WELL-POSEDNESS AND NUMERICAL APPROXIMATION OF A FRACTIONAL DIFFUSION EQUATION WITH A NONLINEAR VARIABLE ORDER

Buyang Li[1,*], Hong Wang[2] and Jilu Wang[3]

Abstract. We prove well-posedness and regularity of solutions to a fractional diffusion porous media equation with a variable fractional order that may depend on the unknown solution. We present a linearly implicit time-stepping method to linearize and discretize the equation in time, and present rigorous analysis for the convergence of numerical solutions based on proved regularity results.

Mathematics Subject Classification. 65M15, 65M60, 65M12, 45K05.

1. Introduction

Fractional diffusion partial differential equations (PDEs) were shown to model anomalously diffusive transport (*e.g.*, of solutes through heterogeneous aquifers) more accurately than integer-order Fickian diffusion PDEs do [38, 39]. The fundamental reason is that the latter, which were obtained if the underlying particle jumps have (i) a mean free path and (ii) a mean waiting time [38, 39], are characterized by solutions with Gaussian type symmetric and exponentially decaying tails. Assumptions (i) and (ii) hold for diffusive transport of solutes in homogeneous aquifers when the solute plumes were observed to exhibit the same tail behavior [4]. However, field tests showed that the diffusive transport of solutes in heterogeneous aquifers often exhibits highly skewed power-law decaying tails [6,37,38]. A traditional practice to model diffusive transport of solutes in heterogeneous aquifers is to tweak the variable parameters that multiply the pre-set integer-order diffusion PDEs to fit the training data, which tends to recover a rapidly varying diffusivity and may overfit the training data but yield less accurate prediction on testing data [47].

The time-fractional PDE (tFPDE)

$$\partial_t^\alpha u(x,t) - \Delta u(x,t) = f(x,t), \quad \text{with} \quad 0 < \alpha < 1, \tag{1.1}$$

was derived *via* the continuous time random walk to model the subdiffusive transport of solutes [37,38] assuming that the mean waiting time has a power-law decaying tail [37,38]. Consequently, (1.1) more accurately models the subdiffusive transport of solutes in heterogeneous aquifers. However, (1.1) was derived as the diffusion limit

[1] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong.
[2] Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA.
[3] Beijing Computational Science Research Center, Beijing 100193, P.R. China.
*Corresponding author: `buyang.li@polyu.edu.hk`, `libuyang@gmail.com`

of a continuous time random walk in the phase space as the number of particle jumps tends to infinity and hence holds for large time $t > 0$ instead of all the way up to the time $t = 0$ as often assumed in the literature for simplicity [37, 38]. To resolve this problem, a mobile-immobile tFPDE model

$$\partial_t u(x,t) + b(x,t)\partial_t^\alpha u(x,t) - \Delta u(x,t) = f(x,t), \quad \text{with} \quad 0 < \alpha < 1, \tag{1.2}$$

was derived in [47, 56] to model subdiffusive transport of solutes in heterogeneous aquifers. In this model, a $b/(1 + b)$ portion of total solute mass in the aqueous phase may get absorbed to the aquifers and forms an immobile phase, and then gets released at a much later time. The significantly increased waiting time leads to a power-law decaying tail of the solute and thus undergoes subdiffusive transport. The subdiffusive transport of the adsorbed particles is modeled by the time-fractional derivative $b\partial_t^\alpha u$. The remaining $1/(1 + b)$ portion of the solute mass in the bulk fluid phase forms a mobile phase that undergoes a Brownian motion, and hence is modeled by the $\partial_t u$ term.

In nonconventional gas and oil recovery, the reservoirs often have insufficient permeability due to the existence of micropores, resulting in a large amount of adsorbed gas or oil mass and significantly decreased flow rate to the wellbore. Hydraulic fracturing is often used to increase the permeability of the porous media [15, 22]. The change of structure of porous media results in change of fractal dimension of media [13, 37] *via* the Hurst index, which in turn leads to change of order in tFPDEs [37]. This can be modeled by variable-order tFPDEs with α depending on x and t; see [25, 29, 50, 51]. In another application, the clogging of pore space in soil by microbial biomass, their body and their byproducts, *i.e.*, bioclogging, reduced the rate of infiltration of water remarkably and has significant impact on water, solute, and energy exchanges between stream water and groundwater as well as the environment [3, 55]. In this case, the fractional order α also depends on the unknown solution u. Motivated by these discussions, we consider the following tFPDE:

$$\begin{cases} \partial_t u(x,t) + b(x,t)\partial_t^{\alpha(x,t,u(x,t))} u(x,t) - \Delta u(x,t) = f(x,t) & (x,t) \in \Omega \times (0,T], \\ u(x,t) = 0 & (x,t) \in \partial\Omega \times (0,T], \\ u(x,0) = u_0(x) & x \in \Omega, \end{cases} \tag{1.3}$$

where Ω is a bounded convex domain, f and u_0 are given source term and initial value, respectively, and $\partial_t^{\alpha(x,t,u(x,t))} u(x,t)$ denotes the Caputo fractional-order partial derivative in time, defined by

$$\partial_t^{\alpha(x,t,u(x,t))} u(x,t) := \frac{1}{\Gamma(1 - \alpha(x,t,u(x,t)))} \int_0^t (t-s)^{-\alpha(x,t,u(x,t))} \partial_s u(x,s)\mathrm{d}s, \tag{1.4}$$

with a variable order $\alpha(x,t,u(x,t))$ which may depend on the unknown solution u, with $\Gamma(s) := \int_0^\infty t^{s-1}e^{-t}\mathrm{d}t$ denoting the Euler gamma function. Such fractional-order time derivatives in PDE models often present mathematical bottlenecks that were not encountered in the context of standard diffusion PDEs in analysis of regularity of solutions and convergence of numerical solutions. Correspondingly, many efforts have been made in developing efficient numerical methods and error analysis for such tFPDEs.

One popular approach is to discretize fractional-order derivatives by using the convolution quadrature (CQ) introduced in [30]. This approach has been used to solve (1.1) and related problems in [2, 11, 17–19, 32]. An advantage of this approach is the existence of a unified framework [31] for constructing and analyzing high-order methods for tFPDEs with constant coefficients, with possible weak singularities at $t = 0$, including CQ generated by the Crank–Nicolson method [19], BDF [18] and Runge–Kutta methods [2]. As explained in Appendix A of [18], error analysis of CQ for tFPDEs with constant coefficients can be equivalently viewed as error analysis of CQ for discretizing an operator-valued convolution integral. Other popular approaches for discretizing time-fractional PDEs include finite difference methods [12, 28, 35, 49] and discontinuous Galerkin methods [36, 40, 42], which approximate the fractional derivative at grid points and in a weak form, respectively, based on piecewise polynomial approximation of the solution. These approaches are flexible for variable stepsizes to improve the accuracy of numerical solutions; see [23, 27, 43–45, 48]. More recently, numerical approximation

of fractional derivatives based on deconvolution was proposed in [26], which has application in discretizing fractional stochastic differential equations and fractional gradient flows.

For all approaches mentioned above, numerical analysis in the literature mainly focuses on tFPDEs with a constant fractional order. When the fractional order α depends only on x, independent of t, convolution quadrature and Laplace transform techniques can still be used similarly as for the fractional substantial derivative in [21, 52]. When the fractional order α depends only on t, independent of x, well-posedness of variable-order tFPDEs was proved in [53] based on eigenfunction decomposition, which reduces the problem to decoupled fractional ordinary differential equations. However, when the fractional order depends on both x and t, or the unknown solution u, many existing techniques are not directly applicable, including the Laplace transform techniques and separation of variables. In this case, analysis of regularity of solutions and convergence of numerical methods is still challenging. This is similar to tFPDEs with time-dependent diffusion coefficients, for which the Laplace transform techniques and separation of variables cannot be applied directly either. For tFPDEs with time-dependent diffusion coefficients, convergence of a semidiscrete finite element method was proved in [41] based on a novel technique which extends the argument for standard parabolic problems in [34] to the fractional case. Existence, uniqueness and regularity of solutions, as well as convergence of a fully discrete numerical method were proved in [20] based on a perturbation argument; also see [24] for new regularity results.

The objective of this paper is to develop techniques to analyze well-posedness and accuracy of numerical solutions for such complicated nonlinear fractional evolution problems as (1.3) under a more general setting such that the variable order $\alpha(x, t, u)$ can depend on both x and t as well as the unknown solution u, and is possibly discontinuous in x. In addition to the time-dependency of the variable order, the nonlinearity brings in new difficulties for regularity and numerical analysis. In particular, the boundedness of solution in $L^\infty(0, T; L^\infty(\Omega))$ becomes essential in proving well-posedness and regularity of the solutions, as well as convergence of numerical solutions, while the Laplace transform techniques cannot be used directly in establishing the regularity of solutions. We present new techniques, based on maximal L^p-regularity of parabolic equations and Schaefer's fixed point theorem in the Banach space $L^\infty(0, T; L^\infty(\Omega))$, to prove the existence and uniqueness of solutions to the nonlinear problem (1.3) under condition (2.1). This condition allows discontinuity of $\alpha(x, t, u)$ in x to fit practical applications in which different physical subdomains of the porous media may have different structures and properties that lead to different fractional orders in the model. Under a slightly stronger condition (3.1), we prove some regularity results for further numerical approximation and error analysis.

We propose a linearly implicit CQ time-stepping method to linearize and discretize the nonlinear problem (1.3), and prove convergence of the proposed method based on the natural regularity of solutions proved in this paper for a given source term $f \in W^{1,1}(0, T; L^2(\Omega))$ and initial value $u_0 \in \dot{H}^2(\Omega)$. Again, the Laplace transform technique cannot be directly used for error analysis due to the nonlinearity of the problem. To overcome the difficulty caused by the time-dependent and nonlinear variable order, we present error analysis by using a perturbation technique which freezes the nonlinear variable order at a fixed time level and estimates the defect based on the natural regularity proved for the solution and a generalized discrete Gronwall's inequality introduced in this paper (see Appendix D). Numerical results are provided to support the theoretical analysis on the convergence of the numerical solutions.

The techniques developed in this paper could also be applicable to other related nonlinear fractional-order equations possibly with time-dependent and nonlinear variable order, and time-dependent diffusion coefficients.

2. Existence and uniqueness of solutions

In this section, we prove existence and uniqueness of solutions to (1.3) under the following conditions:

$$\alpha : \Omega \times [0, T] \times \mathbb{R} \to [0, \alpha_*] \text{ is Lipschitz continuous in the third argument,}$$

$$b \in L^\infty(0, T; L^\infty(\Omega)),$$

$$u_0 \in \dot{H}^2(\Omega), \quad f \in L^p(0, T; L^2(\Omega)) \quad \text{for some} \quad p > \frac{4}{4-d},$$

$$(2.1)$$
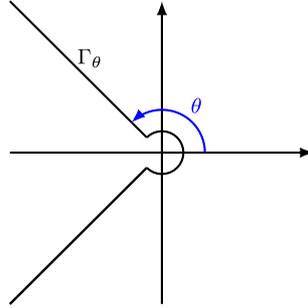
FIGURE 1. The contour Γ_θ on the complex plane.

where $\alpha_* \in (0,1)$ is a fixed constant and $d \in \{1,2,3\}$ is the dimension of space.

2.1. Notation

For simplicity of notation, we define $\dot{H}^2(\Omega) = H^2(\Omega) \cap H_0^1(\Omega)$ and denote by $\dot{H}^s(\Omega) = (L^2(\Omega), \dot{H}^2(\Omega))_{[s/2]}$ the complex interpolation space between $L^2(\Omega)$ and $\dot{H}^2(\Omega)$ for $s \in [0,2]$; see [7]. Thus $\dot{H}^0(\Omega) = L^2(\Omega)$ and $\dot{H}^1(\Omega) = H_0^1(\Omega)$. For $0 < \gamma < 1$, we denote by $C^\gamma(\overline{\Omega})$ the conventional space of Hölder continuous functions on $\overline{\Omega}$ with exponent γ. For a Banach space X (*e.g.*, $X = L^2(\Omega)$), we denote by $C^\gamma([0,T];X)$ the space of Hölder continuous functions from $[0,T]$ to X. Furthermore, we define the following function spaces:

$$
\begin{aligned}
X &= \{w \in W^{1,p}(0,T;L^2(\Omega)) : w(0) = 0\}, \\
Y &= \{w \in W^{1,p}(0,T;L^2(\Omega)) \cap L^p(0,T;\dot{H}^2(\Omega)) : w(0) = 0\}, \\
Z &= L^\infty(0,T;L^\infty(\Omega)).
\end{aligned}
\tag{2.2}
$$

We denote by $(\cdot,\cdot)$ and $\|\cdot\|_{L^2(\Omega)}$ the inner product and norm of $L^2(\Omega)$, and denote by $\|\cdot\|$ the operator norm from $L^2(\Omega)$ to $L^2(\Omega)$.

