A unified framework for the regularization of final value time-fractional diffusion equation.

Walter C. SIMO TAO LEE*

Abstract

This paper focuses on the regularization of backward time-fractional diffusion problem on unbounded domain. This problem is well-known to be ill-posed, whence the need of a regularization method in order to recover stable approximate solution. For the problem under consideration, we present a unified framework of regularization which covers some techniques such as Fourier regularization [19], mollification [12] and approximate-inverse [7]. We investigate a regularization technique with two major advantages: the simplicity of computation of the regularized solution and the avoid of truncation of high frequency components (so as to avoid undesirable oscillation on the resulting approximate-solution). Under classical Sobolev-smoothness conditions, we derive order-optimal error estimates between the approximate solution and the exact solution in the case where both the data and the model are only approximately known. In addition, an order-optimal a-posteriori parameter choice rule based on the Morozov principle is given. Finally, via some numerical experiments in two-dimensional space, we illustrate the efficiency of our regularization approach and we numerically confirm the theoretical convergence rates established in the paper.

Keywords: Backward time-fractional diffusion, sub-diffusion, mollification, regularization, error estimates, parameter choice rule.

1 Introduction

In this paper, we study the final value problem
\[
\begin{aligned}
\frac{\partial^\gamma u}{\partial t^\gamma} &= \Delta u & x \in \mathbb{R}^n, \ t \in (0,T) \\
u(x,T) &= g(x) & x \in \mathbb{R}^n,
\end{aligned}
\]
where we aim at recovering the initial distribution \(u(\cdot,0)\) given the final distribution \(u(\cdot,T)\). In (1), \(\gamma \in (0,1)\) and \(\frac{\partial^\gamma}{\partial t^\gamma}\) denotes the Caputo fractional derivative [5] defined as
\[
\frac{\partial^\gamma u}{\partial t^\gamma} = \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{u'(s)}{(t-s)^\gamma} \, ds,
\]
where \(\Gamma(\cdot)\) is nothing but the Gamma function: \(\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} \, dt\).

Time-fractional diffusion equations usually model sub-diffusion processes such as slow and anomalous diffusion processes which failed to be described by classical diffusion models [2, 3, 9, 10]. Due to the high diversity of such phenomena which are not properly modeled by classical diffusion,
time-fractional diffusion problems have gained much attention in last decades. Beyond applications in diffusion processes, time-fractional equation (1) has also been applied to image de-blurring [15] where the time-fractional derivative allows to capture the memory effect in image blurring.

It is well-known that the ill-posedness of equation (1) comes from the irreversibility of time of the diffusion equation, which is caused by the very smoothing property of the forward diffusion. As a result, very small perturbation of the final distribution \( u(\cdot, T) \) may cause arbitrary large error in the initial distribution \( u(\cdot, 0) \). Hence, a regularization method is crucial in order to recover stable approximate of the initial distribution \( u(\cdot, 0) \). In this regard, many regularization methods have been applied to final value time-fractional diffusion equation. Let us mention the mollification method [12, 17], Fourier regularization [16, 19], the method of quasi-reversibility [6], Tikhonov method [13], total variation regularization [15], boundary condition regularization [18], non-local boundary value method [4], truncation method [14]. Yet, the set of regularization methods applied to backward time-fractional diffusion equation still presents some sparsity, especially compared to the set regularization methods for backward classical diffusion problems.

In this paper, we describe how in the context of regularization of the final value time-fractional diffusion equation (1), the Fourier regularization [19] and mollification [12] are nothing but examples of approximate-inverse [7] regularization. Next, we investigate a regularization technique which yields a better trade-off between stability and accuracy compared to the Fourier regularization [19] and the mollification technique of Van Duc N. et al [12]. We consider noisy setting where \( u(\cdot, T) \) is approximated by a noisy data \( g^\delta \) satisfying

\[
||u(\cdot, T) - g^\delta||_{L^2(\mathbb{R}^n)} \leq \delta,
\]

and we derive order-optimal convergence rates between our approximate solution and the exact solution \( u(\cdot, 0) \) under classical Sobolev smoothness condition

\[
u(\cdot, 0) \in H^p(\mathbb{R}^n) \quad \text{with} \quad ||u(\cdot, 0)||_{H^0(\mathbb{R}^n)} \leq E, \quad p > 0.
\]

We also provide error estimates under the more realistic setting where both the data \( g \) and the forward diffusion operator are only approximately known. The motivation here being that, in practice, the Mittag-Leffler function which plays a major role in the resolution of equation (1) can only be approximated in practice.

The outline of this article is as follows:

In Section 2, we discuss existence of solution of equation (1) and reformulate the equation into an operator equation of the form \( Au(\cdot, 0) = g \) in which \( A \) is a bounded linear operator on \( L^2(\mathbb{R}^n) \). We present key estimates necessary for the regularization analysis and illustrate the ill-posedness of recovering \( u(\cdot, 0) \) from \( g \). Next, we introduce the framework of regularization which includes Fourier regularization, mollification, and approximate inverse. At last, we introduce our regularization approach.

Section 3 deals with error estimates and order-optimality of our regularization technique under the smoothness condition (2). In this section, we derive error estimates between the approximate solution and the exact solution \( u(\cdot, 0) \) in Sobolev spaces \( H^l(\mathbb{R}^n) \) with \( l \geq 0 \). We also give error estimates for the approximation of early distribution \( u(\cdot, t) \) with \( t \in (0, T) \). We end the section by presenting analogous error estimates for the case where both the data \( g \) and the forward diffusion operator \( A \) are only approximately known.

Section 4 is devoted to parameter selection rules which is a critical step in the application of a regularization method. Here we propose a Morozov-like a-posteriori parameter choice rule leading
Moreover, we can also derive the following relation for early distribution \( u \)

\[
\text{Proposition 1. Let us start by the following result about the existence and uniqueness of solution of equation (1).}
\]

2 Regularization

From (4), we can deduce the following relation between the solution \( u \)

\[
\text{Fourier transform of the function} \ f
\]

theoretical convergence rates given in Section 3. Moreover, in this Section, we also carry out a numerical convergence rates analysis in order to confirm the theoretical convergence rates given in Section 3.

In the sequel, \( ||f|| \) or \( ||f||_{L^2} \) always refers to the \( L^2 \)-norm of the function \( f \) on \( \mathbb{R}^n \), \( ||f||_{H^p} \) denotes the Sobolev norm of \( f \) on \( \mathbb{R}^n \) and \( || \cdot || || \) denotes operator norm of a bounded linear mapping. Throughout the paper, \( \hat{f} \) or \( \mathcal{F}(f) \) (resp. \( \mathcal{F}^{-1}(f) \)) denotes the Fourier (resp. inverse Fourier) transform of the function \( f \) defined as

\[
\hat{f}(\xi) = \mathcal{F}(f)(\xi) = \frac{1}{\sqrt{2\pi n}} \int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi} \, dx, \quad \mathcal{F}^{-1}(f)(x) = \frac{1}{\sqrt{2\pi n}} \int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi} \, d\xi, \quad \xi, x \in \mathbb{R}^n.
\]

2 Regularization

Let us start by the following result about the existence and uniqueness of solution of equation (1).

Proposition 1. For all \( \gamma \in (0, 1) \) and \( g \in H^2(\mathbb{R}^n) \), Problem (1) admits a unique weak solution \( u \in C([0, T], L^2(\mathbb{R}^n)) \cap C((0, T], H^2(\mathbb{R}^n)) \). That is, the first equation in (1) holds in \( L^2(\mathbb{R}^n) \) for all \( t \in (0, T) \) and \( u(\cdot, t) \in H^2(\mathbb{R}^n) \) for all \( t \in (0, T) \) with

\[
\lim_{t \to T} ||u(\cdot, t) - g||_{H^2} = 0.
\]

Proposition 1 is merely generalization of [19, Lemma 2.2] where only the case \( n = 1 \) is considered. The idea of the proof is merely to check that the formal solution defined by (5) is the weak solution.

Now, let us define the framework that we will consider for the regularization of problem (1). Consider a data \( g \in H^2(\mathbb{R}^n) \), by applying the Fourier transform in (1) with respect to variable \( x \), we get

\[
\begin{align*}
\frac{\partial^\gamma \hat{u}(\xi, t)}{\partial t^{\gamma}} &= -|\xi|^2 \hat{u}(\xi, t) \quad \xi \in \mathbb{R}^n, \ t \in (0, T) \\
\hat{u}(\xi, T) &= \hat{g}(\xi) \quad \xi \in \mathbb{R}^n.
\end{align*}
\]

By applying the Laplace transform with respect to variable \( t \) in (3), one gets

\[
\begin{align*}
\hat{u}(\xi, t) &= \hat{u}(\xi, 0)E_{\gamma, 1}(-|\xi|^2 t^\gamma) \quad \xi \in \mathbb{R}^n \\
\hat{u}(\xi, T) &= \hat{g}(\xi) \quad \xi \in \mathbb{R}^n,
\end{align*}
\]

where \( E_{\gamma, 1} \) is the Mittag Leffler function [10] defined as

\[
E_{\gamma, 1}(z) = \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(\gamma k + 1)}, \quad z \in \mathbb{C}.
\]

From (4), we can deduce the following relation between the solution \( u(\cdot, 0) \) and the data \( g \) in the frequency domain:

\[
\hat{u}(\xi, 0) = \frac{\hat{g}(\xi)}{E_{\gamma, 1}(-|\xi|^2 T^\gamma)}.
\]

Moreover, we can also derive the following relation for early distribution \( u(\cdot, t) \) with \( t \in (0, T) \)

\[
\forall t \in (0, T), \quad \hat{u}(\xi, t) = \frac{E_{\gamma, 1}(-|\xi|^2 t^\gamma)}{E_{\gamma, 1}(-|\xi|^2 T^\gamma)} \hat{g}(\xi).
\]

end
From equations (5) and (6), we can see that the Mittag Leffler function $E_{\gamma,1}$ plays an important role in time-fractional diffusion equation (1). Hence, let us recall some key estimates about the function $E_{\gamma,1}$ that will be repeatedly used in the sequel.

