program provided by Prof. Podlubny to compute this function (http://www.mathworks.com/matlabcentral/fileexchange/8738).

The reconstructions in case 1 are given in Fig. 8. The exact initial value and the measurement data are shown in the top line of Fig. 8. Since we take $T = 1$ (not so small), the information about the initial distribution cannot be seen directly in this measurement. The reconstructions for different $M$ are presented in the bottom line of Fig. 8. Since the initial value consists of only the first four frequency components, it is reasonable that the reconstructions take optimal performance for $M = 3$. The error distribution with respect to the number of truncation terms for different noise level $\delta$ is presented in Fig. 9.

Next we consider case 2 where $u_0(x, y)$ consists of all frequency components $\lambda_j : j = 1, 2, \ldots$ in a general domain $\Omega$, and therefore any expansion using a finite number of eigenfunctions in our reconstruction scheme causes truncation error. We present the noisy measurement data with $\delta = 0.003$ and the reconstructions in Fig. 10. It can be seen that, in this very general configuration, the reconstruction with $M = 5$ is still acceptable. The error distribution in terms of the number of
Fig. 10. Reconstruction results at $t = 0$ from noisy measurement data for fixed $\delta = 0.003$ for case 2 with different regularizing parameters $M$ for Example 4.

Fig. 11. Error curves in case 2 with respect to truncation term $M$ for different noise level $\delta$ for Example 4.

truncation terms $M$ for different noise levels is shown in Fig. 11. It is worthwhile to point out that the reconstruction error comes from both the noisy measurement data and the finite terms truncations of the inversion scheme.

Acknowledgments

This work is supported by the NSFC (Nos. 11071039, 61179035, 61127002), the Jiangsu Natural Science Foundation (BK2011584) and Southeast University (No. 3207011102). The authors would like to thank two anonymous referees for their valuable comments and suggestions, which made this paper much improved.

References