It is well known that the Dirichlet Laplacian $\Delta : \dot{H}^2(\Omega) \to L^2(\Omega)$ generates a bounded analytic semigroup on BOTH $L^2(\Omega)$ and $C_0(\overline{\Omega})$ with angle $\frac{\pi}{2}$; see [46]. Equivalently, $z - \Delta$ is invertible for $z \in \mathbb{C}$ such that $|\arg(z)| < \pi$, and the following resolvent estimate holds:

$$
\|(z-\Delta)^{-1}\| \le C_\theta |z|^{-1}, \quad \forall z \in \Sigma_\theta \backslash \{0\}, \quad \forall \theta \in (\tfrac{\pi}{2}, \pi),
\tag{2.3}
$$

$$
\|(z-\Delta)^{-1}\|_{L^\infty(\Omega) \to L^\infty(\Omega)} \le C_\theta |z|^{-1}, \quad \forall z \in \Sigma_\theta \backslash \{0\}, \quad \forall \theta \in (\tfrac{\pi}{2}, \pi),
\tag{2.4}
$$

where $\Sigma_\theta = \{z \in \mathbb{C} : |\arg(z)| \le \theta\}$ is a sector on the complex plane.

We denote by $\theta \in (\frac{\pi}{2}, \pi)$ a fixed number and denote by Γ_θ a contour on the complex plane (contained in the sector $\overline{\Sigma}_\theta$), defined by

$$
\Gamma_\theta = \{z \in \mathbb{C} : |\arg(z)| = \theta, |z| \ge 1/T\} \cup \{z \in \mathbb{C} : |\arg(z)| \le \theta, |z| = 1/T\},
\tag{2.5}
$$

as shown in Figure 1. Then $z^{\alpha(x,t)} \in \Sigma_\theta$ for $(x,t) \in \Omega \times [0,T]$ and $z \in \Sigma_\theta \supset \Gamma_\theta$.

Let ϵ_* be a fixed number satisfying $0 < \epsilon_* < \min(1 - \alpha_*, 1 - d/4)$, where $d \in \{1,2,3\}$ is the dimension of space. Related to the number ϵ_*, we define a function

$$
\eta(t) = t^{-\alpha_* - \epsilon_*},
\tag{2.6}
$$

which is integrable in time and has the following property:

$$
|\partial_t^{\alpha(x,t,u(x,t))} v(t)| \le C \int_0^t \eta(t-s) |\partial_t v(s)| \mathrm{d}s =: \eta * |\partial_t v|(t),
$$

where C is some constant and $\eta * |\partial_t v|$ denotes convolution in time between the two functions. This property is a direct consequence of definition (1.4).

The Laplace transform of a function $f(x,t)$ is denoted by $\mathcal{L}f(x,z) = \int_0^\infty f(x,t)e^{-tz}\mathrm{d}t$. For functions of several temporal variables, we denote by $\mathcal{L}_s[f(x,s,t)]$ the Laplace transform in the s-variable and denote by $\mathcal{L}_z^{-1}[\hat{f}(x,z,t)]$ the inverse Laplace transform of $\hat{f}$ in the z-variable. The definition of fractional-order derivative in (1.4) is related to the Laplace transform through the following identity (*cf.* [30]):

$$
\begin{aligned}
&\partial_t^{\alpha(x,t,u(x,t))} v(x,t) \\
&= \mathcal{L}_z^{-1}\big[z^{\alpha(x,t,u(x,t))}\mathcal{L}_t[\widetilde{v}(x,t) - v(x,0)](z)\big](t) \\
&= \frac{1}{2\pi i}\int_{\Gamma_\theta} z^{\alpha(x,t,u(x,t))}\mathcal{L}_t[\widetilde{v}(x,t) - v(x,0)](z)e^{tz}\mathrm{d}z \quad \forall v \in W^{1,p}(0,T;L^2(\Omega)),
\end{aligned}
\tag{2.7}
$$

where $\widetilde{v} \in W^{1,p}(\mathbb{R}_+;L^2(\Omega))$ denotes an extension of v with compact support in t, and this identity is independent of the choice of extension. The following inequalities are often used in literature in analysis of tFPDEs (for example, see [18,30,32]):

$$
\int_{\Gamma_\theta} |z|^{\gamma-1}|e^{tz}|\,|\mathrm{d}z| \le C_\gamma t^{-\gamma} \quad \text{and} \quad \left\|\int_{\Gamma_\theta} z^\gamma(z-\Delta)^{-1}e^{tz}\,\mathrm{d}z\right\| \le C_\gamma t^{-\gamma} \quad \forall t \in (0,T],
\tag{2.8}
$$

where $|\mathrm{d}z|$ denotes the arc length element on the contour Γ_θ, and C_γ is some constant depending on the parameter γ. These two inequalities will be often used in this paper, in analyzing both regularity of solutions to (1.3) and convergence of numerical solutions.

Throughout this paper, we use the abbreviation $v(t) = v(\cdot,t)$ for any function v defined on $\Omega \times (0,T]$, and we denote by $\alpha(t,v(t))$ or $\alpha(t,v)$ the spatial function $\alpha(\cdot,t,v(\cdot,t))$. We denote by C a generic positive constant (independent of τ, n and N) that could be different at different occurrences.

2.2. The linear problem

Before studying the nonlinear problem (1.3), we first consider the following linear problem with a given variable order $\alpha(x,t)$:

$$
\begin{cases}
\partial_t u(x,t) + b(x,t)\partial_t^{\alpha(x,t)}u(x,t) - \Delta u(x,t) = f(x,t) & (x,t) \in \Omega \times (0,T], \\
u(x,t) = 0 & (x,t) \in \partial\Omega \times (0,T], \\
u(x,0) = u_0(x) & x \in \Omega,
\end{cases}
\tag{2.9}
$$

where $\alpha(x,t)$ and $b(x,t)$ are assumed to satisfy the following conditions:

$$
\begin{aligned}
&0 \le \alpha(x,t) \le \alpha_* < 1 \quad \text{for} \quad (x,t) \in \Omega \times [0,T], \\
&b \in L^\infty(0,T;L^\infty(\Omega)), \\
&u_0 \in \dot{H}^2(\Omega), \quad f \in L^p(0,T;L^2(\Omega)) \quad \text{for some} \quad p > 1.
\end{aligned}
\tag{2.10}
$$

In this section, we prove well-posedness of the linear problem (2.9) by utilizing the following three technical tools.

Lemma 2.1 (Maximal L^p-regularity, *cf.* [54], Cor. d or [1], Lem. 8.3)**.** *For any given $1 < p < \infty$ and $f \in L^p(0,T;L^2(\Omega))$, the heat equation*

$$
\begin{cases}
\partial_t u(x,t) - \Delta u(x,t) = f(x,t) & (x,t) \in \Omega \times (0,T], \\
u(x,t) = 0 & (x,t) \in \partial\Omega \times (0,T], \\
u(x,0) = 0 & x \in \Omega,
\end{cases}
\tag{2.11}
$$

has a unique solution $u \in W^{1,p}(0,T;L^2(\Omega)) \cap L^p(0,T;\dot{H}^2(\Omega))$, which satisfies the following estimate:

$$\|u\|_{W^{1,p}(0,T;L^2(\Omega))} + \|u\|_{L^p(0,T;\dot{H}^2(\Omega))} \le C\|f\|_{L^p(0,T;L^2(\Omega))} \quad \forall f \in L^p(0,T;L^2(\Omega)), \tag{2.12}$$

where the constant C is independent of f and T, but may depend on p.

Lemma 2.2 (Young's convolution inequality, *cf.* [8], Thm. 3.9.4). *Let* $g * v(t) := \int_0^t g(t-s)v(s)\mathrm{d}s$ *denote the convolution of* $g \in L^1(0,T)$ *with* $v \in L^p(0,T)$. *Then*

$$\|g * v\|_{L^p(0,t_*)} \le \|g\|_{L^1(0,t_*)}\|v\|_{L^p(0,t_*)}, \quad \forall t_* \in (0,T], \quad \forall 1 \le p \le \infty. \tag{2.13}$$

The following result can be proved by using Lemma 2.2.

Lemma 2.3 (A weighted L^p estimate for fractional-order derivatives). *For* $v \in W^{1,p}(0,T;L^2(\Omega))$ *and* $\lambda \ge 1$, *the following estimate holds:*

$$\|e^{-\lambda t}\partial_t^{\alpha(x,t)}v\|_{L^p(0,t_*;L^2(\Omega))} \le C\lambda^{\alpha_*-1}\|e^{-\lambda t}\partial_t v\|_{L^p(0,t_*;L^2(\Omega))} \quad \forall t_* \in (0,T], \tag{2.14}$$

where the constants C is independent of λ and t_* (but may depend on T).

Proof. Based on the definition (1.4), we have

$$
\begin{aligned}
|\partial_t^{\alpha(x,t)}v(x,t)| &= \frac{1}{\Gamma(1-\alpha(x,t))}\int_0^t (t-s)^{-\alpha(x,t)}\partial_s v(x,s)\mathrm{d}s \\
&\le CT^{-\alpha(x,t)}\int_0^t \left(\frac{t-s}{T}\right)^{-\alpha(x,t)}|\partial_s v(x,s)|\mathrm{d}s \\
&\le C\min(T,1)^{-\alpha_*}T^{\alpha_*}\int_0^t (t-s)^{-\alpha_*}|\partial_s v(x,s)|\mathrm{d}s \quad \forall t \in (0,T],
\end{aligned}
$$

where we have used the inequality $\left(\frac{t-s}{T}\right)^{-\alpha(x,t)} \le \left(\frac{t-s}{T}\right)^{-\alpha_*}$ and $T^{-\alpha(x,t)} \le \min(T,1)^{-\alpha_*}$. Multiplying the above inequality by $e^{-\lambda t}$, we obtain

$$e^{-\lambda t}\partial_t^{\alpha(x,t)}v(x,t) \le C\min(T,1)^{-\alpha_*}T^{\alpha_*}\int_0^t e^{-\lambda(t-s)}(t-s)^{-\alpha_*}|e^{-\lambda s}\partial_s v(x,s)|\mathrm{d}s.$$

For $t_* \in (0,T]$, the inequality above implies

$$
\begin{aligned}
\|e^{-\lambda t}\partial_t^{\alpha(x,t)}v\|_{L^p(0,t_*;L^2(\Omega))} &= C\min(T,1)^{-\alpha_*}T^{\alpha_*}\|(e^{-\lambda t}t^{-\alpha_*})*(e^{-\lambda t}\partial_t v)\|_{L^p(0,t_*;L^2(\Omega))} \\
&\le C\min(T,1)^{-\alpha_*}T^{\alpha_*}\|e^{-\lambda t}t^{-\alpha_*}\|_{L^1(0,t_*)}\|e^{-\lambda t}\partial_t v\|_{L^p(0,t_*;L^2(\Omega))},
\end{aligned}
$$

where we have used Lemma 2.2 in the last inequality. Then we note that

$$\|e^{-\lambda t}t^{-\alpha_*}\|_{L^1(0,t_*)} = \int_0^{t_*} e^{-\lambda t}t^{-\alpha_*}\mathrm{d}t \le \lambda^{\alpha_*-1}\int_0^\infty e^{-s}s^{-\alpha_*}\mathrm{d}s \le C\lambda^{\alpha_*-1},$$

with a constant C independent of $\lambda \ge 1$. Combining the two estimates above, we obtain the desired result in Lemma 2.3. $\qquad\square$

By using Lemma 2.3, we can prove well-posedness of the linear problem (2.9).

Theorem 2.4 (Existence and uniqueness of solutions to the linear problem). *Under assumption* (2.10)*, problem* (2.9) *has a unique solution*

$$u \in W^{1,p}(0,T; L^2(\Omega)) \cap L^p(0,T; \dot{H}^2(\Omega)) \hookrightarrow C([0,T]; L^2(\Omega)), \tag{2.15}$$

and the solution satisfies the following estimate:

$$\|u\|_{W^{1,p}(0,t_*;L^2(\Omega))} + \|u\|_{L^p(0,t_*;\dot{H}^2(\Omega))} \le C \left(\|f\|_{L^p(0,t_*;L^2(\Omega))} + \|u_0\|_{\dot{H}^2(\Omega)} \right) \quad \forall t_* \in (0,T], \tag{2.16}$$

where the constant C is independent of $\alpha(x,t)$ and $t_* \in (0,T]$ (but may depend on α_* and T).

Proof. By considering the equation of $u - u_0$, we can reduce problem (2.9) to the case with zero initial value (with f replaced by $f + \Delta u_0$). Hence, we focus on the case $u_0 = 0$.

First, we denote by X_λ the function space X defined in (2.2), equipped with the equivalent norm

$$\|w\|_{X_\lambda} := \|e^{-\lambda t}\partial_t w\|_{L^p(0,T;L^2(\Omega))},$$

where $\lambda \ge 1$ is a parameter to be chosen later. We define a map $M : X_\lambda \to X_\lambda$ in the following way: for any $v \in X_\lambda$, we define $w = Mv$ to be the solution of

$$\begin{cases} \partial_t w(x,t) - \Delta w(x,t) = f(x,t) - b(x,t)\partial_t^{\alpha(x,t)}v(x,t) & (x,t) \in \Omega \times (0,T], \\ w(x,t) = 0 & (x,t) \in \partial\Omega \times (0,T], \\ w(x,0) = 0 & x \in \Omega. \end{cases} \tag{2.17}$$

Since $v \in X_\lambda$ implies $b\partial_t^{\alpha(x,t)}v \in L^p(0,T;L^2(\Omega))$ (this was proved in Lem. 2.3), it follows from Lemma 2.1 that (2.17) has a unique solution $w \in Y \hookrightarrow X_\lambda$, where Y is defined in (2.2). Hence the map $M : X_\lambda \to X_\lambda$ is well defined, and any fixed point of M is a solution of (2.9) in Y. We shall prove existence and uniqueness of a fixed point of M by using the Banach fixed point theorem.