**Lemma 1.** Let $\gamma \in [\gamma_0, \gamma_1] \subset (0,1)$, there exists constants $C_1$ and $C_2$ depending only on $\gamma_0$ and $\gamma_1$ such that

$$\forall x \leq 0, \quad \frac{C_1}{\Gamma(1 - \gamma)} \frac{1}{1 - x} \leq E_{\gamma,1}(x) \leq \frac{C_2}{\Gamma(1 - \gamma)} \frac{1}{1 - x}. \quad (7)$$

For a proof of Lemma 1, see [19, Lemma 2.1]. From Lemma 1, we can easily derive the next Lemma.

**Lemma 2.** Assume $\gamma \in [\gamma_0, \gamma_1] \subset (0,1)$, then for every $\xi \in \mathbb{R}^n$ and $t \in (0, T]$,

$$\frac{1}{(1 \lor t^\gamma)} \left( \frac{C_1}{\Gamma(1 - \gamma)} \frac{1}{1 + |\xi|^2} \right) \leq E_{\gamma,1}(-|\xi|^2 t^\gamma) \leq \frac{1}{(1 \land t^\gamma)} \left( \frac{C_2}{\Gamma(1 - \gamma)} \frac{1}{1 + |\xi|^2} \right). \quad (8)$$

and

$$\frac{C_1}{C_2} \leq \frac{E_{\gamma,1}(-|\xi|^2 t^\gamma)}{E_{\gamma,1}(-|\xi|^2 T^\gamma)} \leq \frac{C_2}{C_1} \left( \frac{T}{t} \right)^\gamma. \quad (9)$$

In (8), $\lor$ denotes the maximum while $\land$ denotes the minimum, that is, $1 \lor t^\gamma = \max \{1, t^\gamma\}$ and $1 \land t^\gamma = \min \{1, t^\gamma\}$.

From (6) and (9), we get that for every $t \in (0, T)$, $||u(\cdot, t)||_L^2 \leq (C_2/C_1) (T/t)^\gamma ||g||_L^2$, which implies that the problem of recovering $u(\cdot, t)$ from $g$ is actually well posed. However, that of recovering $u(\cdot, 0)$ is ill-posed. Indeed, from (5) and (8), we get that

$$||u(\xi, 0)|| \geq C_\gamma (1 + |\xi|^2) \hat{g}(\xi), \quad \text{with} \quad C_\gamma = \Gamma(1 - \gamma)(1 \land T^\gamma)/C_2. \quad (10)$$

From (10), we see that very small perturbations in high frequencies in the data $g$ leads to arbitrary large errors in the solution $u(\cdot, 0)$. Therefore, one needs a regularization method to recover stable estimates of $u(\cdot, 0)$.

**Remark 1.** From (5) and (8), we can also derive that

$$||u(\xi, t)||_L^2 \leq D_\gamma (1 + |\xi|^2) \hat{g}(\xi), \quad \text{with} \quad D_\gamma = \Gamma(1 - \gamma)(1 \lor T^\gamma)/C_1. \quad (11)$$

Estimate (11) illustrates the fact that the backward time-fractional diffusion equation is less ill-posed (mildly ill-posed) on the contrary to the classical backward diffusion equation which is exponentially ill-posed. This is actually due to the asymptotic slow decay of the Mittag Leffler function $E_{\gamma,1}(-|\xi|^2)$ compared to $\exp(-|\xi|^2) = E_{1,1}(-|\xi|^2)$.

From (5), we can reformulate equation (1) into an operator equation

$$Au(\cdot, 0) = g. \quad (12)$$

where $A : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ is the linear forward diffusion operator which maps the initial distribution $u(\cdot, 0)$ to the final distribution $g$, that is,

$$A = \mathcal{F}^{-1} \left( E_{\gamma,1}(-|\xi|^2 T^\gamma) \right) \mathcal{F}. \quad (13)$$

In the sequel, given a data $g \in L^2(\mathbb{R}^n)$, we aim at recovering $u(\cdot, 0) \in L^2(\mathbb{R}^n)$. Let $\varphi$ be a smooth real-valued function in $L^1(\mathbb{R}^n)$ satisfying $\int_{\mathbb{R}^n} \varphi(x) \, dx = 1$. It is well-known that the family of functions $(\varphi_\alpha)_{\alpha > 0}$ defined by

$$\forall x \in \mathbb{R}^n, \quad \varphi_\alpha(x) := \frac{1}{\alpha^n} \varphi \left( \frac{x}{\alpha} \right),$$
satisfies
\[
\forall f \in L^2(\mathbb{R}^n), \quad \varphi_\alpha * f \to f \quad \text{in} \quad L^2(\mathbb{R}^n) \quad \text{as} \quad \alpha \downarrow 0, \tag{14}
\]
where \( \varphi_\alpha * f \) is nothing but the convolution of the functions \( \varphi_\alpha \) and \( f \) defined as \( (\varphi_\alpha * f)(x) = \int_{\mathbb{R}^n} \varphi_\alpha(x-y)f(y) \, dy \). For \( \alpha > 0 \), let \( M_\alpha \) be the mollifier operator defined by
\[
\forall \alpha > 0, \quad \forall f \in L^2(\mathbb{R}^n), \quad M_\alpha f = \varphi_\alpha * f. \tag{15}
\]
From (14), we see that the family of operators \( (M_\alpha)_{\alpha > 0} \) is an approximation of unity in \( L^2(\mathbb{R}^n) \), that is,
\[
\forall f \in L^2(\mathbb{R}^n), \quad M_\alpha f \to f \quad \text{in} \quad L^2(\mathbb{R}^n) \quad \text{as} \quad \alpha \downarrow 0. \tag{16}
\]
Let \( u_\alpha \) be the solution of the equation
\[
\begin{cases}
\frac{\partial u_\alpha}{\partial t} = \Delta u & x \in \mathbb{R}^n, \quad t \in (0, T) \\
u(x, T) = (M_\alpha g)(x) & x \in \mathbb{R}^n. \tag{17}
\end{cases}
\]
From (5) and (6), by replacing \( g \) by \( M_\alpha g \), and using the fact that the \( \widehat{M_\alpha g}(\xi) = \sqrt{2\pi^n} \varphi_\alpha(\xi) \widehat{g}(\xi) \), one gets
\[
\begin{cases}
\widehat{u_\alpha}(\xi, 0) = \frac{\sqrt{2\pi^n}}{E_{\gamma,1}(|\xi|^2T^n)} \widehat{\varphi_\alpha}(\xi) \widehat{g}(\xi) = \sqrt{2\pi^n} \varphi_\alpha(\xi) \widehat{u}(\xi, 0) \\
\widehat{u_\alpha}(\xi, t) = \frac{\sqrt{2\pi^n}}{E_{\gamma,1}(|\xi|^2T^n)} \widehat{\varphi_\alpha}(\xi) \widehat{g}(\xi) = \sqrt{2\pi^n} \varphi_\alpha(\xi) \widehat{u}(\xi, t), \quad t \in (0, T). \tag{18}
\end{cases}
\]
which yields that
\[
\forall t \in [0, T] \quad u_\alpha(\cdot, t) = M_\alpha u(\cdot, t). \tag{19}
\]

**Proposition 2.** Assume that there exists a function \( g : \mathbb{R}^*_+ \to \mathbb{R}_+ \) such that the mollifier kernel \( \varphi \) verifies
\[
\forall \alpha > 0, \quad \frac{\widehat{\varphi}(\alpha \xi)}{E_{\gamma,1}(-|\xi|^2T^n)} \leq g(\alpha), \tag{20}
\]
Then the family \( (u_\alpha)_{\alpha > 0} \) solution of equation (17) defines a regularization method for problem (1) in \( L^2(\mathbb{R}^n) \).

**Proof.** From (18), noticing that \( \widehat{\varphi_\alpha}(\xi) = \widehat{\varphi}(\alpha \xi) \) and using the fact that the function \( E_{\gamma,1} \) is increasing on \( \mathbb{R}_- \) with \( E_{\gamma,1}(0) = 1 \), we can see that (20) implies that for every \( t \in [0, T] \) and \( \alpha > 0 \), the mapping \( R_{\alpha, t} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n) \) which maps the data \( g \) to \( u_\alpha(\cdot, t) \) is bounded with \( ||R_{\alpha, t}|| \leq \sqrt{2\pi^n} g(\alpha) \). Moreover, from (16) and (19), we deduce that for every \( t \in [0, T] \), \( u_\alpha(\cdot, t) \) converges to \( u(\cdot, t) \) in \( L^2(\mathbb{R}^n) \) as \( \alpha \) goes to 0. \( \square \)

From Proposition 2, we can see that, choosing a kernel \( \varphi \) which satisfies condition (20) allows to defines a regularization method for equation (1). Now, let us show that the family of regularization methods defined in this way actually coincides with approximate-inverse introduced by Louis and Maass [7].

From the first equation in (18), we can derive that
\[
u_\alpha(x, 0) = \int_{\mathbb{R}^n} e^{ix\xi} \frac{\widehat{\varphi_\alpha}(\xi) \widehat{g}(\xi)}{E_{\gamma,1}(-|\xi|^2T^n)} \, d\xi. \tag{21}
\]
From (19) and (21), we can reformulate \( u_\alpha(x, 0) \) as follows
\[
\begin{cases}
\forall x \in \mathbb{R}^n, \quad u_\alpha(x, 0) = \langle c_\alpha(x, \cdot), u(\cdot, 0) \rangle_{L^2}, \quad \text{with} \quad c_\alpha(x, y) = \varphi_\alpha(x-y) \\
\forall x \in \mathbb{R}^n, \quad u_\alpha(x, 0) = \langle v_{x, \alpha}, g \rangle_{L^2}, \quad \text{with} \quad v_{x, \alpha} = F^{-1} \left( \frac{e^{ix\xi} \widehat{\varphi_\alpha}(\xi)}{E_{\gamma,1}(-|\xi|^2T^n)} \right). \tag{22}
\end{cases}
\]
Moreover, one can easily check that $v_{x,\alpha}$ is nothing but the solution of the adjoint equation

$$A^*f = e_\alpha(x,\cdot).$$

Hence, we deduce that the solution $u_\alpha$ of equation (17) actually corresponds to the approximate-inverse [7] regularized solution of equation (12), the pair $(v_{x,\alpha}, e_\alpha)$ being what is usually called reconstruction kernel - mollifier.

This setting of regularization we just described encompasses many regularization methods that appears distinctively in the literature of the regularization of the final value time-fractional diffusion equation. Each regularization method being a particular choice of the mollifier kernel $\varphi$.

For the Fourier regularization [19] (where $n = 1$), we have

$$\hat{u}_{\xi_{max}}(\xi, t) = \chi_{[-\xi_{max}, \xi_{max}]}(\xi) \frac{E_{\gamma,1}(-|\xi|^2T^\gamma)}{E_{\gamma,1}(-|\xi|^2T^\gamma)} \hat{g}(\xi),$$

(23)

where $\chi_{\Omega}$ denotes the characteristic function of the set $\Omega$ equal to 1 on $\Omega$ and 0 elsewhere. By comparing (23) and (18), we readily get that

$$\alpha = 1/\xi_{max}, \quad \text{and} \quad \varphi(x) = \mathcal{F}^{-1}\left(\frac{1}{\sqrt{2\pi}} \chi_{[-1,1]}\right) = \frac{\sin(x)}{\pi x}. \quad (24)$$

In this case, condition (20) merely reads

$$\frac{\chi_{[-1,1]}(\xi/\xi_{max})}{E_{\gamma,1}(-|\xi|^2T^\gamma)} \leq \varrho\left(\frac{1}{\xi_{max}}\right).$$

Using (8), we can see that this condition is satisfies with $\varrho(\alpha) = C(1 + 1/\alpha^2)$ where $C = (1 \lor T^\gamma)\Gamma(1 - \gamma)/C_1$.

For the mollification method of N. Van Duc et al. [12], the mollifier operator $M_\alpha$ is denoted $S_\nu$ where $S_\nu$ is the convolution by the so called Dirichlet kernel $D_\nu$ defined as

$$D_\nu(x) = \frac{1}{\pi^n} \prod_{j=1}^n \frac{\sin(\nu x_j)}{x_j}, \quad \text{with} \quad \nu > 0 \text{ and } x \in \mathbb{R}^n.$$  

Hence we can deduce that, this merely corresponds to

$$\alpha = 1/\nu, \quad \text{and} \quad \varphi(x) = \prod_{j=1}^n \frac{\sin(x_j)}{\pi x_j}. \quad (25)$$

Given that the Fourier transform of the kernel $D_\nu$ is given by

$$\mathcal{F}(D_\nu)(\xi) = \chi_{\Lambda}(\xi), \quad \text{with} \quad \Lambda = \{x \in \mathbb{R}^n : |x_j| \leq \nu, \ j = 1, ..., n\}, \quad (26)$$

we see that condition (20) merely reads

$$\frac{\chi_{\Lambda}(\xi/\nu)}{E_{\gamma,1}(-|\xi|^2T^\gamma)} \leq \varrho\left(\frac{1}{\nu}\right),$$

which is fulfilled with $\varrho(\alpha) = C(1 + \sqrt{n}/\alpha^2)$ where $C = (1 \lor T^\gamma)\Gamma(1 - \gamma)/C_1$.

By the way, from (24) and (25), we can see that the Fourier regularization and the mollification approach of N. Van Duc et al. actually coincides, the latter approach being a generalization of the
former to $n-$dimensional case. From (26), we can conclude that both regularization approaches are nothing but truncation methods. That is, the regularization is done by merely throwing away high frequency components of the data, which are responsible of the ill-posedness, and conserving unchanged the remaining frequency components. In order word, these two methods can be regarded as spectral cut-off methods. It is important to notice that though high frequency components are responsible of ill-posedness, nevertheless, the still carry non-negligible information on the sought solution. Therefore, it is desirable to apply a regularization which do not suppress high frequency components but which applies much regularization to those components compared to low frequency components. Let us point out that mere truncation of high frequency components usually entails Gibbs phenomena and oscillation of the approximate solution which should be avoided as far as possible. This is actually possible by choosing a kernel $\varphi$ whose Fourier transform is supported on the whole domain $\mathbb{R}^n$.

Now on, let us consider a mollifier kernel $\varphi$ defined by

$$\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \exp(-\tau|\xi|^s), \quad \tau > 0, \quad s > 0,$$

i.e.

$$\varphi = \frac{1}{\sqrt{2\pi}^n} F^{-1}(\exp(-\tau|\xi|^s)), \quad (27)$$

where $\tau$ and $s$ are two free positive parameters. From (27), we can see that $\varphi \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and satisfies

$$\int_{\mathbb{R}^n} \varphi(x) \, dx = \sqrt{2\pi}^n \hat{\varphi}(0) = 1.$$

**Lemma 3.** Let $b$ and $d$ be two positive numbers, consider the function $f_{b,d}(x) = (1 + x)e^{-bx^d}$. Then there exists a constant $C$ depending only on $d$ such that

$$\sup_{x \geq 0} f_{b,d}(x) \leq \frac{C}{b^{1/d}} \quad \text{as} \quad b \downarrow 0. \quad (28)$$

The proof of Lemma 3 is deferred to appendix. Lemma 3 will help us to prove that the kernel $\varphi$ given by (27) allows us to define a regularization method for equation (1).

**Proposition 3.** Let $M_\alpha$ be the mollifier operator defined by (15) with the kernel $\varphi$ given in (27) with $\tau$ and $s$ being two positive numbers. Then the family $(u_\alpha)_{\alpha > 0}$ of function $u_\alpha$ solution of equation (17) defines a regularization method for equation (1).

**Proof.** In view of Proposition 2, it suffices to prove that the kernel $\varphi$ given in (27) verifies (20). By considering (27) and estimate (8), we have

$$\frac{\hat{\varphi}(\alpha \xi)}{E_{\gamma,1}(\xi^2 T^\gamma)} \leq C(1 + |\xi|^2) \exp(-\tau \alpha^s |\xi|^s), \quad \text{with} \quad C = \Gamma(1 - \gamma)(1 \vee T^\gamma)/\sqrt{2\pi}^n C_1. \quad (29)$$

The right hand side in (29) is nothing but $C f_{b,d}(|\xi|^2)$ with $b = \tau \alpha^s$ and $d = s/2$. Hence from (28), we deduce that there exists a constant $C$ independent on $\alpha$ such that

$$\forall \xi \in \mathbb{R}^n, \quad \frac{\hat{\varphi}(\alpha \xi)}{E_{\gamma,1}(\xi^2 T^\gamma)} \leq \frac{C}{\alpha^2} \quad \text{as} \quad \alpha \to 0,$$

whence (20) with $\varphi(\alpha) = C/\alpha^2$. \qed

**Remark 2.** By defining the mollifier kernel $\varphi$ as in (27), we can see that the regularization technique induces a more suitable treatment of frequency components. Indeed, with our choice of mollifier kernel, the amount of regularization smoothly depends on the magnitude of the frequency components.
The higher the frequency, the stronger the regularization applied, and similarly, the lower the frequency, the lower the regularization applied. This is actually desirable for a regularization method given that as the frequency gets higher, the noise in the frequency components gets much more amplified, and as the frequency gets lower, the noise in the frequency components gets less and less amplified.

Remark 3. From Proposition 2, we can see that the family of function $(\varphi_s)_{s>0}$ defined by (27) allows to define a family of regularization methods, each regularization method being determined by the choice of the free parameter $s > 0$. For instance, the choice $s = 1$ means considering a Cauchy convolution kernel while $s = 2$ means taking a Gaussian convolution kernel.

Let us end this section by the following Lemma which gives rates of convergence of the mollifier operator $M_\alpha$ corresponding to kernel $\varphi$ defined by (27) on Sobolev spaces $H^p(\mathbb{R}^n)$ with $p > 0$.

**Lemma 4.** Let $p > 0$ and $M_\alpha$ be the mollifier operator defined by (15) with the mollifier kernel $\varphi$ given by (27). Then

$$\forall f \in H^p(\mathbb{R}^n), \quad \|f - M_\alpha f\|_{L^2} \leq \tau \frac{p/s}{\alpha} \|f\|_{H^p}. \quad (30)$$

**Proof.** Let $f \in H^p(\mathbb{R}^n)$. If $p < s$, using Parseval identity, we have

$$\|f - M_\alpha f\|_{L^2} = \left\|1 - \sqrt{2\pi} \hat{\varphi}(\alpha \xi) \hat{f}(\xi)\right\|_{L^2}$$

$$= \left\|[1 - \exp(-\tau(\alpha|\xi|^s))^{p/s} \hat{f}(\xi) \times [1 - \exp(-\tau(\alpha|\xi|^s))]^{1-p/s}\right\|_{L^2}$$

$$\leq \left\|[1 - \exp(-\tau(\alpha|\xi|^s))^{p/s} \hat{f}(\xi)\right\|_{L^2}$$

$$\leq (\tau \alpha^s)^{p/s} \left\|\hat{f}(\xi)\right\|_{L^2} \leq \tau \alpha^s \|f\|_{H^p}.$$ 

If $p \geq s$, $\|f - M_\alpha f\|_{L^2} = \left\|[1 - e^{-\tau(\alpha|\xi|^s)}] \hat{f}(\xi)\right\|_{L^2} \leq \tau \alpha^s \left\|\xi|\hat{f}(\xi)\right\|_{L^2} \leq \tau \alpha^s \|f\|_{H^s} \leq \tau \alpha^s \|f\|_{H^p}. \quad \square$

### 3 Error estimates

Henceforth, $\varphi$ denotes the mollifier kernel defined by (27) and $g^\delta \in L^2(\mathbb{R}^n)$ denotes a noisy data satisfying the noise level condition

$$\|g - g^\delta\|_{L^2} \leq \delta, \quad (31)$$

where $g = u(\cdot, T)$ is the exact final distribution. Let us introduce the regularized solution $u_\alpha^\delta$ corresponding to the noisy data $g^\delta$ as the solution of equation

$$\begin{cases}
\frac{\partial^n u}{\partial t^n} = \Delta u, & x \in \mathbb{R}^n, \quad t \in (0, T) \\
u(x, T) = (M_\alpha g^\delta)(x) & x \in \mathbb{R}^n.
\end{cases} \quad (32)$$

Equivalently, we can define $u_\alpha^\delta$ in the frequency domain by

$$\hat{u}_\alpha^\delta(\xi, t) = \sqrt{2\pi} \left\langle \frac{E_{\gamma,1}(-|\xi|^2 t)}{E_{\gamma,1}(-|\xi|^2 T)} \hat{\varphi}(\alpha \xi) \hat{g}(\xi), \quad t \in [0, T]. \quad (33)$$

It is well known that without assuming a smoothness condition on the exact solution $u(\cdot, 0)$ (or on the exact data $g$), it is impossible to exhibit a rate of convergence of regularized solution towards the exact solution [11]. Henceforth, we consider the following classical Sobolev smoothness condition:

$$u(\cdot, 0) \in H^p(\mathbb{R}^n), \quad \|u(\cdot, 0)\|_{H^p} \leq E, \quad \text{with} \quad p > 0, \quad E > 0. \quad (34)$$
Before presenting the main results of this section, let us state some lemmas which will be useful in the sequel.