Second, we denote $\widetilde{f}(x,t) = f(x,t) - b(x,t)\partial_t^{\alpha(x,t)}v(x,t)$ and consider the Laplace transform of (2.17) in time. Then we obtain $(z - \Delta)\mathcal{L}w(z) = \mathcal{L}\widetilde{f}(z)$, which furthermore implies $z\mathcal{L}w(z) = z(z - \Delta)^{-1}\mathcal{L}\widetilde{f}(z)$. Since the Laplace transform maps $e^{-\lambda t}\partial_t w$ to $(z + \lambda)\mathcal{L}w(z + \lambda)$ and $e^{-\lambda t}\widetilde{f}$ to $\mathcal{L}\widetilde{f}(z + \lambda)$, respectively, it follows that

$$\begin{aligned} e^{-\lambda t}\partial_t w = \mathcal{L}^{-1}[(z + \lambda)\mathcal{L}w(z + \lambda)] &= \mathcal{L}^{-1}[(z + \lambda)(z + \lambda - \Delta)^{-1}\mathcal{L}\widetilde{f}(z + \lambda)] \\ &= \mathcal{L}^{-1}[(z + \lambda)(z + \lambda - \Delta)^{-1}\mathcal{L}(e^{-\lambda t}\widetilde{f})(z)]. \end{aligned}$$

Setting $z = i\xi$ in the equality above and using the relation between Laplace and Fourier transforms (after extending $\widetilde{f}$ to be zero for $t < 0$), we obtain

$$e^{-\lambda t}\partial_t w = \mathcal{F}^{-1}[(i\xi + \lambda)(i\xi + \lambda - \Delta)^{-1}\mathcal{F}(e^{-\lambda t}\widetilde{f})(\xi)].$$

Since the operators

$$m(\xi) = (i\xi + \lambda)(i\xi + \lambda - \Delta)^{-1} \quad \text{and} \quad \xi m'(\xi) = i\xi(i\xi + \lambda - \Delta)^{-1} + i\xi(i\xi + \lambda)(i\xi + \lambda - \Delta)^{-2}$$

are bounded on $L^2(\Omega)$ for all $\xi \in \mathbb{R}$ and $\lambda > 0$, it follows that $m(\xi)$ is a Hörmander–Mihlin multiplier on $L^p(0,T;L^2(\Omega))$ for all $1 < p < \infty$ (*cf.* [16], Thm. 5.2.7). This means that

$$\begin{aligned} \|e^{-\lambda t}\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} &\le C\|e^{-\lambda t}\widetilde{f}\|_{L^p(0,t_*;L^2(\Omega))} \\ &\le C\|e^{-\lambda t}f\|_{L^p(0,t_*;L^2(\Omega))} + C\|e^{-\lambda t}\partial_t^{\alpha(x,t)}v\|_{L^p(0,t_*;L^2(\Omega))} \\ &\le C\|e^{-\lambda t}f\|_{L^p(0,t_*;L^2(\Omega))} + C\lambda^{\alpha_* - 1}\|e^{-\lambda t}\partial_t v\|_{L^p(0,t_*;L^2(\Omega))}, \end{aligned} \tag{2.18}$$

where we have used Lemma 2.3 in the last inequality, with a constant C independent of $\lambda > 0$ and $t_* \in (0, T]$.

Third, if $v_1, v_2 \in X_\lambda$ and $w_1 = Mv_1$ and $w_2 = Mv_2$, then $w = w_1 - w_2$ is the solution of the equation

$$\begin{cases} \partial_t w(x,t) - \Delta w(x,t) = b(x,t)\partial_t^{\alpha(x,t)}(v_2(x,t) - v_1(x,t)) & (x,t) \in \Omega \times (0,T], \\ w(x,t) = 0 & (x,t) \in \partial\Omega \times (0,T], \\ w(x,t) = 0 & x \in \Omega. \end{cases} \tag{2.19}$$

By applying (2.18) to (2.19) with $t_* = T$, we have

$$\|w_1 - w_2\|_{X_\lambda} \le C\lambda^{\alpha_*-1}\|v_1 - v_2\|_{X_\lambda}.$$

Choosing a sufficiently large λ guarantees that the map $M : X_\lambda \to X_\lambda$ is a contraction. According to the Banach fixed point theorem, M has a unique fixed point $w \in X_\lambda$ (thus $w = Mw \in Y$), i.e., (2.9) has a unique solution in Y when $u_0 = 0$.

Finally, by setting $v = w$ in (2.18) and choosing a sufficiently large λ (the largeness is independent of $t_* \in (0, T]$), we obtain the desired estimate (2.16). This completes the proof of Theorem 2.4. □

Remark 2.5. In Theorem 2.4 we have proved existence and uniqueness by using the Banach fixed point theorem with exponentially weighted norms in time. This trick is already known in the analysis of weakly singular Volterra integral equations – a class of problems that is closely related to the problem studied here; see [5, 9].

2.3. The nonlinear problem

In this section, we prove well-posedness of the nonlinear problem (1.3) by utilizing Theorem 2.4 and the following Schaefer's fixed point theorem.

Lemma 2.6 (Schaefer's fixed point theorem ([14], Chap. 9.2, Thm. 4)). *For a Banach space Z, suppose that $M : Z \to Z$ is a continuous and compact map. If the set*

$$\bigcup_{\theta \in [0,1]} \{v \in Z : v = \theta\, Mv\} \tag{2.20}$$

is bounded in Z, then the map M has at least one fixed point.

The main result of this section is the following theorem.

Theorem 2.7 (Existence and uniqueness of solutions). *Under the assumption (2.1), problem (1.3) has a unique solution*

$$u \in W^{1,p}(0,T;L^2(\Omega)) \cap L^p(0,T;\dot{H}^2(\Omega)) \hookrightarrow C([0,T];L^2(\Omega)). \tag{2.21}$$

Proof. Similarly as the linear problem, by considering the equation of $u - u_0$ we can reduce the problem to zero initial value (replacing f by $f + \Delta u_0$ and modifying the definition of α). Therefore, we only need to focus on the case $u_0 = 0$.

To prove existence of solutions for the nonlinear problem (1.3), we define a nonlinear map $M : Z \to Z$ by setting $w = Mv$ to be the solution of

$$\partial_t w + b(t)\partial_t^{\alpha(t,v)}w - \Delta w = f \tag{2.22}$$

with zero boundary and initial conditions. Theorem 2.4 guarantees that the map $M : Z \to Y$ is well defined.

To prove continuity of the map $M : Z \to Y$, we consider $v_1, v_2 \in Z$ and denote $w_1 = Mv_1$, $w_2 = Mv_2$. Then $w = w_1 - w_2$ is the solution of the equation

$$\partial_t w - \Delta w + b(t)\partial_t^{\alpha(t,v_1)}w = -b(t)\left(\partial_t^{\alpha(t,v_1)}w_2 - \partial_t^{\alpha(t,v_2)}w_2\right) \tag{2.23}$$

with zero boundary and initial conditions. By applying estimate (2.16) to equation (2.23), we obtain

$$
\begin{aligned}
\|\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} + \|w\|_{L^p(0,t_*;\dot{H}^2(\Omega))} &\leq C\|b(t)\left(\partial_t^{\alpha(t,v_1)}w_2 - \partial_t^{\alpha(t,v_2)}w_2\right)\|_{L^p(0,t_*;L^2(\Omega))} \\
&\leq C\|v_1 - v_2\|_{L^\infty(0,t_*;L^\infty(\Omega))}\|\partial_t w_2\|_{L^p(0,t_*;L^2(\Omega))},
\end{aligned}
\tag{2.24}
$$

where the last inequality follows from Lemma 2.8 (to be proved below).

Since the regularity estimate (2.16) implies $\|\partial_t w_2\|_{L^p(0,T;L^2(\Omega))} \leq C$, with a constant C independent of v_1 and v_2, choosing $t_* = T$ in (2.24) yields

$$
\|w\|_Y \leq C\|v_1 - v_2\|_Z.
$$

This proves the continuity of $M : Z \to Y$. Since Y is embedded into Z compactly when $p > \frac{4}{4-d}$ and $d \in \{1,2,3\}$ (see Appendix A), it follows that the map $M : Z \to Z$ is compact.

It is easy to see that any fixed point of M is a solution of (1.3) in Y. We now prove existence of a fixed point by using Schaefer's fixed point theorem (Lem. 2.6). If $v = \theta M v$ for some $\theta \in [0,1]$ then we denote $w = Mv$ with $v = \theta w$. Then w is the solution of the equation

$$
\partial_t w + b(t)\partial_t^{\alpha(t,\theta w)}w - \Delta w = f
\tag{2.25}
$$

with zero boundary and initial conditions. Applying the estimate (2.16) of Theorem 2.4, we immediately obtain

$$
\|w\|_{W^{1,p}(0,T;L^2(\Omega))} + \|w\|_{L^p(0,T;\dot{H}^2(\Omega))} \leq C\left(\|f\|_{L^p(0,T;L^2(\Omega))} + \|u_0\|_{\dot{H}^2(\Omega)}\right),
\tag{2.26}
$$

where the right-hand side is independent of $\theta \in [0,1]$. Therefore, by Schaefer's fixed point theorem, the map M has at least one fixed point u, which is a solution of the nonlinear problem (2.9), and this solution has regularity (2.21) because M maps Z into Y.

It remains to prove uniqueness of the solution. In fact, if w_1 and w_2 are two solutions of (2.9), then $w = w_1 - w_2$ is the solution of (2.23) with $v_1 = w_1$ and $v_2 = w_2$ therein. Then estimate (2.24) implies

$$
\begin{aligned}
\|\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} &+ \|w\|_{L^p(0,t_*;\dot{H}^2(\Omega))} \\
&\leq C\|w\|_{L^\infty(0,t_*;L^\infty(\Omega))} \\
&\leq C_\varepsilon\|w\|_{L^1(0,t_*;L^2(\Omega))} + \varepsilon\left(\|\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} + \|w\|_{L^p(0,t_*;\dot{H}^2(\Omega))}\right),
\end{aligned}
$$

where we have used Lemma B.1 (see Appendix B) in deriving the last inequality, in which ε can be arbitrarily small at the expense of enlarging the constant C_ε. By choosing a sufficiently small ε and using inequality

$$
\|w(t_*)\|_{L^2(\Omega)} \leq \|\partial_t w\|_{L^1(0,t_*;L^2(\Omega))} \leq T^{1-\frac{1}{p}}\left(\|\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} + \|w\|_{L^p(0,t_*;\dot{H}^2(\Omega))}\right),
$$

we obtain

$$
\|w(t_*)\|_{L^2(\Omega)} \leq C\|w\|_{L^1(0,t_*;L^2(\Omega))} \quad \forall t_* \in (0,T].
$$

Then Gronwall's inequality yields $w = 0$. This proves uniqueness of the solution. $\square$

In the proof of Theorem 2.7 we have used the following result.

Lemma 2.8. *For $v_1, v_2 \in Z$ and $v \in X$, the following estimates hold:*

$$
\|\partial_t^{\alpha(t,v_1)}v - \partial_t^{\alpha(t,v_2)}v\|_{L^p(0,t_*;L^2(\Omega))} \leq C\|v_1 - v_2\|_{L^\infty(0,t_*;L^\infty(\Omega))}\|\partial_t v\|_{L^p(0,t_*;L^2(\Omega))} \quad \forall t_* \in (0,T],
$$

where the constant C is independent of $t_* \in (0,T]$ (but may depend on T).