**Lemma 5.** Let \( p \geq 0 \), and \( v \) be a solution of equation \( \frac{\partial^p v}{\partial t^p} = \Delta v \) on \( \mathbb{R}^n \). If \( v(\cdot, 0) \in H^p(\mathbb{R}^n) \), then for every \( t \in (0, T] \), \( v(\cdot, t) \in H^{p+2}(\mathbb{R}^n) \) and

\[
\|v(\cdot, t)\|_{H^{p+2}} \leq \frac{C}{1 \wedge t^\gamma} \|v(\cdot, 0)\|_{H^p}, \quad \text{with} \quad C = \frac{C_2}{\Gamma(1 - \gamma)}. \tag{35}
\]

**Proof.** The proof follows readily by applying Parseval identity and estimate (8) to equation \( \hat{v}(\xi, t) = \hat{v}(\xi, 0)E_{\gamma, 1}(-|\xi|^2 t^\gamma) \) which links \( v(\cdot, 0) \) and \( v(\cdot, t) \) in the frequency domain. \( \square \)

The next lemma illustrates the fact that the Sobolev smoothness condition (34) is nothing but a Hölder source condition.

**Lemma 6.** Let \( p > 0 \), \( u \) be the solution of problem (1) and \( A \) being the forward diffusion operator defined in (13). The smoothness condition \( u(\cdot, 0) \in H^p(\mathbb{R}^n) \) is equivalent to the Hölder source condition \( u(\cdot, 0) = (A^*A)^{p/4}w \) with \( w \in L^2(\mathbb{R}^n) \), satisfying

\[
\left( \frac{\Gamma(1 - \gamma)(1 \wedge T^\gamma)}{C_2} \right)^{p/2} \|u(\cdot, 0)\|_{H^p} \leq \|w\|_{L^2} \leq \left( \frac{\Gamma(1 - \gamma)(1 \wedge T^\gamma)}{C_1} \right)^{p/2} \|u(\cdot, 0)\|_{H^p} \tag{36}
\]

**Proof.** For \( u(\cdot, 0) \in H^p(\mathbb{R}^n) \), formally define \( w \) in the frequency domain by

\[
\hat{w} (\xi) = E_{\gamma, 1}(-|\xi|^2 T^\gamma)^{-p/2} \hat{u} (\xi, 0), \quad \iff \quad \hat{u} (\xi, 0) = (E_{\gamma, 1}(-|\xi|^2 T^\gamma)^{2})^{p/4} \hat{w} (\xi)
\]

From (13), we can verify that the above definition of \( w \) from \( u(\cdot, 0) \) is merely reformulation of the equation \( u(\cdot, 0) = (A^*A)^{p/4}w \) in the frequency domain. Next, we can check that \( w \) is well defined and belongs to \( L^2(\mathbb{R}^n) \). Finally, estimate (36) is deduced from (8). \( \square \)

**Remark 4.** From Lemma 6, we can deduce that the order optimal convergence rate under smoothness condition (34) is nothing but \( C E \frac{\gamma}{p + 2} \delta \frac{\gamma}{p + 2} \) with \( C \geq 1 \) independent of \( E \) and \( \delta \).

The next Lemma which generalizes [12, Lemma 3] will be useful in the sequel for establishing Sobolev norm error estimates.

**Lemma 7.** Let \( p \geq 0 \) and \( v \) be a solution of equation \( \frac{\partial^p v}{\partial t^p} = \Delta v \) on \( \mathbb{R}^n \). If \( v(\cdot, 0) \in H^p(\mathbb{R}^n) \) then

\[
\forall t \in [0, p], \quad \forall t \in (0, T], \quad \|v(\cdot, t)\|_{H^{p+2}} \leq \frac{C(\gamma)}{(1 \wedge t^\gamma)} \|v(\cdot, 0)\|_{H^p} \|v(\cdot, t)\|_{H^p}^{\frac{2+\gamma}{p+2}}, \tag{37}
\]

where \( C(\gamma) = (1 \wedge T^\gamma)C_2/C_1 \). Moreover,

\[
\forall t \in [0, p], \quad \forall t \in [0, T], \quad \|v(\cdot, t)\|_{H^{l}} \leq \bar{C}(\gamma) \|v(\cdot, 0)\|_{H^p} \|v(\cdot, T)\|_{H^p}^{\frac{p-l}{p+2}}, \tag{38}
\]

where \( \bar{C}(\gamma) = \Gamma(1 - \gamma)(1 \wedge T^\gamma)/C_1 \).
Theorem 1. Assume that the solution $u$ of problem (1) satisfies the smoothness condition (34). Consider a noisy approximation $g^\delta$ satisfying (31) and let $u^\delta_\alpha$ be the regularized solution defined by (33). Then for the a-priori selection rule $\alpha(\delta) = (\delta/E)^{1/p+2}$, we have

$$
\forall l \in [0, p], \quad \text{such that } \ p - l \leq s, \quad \|u(\cdot, 0) - u^\delta_\alpha(\cdot, 0)\|_{H^l} \leq CE^{2+1/p+2} \delta^{p-l/2},
$$

where $C$ is a constant independent of $\delta$ and $E$.

Proof. Using the Parseval identity, we have

$$
\|u(\cdot, 0) - u_\alpha(\cdot, 0)\|_{H^l} = \left\|1 - \sqrt{2\pi}^n \tilde{\varphi}(\alpha \xi) \left(1 + |\xi|^2\right)^{l/2} \hat{\varphi}(\xi, 0)\right\|_{L^2} = \|\tilde{u} - M_\alpha \tilde{u}\|_{L^2},
$$

where $\tilde{u} = \mathcal{F}^{-1} \left((1 + |\xi|^2)^{l/2} \hat{\varphi}(\xi, 0)\right)$. Since $u(\cdot, 0) \in H^p(\mathbb{R}^n)$, $\tilde{u} \in H^{p-l}(\mathbb{R}^n)$, then applying (30) to (41), we deduce that if $p - l \leq s$, then

$$
\|u(\cdot, 0) - u_\alpha(\cdot, 0)\|_{H^l} \leq \tau^{p-l} S^{p-l} \|\tilde{u}\|_{H^{p-l}} = \tau^{p-l} \alpha^{p-l} \|u(\cdot, 0)\|_{H^p} \leq \tau^{p-l} \alpha^{p-l} E.
$$
On the other hand, using (8) and (31), we have

$$\left\| u_\alpha (\cdot,0) - u^\delta_\alpha (\cdot,0) \right\|_{H^l} = \left\| (1 + |\xi|^2)^{1/2} \sqrt{2\pi^n} \hat{\varphi}(\alpha \xi) \frac{\hat{g}(\xi) - \hat{g}^\delta(\xi)}{E_{\gamma,1}(-|\xi|^2 t^\gamma)} \right\|_{L^2} \leq \delta (1 + T^\gamma) \frac{\Gamma(1 - \gamma)}{C_1} \left\| (1 + |\xi|^2)^{1/2} \exp(-\tau(\alpha |\xi|)^s) \right\|_\infty.$$  \hspace{1cm} (43)

But \((1 + |\xi|^2)^{1/2} \exp(-\tau(\alpha |\xi|)^s) = (f_{b,d}(|\xi|^2))^{1/2} \) with \( d = s/2 \) and \( b = \frac{\tau(\alpha |\xi|)^s}{1 + T^\gamma} \rightarrow 0 \) as \( \alpha \rightarrow 0 \).

Hence applying Lemma 3, we get that there exists a constant \( C \) independent of \( \alpha \), such that

$$\forall \xi \in \mathbb{R}^n, \quad (1 + |\xi|^2)^{1/2} \exp(-\tau(\alpha |\xi|)^s) \leq (C/\alpha^2)^{1/2}. \hspace{1cm} (44)$$

Applying (44) to (43), we deduce the existence of a constant \( \tilde{C} \) independent of \( \alpha, \delta \) and \( E \) such that

$$\left\| u_\alpha (\cdot,0) - u^\delta_\alpha (\cdot,0) \right\|_{H^l} \leq \frac{\tilde{C}}{\alpha^{s+l}}. \hspace{1cm} (45)$$

Finally from (42) and (45), we deduce that

$$\left\| u(\cdot,0) - u^\delta_\alpha (\cdot,0) \right\|_{H^l} \leq \frac{\tau p_{s+l}}{\alpha^s} \alpha^{p-l} E + \frac{\delta}{\alpha^{s+l}}. \hspace{1cm} (46)$$

By choosing \( \alpha(\delta) = (\delta/E)^{\frac{1}{p + l}} \) in (46), we get (40).

\begin{remark}
By considering \( l = 0 \) in Theorem 1, we get that if \( p \leq s \), then

$$\left\| u(\cdot,0) - u^\delta_\alpha (\cdot,0) \right\|_{L^2} \leq C E \frac{\delta}{\alpha^{p+2}}.$$

Hence, our regularization method is order-optimal under the classical smoothness condition (34). Notice that the condition \( p \leq s \) is not at all restrictive since the parameter \( s \) is freely chosen in \( \mathbb{R}^*_+ \).