Proof. We use the following mean value result:

$$\frac{(t-s)^{-\alpha(t,v_1(t))}}{\Gamma(1-\alpha(t,v_1))} - \frac{(t-s)^{-\alpha(t,v_2(t))}}{\Gamma(1-\alpha(t,v_2))}$$

$$= -[\alpha(t,v_1(t)) - \alpha(t,v_2(t))] \int_0^1 \frac{(t-s)^{-\kappa_\omega(t)}\ln(t-s)}{\Gamma(1-\kappa_\omega(t))}\mathrm{d}\omega$$

$$+ [\alpha(t,v_1(t)) - \alpha(t,v_2(t))] \int_0^1 \frac{(t-s)^{-\kappa_\omega(t)}\Gamma'(1-\kappa_\omega(t))}{\Gamma(1-\kappa_\omega(t))^2}\mathrm{d}\omega,$$

where $\kappa_\omega(t) = (1-\omega)\alpha(t,v_1(t)) + \omega\alpha(t,v_2(t))$. Using definition (1.4) and the identity above, we have

$$|\partial_t^{\alpha(t,v_1(t))}v(t) - \partial_t^{\alpha(t,v_2(t))}v(t)| \le C|\alpha(t,v_1(t)) - \alpha(t,v_2(t))| \int_0^t (t-s)^{-\alpha_*-\epsilon_*}|\partial_s v(s)|\mathrm{d}s.$$

By taking the $L^p(0,t_*;L^2(\Omega))$ norm of this inequality, we obtain

$$\|\partial_t^{\alpha(t,v_1)}v - \partial_t^{\alpha(t,v_2)}v\|_{L^p(0,t_*;L^2(\Omega))}$$

$$\le C\|v_1 - v_2\|_{L^\infty(0,t_*;L^\infty(\Omega))}\left(\int_0^{t_*} t^{-\alpha_*-\epsilon_*}\mathrm{d}t\right)\|\partial_t v\|_{L^p(0,t_*;L^2(\Omega))},$$

where we have used Lipschitz continuity of α with respect to v. This proves the desired result. $\square$

3. Regularity of solutions

In this section, we present some regularity results for solutions of the nonlinear problem (1.3). These results will be used in Section 4 in analyzing convergence of numerical solutions. In addition to (2.1), we assume that the following conditions hold:

$$\begin{aligned}&\alpha : \Omega \times [0,T] \times \mathbb{R} \to [0,\alpha_*] \text{ is Lipschitz continuous in both } t \text{ and } u,\\ &b \in W^{1,\infty}(0,T;L^\infty(\Omega)),\\ &u_0 \in \dot{H}^2(\Omega) \quad \text{and} \quad f \in W^{1,1}(0,T;L^2(\Omega)),\end{aligned} \tag{3.1}$$

where $\alpha_* \in (0,1)$ is a fixed constant and $d \in \{1,2,3\}$ is the dimension of space.

When the function $\alpha(x,t,u)$ is Lipschitz continuous in both t and u, as assumed in condition (3.1), we have the following result as an extension of Lemma 2.8.

Lemma 3.1. *For $v_1, v_2 \in Z$ and $v \in X$, the following estimate holds:*

$$|\partial_s^{\alpha(s,v_1(s))}v(s) - \partial_s^{\alpha(t,v_2(t))}v(s)| \le C\int_0^s (|s-t| + |v_1(s) - v_2(t)|)(s-\sigma)^{-\alpha_*-\epsilon_*}|\partial_\sigma v(\sigma)|\mathrm{d}\sigma.$$

Proof. We use the following mean value result:

$$\frac{(s-\sigma)^{-\alpha(s,v_1(s))}}{\Gamma(1-\alpha(s,v_1(s)))} - \frac{(s-\sigma)^{-\alpha(t,v_2(t))}}{\Gamma(1-\alpha(t,v_2(t)))}$$

$$= -[\alpha(s,v_1(s)) - \alpha(t,v_2(t))] \int_0^1 \frac{(s-\sigma)^{-\kappa_\omega(s,t)}\ln(s-\sigma)}{\Gamma(1-\kappa_\omega(s,t))}\mathrm{d}\omega$$

$$+ [\alpha(s,v_1(s)) - \alpha(t,v_2(t))] \int_0^1 \frac{(s-\sigma)^{-\kappa_\omega(s,t)}\Gamma'(1-\kappa_\omega(s,t))}{\Gamma(1-\kappa_\omega(s,t))^2}\mathrm{d}\omega,$$

where $\kappa_\omega(s,t) = (1-\omega)\alpha(s, v_1(s)) + \omega\alpha(t, v_2(t))$. Then, using definition (1.4) and the identity above, we obtain the desired inequality by using the Lipschitz continuity

$$|\alpha(s, v_1(s)) - \alpha(t, v_2(t))| \leq C(|s-t| + |v_1(s) - v_2(t)|).$$

$\square$

The main result of this section is the following theorem.

Theorem 3.2. *Let* $1 < p < \frac{4}{d}$, $0 < \gamma < 1 - \frac{d}{4}$ *and* $\alpha_* < \beta < 1$. *Then, under condition* (3.1), *the solution of the nonlinear problem* (1.3) *satisfies*

$$\partial_t u \in L^p(0, T; L^\infty(\Omega)) \cap C^\gamma([0,T]; L^\infty(\Omega)), \tag{3.2}$$

$$u \in W^{1,\infty}(0, T; L^2(\Omega)) \cap L^\infty(0, T; \dot{H}^2(\Omega)), \tag{3.3}$$

$${}^R\partial_t^\beta \partial_t u \in L^1(0, T; L^2(\Omega)), \tag{3.4}$$

where ${}^R\partial_t^\beta \partial_t u$ *denotes the Riemann–Liouville fractional derivative of* $\partial_t u$, *defined in terms of the Laplace transform by* ${}^R\partial_t^\beta \partial_t u = \mathcal{L}_z^{-1}[z^\beta \mathcal{L}[\partial_t u](z)]$.

Proof. To prove the regularity results in (3.2)–(3.4), we denote $w(t) = u(t) - u_0$ and rewrite (1.3) as

$$\begin{cases} \partial_t w(t) - \Delta w(t) = (\Delta u_0 + f(t)) - b(t)\partial_t^{\alpha(t, u(t))} w(t) & \text{in} \quad \Omega \quad \text{for} \quad t \in (0,T], \\ w(t) = 0 & \text{on} \quad \partial\Omega \quad \text{for} \quad t \in (0,T], \\ w(0) = 0 & \text{in} \quad \Omega. \end{cases} \tag{3.5}$$

The solution of (3.5) can be expressed by

$$w(t) = \int_0^t E(t-s)(\Delta u_0 + f(s))\mathrm{d}s - \int_0^t E(t-s)b(s)\partial_s^{\alpha(s, u(s))} w(s)\mathrm{d}s =: I(t) - J(t), \tag{3.6}$$

where $E(t) = e^{t\Delta}$, with $t \geq 0$, is the semigroup of operators generated by the Dirichlet Laplacian, given by

$$E(t)v := \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\theta} e^{zt}(z-\Delta)^{-1}v \, \mathrm{d}z \quad \forall v \in L^1(\Omega), \tag{3.7}$$

where Γ_θ is the contour defined in (2.5). It is well known that the semigroup generated by the Dirichlet Laplacian is bounded and analytic in the sector Σ_θ (see [46]), equivalently the following estimate holds:

$$\|E(t)\|_{L^p(\Omega)\to L^p(\Omega)} + t\|E'(t)\|_{L^p(\Omega)\to L^p(\Omega)} + t\|\Delta E(t)\|_{L^p(\Omega)\to L^p(\Omega)} \leq C, \quad \forall t > 0, \quad \forall 1 \leq p \leq \infty. \tag{3.8}$$

Identity (3.7) implies that $E(t)$ is the inverse Laplace transform of the operator $(z-\Delta)^{-1}$. Therefore,

$${}^R\partial_t^\beta E(t) \text{ is the inverse Laplace transform of } z^\beta(z-\Delta)^{-1}. \tag{3.9}$$

This result is used in the subsequent analysis, which is divided into three parts.

3.1. Part I: proof of (3.2)

Note that $f \in W^{1,1}(0, T; L^2(\Omega)) \hookrightarrow L^{p_0}(0, T; L^2(\Omega))$ for all $1 < p_0 < \infty$. For any $p_0 > \frac{4}{4-d}$ (or equivalently $1 - \frac{d}{4} - \frac{1}{p_0} > 0$), Theorem 2.7 implies that the solution has regularity $u \in W^{1,p_0}(0, T; L^2(\Omega)) \cap L^{p_0}(0, T; \dot{H}^2(\Omega))$,

which is embedded into $C^\gamma([0,T]; L^\infty(\Omega))$ for $0 < \gamma < 1 - \frac{d}{4} - \frac{1}{p_0}$; see Appendix A. Since p_0 can be arbitrarily large, it follows that

$$u \in C^\gamma([0,T]; L^\infty(\Omega)) \quad \text{for any} \quad 0 < \gamma < 1 - \frac{d}{4}. \tag{3.10}$$

To prove $\partial_t u = \partial_t w \in L^p(0,T; L^\infty(\Omega))$ for $1 < p < \frac{4}{d}$, we estimate $\partial_t I(t)$ and $\partial_t J(t)$ separately, which are defined in (3.6).

First, we see that

$$\partial_t I(t) = \partial_t \int_0^t E(s)(\Delta u_0 + f(t-s)) \mathrm{d}s = E(t)(\Delta u_0 + f(0)) + \int_0^t E(t-s) \partial_s f(s) \mathrm{d}s.$$

Since $\|v\|_{L^\infty(\Omega)} \le C \|v\|_{L^2(\Omega)}^{1-\frac{d}{4}} \|v\|_{\dot{H}^2(\Omega)}^{\frac{d}{4}}$ for $d \in \{1,2,3\}$, it follows from (3.7) that

$$\begin{aligned}
\|E(t)v\|_{L^\infty(\Omega)} &\le C \int_{\Gamma_\theta} |e^{tz}| \|(z-\Delta)^{-1}v\|_{L^\infty(\Omega)} |\mathrm{d}z| \\
&\le C \int_{\Gamma_\theta} |e^{tz}| \|(z-\Delta)^{-1}v\|_{L^2(\Omega)}^{1-\frac{d}{4}} \|(z-\Delta)^{-1}v\|_{\dot{H}^2(\Omega)}^{\frac{d}{4}} |\mathrm{d}z| \\
&\le C \int_{\Gamma_\theta} |e^{tz}| (|z|^{-1}\|v\|_{L^2(\Omega)})^{1-\frac{d}{4}} \|v\|_{L^2(\Omega)}^{\frac{d}{4}} |\mathrm{d}z| \qquad \text{(here (2.3) is used)} \\
&\le C \int_{\Gamma_\theta} |e^{tz}| |z|^{-(1-\frac{d}{4})} \|v\|_{L^2(\Omega)} |\mathrm{d}z| \\
&\le C t^{-\frac{d}{4}} \|v\|_{L^2(\Omega)} \quad \forall v \in L^2(\Omega), \qquad \text{(here (2.8) is used)}
\end{aligned}$$

it follows that

$$\|\partial_t I(t)\|_{L^\infty(\Omega)} \le C t^{-\frac{d}{4}} \|\Delta u_0 + f(0)\|_{L^2(\Omega)} + C \int_0^t (t-s)^{-\frac{d}{4}} \|\partial_s f(s)\|_{L^2(\Omega)} \mathrm{d}s,$$

which furthermore implies, for $1 < p < \frac{4}{d}$,

$$\begin{aligned}
\|\partial_t I(t)\|_{L^p(0,T;L^\infty(\Omega))} &\le C \left(\int_0^T t^{-p\frac{d}{4}} \mathrm{d}t \right)^{\frac{1}{p}} \left(\|\Delta u_0 + f(0)\|_{L^2(\Omega)} + \|\partial_s f\|_{L^1(0,T;L^2(\Omega))} \right) \\
&\le C \left(\|\Delta u_0 + f(0)\|_{L^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))} \right). \tag{3.11}
\end{aligned}$$

Second, we consider

$$\begin{aligned}
\partial_t J(t) &= \partial_t \int_0^t E(t-s) b(s) \partial_s^{\alpha(s,u(s))} w(s) \mathrm{d}s \\
&= \int_0^t E'(t-s) b(s) \partial_s^{\alpha(s,u(s))} w(s) \mathrm{d}s + E(0) b(t) \partial_t^{\alpha(t,u(t))} w(t) \\
&= \int_0^t E'(t-s) [b(s) \partial_s^{\alpha(s,u(s))} w(s) - b(t) \partial_s^{\alpha(t,u(t))} w(s)] \mathrm{d}s \\
&\quad + \int_0^t E'(t-s) b(t) \partial_s^{\alpha(t,u(t))} w(s) \mathrm{d}s + b(t) \partial_t^{\alpha(t,u(t))} w(t) \\
&= \int_0^t E'(t-s) [b(s) \partial_s^{\alpha(s,u(s))} w(s) - b(t) \partial_s^{\alpha(t,u(t))} w(s)] \mathrm{d}s \\
&\quad + b(t) \left[\partial_t \int_0^t E(t-s) \partial_s^{\alpha(t_*,u(t_*))} w(s) \mathrm{d}s \right]\bigg|_{t_*=t} \\
&=: J_1(t) + J_2(t). \tag{3.12}
\end{aligned}$$

By using estimate (3.8), we have

$$
\begin{aligned}
\|J_1(t)\|_{L^\infty(\Omega)} & \\
\leq C \int_0^t & (t-s)^{-1} \|b(s) - b(t)\|_{L^\infty(\Omega)} \|\partial_s^{\alpha(s,u(s))} w(s)\|_{L^\infty(\Omega)} \mathrm{d}s \\
+ C \int_0^t & (t-s)^{-1} \|b(t)\|_{L^\infty(\Omega)} \|\partial_s^{\alpha(s,u(s))} w(s) - \partial_s^{\alpha(t,u(t))} w(s)\|_{L^\infty(\Omega)} \mathrm{d}s \\
\leq C \int_0^t & \|\partial_s^{\alpha(s,u(s))} w(s)\|_{L^\infty(\Omega)} \mathrm{d}s \quad \text{(here Lem. 3.1 is used)} \\
+ C \int_0^t & (t-s)^{-1} (|s-t| + \|u(s) - u(t)\|_{L^\infty(\Omega)}) \int_0^s (s-\sigma)^{-\alpha_* - \epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
\leq C \int_0^t \int_0^s & (s-\sigma)^{-\alpha_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
+ C(1 + \|u\|_{C^{\epsilon_*}([0,T];L^\infty(\Omega))}) & \int_0^t (t-s)^{-(1-\epsilon_*)} \int_0^s (s-\sigma)^{-\alpha_* - \epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
\leq C \int_0^t & (t-\sigma)^{-\alpha_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma,
\end{aligned}
\tag{3.13}
$$

where we have used the Hölder regularity $u \in C^{\epsilon_*}([0,T]; L^\infty(\Omega))$ shown in (3.10) in deriving the second to last inequality, and changed in the order of integration in the last inequality. Multiplying the above inequality by $e^{-\lambda t}$, we obtain

$$
\|e^{-\lambda t} J_1(t)\|_{L^\infty(\Omega)} \leq C \int_0^t e^{-\lambda(t-\sigma)} (t-\sigma)^{-\alpha_*} e^{-\lambda\sigma} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma.
$$