The next theorem shows rate of convergence of the error \( u(\cdot,t) - u^\delta_\alpha (\cdot,t) \) when approximating earlier distribution \( u(\cdot,t), t \in (0,T) \) using the regularized solution \( u^\delta_\alpha \).

\begin{theorem}
Consider the setting of Theorem 1. By considering \( \alpha(\delta) = (\delta/E)^{\frac{1}{p + l}} \), we have

$$\forall t \in (0,T], \forall l \in [0,p+2], \quad s.t. \quad p + 2 - l \leq s, \quad \left\| u(\cdot,t) - u^\delta_\alpha(\cdot,t) \right\|_{H^l} \leq \frac{C}{t^{\gamma}} E \frac{\delta}{\alpha^{p+2}} \frac{1}{\alpha^{s+l}}, \hspace{1cm} (47)$$

where \( C \) is a constant independent of \( \delta \) and \( E \). Moreover,

$$\forall t \in (0,T], \forall l \in [0,p], \quad s.t. \quad p - l \leq s, \quad \left\| u(\cdot,t) - u^\delta_\alpha(\cdot,t) \right\|_{H^l} \leq \tilde{C} E \frac{\delta}{\alpha^{p+2}} \frac{1}{\alpha^{s+l}}, \hspace{1cm} (48)$$

where \( \tilde{C} \) is a constant independent of \( \delta, E \) and \( t \).

\begin{proof}
By noticing that \( |u(\xi,t) - u^\delta_\alpha(\xi,t)| = E_{\gamma,1}(-|\xi|^2 t^\gamma) |u(\xi, 0) - u^\delta_\alpha(\xi, 0)| \leq |u(\xi, 0) - u^\delta_\alpha(\xi, 0)| \), we can deduce (48) from (40). Let us prove (47). We have

$$\left\| u(\cdot,t) - u_\alpha(\cdot, t) \right\|_{H^l} = \left\| \left( 1 - \sqrt{2\pi^n} \hat{\varphi}(\alpha \xi) \right) \left( 1 + |\xi|^2)^{1/2} \hat{u}(\xi, t) \right\|_{L^2} = \left\| \tilde{u} - M_\alpha \tilde{u} \right\|_{L^2},$$

where \( \tilde{u} = F^{-1} \left( (1 + |\xi|^2)^{1/2} \hat{u}(\xi, t) \right) \). Since \( u(\cdot,0) \in H^p(\mathbb{R}^n) \), from Lemma 5, we get \( u(\cdot,t) \in H^{p+2}(\mathbb{R}^n) \) which implies that \( \tilde{u} \in H^{p+2-l}(\mathbb{R}^n) \), then applying (30) and (35), we get that if \( p + 2 - l \leq s \), then

$$\left\| u(\cdot,t) - u_\alpha(\cdot, t) \right\|_{H^l} \leq \frac{\tau p_{s+2}^{s+l}}{\alpha^{p+2-l}} \alpha^{p-2-l} \left\| \tilde{u} \right\|_{H^{p+2-l}} = \frac{\tau p_{s+2}^{s+l}}{\alpha^{p+2-l}} \frac{\tau p_{s+2}^{s+l}}{\alpha^{p-2-l}} \left\| u(\cdot,t) \right\|_{H^{p+2}} \leq \frac{C_2 \tau p_{s+2}^{s+l}}{(1 \wedge t^\gamma) \Gamma(1 - \gamma)} \alpha^{p+2-l} E. \hspace{1cm} (49)$$

\end{proof}

\end{theorem}
On the other hand, using (9), we have

$$
\left\| u_\alpha(\cdot, t) - u_\alpha^\delta(\cdot, t) \right\|_{H^l} = \left\| (1 + |\xi|^2)^{l/2} \sqrt{2\pi^n} \widehat{\varphi}(\alpha \xi) \frac{E_{\gamma, 1}(-|\xi|^2 T^\gamma)}{E_{\gamma, 1}(-|\xi|^2 T^\gamma)} \left| g(\xi) - g^\delta(\xi) \right| \right\|_{L^2} \\
\leq \frac{\delta C_2}{C_1} \left( \frac{T}{t} \right)^\gamma \left( 1 + |\xi|^2 \right)^{l/2} \exp(-\tau(\alpha |\xi|)^s) \right\|_{L^\infty}. \quad (50)
$$

When \( l = 0 \), we can easily see from (50) that

$$
\left\| u_\alpha(\cdot, t) - u_\alpha^\delta(\cdot, t) \right\|_{H^l} \leq \frac{\delta C_2}{C_1} \left( \frac{T}{t} \right)^\gamma. \quad (51)
$$

When \( l > 0 \), using similar reasoning yielding to (44), we have

$$
\forall \xi \in \mathbb{R}^n, \quad (1 + |\xi|^2)^{l/2} \exp(-\tau(\alpha |\xi|)^s) \leq (C/\alpha^2)^{l/2} \quad (52)
$$

From (50), (51) and (52), we deduce that for every \( l \geq 0 \) the exists a constant \( C \) independent of \( \alpha, \delta \) and \( t \) such that

$$
\left\| u_\alpha(\cdot, t) - u_\alpha^\delta(\cdot, t) \right\|_{H^l} \leq \frac{C}{\pi^\gamma} \frac{\delta}{t^l} \quad (53)
$$

Lastly from (49) and (53), we get that if \( p + 2 - l \leq s \), then

$$
\left\| u(\cdot, t) - u_\alpha^\delta(\cdot, t) \right\|_{L^2} \leq \frac{\bar{C}}{1 \wedge T^\gamma} \alpha^{p+2-l} E + \frac{C}{t^l} \frac{\delta}{\alpha^l}, \quad (54)
$$

where \( C \) and \( \bar{C} \) are constants independent of \( \alpha, \delta, E \) and \( t \). For \( \alpha(\delta) = (\delta/E)^{p+2} \), we get (47). \( \square \)

**Remark 6.** In Theorem 2, in the estimate (48), the rate is lower than the rate in (47), however, notice that the factor \( 1/t^l \) in (47) blows up as \( t \) decreases to 0. Notice that the rate in (48) cannot be improved without multiplying by a factor which blows up as \( t \) goes to 0.

**Remark 7.** By choosing \( l = 0 \) in Theorem 2, we get the rate \( \| u(\cdot, t) - u_\alpha^\delta(\cdot, t) \|_{L^2} \leq \frac{C}{\pi^\gamma} \delta \), which means we can also recover earlier distribution \( u(\cdot, t) \) with the best possible rate \( \delta \).

Now let us study error estimates when both the data \( g \) and the operator \( A \) are only approximately known. Indeed, though the operator \( A \) is explicitly known as (13), in practical implementation, this operator is only approximated given that the Mittag Leffler function can only be approximated though with desired accuracy [10].

In the sequel we assume that \( \psi_h \) is a positive function defined on \( \mathbb{R}_+ \times (0, T] \) and satisfying

$$
\forall t \in (0, T], \quad \left\| \psi_h(\xi, t) - E_{\gamma, 1}(-|\xi|^2 T^\gamma) \right\|_{L^\infty} \leq h. \quad (55)
$$

With the function \( \psi_h \), we can approximate operator \( A \) by the operator \( A_h \) defined by

$$
A_h = \mathcal{F}^{-1} \psi_h(\xi, T) \mathcal{F} \quad \text{i.e.} \quad \widehat{A_h} f(\xi) = \psi_h(|\xi|, T) \hat{f}(\xi). \quad (56)
$$

From (55), given that for every \( t > 0 \), and \( \xi \in \mathbb{R}^n, \quad E_{\gamma, 1}(-|\xi|^2 T^\gamma) \in (0, 1] \), we can deduce that \( ||A - A_h|| \leq h \). Let \( u_\alpha^h \) be the regularized solution defined in the frequency domain by

$$
\begin{align*}
\hat{u}_\alpha^h(\xi, 0) &= \sqrt{2\pi^n} \hat{\varphi}(\alpha \xi) \frac{\hat{g}(\xi)}{\psi_h(|\xi|, T)} \\
\hat{u}_\alpha^h(\xi, t) &= \sqrt{2\pi^n} \hat{\varphi}(\alpha \xi) \frac{\psi_h(|\xi|, T)}{\psi_h(|\xi|, T)} \hat{g}(\xi) \quad \text{for} \quad t \in (0, T].
\end{align*} \quad (57)
$$

The next theorem provides error estimates in the approximation of the initial distribution \( u(\cdot, 0) \) under the practical setting where both the data \( g = u(\cdot, T) \) and the forward diffusion operator are only approximately known.
Theorem 3. Consider the setting of Theorem 1. Assume that \( h \leq 1/2 \), let \( \psi_h \) be a function satisfying (55) and \( u_{\alpha h} \) be the approximate solution defined in (57). Then for the a-priori selection rule \( \alpha(\delta, h) = (h + \delta/E)^{1/2} \), the following convergence rate holds:

\[
\| u(\cdot, 0) - u_{\alpha h}(\cdot, 0) \|_{H^1} \leq C \left( h + \delta/E \right)^{1/2} \tag{58}
\]

where \( C \) is a constant independent of \( \delta, h \) and \( E \).

PROOF. For simplicity of notation, we set \( \psi(|\xi|, t) = E_{\gamma, 1}(-|\xi|^2 t^\gamma) \) for all \( t \in [0, T] \). For every \( t \in (0, T] \) and \( \xi \in \mathbb{R}^n \) we have

\[
\frac{E_{\gamma, 1}(-|\xi|^2 t^\gamma)}{\psi_h(|\xi|, t)} \leq \frac{\psi(|\xi|, t)}{\psi(|\xi|, t) - \psi_h(|\xi|, t)} = \frac{1}{1 - \frac{\psi(|\xi|, t) - \psi_h(|\xi|, t)}{\psi(|\xi|, t)}} \leq \frac{1}{1 - h} \leq 1 + 2h \leq 2. \tag{59}
\]

The first inequality in (59) is due to the fact that \( \psi(|\xi|, t) \leq |\psi(|\xi|, t) - \psi_h(|\xi|, t)| + \psi_h(|\xi|, t) \), the second inequality comes from (55) and the last two inequalities are due to the fact that \( h \leq 1/2 \). Let \( l \in [0, p] \) such that \( p - l \leq s \), we recall that from (42), we have

\[
\| u(\cdot, 0) - M_{\alpha} u(\cdot, 0) \|_{H^1} \leq \tau^\frac{p - l}{4} \alpha^{p - l} E. \tag{60}
\]

By noticing that

\[
\left| \frac{\psi(|\xi|, T)}{\psi_h(|\xi|, T)} g^\delta(\xi) \right| \leq \left| \left[ 1 - \frac{\psi(|\xi|, T)}{\psi_h(|\xi|, T)} \right] g^\delta(\xi) \right| \leq \left[ \frac{\psi(|\xi|, T) - \psi_h(|\xi|, T)}{\psi(|\xi|, T)} \right] g^\delta(\xi) \leq 2h |g^\delta(\xi)| + 2 \left| \frac{\psi(|\xi|, T) - \psi_h(|\xi|, T)}{\psi(|\xi|, T)} \right| g^\delta(\xi) \tag{61}
\]

we deduce that

\[
\| M_{\alpha} u(\cdot, 0) - u_{\alpha h}(\cdot, 0) \|_{H^1} = \left\| \sqrt{2\pi^2} \varphi(\alpha \xi)(1 + |\xi|^2)^{1/2} \left[ \frac{\hat{g}(\xi)}{\psi(|\xi|, T)} - \frac{\hat{g}^\delta(\xi)}{\psi(|\xi|, T)} \right] \right\|_{L^2} \leq (2h \| g \|_{L^2} + 2\delta) \left\| \frac{\exp(-\tau(\alpha \xi)^s)(1 + |\xi|^2)^{1/2}}{E_{\gamma, 1}(-|\xi|^2 T^\gamma)} \right\|_{\infty} \tag{62}
\]

using (61) and (31)

\[
\leq (2hE + 2\delta) C \left( 1 + |\xi|^2 \right)^{1/2} \| \exp(-\tau(\alpha|\xi|)^s) \|_{\infty} \tag{63}
\]

\[
\leq C \frac{hE + \delta}{\alpha^{2+l}} \tag{64}
\]

we get

\[
\| u(\cdot, 0) - u_{\alpha h} \|_{H^1} \leq \tau^\frac{p - l}{4} \alpha^{p - l} E + C E \frac{h + \delta/E}{\alpha^{2+l}} \tag{65}
\]

from which (58) follows by choosing \( \alpha(\delta, h) = (h + \delta/E)^{1/2} \). □
Remark 8. From Theorem 3, by choosing \( l = 0 \) in (58), we get the rate
\[
\left\| u(\cdot, 0) - u^{\delta, h}_\alpha \right\|_{L^2} \leq C E \frac{\sqrt{\varepsilon}}{t^{\gamma/2}} (\delta + h E)^{\frac{1}{2} - \frac{\gamma}{2}}.
\]
Hence in the practical setting where both the data and the operator are approximated, we are able to derive order-optimal convergence rates.

The next theorem exhibits error estimates when approximating \( u(\cdot, t) \) with \( t \in (0, T] \).

**Theorem 4.** Consider the setting of Theorem 3. Then for the a-priori selection rule \( \alpha(\delta, h) = (h + \delta/E)^{\frac{1}{2}} \), the following convergence rate holds:
\[
\forall t \in (0, T], \forall l \in [0, p+2], \ s. \ t. \ p+2-l \leq s, \quad \left\| u(\cdot, t) - u^{\delta, h}_\alpha(\cdot, t) \right\|_{H^l} \leq \frac{C}{t^l} E \frac{\sqrt{\varepsilon}}{t^{\gamma/2}} (\delta + h E)^{1 - \frac{\gamma}{2}}, \quad (63)
\]
where \( C \) is a constant independent of \( \delta, h, E \) and \( t \).

**Proof.** Let \( u \) satisfies (34) and \( t \in (0, T] \). For \( l \in [0, p+2] \) such that \( p+2-l \leq s \), recall that from (49), we have
\[
\| u(\cdot, t) - M_\alpha u(\cdot, t) \|_{H^l} \leq \frac{C_2 \tau^{p+2-l}}{(1 \land t^\gamma) \Gamma(1 - \gamma)} \alpha^{p+2-l} E. \quad (64)
\]
On the other hand, using (55), we have
\[
\left| \begin{array}{c}
g(\xi) - \psi_h(\xi, t) \frac{\psi(\xi, t)}{\psi(\xi, T)} \hat{g}(\xi) \\
\phi_h(\xi, t) \frac{\phi(\xi, t)}{\phi(\xi, T)} \hat{\phi}(\xi)
\end{array} \right| \leq \left| \begin{array}{c}
1 - \frac{\psi_h(\xi, t)}{\psi(\xi, t)} \frac{\psi(\xi, T)}{\phi_h(\xi, T)} \hat{g}(\xi) \\
\frac{\phi_h(\xi, t)}{\phi(\xi, t)} \frac{\phi(\xi, T)}{\phi_h(\xi, T)} \hat{\phi}(\xi)
\end{array} \right| \left| \begin{array}{c}
\frac{\psi(\xi, t)}{\psi(\xi, T)} \hat{g}(\xi) \\
\frac{\phi(\xi, t)}{\phi(\xi, T)} \hat{\phi}(\xi)
\end{array} \right|
\leq \left| \begin{array}{c}
\frac{\psi(\xi, t)}{\psi(\xi, T)} \hat{g}(\xi) \\
\frac{\phi(\xi, t)}{\phi(\xi, T)} \hat{\phi}(\xi)
\end{array} \right| \left( 1 + h \right) \left( \frac{\psi_h(\xi, T)}{\psi_h(\xi, T)} \right) \hat{g}(\xi) - \frac{\psi(\xi, T)}{\psi(\xi, T)} \hat{\phi}(\xi)
\leq h |\hat{g}(\xi)| + (3/2) \left( 2 h |\hat{g}(\xi)| + 2 |\hat{g}(\xi) - \hat{\phi}(\xi)| \right) \text{ using (61)}
\leq 4 h |\hat{g}(\xi)| + 3 |\hat{g}(\xi) - \hat{\phi}(\xi)|. \quad (65)
\]
From (65), we deduce that
\[
\left\| M_\alpha u(\cdot, t) - u^{\delta, h}_\alpha(\cdot, t) \right\|_{H^l} = \left\| \sqrt{2 \pi} \hat{\phi}(\alpha \xi)(1 + |\xi|^2)/2 \left[ \frac{\psi(\xi, t)}{\psi(\xi, T)} \hat{g}(\xi) - \frac{\psi_h(\xi, t)}{\psi_h(\xi, T)} \hat{\phi}(\xi) \right] \right\|_{L^2}
\leq \left\| \sqrt{2 \pi} \hat{\phi}(\alpha \xi)(1 + |\xi|^2)/2 \left[ \frac{\psi(\xi, t)}{\psi(\xi, T)} \hat{g}(\xi) - \frac{\psi_h(\xi, t)}{\psi_h(\xi, T)} \hat{\phi}(\xi) \right] \right\|_{L^2}
\leq \frac{C_2 \tau^{p+2-l}}{(1 \land t^\gamma) \Gamma(1 - \gamma)} \alpha^{p+2-l} E \quad \text{using (52)},
\]
where \( C \) is a constant independent of \( \alpha, \delta, E \) and \( t \). Finally, from (64) and (66), we get
\[
\left\| u(\cdot, t) - u^{\delta, h}_\alpha(\cdot, t) \right\|_{H^l} \leq \frac{C_2 \tau^{p+2-l}}{(1 \land t^\gamma) \Gamma(1 - \gamma)} \alpha^{p+2-l} E + \frac{C}{t^l} E \frac{\sqrt{\varepsilon}}{t^{\gamma/2}} (\delta + h E)^{1 - \frac{\gamma}{2}}
\]
from which (63) follows by choosing \( \alpha(\delta) = (h + \delta/E)^{\frac{1}{2}} \).

**Remark 9.** By choosing \( l = 0 \) in Theorem 4, we recover the best possible rate
\[
\left\| u(\cdot, t) - u^{\delta, h}_\alpha(\cdot, t) \right\|_{L^2} \leq \frac{C}{t^\gamma}(\delta + h E).
\]
Now let us focus on the choice of the regularization parameter \( \alpha \) when we don’t have precise a-priori information about the smoothness of the sought solution \( u(\cdot, 0) \).
4 A-posteriori parameter choice rule

The choice of the regularization parameter is a crucial step for any regularization method. As a matter of fact, no matter the regularization method considered, a bad choice of the regularization parameter results in poor approximate solution.

Let us consider the following a-posteriori parameter choice rule based on the Morozov principle [8].

\[
\alpha(\delta, g^\delta) = \sup \left\{ \alpha > 0, \text{ s.t. } ||g^\delta - M_\alpha g^\delta||_{L^2} < \theta \delta \right\},
\]

(67)

where \( \theta > 1 \) is a free real parameter.

**Proposition 4.** Assume that the noise level \( \delta \) and the noisy data \( g^\delta \) satisfies

\[
0 < \theta \delta < ||g^\delta||,
\]

(68)

then the parameter \( \alpha(\delta, g^\delta) \) expressed in (67) is well defined and satisfies

\[
||u^\delta_{\alpha(\delta, g^\delta)}(\cdot, T) - g^\delta|| = \theta \delta \quad \text{i.e.} \quad \left||1 - e^{-\tau(\alpha|\xi|)^*}\delta\hat{g}(\xi)\right||_{L^2} = \theta \delta.
\]

(69)

**Proof.** Let the function \( v : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) be defined by \( v(\alpha) := ||g^\delta - M_\alpha g^\delta||_{L^2}^2 \). By Parseval identity, we get that \( v(\alpha) = \left||1 - e^{-\tau(\alpha|\xi|)^*}\delta\hat{g}(\xi)\right||_{L^2}^2 \). Using the dominated convergence theorem, we can readily check that \( \lim_{\alpha \rightarrow 0} v(\alpha) = 0 \) and \( \lim_{\alpha \rightarrow \infty} v(\alpha) = \left||\delta\hat{g}(\xi)\right||_{L^2}^2 \). Moreover, we can also check using derivation under integral sign that \( v \) is differentiable with \( v'(\alpha) > 0 \) for all \( \alpha > 0 \) which implies that \( v \) is strictly increasing. Hence if (68) is satisfied, then \( \theta \delta \) is in the range of the one-to-one function \( v \), whence the existence and uniqueness of \( \alpha(\delta, g^\delta) \) defined in (67) which is characterized by (69).

\[\square\]

**Remark 10.** The condition (68) is quite reasonable as we do not expect to recover reasonable approximate solution if the data is dominated by noise.

**Theorem 5.** Consider the setting of Theorem 1. Assume that (68) is satisfied and let \( \alpha(\delta, g^\delta) \) satisfying (69). If \( p + 2 \leq s \), then the following holds:

\[
\forall l \in [0, p], \quad \forall t \in [0, T], \quad \left\| u(\cdot, t) - u^\delta_{\alpha(\delta, g^\delta)}(\cdot, t) \right\|_{H^l} \leq C E_{p+2}^{2l+2} \delta^{p-l} \] 

(70)

where \( C \) is a constant independent of \( \delta, E \) and \( t \).