Then, by considering the $L^p(0,T)$ norm of the above inequality and using Lemma 2.2, we obtain

$$
\|e^{-\lambda t} J_1\|_{L^p(0,T;L^\infty(\Omega))} \leq C\lambda^{\alpha_* - 1} \|e^{-\lambda t} \partial_t w\|_{L^p(0,T;L^\infty(\Omega))}.
\tag{3.14}
$$

Third, we note that $J_2(t) = b(t) \partial_t J_3(t, t_*)\big|_{t_* = t}$ with

$$
J_3(t, t_*) = \int_0^t E(t-s) \partial_s^{\alpha(t_*, u(t_*))} w(s) \mathrm{d}s.
\tag{3.15}
$$

Let $\phi(s, t_*) = \partial_s^{\alpha(t_*, u(t_*))} w(s)$. Then for a fixed $t_* \in (0,T]$ we can express $\partial_t J_3(t, t_*)$ by using inverse Laplace transform, *i.e.*,

$$
\begin{aligned}
\partial_t J_3(t, t_*) &= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta} e^{zt} (z - \Delta)^{-1} z \, \widehat{\phi}(z, t_*) \, \mathrm{d}z \\
&= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta} e^{zt} (z - \Delta)^{-1} z^{\alpha_* + \epsilon_*} z^{1 - \alpha_* - \epsilon_*} \widehat{\phi}(z, t_*) \, \mathrm{d}z \\
&= \int_0^t F(t-s) \partial_s^{1 - \alpha_* - \epsilon} \phi(s, t_*) \mathrm{d}s,
\end{aligned}
$$

where

$$
F(t) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta} e^{zt} (z - \Delta)^{-1} z^{\alpha_* + \epsilon_*} \, \mathrm{d}z,
$$

which satisfies the following estimate (using (2.8) again):

$$\|F(t)\|_{L^\infty(\Omega)\to L^\infty(\Omega)} = \left\|\frac{1}{2\pi i}\int_{\Gamma_\theta} e^{zt}(z-\Delta)^{-1}z^{\alpha_*+\epsilon_*}\,dz\right\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \le Ct^{-(\alpha_*+\epsilon_*)}.$$

Therefore,

$$e^{-\lambda t}J_2(t) = e^{-\lambda t}b(t)\partial_t J_3(t,t) = b(t)\int_0^t e^{-\lambda(t-s)}F(t-s)[e^{-\lambda s}\partial_s^{1-\alpha_*-\epsilon_*}\phi(s,t)]ds. \tag{3.16}$$

Since

$$
\begin{aligned}
e^{-\lambda s}\|\partial_s^{1-\alpha_*-\epsilon_*}\phi(s,t)\|_{L^\infty(\Omega)} &= e^{-\lambda s}\|\partial_s^{1-\alpha_*-\epsilon_*}[\partial_s^{\alpha(t,u(t))-1}\partial_s w(s)]\|_{L^\infty(\Omega)}\\
&= e^{-\lambda s}\|\partial_s^{\alpha(t,u(t))-\alpha_*-\epsilon_*}\partial_s w(s)\|_{L^\infty(\Omega)}\\
&= e^{-\lambda s}\|K*\partial_s w(s)\|_{L^\infty(\Omega)}\\
&= \|(e^{-\lambda s}K(s))*(e^{-\lambda s}|\partial_t w(s)|)\|_{L^\infty(\Omega)}\\
&\le (e^{-\lambda s}\|K(s)\|_{L^\infty(\Omega)})*(e^{-\lambda s}\|\partial_t w(s)\|_{L^\infty(\Omega)})\\
&\le (e^{-\lambda s}s^{-(1-\epsilon_*)})*(e^{-\lambda s}\|\partial_t w(s)\|_{L^\infty(\Omega)}),
\end{aligned}\tag{3.17}
$$

where we have used the following estimate:

$$K(s) = \frac{s^{\alpha_*+\epsilon_*-\alpha(t,u(t))-1}}{\Gamma(\alpha_*+\epsilon_*-\alpha(t,u(t)))} \le Cs^{-(1-\epsilon_*)},$$

substituting (3.17) into (3.16) yields

$$
\begin{aligned}
e^{-\lambda t}\|J_2(t)\|_{L^\infty(\Omega)} &\le \|b(t)\|_{L^\infty(\Omega)}\int_0^t e^{-\lambda(t-s)}\|F(t-s)\|_{L^\infty(\Omega)\to L^\infty(\Omega)}e^{-\lambda s}\|\partial_s^{1-\alpha_*-\epsilon_*}\phi(s,t)\|_{L^\infty(\Omega)}ds\\
&\le C\int_0^t e^{-\lambda(t-s)}(t-s)^{-(\alpha_*+\epsilon_*)}[(e^{-\lambda s}s^{-(1-\epsilon_*)})*(e^{-\lambda s}\|\partial_t w(s)\|_{L^\infty(\Omega)})]ds.
\end{aligned}
$$

Therefore, by using Lemma 2.2,

$$
\begin{aligned}
\|e^{-\lambda t}J_2\|_{L^p(0,T;L^\infty(\Omega))} &\le C\lambda^{\alpha_*+\epsilon_*-1}\|(e^{-\lambda s}s^{-(1-\epsilon_*)})*(e^{-\lambda s}\|\partial_t w(s)\|_{L^\infty(\Omega)})\|_{L^p(0,T)}\\
&\le C\lambda^{\alpha_*+\epsilon_*-1}\|(e^{-\lambda s}s^{-(1-\epsilon_*)})\|_{L^1(0,T)}\|e^{-\lambda t}\partial_t w\|_{L^p(0,T;L^\infty(\Omega))}\\
&\le C\lambda^{\alpha_*-1}\|e^{-\lambda t}\partial_t w\|_{L^p(0,T;L^\infty(\Omega))}.
\end{aligned}\tag{3.18}
$$

Substituting estimates (3.14) and (3.18) into (3.12), we obtain

$$\|e^{-\lambda t}\partial_t J\|_{L^p(0,T;L^\infty(\Omega))} \le C\lambda^{\alpha_*-1}\|e^{-\lambda t}\partial_t w\|_{L^p(0,T;L^\infty(\Omega))}. \tag{3.19}$$

Then substituting (3.11) and (3.19) into (3.6) yields

$$
\begin{aligned}
&\|e^{-\lambda t}\partial_t w\|_{L^p(0,T;L^\infty(\Omega))}\\
&\le \|e^{-\lambda t}\partial_t I\|_{L^p(0,T;L^\infty(\Omega))} + \|e^{-\lambda t}\partial_t J\|_{L^p(0,T;L^\infty(\Omega))}\\
&\le \|\partial_t I\|_{L^p(0,T;L^\infty(\Omega))} + \|e^{-\lambda t}\partial_t J\|_{L^p(0,T;L^\infty(\Omega))}\\
&\le C\big(\|\Delta u_0 + f(0)\|_{L^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))}\big) + C\lambda^{\alpha_*-1}\|e^{-\lambda t}\partial_t w\|_{L^p(0,T;L^\infty(\Omega))}.
\end{aligned}
$$

By choosing a sufficiently large λ, the last term on the right-hand side can be absorbed by the left-hand side. As a result, we obtain

$$\|\partial_t w\|_{L^p(0,T;L^\infty(\Omega))} \le C\big(\|\Delta u_0\|_{L^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))}\big).$$

This proves (3.2) in view of $\partial_t u = \partial_t w$.

3.2. Part II: proof of (3.3)

By using expressions (3.6) and (3.12), we have

$$\partial_t w(t) = \partial_t I(t) - J_1(t) - J_2(t). \tag{3.20}$$

We use this expression to prove (3.3).

First, by using the semigroup estimate (3.8),

$$\|\partial_t I(t)\|_{L^2(\Omega)} = \left\| E(t)(\Delta u_0 + f(0)) + \int_0^t E(s)\partial_t f(t-s)\mathrm{d}s \right\|_{L^2(\Omega)}$$

$$\leq \|E(t)\|\|\Delta u_0 + f(0)\|_{L^2(\Omega)} + \int_0^t \|E(s)\|\|\partial_t f(t-s)\|_{L^2(\Omega)}\mathrm{d}s$$

$$\leq C\left(\|u_0\|_{\dot{H}^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))} \right). \tag{3.21}$$

Second, similarly as (3.13), by using Lemma 3.1 we have

$$\|J_1(t)\|_{L^2(\Omega)} \leq C\int_0^t (t-s)^{-1}\|b(s)-b(t)\|_{L^\infty(\Omega)}\|\partial_s^{\alpha(s,u(s))}w(s)\|_{L^2(\Omega)}\mathrm{d}s$$

$$+ C\int_0^t (t-s)^{-1}\|b(t)\|_{L^\infty(\Omega)}\|\partial_s^{\alpha(s,u(s))}w(s) - \partial_s^{\alpha(t,u(t))}w(s)\|_{L^2(\Omega)}\mathrm{d}s$$

$$\leq C\int_0^t \int_0^s (s-\sigma)^{-\alpha_*}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma\mathrm{d}s \quad \text{(Lem. 3.1 is used)}$$

$$+ C\int_0^t (t-s)^{-1}(|t-s| + \|u(t)-u(s)\|_{L^\infty(\Omega)})\int_0^s (s-\sigma)^{-\alpha_*-\epsilon_*}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma\mathrm{d}s$$

$$\leq C\int_0^t \int_0^s (s-\sigma)^{-\alpha_*}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma\mathrm{d}s \quad \text{(Hölder regularity (3.10) is used)}$$

$$+ C\int_0^t (t-s)^{-(1-\epsilon_*)}\int_0^s (s-\sigma)^{-\alpha_*-\epsilon_*}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma\mathrm{d}s$$

$$\leq C\int_0^t (t-\sigma)^{-\alpha_*}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma, \tag{3.22}$$

where we have changed the order of integration in the deriving the last inequality.

Third, by using expression $J_2(t) = b(t)\partial_t J_3(t,t_*)\big|_{t_*=t}$ in (3.15) and estimates (3.16) and (3.17) with $\lambda = 0$, we have

$$\|J_2(t)\|_{L^2(\Omega)} \leq \|b(t)\|_{L^\infty(\Omega)}\int_0^t \|F(t-s)\|\|\partial_s^{1-\alpha_*-\epsilon_*}\phi(s,t)\|_{L^2(\Omega)}\mathrm{d}s$$

$$\leq C\int_0^t (t-s)^{-(\alpha_*+\epsilon_*)}\int_0^s (s-\sigma)^{-(1-\epsilon_*)}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma\mathrm{d}s$$

$$\leq C\int_0^t (t-\sigma)^{-\alpha_*}\|\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma, \tag{3.23}$$

where we have changed the order of integration in deriving the last inequality.

Finally, substituting estimates (3.21)–(3.23) into (3.20), we have

$$\|\partial_t w(t)\|_{L^2(\Omega)} \leq C(\|u_0\|_{\dot{H}^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))}) + C\int_0^t (t-\sigma)^{-\alpha_*}\|\partial_s w(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

By applying Gronwall's inequality with weakly singular kernels (see [10], Lem. 1), we can remove the last term on the right-hand side of the above inequality. This yields $\partial_t w \in L^\infty(0, T; L^2(\Omega))$ and therefore

$$\|\partial_t^{\alpha(t,u(t))} w(t)\|_{L^2(\Omega)} \le C \int_0^t (t-s)^{-\alpha_*} \|\partial_t w(s)\|_{L^2(\Omega)} ds \le C.$$

Since both $\partial_t w$ and $\partial_t^{\alpha(t,u(t))} w$ are in $L^\infty(0, T; L^2(\Omega))$, from equation (3.5) we obtain $\Delta w = \partial_t w + b \partial_t^{\alpha(t,u(t))} w - f \in L^\infty(0, T; L^2(\Omega))$, which implies $w \in L^\infty(0, T; \dot{H}^2(\Omega))$. This proves (3.3).