Proof. For simplicity of notation, let \( \alpha := \alpha(\delta, g^\delta) \) defined by (69) and let \( g_\alpha = M_\alpha g \) where \( g = u(\cdot, T) \) and \( g^\delta_\alpha = M_\alpha g^\delta \). We have

\[
\|g - g_\alpha\| \leq \|g - g^\delta\| + \|g^\delta - g^\delta_\alpha\| + \|g^\delta_\alpha - g_\alpha\| \leq \delta + \theta \delta + \|e^{-\tau(\alpha|\xi|)^*}\delta\hat{g}(\xi)\| \leq (\theta + 2)\delta.
\]

(71)

Let \( p \geq 0 \) such that \( p + 2 \leq s \), using (30) and (35), we have

\[
\theta \delta = \left\|1 - e^{-\tau(\alpha|\xi|)^*}\delta\hat{g}(\xi)\right\| \leq \left\|1 - e^{-\tau(\alpha|\xi|)^*}\delta\hat{g}(\xi)\right\| + \left\|1 - e^{-\tau(\alpha|\xi|)^*}\delta\hat{g}(\xi)\right\| \leq \tau^{(p+2)/s} \alpha^{p+2} ||g||_{H^{p+2}} + \delta \leq C(p, \gamma) \alpha^{p+2} E + \delta
\]

(72)
From (72), we deduce that

\[
(\theta - 1)\delta \leq C(p, \gamma)\alpha^{p+2} E \quad \Rightarrow \quad \frac{1}{\alpha} \leq \left( \frac{C(p, \gamma)}{\theta - 1} \right)^{\frac{1}{p+2}} E^{\frac{1}{p+2}} \delta^{-\frac{1}{p+2}}
\]  

(73)

By noticing that \(\|u(\cdot, 0) - M_\alpha u(\cdot, 0)\|_{H^p} \leq \|u(\cdot, 0)\|_{H^p} \leq E\), and applying (38) of Lemma 7 together with (71), we get

\[
\forall l \in [0, p], \quad \forall t \in [0, T], \quad \|u(\cdot, t) - u_\alpha(\cdot, t)\|_{H^l} \leq C l \alpha^{l+2} \delta^{\frac{l}{p+2}},
\]

(74)

where \(u_\alpha(\cdot, t) = M_\alpha u(\cdot, t)\) is the solution of equation (17) and \(C(\alpha, p, \theta) = \tilde{C}(\gamma)(\theta - 1)^{\frac{p+1}{p+2}}\). On the other hand, from (45) and (73), we have that for every \(l \in [0, p]\) and \(t \in [0, T]\),

\[
\left\| u_\alpha(\cdot, t) - u_\delta(\cdot, t) \right\|_{H^l} \leq \left\| u_\alpha(\cdot, 0) - u_\delta(\cdot, 0) \right\|_{H^l} \leq \tilde{C} \delta l \alpha^{l+2} \leq C E^{\frac{1}{p+2}} \delta^{\frac{l}{p+2}},
\]

(75)

where \(C\) is a constant independent of \(\delta, E\) and \(t\). Estimate (70) follows from (74) and (75). □

**Remark 11.** In Theorem 5, by choosing \(l = 0\) and \(t = 0\) in (70), we obtain the rate

\[
\left\| u(\cdot, t) - u_\alpha(\cdot, t) \right\|_{L^2} \leq C E^{\frac{1}{p+2}} \delta^{\frac{l}{p+2}},
\]

which means that the parameter choice rule (67) leads to order-optimal rate.

Lastly, the next theorem exhibits rates of convergence when approximating early distribution \(u(\cdot, t)\) with \(t \in (0, T)\) with the parameter choice rule (67).

**Theorem 6.** Consider the setting of Theorem 5. The following estimate holds

\[
\forall t \in (0, T), \quad \forall l \in [0, p + 1], \quad \|u(\cdot, t) - u_\alpha(\cdot, t)\|_{H^l} \leq \frac{C}{l^{1/4}} E^{\frac{1/2}{p+2}} \delta^{\frac{l}{p+2}},
\]

(76)

where \(C\) is a constant independent of \(\delta, E\) and \(t\).

**Proof.** Let \(t \in (0, T), \quad l \in [0, p]\) and \(\alpha = \alpha(\delta, g^\delta)\) defined by (69). By applying (37), we get

\[
\left\| u(\cdot, t) - u_\alpha(\cdot, t) \right\|_{H^{l+2}} \leq \frac{C(\gamma)}{1 + t^2} \left\| u(\cdot, 0) - u_\alpha(\cdot, 0) \right\|_{H^p}^{\frac{l+2}{p+2}} \left\| u(\cdot, T) - u_\alpha(\cdot, T) \right\|_{L^2}^{\frac{l+1}{p+2}}
\]

\[
\leq \frac{C(\gamma)}{1 + t^2} E^{\frac{l+2}{p+2}} \left\| g - g_\alpha \right\|_{L^2}^{\frac{l+1}{p+2}}
\]

\[
\leq \frac{C(\gamma)}{1 + t^2} \left( \theta + 2 \right)^{\frac{l+2}{p+2}} E^{\frac{l+2}{p+2}} \delta^{\frac{l+1}{p+2}} \quad \text{from (71)}.
\]

From (77), we deduce that for all \(l \in [0, p + 2]\),

\[
\left\| u(\cdot, t) - u_\alpha(\cdot, t) \right\|_{H^l} \leq \frac{C}{1 + t^2} E^{\frac{l+1/2}{p+2}} \delta^{\frac{l+1/2}{p+2}}.
\]

(78)

On the other hand, from (53), we have

\[
\left\| u_\alpha(\cdot, t) - u_\delta(\cdot, t) \right\|_{H^l} \leq \frac{C \delta}{l^{1/4} \alpha} \leq \frac{C}{l^{1/4} \alpha} \left( \frac{C(p, \gamma)}{\theta - 1} \right)^{\frac{l+1/2}{p+2}} E^{\frac{l+1/2}{p+2}} \delta^{1-\frac{l}{p+2}} \quad \text{using (73)}.
\]

Estimate (76) follows readily from (78) and (79). □

**Remark 12.** In Theorem 6, taking \(l = 0\), we get the best possible rate \(\|u(\cdot, t) - u_\delta(\cdot, g^\delta)(\cdot, t)\|_{L^2} \leq C \frac{\delta}{\tau^2}\) for the posteriori parameter choice rule (67).

Let us end this section with the following algorithm for approximating the regularization parameter \(\alpha(\delta, g^\delta)\).
Algorithm 1
1: Set $\alpha_0 \gg 1$ and $q \in (0, 1)$
2: Set $\alpha = \alpha_0$ (initial guess)
3: while $\|1 - e^{-\tau(\alpha \xi)_{p}} \tilde{g}^\delta(\xi)\| > \theta \delta$ do
4: \hspace{1em} $\alpha = q \times \alpha$
5: end while

5 Numerical experiments

In order to illustrate the effectiveness of our regularization approach, we consider four numerical examples in two-dimension space where we invariably set $T = 1$ and $\gamma = 0.8$.

**Example 1**: $u(x, 0) = e^{-x_1^2 - x_2^2}$.

**Example 2**: $u(x, 0) = e^{-|x_1| - |x_2|}$.

**Example 3**: $u(x, 0) = v(x_1)v(x_2)$ where $v$ is triangle impulse set as $v(\lambda) = \begin{cases} 
1 + \lambda/3 & \text{if } \lambda \in [-3, 0] \\
1 - \lambda/3 & \text{if } \lambda \in (0, 3] \\
0 & \text{otherwise}.
\end{cases}$

**Example 4**: $u(x, 0) = \begin{cases} 
1 & \text{if } x \in [-5, 5]^2 \\
0 & \text{otherwise}.
\end{cases}$

Notice that, in Examples 1 2 and 3, $u$ is continuous on the contrary to Example 4. Moreover, in Example 1, $u(\cdot, 0) \in H^p(\mathbb{R}^2)$ for all $p > 0$; in Example 2, $u(\cdot, 0) \in H^p(\mathbb{R}^2)$ for $p < 1$; in Example 3, $u(\cdot, 0) \in H^1(\mathbb{R}^2)$; and in Example 4, $u(\cdot, 0) \in H^p(\mathbb{R}^2)$ for $p < 1/2$.

In the four examples, the support of $u(\cdot, 0)$ is considered $[-L, L]^2$ which is uniformly discretized as $(x_1(i), x_2(j))$ where $x_1(i) = x_2(i) = -L + (i - 0.5)\kappa$ with $\kappa = 2L/N$ and $i, j = 1, ..., N$. In all the simulations, we set $L = 10$ and $N = 256$.

The noisy data $g^\delta$ is generated as $g^\delta(x) = u(x, T) + \eta \epsilon(x)$, where $\epsilon(x)$ is a random number drawn from the standard normal distribution and $\eta$ is a parameter allowing to control to amount to noise added to the data. The noise level $\delta$ is nothing but $\eta E \|\epsilon\|_2$ and the percentage of noise $\text{perc\_noise}$ is nothing but

$$\text{perc\_noise} = \frac{100 \times \eta E \|\epsilon\|_2}{\|u(\cdot, T)\|_2} \%.$$

For the mollifier kernel $\varphi$ defined in (27), we invariably choose $s = 4$ and $\tau = 1/2$. The Fourier transform and inverse Fourier transform involved in the computation of the reconstructed solution $u_\alpha^\delta$ are quite rapidly evaluated with the numerical procedure from [1] based on fast Fourier transform (FFT) algorithm. From the Shannon-Nyquist principle, we set the frequency domain to $[-\Omega, \Omega]^2$ with $\Omega = \pi N/2L$.

For all the results in this section, we consider the parameter choice rule $\alpha(\delta, g^\delta)$ defined by (67) with $\theta = 1.01$. We compute this parameter via Algorithm 1 with $q = 0.99$ and $\alpha_0 = 10$. On Figures 1 and 2, we illustrate the approximate solution for $\text{perc\_noise} = 1\%$ in each example.

The reconstructed solution $u_\alpha^\delta(\delta, g^\delta)$ is computed via formula (33) (with $t = 0$) followed by inverse Fourier transform. For each approximate solution $u_\alpha^\delta(\delta, g^\delta)$, we assess the relative error

$$\text{rel\_err} = \frac{\|u_\alpha^\delta(\delta, g^\delta) - u(\cdot, 0)\|_2}{\|u(\cdot, 0)\|_2}.$$
In order to numerically confirm the theoretical rates of the reconstruction error, on Figure 3, we plot $\ln(\text{rel.err})$ versus $\ln(\delta)$ for various values of $\delta$. We recall that if the reconstruction error is of order $O(\delta^r)$ as $\delta \to 0$, then the curve $(\ln(\delta), \ln(\text{rel.err}))$ should exhibit a line shape with slope equal to $r$. From Figure 3, we can see that for each example, the plot clearly exhibits a line shape, confirming thus the power rate in the reconstruction error. Though the numerical order $r_{\text{num}}$ observed is different from the theoretical rate $r = p/(p+2)$, however, the numerical order does increase as the solution gets smoother. In fact, the higher order (0.5915) is achieved for example 1 which corresponds to the smoothest case; next in example 2 and 3, where the solution $u(\cdot, 0)$ has approximated the same regularity, the numerical orders in this cases are closed; lastly, the lowest numerical order (0.16) is achieved in Example 4 where the solution is the least regular.

Fig. 1: Illustration reconstructed solution $u^\delta_{a(\delta, g^\delta)}$ for perc噪音 = 1% in Example 1 (first row) and Example 2 (second row).

Fig. 2: Illustration reconstructed solution $u^\delta_{a(\delta, g^\delta)}$ for perc噪音 = 1% in Example 3 (first row) and Example 4 (second row).

Finally, in order to assess the numerical stability and convergence of our method, we run a Monte Carlo simulations of 200 replications of noise term for each example with various percentage of noise. The results are summarized in Tables 1 and 2. From Tables 1 and 2, we can see that:
6 Conclusion

In this paper, we focused on the regularization of final value time-fractional diffusion equation on unbounded domain. We presented a broad class of regularization methods which encompasses some regularization methods appearing distinctly in literature for this problem. We proposed a simple regularization method which smoothly regularizes the problem without truncating high frequency components. We proved order-optimality of our regularization approach in the practical setting where not only the data are approximated but also the forward diffusion operator. We also showed that with our regularization approach, we can also approximate early distribution $u(\cdot, t)$ with $t \in (0, T)$ with the best possible rate. For full applicability of the method, based on the Morozov principle, we provided a parameter choice rule leading to order-optimal rates. Finally, the soundness and efficiency of the regularization technique couples with the parameter choice rule prescribed is illustrated through some numerical simulations in two-dimensional space.

![Fig. 3: Illustration of numerical rates of convergence for the a-posteriori rule (67).](image)

| perc_noise | Example 1 | \(10\%\) | \(5\%\) | \(1\%\) | \(0.5\%\) | Example 2 | \(10\%\) | \(5\%\) | \(1\%\) | \(0.5\%\) |
|------------|-----------|--------|--------|--------|----------|-----------|--------|--------|--------|----------|
| mean(rel_err) | 1.585e-1  | 9.617e-2 | 3.457e-2 | 2.515e-2 | 2.203e-1 | 1.691e-1 | 9.178e-2 | 6.975e-2 |
| var(rel_err)  | 7.02e-5   | 2.145e-5 | 4.343e-7 | 5.489e-7 | 6.242e-5 | 2.864e-5 | 4.25e-6  | 1.43e-6  |
| mean(α(δ, gδ)) | 0.993964 | 0.825553 | 0.50099 | 0.468599 | 1.1373 | 0.929768 | 0.603317 | 0.497485 |

Tab. 1: Summary Monte Carlo simulation for Example 1 and 2 with 200 sample size.

| perc_noise | Example 3 | \(10\%\) | \(5\%\) | \(1\%\) | \(0.5\%\) | Example 4 | \(10\%\) | \(5\%\) | \(1\%\) | \(0.5\%\) |
|------------|-----------|--------|--------|--------|----------|-----------|--------|--------|--------|----------|
| mean(rel_err) | 9.438e-2  | 7.273e-2 | 3.962e-2 | 3.108e-2 | 2.605e-1 | 2.3e-1   | 1.768e-1 | 1.587e-1 |
| var(rel_err)  | 6.625e-6  | 3.439e-6 | 6.711e-7 | 2.762e-7 | 1.722e-5 | 1.071e-5 | 4.527e-6 | 2.666e-6 |
| mean(α(δ, gδ)) | 1.3899   | 1.1295  | 0.749755 | 0.632783 | 1.5956  | 1.24342 | 0.731601 | 0.591123 |

Tab. 2: Summary Monte Carlo simulation for Example 3 and 4 with 200 sample size.

- The reconstruction error and the regularization parameter $\alpha(\delta, g^\delta)$ both decrease as the noise level decreases. This indicates the numerical convergence of the regularization method coupled with the parameter choice rule (67).

- The variance of the reconstruction error in all cases is not larger than $10^{-4}$. This is a significant indicator of the high numerical stability of the regularization technique coupled with parameter choice rule (67).

- Lastly, the reconstruction error are much smaller in Example 1 and 3 compared to Examples 2 and 4. This observation confirms the following known fact: the smoother the sought solution, the better the reconstruction is expected. Indeed, the solution in Examples 1 and 3 are much smoother than the solution in Examples 2 and 4.
Appendix

Proof of Lemma 3: Let the function $f(x) = (1 + x)e^{-bx^d}$ with $b, d \in \mathbb{R}^*_+$. The function $f$ is continuous on $[0, +\infty)$ with $f(0) = 1$ and $\lim_{x \to +\infty} f(x) = 0$. Moreover, $f$ is smooth on $(0, \infty)$, hence, $f$ admits a global maximum on $\mathbb{R}_+$ attained at a critical point. Now, notice that $f'(x) = p(x)e^{-bx^d}$ with $p(x) = 1 - bdx^{d-1}(1 + x)$. Hence $f'(x) = 0$ implies that $p(x) = 0$. But

$$p(x) = 0 \Rightarrow x^{d-1}(1 + x) = \frac{1}{bd} \to +\infty \quad \text{as} \quad b \downarrow 0. \quad (80)$$

If $d \geq 1$, we can verify that $p$ has a unique root $\bar{x}$ on $\mathbb{R}_+$ which maximizes $f$. Moreover, $\bar{x} \to +\infty$ as $b \downarrow 0$. Since $\bar{x} \to +\infty$ as $b \downarrow 0$, from (80), we deduce that $x^d \sim 1/bd$. Since $p(\bar{x}) = 0$, then, $(1 + \bar{x}) = (1/bd)\bar{x}^{1-d}$ so that

$$\sup_{x \geq 0} f(x) = f(\bar{x}) \sim \frac{1}{bd} \bar{x}^{1-d} \exp(-b(\bar{x})^{1-d-1}) \sim \frac{1}{bd} (bd)^{-1-d} \exp(-1/d) = \frac{(d \exp(1))^{-1/d}}{b^{1/d}}, \quad (81)$$

whence the desired estimate (28).

If $d < 1$, we can verify that $p$ has two roots $x_0$ and $x_1$ on $\mathbb{R}_+$ with $x_0 < \bar{x}$ satisfying $x_0 \to 0$ as $b \downarrow 0$ and $x_1 \to +\infty$ as $b \downarrow 0$. Moreover we can check that $f$ admits a local minimum at $x_0$ and local maximum at $x_1$, so that the global maximum of $f$ on $[0, +\infty)$ is $\max(f(0), f(x_1)) = f(x_1)$ when $b \downarrow 0$. Similarly to the case $d \geq 1$, the roots $x_1$ satisfies $x_1^d \sim 1/bd$ and estimate (81) is also valid for $\bar{x} = x_1$ whence (28).

References

[1] D. H. Bailey and P. N. Swarztrauber. A fast method for the numerical evaluation of continuous fourier and laplace transforms. *SIAM Journal on Scientific Computing*, 15(5):1105–1110, 1994.

[2] V. Balakrishnan. Anomalous diffusion in one dimension. *Physica A: Statistical Mechanics and its applications*, 132(2-3):569–580, 1985.

[3] A. V. Chechkin, R. Gorenflo, and I. M. Sokolov. Fractional diffusion in inhomogeneous media. *Journal of Physics A: Mathematical and General*, 38(42):L679, 2005.

[4] D. N. Hao, J. Liu, N. Van Duc, and N. Van Thang. Stability results for backward time-fractional parabolic equations. *Inverse Problems*, 35(12):125006, 2019.

[5] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and applications of fractional differential equations*, volume 204. elsevier, 2006.

[6] J. Liu and M. Yamamoto. A backward problem for the time-fractional diffusion equation. *Applicable Analysis*, 89(11):1769–1788, 2010.

[7] A. K. Louis and P. Maass. A mollifier method for linear operator equations of the first kind. *Inverse problems*, 6(3):427, 1990.

[8] P. Mathé and B. Hofmann. How general are general source conditions? *Inverse Problems*, 24(1):015009, 2008.

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

[10] I. Podlubnv. Fractional differential equations academic press. *San Diego, Boston*, 1999.
[11] E. Schock. Approximate solution of ill-posed equations: arbitrarily slow convergence vs. super-convergence. In Constructive methods for the practical treatment of integral equations, pages 234–243. Springer, 1985.

[12] N. Van Duc, P. Q. Muoi, and N. Van Thang. A mollification method for backward time-fractional heat equation. Acta Mathematica Vietnamica, 45(3):749–766, 2020.

[13] J.-G. Wang, T. Wei, and Y.-B. Zhou. Optimal error bound and simplified tikhonov regularization method for a backward problem for the time-fractional diffusion equation. Journal of computational and applied mathematics, 279:277–292, 2015.

[14] L. Wang and J. Liu. Data regularization for a backward time-fractional diffusion problem. Computers & Mathematics with Applications, 64(11):3613–3626, 2012.

[15] L. Wang and J. Liu. Total variation regularization for a backward time-fractional diffusion problem. Inverse problems, 29(11):115013, 2013.

[16] X. Xiong, Q. Zhou, and Y. Hon. An inverse problem for fractional diffusion equation in 2-dimensional case: Stability analysis and regularization. Journal of Mathematical Analysis and Applications, 393(1):185–199, 2012.

[17] F. Yang, C.-L. Fu, and X.-X. Li. A mollification regularization method for unknown source in time-fractional diffusion equation. International Journal of Computer Mathematics, 91(7):1516–1534, 2014.

[18] M. Yang and J. Liu. Solving a final value fractional diffusion problem by boundary condition regularization. Applied Numerical Mathematics, 66:45–58, 2013.

[19] M. Yang and J. Liu. Fourier regularization for a final value time-fractional diffusion problem. Applicable Analysis, 94(7):1508–1526, 2015.