3.3. Part III: proof of (3.4)

By using expression (3.6), we have

$$^R\partial_t^\beta \partial_t w(t) = {}^R\partial_t^\beta \partial_t I(t) - {}^R\partial_t^\beta \partial_t J(t), \tag{3.24}$$

with

$$^R\partial_t^\beta \partial_t I(t) = {}^R\partial_t^\beta E(t)(\Delta u_0 + f(0)) + {}^R\partial_t^\beta \int_0^t E(s)\partial_t f(t-s)ds,$$

$$^R\partial_t^\beta \partial_t J(t) = {}^R\partial_t^\beta \partial_t \int_0^t E(t-s)b(s)\partial_s^{\alpha(s,u(s))} w(s)ds$$

$$= \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^{1+\beta}(z-\Delta)^{-1} \mathcal{L}_s\big[b(s)\partial_s^{\alpha(s,u(s))} w(s)\big](z)dz,$$

where we have used (3.9) in the two identities above.

The two terms in the expression of $^R\partial_t^\beta \partial_t I(t)$ can be estimated by using (3.9) and (2.8) directly:

$$\| {}^R\partial_t^\beta E(t)(\Delta u_0 + f(0))\|_{L^2(\Omega)} = \left\| \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^\beta (z-\Delta)^{-1}(\Delta u_0 + f(0))\,dz \right\|_{L^2(\Omega)}$$

$$\le C \int_{\Gamma_\theta} |e^{zt}||z|^{\beta-1}|dz| \|\Delta u_0 + f(0)\|_{L^2(\Omega)}$$

$$\le C t^{-\beta} \|\Delta u_0 + f(0)\|_{L^2(\Omega)},$$

and

$$\left\| {}^R\partial_t^\beta \int_0^t E(s)\partial_t f(t-s)ds \right\|_{L^2(\Omega)} = \left\| \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^\beta (z-\Delta)^{-1} \mathcal{L}[\partial_t f](z)\,dz \right\|_{L^2(\Omega)}$$

$$= \left\| \int_0^t {}^R\partial_t^\beta E(s)\partial_t f(t-s)ds \right\|_{L^2(\Omega)}$$

$$\le C \int_0^t (t-s)^{-\beta} \|\partial_s f(s)\|_{L^2(\Omega)}ds$$

$$= C t^{-\beta} * \|\partial_t f(t)\|_{L^2(\Omega)},$$

where we have used the following estimate, as a result of (2.8):

$$\| {}^R\partial_t^\beta E(s)\| = \left\| \frac{1}{2\pi i} \int_\Gamma e^{zs} z^\beta (z-\Delta)^{-1}dz \right\| \le C s^{-\beta}.$$

Since $t^{-\beta}$ is integrable in time, it follows that

$$\|t^{-\beta}\|\Delta u_0 + f(0)\|_{L^2(\Omega)}\|_{L^1(0,T)} \leq C\|\Delta u_0 + f(0)\|_{L^2(\Omega)},$$
$$\|t^{-\beta} * \|\partial_t f\|_{L^2(\Omega)}\|_{L^1(0,T)} \leq C\|\partial_t f\|_{L^1(0,T;L^2(\Omega))}.$$

This implies

$$\| {}^R\partial_t^\beta \partial_t I(t)\|_{L^1(0,T;L^2(\Omega))} \leq C\left(\|u_0\|_{\dot{H}^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))}\right). \tag{3.25}$$

To estimate ${}^R\partial_t^\beta \partial_t J(t)$, we decompose it in the following way:

$$
\begin{aligned}
{}^R\partial_t^\beta \partial_t J(t) &= \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^{1+\beta} (z-\Delta)^{-1} \mathcal{L}_s\Big[b(s)\partial_s^{\alpha(s,u(s))} w(s) - b(t)\partial_s^{\alpha(t,u(t))} w(s)\Big](z)\mathrm{d}z \\
&\quad + \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^{1+\beta} (z-\Delta)^{-1} \mathcal{L}_s\Big[b(t)\partial_s^{\alpha(t,u(t))} w(s)\Big](z)\mathrm{d}z \\
&=: \widetilde{J}_1(t) + \widetilde{J}_2(t).
\end{aligned}
$$

If we denote

$$F_2(t) = \frac{1}{2\pi i} \int_{\Gamma_\theta} z^\beta (z-\Delta)^{-1} e^{zt}\mathrm{d}z, \quad \text{with} \quad \|F_2(t)\| \leq Ct^{-\beta} \quad \text{and} \quad \|F_2'(t)\| \leq Ct^{-1-\beta},$$

then

$$
\begin{aligned}
\widetilde{J}_1(t) &= \partial_t \int_0^t F_2(t-s)\big[b(s)\partial_s^{\alpha(s,u(s))} w(s) - b(t)\partial_s^{\alpha(t,u(t))} w(s)\big]\mathrm{d}s \\
&= \partial_t \int_0^t F_2(t-s)(b(s)-b(t))\partial_s^{\alpha(s,u(s))} w(s)\mathrm{d}s \\
&\quad + \partial_t \int_0^t F_2(t-s)b(t)[\partial_s^{\alpha(s,u(s))} w(s) - \partial_s^{\alpha(t,u(t))} w(s)]\mathrm{d}s \\
&= \int_0^t F_2'(t-s)(b(s)-b(t))\partial_s^{\alpha(s,u(s))} w(s)\mathrm{d}s \\
&\quad - \partial_t b(t) \int_0^t F_2(t-s)\partial_s^{\alpha(t,u(t))} w(s)\mathrm{d}s \\
&\quad + b(t) \int_0^t F_2'(t-s)[\partial_s^{\alpha(s,u(s))} w(s) - \partial_s^{\alpha(t,u(t))} w(s)]\mathrm{d}s \\
&\quad - b(t) \int_0^t F_2(t-s)\partial_t\partial_s^{\alpha(t,u(t))} w(s)\mathrm{d}s \\
\widetilde{J}_2(t) &= \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^{1+\beta} (z-\Delta)^{-1} b(t) z^{\alpha(t,u(t))} \mathcal{L}_s\big[w(s)\big](z)\mathrm{d}z \\
&= \frac{1}{2\pi i} \int_{\Gamma_\theta} e^{zt} z^\beta (z-\Delta)^{-1} b(t) z^{\alpha(t,u(t))} \mathcal{L}_s\big[\partial_s w(s)\big](z)\mathrm{d}z \\
&= b(t) \int_0^t F_2(t-s) {}^R\partial_s^{\alpha(t,u(t))} \partial_s w(s)\mathrm{d}s,
\end{aligned}
$$

where we have used the identity $\mathcal{L}_s[{}^R\partial_s^{\alpha(t,u(t))}\partial_s w(s)] = z^{\alpha(t,u(t))}\mathcal{L}_s[\partial_s w(s)]$.

By using Lemma 3.1, we have

$$
\begin{aligned}
\|\widetilde{J}_1(t)&\|_{L^2(\Omega)} \\
&\le C \int_0^t (t-s)^{-1-\beta} \|b(s)-b(t)\|_{L^2(\Omega)} \int_0^s (s-\sigma)^{-\alpha_*-\epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
&\quad + C \int_0^t (t-s)^{-\beta} \int_0^s (s-\sigma)^{-\alpha_*-\epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
&\quad + C \int_0^t (t-s)^{-1-\beta} (|t-s| + \|u(s)-u(t)\|_{L^2(\Omega)}) \int_0^s (s-\sigma)^{-\alpha_*-\epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
&\quad + C \int_0^t (t-s)^{-\beta} (1+\|\partial_t u\|_{L^2(\Omega)}) \int_0^s (s-\sigma)^{-\alpha_*-\epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
&\le C \int_0^t (t-\sigma)^{1-\beta-\alpha_*-\epsilon_*} \|\partial_\sigma w(\sigma)\|_{L^\infty(\Omega)} \mathrm{d}\sigma
\end{aligned}
$$

where we have used the regularity $\partial_t u \in C([0,T]; L^2(\Omega))$ proved in Part II of the proof, and changed the order of integration in deriving the last inequality. Since $1 - \beta - \alpha_* - \epsilon_* > -1$, it follows that $t^{1-\beta-\alpha_*-\epsilon_*}$ is integrable in time. Therefore,

$$
\|\widetilde{J}_1\|_{L^1(0,T;L^2(\Omega))} \le C \|t^{1-\beta-\alpha_*-\epsilon_*}\|_{L^1(0,T)} \|\partial_t w\|_{L^1(0,T;L^\infty(\Omega))} \le C, \tag{3.26}
$$

where we have used the regularity (3.2) proved in Part I of the proof.

Since $\beta > \alpha_*$, it follows that $\alpha(t,u(t)) - \beta \le \alpha_* - \beta < 0$ and therefore

$$
\begin{aligned}
\|\widetilde{J}_2(t)\|_{L^2(\Omega)} &= \int_0^t F_2(t-s) \, {}^R\partial_s^{\alpha(t,u(t))} \partial_s w(s) \mathrm{d}s \\
&\le C \int_0^t (t-s)^{-\beta} \| {}^R\partial_s^{\alpha(t,u(t))-\beta} \, {}^R\partial_s^\beta \partial_s w(s)\|_{L^2(\Omega)} \mathrm{d}s \\
&\le C \int_0^t (t-s)^{-\beta} \int_0^s (s-\sigma)^{-(1+\alpha_*-\beta)} \| {}^R\partial_\sigma^\beta \partial_\sigma w(\sigma)\|_{L^2(\Omega)} \mathrm{d}\sigma \mathrm{d}s \\
&\le C \int_0^t (t-\sigma)^{-\alpha_*} \| {}^R\partial_\sigma^\beta \partial_\sigma w(\sigma)\|_{L^2(\Omega)} \mathrm{d}\sigma,
\end{aligned}
$$

where we have changed the order of integration in deriving the last inequality.

Since ${}^R\partial_t^\beta \partial_t J(t) = \widetilde{J}_1(t) + \widetilde{J}_2(t)$, substituting the estimate of $\widetilde{J}_2(t)$ into (3.24) yields

$$
\begin{aligned}
\| {}^R\partial_t^\beta \partial_t w(t)\|_{L^2(\Omega)} &\le \| {}^R\partial_t^\beta \partial_t I(t)\|_{L^2(\Omega)} + \| {}^R\partial_t^\beta \partial_t J(t)\|_{L^2(\Omega)} \\
&\le \| {}^R\partial_t^\beta \partial_t I(t)\|_{L^2(\Omega)} + \|\widetilde{J}_1(t)\|_{L^2(\Omega)} \\
&\quad + C \int_0^t (t-\sigma)^{-\alpha_*} \| {}^R\partial_\sigma^\beta \partial_\sigma w(\sigma)\|_{L^2(\Omega)} \mathrm{d}\sigma.
\end{aligned}
$$

Multiplying the inequality above by $e^{-\lambda t}$ and considering the $L^1(0,T)$ norm, we obtain

$$
\begin{aligned}
\|e^{-\lambda t}\|^R\partial_t^\beta\partial_t w\|_{L^2(\Omega)}\|_{L^1(0,T)} \\
\leq \|\,^R\partial_t^\beta\partial_t I\|_{L^1(0,T;L^2(\Omega))} + \|\widetilde{J}_1\|_{L^1(0,T;L^2(\Omega))} \\
+ C\left\|\int_0^t e^{-\lambda(t-\sigma)}(t-\sigma)^{-\alpha_*}e^{-\lambda\sigma}\|^R\partial_\sigma^\beta\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma\right\|_{L^1(0,T)} \\
\leq C + C\left(\int_0^T e^{-\lambda t}t^{-\alpha_*}\mathrm{d}t\right)\int_0^T e^{-\lambda\sigma}\|^R\partial_\sigma^\beta\partial_\sigma w(\sigma)\|_{L^2(\Omega)}\mathrm{d}\sigma \\
\leq C + C\lambda^{\alpha_*-1}\|e^{-\lambda t}\|^R\partial_t^\beta\partial_t w\|_{L^2(\Omega)}\|_{L^1(0,T)},
\end{aligned}
$$

where we have used (3.25) and (3.26), and Lemma 2.2, in deriving the second to last inequality. By choosing a sufficiently large λ, the second term on the right-hand side can be absorbed by the left-hand side. Then we obtain (3.4).

The proof of Theorem 3.2 is complete.   □

4. Numerical approximation

In this section, we propose a linearly implicit time-stepping method for discretizing (1.3) in time, and present an error estimate based on the regularity results proved in Theorem 3.2.

4.1. A linearly implicit method

Let $t_n = n\tau$, $n = 0,1,\dots,N$, be a uniform partition of the time interval $[0,T]$ with stepsize $\tau = T/N$. We consider the following time-stepping method: for given functions $u_j \in \dot{H}^2(\Omega)$, $j = 0,1,\dots,n-1$, find $u_n \in \dot{H}^2(\Omega)$ satisfying the equation

$$
\bar{\partial}_\tau u_n(x) + b(x,t_n)\bar{\partial}_\tau^{\alpha(x,t_n,u_{n-1})}(u_n(x) - u_0(x)) - \Delta u_n(x) = f_n(x), \tag{4.1}
$$

where $f_n(x) = f(x,t_n)$, $\bar{\partial}_\tau u_n(x) = (u_n(x) - u_{n-1}(x))/\tau$, and $\bar{\partial}_\tau^{\alpha(x,t_n,u_{n-1})}(u_n(x) - u_0(x))$ denotes the CQ generated by the backward Euler method, *i.e.*,

$$
\bar{\partial}_\tau^{\alpha(x,t_n,u_{n-1})}(u_n(x) - u_0(x)) = \sum_{j=0}^n \kappa_j(x,t_n,u_{n-1})(u_{n-j}(x) - u_0(x)), \tag{4.2}
$$

with $\kappa_j(x,t_n,u_{n-1})$ being the coefficients in the power series expansion

$$
\left(\frac{1-\zeta}{\tau}\right)^{\alpha(x,t_n,u_{n-1})} = \sum_{j=0}^\infty \kappa_j(x,t_n,u_{n-1})\zeta^j \quad \text{for} \quad \zeta \in \mathbb{C} \quad \text{such that} \quad |\zeta| < 1. \tag{4.3}
$$

We have used an explicit scheme in the nonlinear fractional order to linearize the equation. For abbreviation, we denote

$$
\alpha(t_n,u_{n-1}) = \alpha(\cdot,t_n,u_{n-1}) \quad \text{and} \quad \kappa_j(t_n,u_{n-1}) = \kappa_j(\cdot,t_n,u_{n-1})
$$

in the following error analysis.

By using Cauchy's integral formula, the coefficients in the power series expansion (4.3) can be expressed as

$$\kappa_j(t_n, u_{n-1}) = \frac{1}{2\pi i} \int_{|\zeta|=1} \left(\frac{1-\zeta}{\tau}\right)^{\alpha(t_n, u_{n-1})} \zeta^{-j-1} \mathrm{d}\zeta$$

$$= \frac{\tau}{2\pi i} \int_{\substack{z=i\xi \\ |\xi| \le \frac{\pi}{\tau}}} \left(\frac{1-e^{-\tau z}}{\tau}\right)^{\alpha(t_n, u_{n-1})} e^{t_j z} \mathrm{d}z \quad \text{(change of variable } \zeta = e^{-\tau z})$$

$$= \frac{\tau}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1-e^{-\tau z}}{\tau}\right)^{\alpha(t_n, u_{n-1})} e^{t_j z} \mathrm{d}z,$$

where we have deformed the contour of integration in the last equality to

$$\Gamma_\theta^\tau = \{z \in \Gamma_\theta : |\mathrm{Im}(z)| \le \pi/\tau\}, \tag{4.4}$$

which is a finite part of the contour Γ_θ defined in (2.5). This technique (of deforming the contour to the left half of the complex plane) is often used in analysis of CQ in the literature; see [30, 32]. Similarly as (2.8), contour integrals on the finite part Γ_θ^τ satisfy the following estimates:

$$\int_{\Gamma_\theta^\tau} |z|^{\gamma-1} |e^{tz}| \, |\mathrm{d}z| \le C(\tau+t)^{-\gamma} \quad \text{and} \quad \left\| \int_{\Gamma_\theta^\tau} z^\gamma (z-\Delta)^{-1} e^{tz} \, \mathrm{d}z \right\| \le C(\tau+t)^{-\gamma}. \tag{4.5}$$

The extra τ in $(\tau+t)^{-\gamma}$ can be obtained by using the property $|e^{tz}| \le C|e^{(\tau+t)z}|$, which holds because $\tau|z| \le \pi/\sin(\theta)$ on the contour Γ_θ^τ. We shall frequently use these estimates in the subsequent error analysis.

4.2. Consistency of CQ

The consistency error of CQ has been studied in the literature for sufficiently smooth function. Here we present analysis for the consistency error based on the regularity results proved in Theorem 3.2.

Lemma 4.1. *Under condition* (3.1), *if* $v(0) = 0$ *and* $^R\partial_t^\beta \partial_t v \in L^1(0, T; L^2(\Omega))$ *for some* $\alpha_* < \beta < 1$, *then*

$$\tau \sum_{n=1}^N \left\| \partial_t^{\alpha(t_n, u(t_n))} v(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_n) \right\|_{L^2(\Omega)} \le C\tau \|^R\partial_t^\beta \partial_t v\|_{L^1(0,T;L^2(\Omega))},$$

where $u(t)$ is the solution of the continuous problem (1.3).

Proof. First, we define $K_{\tau,\epsilon}^n(t) = \sum_{j=0}^\infty \kappa_j(t_n, u(t_n)) \delta_{t_j+\epsilon}(t)$ with $0 < \epsilon < \tau$, where $\delta_{t_j+\epsilon}(t)$ is the delta function concentrated at the point $t_j + \epsilon$. Then

$$\bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_n) = \bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_m)|_{t_m=t_n}$$

and

$$\bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_m) = \sum_{j=0}^m \kappa_j(t_n, u(t_n)) v(t_{m-j}) = \lim_{\epsilon \to 0} \int_0^{t_m} K_{\tau,\epsilon}^n(s) v(t_m - s) \mathrm{d}s$$

$$= \lim_{\epsilon \to 0} K_{\tau,\epsilon}^n * v(t_m)$$

$$= \lim_{\epsilon \to 0} (K_{\tau,\epsilon}^n * 1) * \partial_t v(t_m) \quad \text{for} \quad m \ge 1,$$

where we have used the Newton–Leibnitz formula $v(t) = v(0) + \int_0^t \partial_t v(s) \mathrm{d}s = (1 * \partial_t v)(t)$ (when $v(0) = 0$). The Laplace transform of the kernel $\lim_{\epsilon \to 0} K_{\tau,\epsilon}^n * 1$ is

$$\mathcal{L}[\lim_{\epsilon \to 0} K_{\tau,\epsilon}^n * 1](z) = \mathcal{L}[\lim_{\epsilon \to 0} K_{\tau,\epsilon}^n](z)\mathcal{L}[1](z) = \left(\sum_{j=0}^{\infty} \kappa_j(t_n, u(t_n))e^{-j\tau z} \right) z^{-1}$$

$$= \left(\frac{1 - e^{-\tau z}}{\tau} \right)^{\alpha(t_n, u(t_n))} z^{-1}, \tag{4.6}$$

where we have used the definition of κ_j in (4.3).

Second, we note that

$$\partial_t^{\alpha(t_n, u(t_n))} v(t_m) = \int_0^{t_m} E^n(t_m - s)\partial_t v(s)\mathrm{d}s = E^n * \partial_t v(t_m),$$

with

$$E^n(t) = \frac{1}{\Gamma(1 - \alpha(u(t_n)))} t^{-\alpha(t_n, u(t_n))} = \frac{1}{2\pi i} \int_{\Gamma_\theta} z^{\alpha(t_n, u(t_n)) - 1} e^{tz} \mathrm{d}z,$$

$$^R\partial_t^{1-\beta} E^n(t) = \mathcal{L}_z^{-1}[z^{1-\beta} \mathcal{L}E^n(z)] = \frac{1}{2\pi i} \int_{\Gamma_\theta} z^{\alpha(t_n, u(t_n)) - \beta} e^{tz} \mathrm{d}z. \tag{4.7}$$

Therefore,

$$\bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_m) - \partial_t^{\alpha(t_n, u(t_n))} v(t_m) = (\lim_{\epsilon \to 0} K_{\tau,\epsilon}^n * 1 - E^n) * \partial_t v(t_m)$$

$$= \mathcal{L}_z^{-1}\left[\mathcal{L}[\lim_{\epsilon \to 0} K_{\tau,\epsilon}^n * 1 - E^n](z) \, \mathcal{L}[\partial_t v](z) \right](t_m)$$

$$= \mathcal{L}_z^{-1}\left[z^{-\beta} \mathcal{L}[K_{\tau,\epsilon}^n * 1 - E^n](z) \, z^{\beta} \mathcal{L}[\partial_t v](z) \right](t_m)$$

$$= F * (^R\partial_t^\beta \partial_t v)(t_m), \tag{4.8}$$

with

$$F(t) = \mathcal{L}_z^{-1}\left[z^{-\beta} \mathcal{L}[K_{\tau,\epsilon}^n * 1 - E^n](z) \right](t)$$

$$= \frac{1}{2\pi i} \int_{\Gamma_\theta} \left[\left(\frac{1 - e^{-\tau z}}{\tau} \right)^{\alpha(t_n, u(t_n))} - z^{\alpha(t_n, u(t_n))} \right] z^{-\beta - 1} e^{tz} \mathrm{d}z,$$

where we have used (4.6). In estimating $F(t)$, we consider the two cases $t \geq \tau$ and $t \in (0, \tau)$ separately.

If $t \geq \tau$ then

$$|F(t)| \leq \int_{\Gamma_\theta^\tau} C\tau |z|^{\alpha(t_n, u(t_n)) - \beta} e^{t\mathrm{Re}(z)} |\mathrm{d}z| + \int_{\Gamma_\theta \backslash \Gamma_\theta^\tau} C\tau |z|^{\alpha(t_n, u(t_n)) - \beta} e^{(t - \alpha_* \tau)\mathrm{Re}(z)} |\mathrm{d}z|$$

$$\leq C\tau t^{\beta - \alpha(t_n, u(t_n)) - 1}.$$

If $t \in (0, \tau)$ then

$$|F(t)| = |^R\partial_t^{-\beta}(K_{\tau,\epsilon}^n * 1) - {}^R\partial_t^{-\beta} E^n|$$

$$= \left| \frac{1}{\Gamma(1 + \beta)} \tau^{-\alpha(t_n, u(t_n))} t^\beta - \frac{1}{\Gamma(1 + \beta - \alpha(t_n, u(t_n)))} t^{\beta - \alpha(t_n, u(t_n))} \right|$$

$$\leq C\tau^{\beta - \alpha(t_n, u(t_n))}$$

$$\leq C\tau(\tau + t)^{\beta - \alpha(t_n, u(t_n))}.$$

Overall, combining the two cases $t \geq \tau$ and $t \in (0, \tau)$, we have

$$|F(t)| \leq C\tau(\tau + t)^{\beta - \alpha(t_n, u(t_n)) - 1} \leq C\tau(\tau + t)^{\beta - \alpha_* - 1}.$$

Using this estimate in (4.8), we have

$$\|\bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_n) - \partial_t^{\alpha(t_n, u(t_n))} v(t_n)\|_{L^2(\Omega)} = \|F * ({}^R\partial_t^\beta \partial_t v)(t_n)\|_{L^2(\Omega)}$$

$$\leq \int_0^{t_n} \|F(t_n - s)\|_{L^\infty(\Omega)} \|{}^R\partial_t^\beta v(s)\|_{L^2(\Omega)}$$

$$\leq \int_0^{t_n} C\tau(t_{n+1} - s)^{\beta - \alpha_* - 1} \|{}^R\partial_t^\beta v(s)\|_{L^2(\Omega)}.$$

Summing up the inequality above for $n = 1, \ldots, N$, we obtain

$$\tau \sum_{n=1}^N \|\bar{\partial}_\tau^{\alpha(t_n, u(t_n))} v(t_n) - \partial_t^{\alpha(t_n, u(t_n))} v(t_n)\|_{L^2(\Omega)}$$

$$\leq C\tau \left[\tau \sum_{n=0}^N t_{n+1}^{\beta - \alpha_* - 1} \right] \|{}^R\partial_t^\beta v(s)\|_{L^1(0, T; L^2(\Omega))} \leq C\tau \|{}^R\partial_t^\beta v(s)\|_{L^1(0, T; L^2(\Omega))}.$$

This proves the desired estimate in Lemma 4.1.   □

4.3. Convergence of numerical solutions

In Theorem 3.2, we have proved the regularity ${}^R\partial_t^\beta \partial_t u \in L^1(0, T; L^2(\Omega))$ for $\alpha_* < \beta < 1$, which together with Lemma 4.1 imply

$$\tau \sum_{n=1}^N \left\| \partial_t^{\alpha(t_n, u(t_n))} u(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_n))} (u(t_n) - u_0) \right\|_{L^2(\Omega)} \leq C\tau. \tag{4.9}$$

This result is used to prove convergence of numerical solutions in this subsection.

Theorem 4.2 (Error estimate)**.** *Under assumptions* (3.1)*, there exists a positive constant* τ_0*, such that when* $\tau \leq \tau_0$*, the discrete solution given by* (4.1) *satisfies the following error estimate:*

$$\max_{1 \leq n \leq N} \|u(t_n) - u_n\|_{L^2(\Omega)} \leq C\tau.$$

Proof. Let $w_n = u_n - u_0$ and rewrite (4.1) as

$$\bar{\partial}_\tau w_n - \Delta w_n = (\Delta u_0 + f_n) - b(t_n) \bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} w_n. \tag{4.10}$$

Then the solution of (4.10) can be expressed as

$$w_m = \tau \sum_{n=1}^m E_{m-n} (\Delta u_0 + f_n) - \tau \sum_{n=1}^m E_{m-n} b(t_n) \bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} w_n =: I_m - J_m, \tag{4.11}$$

where

$$E_n = \frac{1}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta \right)^{-1} e^{t_n z} dz, \tag{4.12}$$

which satisfies the following standard estimate (as a result of (4.5)):

$$\|E_n\|_{L^2(\Omega)\to L^2(\Omega)} \le C \quad \forall n = 1, 2, \dots \tag{4.13}$$

Similarly, let $w(t) = u(t) - u_0$ and rewrite (1.3) as (3.5). The exact solution of (3.5) can be expressed as (3.6). The difference between (4.11) and (3.6) yields

$$u_m - u(t_m) = w_m - w(t_m) = (I_m - I(t_m)) - (J_m - J(t_m)). \tag{4.14}$$

In the following, we estimate the two part of the error separately. The proof is divided into two parts.

Part I: Estimation for $\|u_m - u(t_m)\|_{L^2(\Omega)}$ in terms of $\tau \sum_{j=1}^{m} \|\bar{\partial}_\tau^\beta (u_j - u(t_j))\|_{L^2(\Omega)}$.

We consider the L^2 norm of the error in (4.14):

$$\|u_m - u(t_m)\|_{L^2(\Omega)} \le \|I_m - I(t_m)\|_{L^2(\Omega)} + \|J_m - J(t_m)\|_{L^2(\Omega)}. \tag{4.15}$$

Since $I_m - I(t_m)$ corresponds to the error of discretizing the heat equation, we have the following standard estimate (*cf.* Lem. C.1 in Appendix C with $\beta = 0$):

$$\|I_m - I(t_m)\|_{L^2(\Omega)} \le C\tau \left(\|u_0\|_{\dot{H}^2} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))} \right). \tag{4.16}$$

To estimate $\|J_m - J(t_m)\|_{L^2(\Omega)}$, we denote

$$g_n = b(t_n)\bar{\partial}_\tau^{\alpha(t_n,u_{n-1})} w_n \quad \text{and} \quad g(t) = b(t)\partial_t^{\alpha(t,u(t))} w(t).$$

Then

$$
\begin{aligned}
J_m &- J(t_m) \\
&= \tau \sum_{n=1}^{m} E_{m-n}(g_n - g(t_n)) + \left[\tau \sum_{n=1}^{m} E_{m-n}g(t_n) - \int_0^{t_m} E(t_m - t)g(t)\mathrm{d}t \right] \\
&=: \mathcal{E}_{m,1} + \mathcal{E}_{m,2}.
\end{aligned}
\tag{4.17}
$$

Since $\mathcal{E}_{m,2}$ can be viewed as the error of discretizing a heat equation with the source term $g(t)$, the following standard error estimate holds similarly as (4.16):

$$\|\mathcal{E}_{m,2}\|_{L^2(\Omega)} \le C\tau\|g\|_{W^{1,1}(0,T;L^2(\Omega))}.$$

Since $\|g\|_{W^{1,1}(0,T;L^2(\Omega))} \le C\|\partial_t\partial_t^{\alpha(t,u(t))}w(t)\|_{L^1(0,T;L^2(\Omega))}$ and

$$
\begin{aligned}
\partial_t &\partial_t^{\alpha(t,u(t))} w(t) \\
&= (\partial_t\alpha(t,u) + \partial_t u\partial_u\alpha(t,u))\frac{\Gamma'(1 - \alpha(t,u))}{\Gamma(1 - \alpha(t,u(t)))^2} \int_0^t s^{-\alpha(t,u(t))}\partial_t w(t - s)\mathrm{d}s \\
&\quad - \frac{\partial_t\alpha(t,u) + \partial_t u\partial_u\alpha(t,u)}{\Gamma(1 - \alpha(t,u))} \int_0^t s^{-\alpha(t,u(t))} \ln(s)\partial_t w(t - s)\mathrm{d}s + {}^R\partial_t^{\alpha(t,u(t))}\partial_t w(t),
\end{aligned}
$$

by the regularity results in Theorem 3.2 we see that $\partial_t\partial_t^{\alpha(t,u(t))}w(t) \in L^1(0,T;L^2(\Omega))$, and therefore

$$g \in W^{1,1}(0,T;L^2(\Omega)), \tag{4.18}$$

which furthermore implies

$$\|\mathcal{E}_{m,2}\|_{L^2(\Omega)} \le C\tau. \tag{4.19}$$

Substituting (4.16)–(4.19) into (4.15), we have

$$\|u_m - u(t_m)\|_{L^2(\Omega)} \le C\tau + \|\mathcal{E}_{m,1}\|_{L^2(\Omega)}. \tag{4.20}$$

It remains to estimate $\|\mathcal{E}_{m,1}\|_{L^2(\Omega)}$. To this end, we consider the decomposition

$$
\begin{aligned}
\mathcal{E}_{m,1} &= \tau \sum_{n=1}^m E_{m-n}(g_n - g_\tau(t_n)) + \tau \sum_{n=1}^m E_{m-n}(g_\tau(t_n) - \widetilde{g}_\tau(t_n)) \\
&\quad + \tau \sum_{n=1}^m E_{m-n}(\widetilde{g}_\tau(t_n) - g_\tau^*(t_n)) + \tau \sum_{n=1}^m E_{m-n}(g_\tau^*(t_n) - g(t_n)) \\
&= \mathcal{E}_{m,3} + \mathcal{E}_{m,4} + \mathcal{E}_{m,5} + \mathcal{E}_{m,6},
\end{aligned} \tag{4.21}
$$

where

$$
\begin{aligned}
g_\tau(t_n) &= b(t_n)\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} w(t_n), \quad \widetilde{g}_\tau(t_n) = b(t_n)\bar{\partial}_\tau^{\alpha(t_n, u(t_{n-1}))} w(t_n), \\
g_\tau^*(t_n) &= b(t_n)\bar{\partial}_\tau^{\alpha(t_n, u(t_n))} w(t_n), \quad g(t_n) = b(t_n)\partial_t^{\alpha(t_n, u(t_n))} w(t_n).
\end{aligned}
$$

First, Lemma 4.1 implies $\tau \sum_{n=1}^N \|g_\tau^*(t_n) - g(t_n)\|_{L^2(\Omega)} \le C\tau$, which together with (4.13) imply the following result:

$$\|\mathcal{E}_{m,6}\|_{L^2(\Omega)} \le C\tau \sum_{n=1}^m \|g_\tau^*(t_n) - g(t_n)\|_{L^2(\Omega)} \le C\tau. \tag{4.22}$$

Second, for any bounded sequence v_n we associate it with a function $\widetilde{v}(\zeta) := \sum_{n=0}^\infty v_n \zeta^n$, which is analytic in the unit disk $|\zeta| < 1$ and is called the generating function of the sequence. The sequence v_n can be determinedThe generating function of the sequence $\bar{\partial}_\tau^{\alpha(t_n, v_1)} v_n$ is simply $\delta(\zeta)^{\alpha(t_n, v_1)} \widetilde{v}(\zeta)$. By expressing $\bar{\partial}_\tau^{\alpha(t_n, v_1)} w(t_n)$ and $\bar{\partial}_\tau^{\alpha(t_n, v_2)} w(t_n)$ in terms of the Cauchy integral formula, with the notation $\delta(\zeta) = \frac{1-\zeta}{\tau}$ we have

$$
\begin{aligned}
&\bar{\partial}_\tau^{\alpha(t_n, v_1)} w(t_n) - \bar{\partial}_\tau^{\alpha(t_n, v_2)} w(t_n) \\
&= \frac{1}{2\pi i} \int_{|\zeta|=1} \big(\delta(\zeta)^{\alpha(t_n, v_1)} - \delta(\zeta)^{\alpha(t_n, v_2)}\big) \widetilde{w}(\zeta) \zeta^{-n-1} \mathrm{d}\zeta \\
&= \big[\alpha(t_n, v_1) - \alpha(t_n, v_2)\big] \int_0^1 \frac{1}{2\pi i} \int_{|\zeta|=1} \big[\delta(\zeta)^{(1-\theta)\alpha(t_n, v_1) + \theta\alpha(t_n, v_2) - 1} \ln\delta(\zeta)\big] \frac{\delta(\zeta)\widetilde{w}(\zeta)}{\zeta^{n+1}} \mathrm{d}\zeta \mathrm{d}\theta \\
&= \big[\alpha(t_n, v_1) - \alpha(t_n, v_2)\big] \int_0^1 \sum_{j=0}^n d_{\theta,j} \bar{\partial}_\tau w(t_{n-j}) \mathrm{d}\theta
\end{aligned} \tag{4.23}
$$

with

$$
\begin{aligned}
|d_{\theta,j}| &= \left| \frac{1}{2\pi i} \int_{|\zeta|=1} \big[\delta(\zeta)^{(1-\theta)\alpha(t_n, v_1) + \theta\alpha(t_n, v_2) - 1} \ln\delta(\zeta)\big] \zeta^{-j-1} \mathrm{d}\zeta \right| \\
&= \left| \frac{\tau}{2\pi i} \int_{\Gamma_\theta^\tau} \big[\delta(e^{-z\tau})^{(1-\theta)\alpha(t_n, v_1) + \theta\alpha(t_n, v_2) - 1} \ln\delta(e^{-z\tau})\big] e^{zt_j} \mathrm{d}z \right| \\
&\le C\tau \int_{\Gamma_\theta^\tau} (|z|^{(1-\theta)\alpha(t_n, v_1) + \theta\alpha(t_n, v_2) - 1} \ln|z| e^{-C|z|t_j} |\mathrm{d}z| \\
&\le C\tau t_{j+1}^{-\alpha_* - \epsilon_*}, \quad \text{(here we have used (4.5))}
\end{aligned} \tag{4.24}
$$

which implies that

$$\|\bar{\partial}_\tau^{\alpha(t_n, u(t_{n-1}))} w(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_n))} w(t_n)\|_{L^2(\Omega)}$$

$$\leq C\|\alpha(t_n, u(t_{n-1})) - \alpha(t_n, u(t_n))\|_{L^2(\Omega)} \sum_{j=0}^{n} \tau t_{j+1}^{-\alpha_* - \epsilon_*} \|\bar{\partial}_\tau w(t_{n-j})\|_{L^\infty(\Omega)} \qquad (4.25)$$

$$\leq C\|u(t_{n-1}) - u(t_n)\|_{L^2(\Omega)} \sum_{j=0}^{n} \tau t_{j+1}^{-\alpha_* - \epsilon_*} \|\bar{\partial}_\tau w(t_{n-j})\|_{L^\infty(\Omega)}$$

$$\leq C\tau \sum_{j=0}^{n} \tau t_{j+1}^{-\alpha_* - \epsilon_*} \|\bar{\partial}_\tau w(t_{n-j})\|_{L^\infty(\Omega)},$$

where we have used $\partial_t u \in L^\infty(0, T; L^2(\Omega))$ (proved in Thm. 3.2) in deriving the last inequality. Therefore,

$$\tau \sum_{n=1}^{N} \|\widetilde{g}_\tau(t_n) - g_\tau^*(t_n)\|_{L^2(\Omega)} \leq C\tau \sum_{n=0}^{N} \tau t_{n+1}^{-\alpha_* - \epsilon_*} \tau \sum_{n=1}^{N} \|\bar{\partial}_\tau w(t_n)\|_{L^\infty(\Omega)}$$

$$\leq C\tau \sum_{n=0}^{N} \tau t_{n+1}^{-\alpha_* - \epsilon_*} \|\partial_t w\|_{L^1(0,T;L^\infty(\Omega))}$$

$$\leq C\tau,$$

where we have used $\partial_t w = \partial_t u \in L^p(0, T; L^\infty(\Omega))$, which is proved in Theorem 3.2. Hence,

$$\|\mathcal{E}_{m,5}\|_{L^2(\Omega)} \leq C\tau \sum_{n=1}^{N} \|\widetilde{g}_\tau(t_n) - g_\tau^*(t_n)\|_{L^2(\Omega)} \leq C\tau. \qquad (4.26)$$

Third, similar as the estimation for $\|\bar{\partial}_\tau^{\alpha(t_n, u(t_{n-1}))} w(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_n))} w(t_n)\|_{L^2(\Omega)}$ in (4.25), by using estimates (4.23) and (4.24) we obtain

$$\|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} w(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_{n-1}))} w(t_n)\|_{L^2(\Omega)}$$

$$\leq C\|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)} \sum_{j=0}^{n} \tau t_{j+1}^{-\alpha_* - \epsilon_*} \|\bar{\partial}_\tau w(t_{n-j})\|_{L^\infty(\Omega)}. \qquad (4.27)$$

For $1 < p < \frac{4}{d}$, we denote by $1 < p' < \infty$ the number satisfying $1 = \frac{1}{p} + \frac{1}{p'}$. Then, by using Hölder's inequality, we have

$$\tau \sum_{n=1}^{m} \|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} w(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_{n-1}))} w(t_n)\|_{L^2(\Omega)}$$

$$\leq C\|(u_{n-1} - u(t_{n-1}))_{n=1}^{m}\|_{L^{p'}(L^2(\Omega))} \left\|\left(\sum_{j=0}^{n} \tau t_{j+1}^{-\alpha_* - \epsilon_*} \|\bar{\partial}_\tau w(t_{n-j})\|_{L^\infty(\Omega)}\right)_{n=1}^{m}\right\|_{L^p}$$

$$\leq C\|(u_{n-1} - u(t_{n-1}))_{n=1}^{m}\|_{L^{p'}(L^2(\Omega))} \|\bar{\partial}_\tau w(t_n)_{n=1}^{m}\|_{L^p(L^\infty(\Omega))}$$

$$\leq C\|(u_{n-1} - u(t_{n-1}))_{n=1}^{m}\|_{L^{p'}(L^2(\Omega))} \|\partial_t w\|_{L^p(0,T;L^\infty(\Omega))}$$

$$(\partial_t w = \partial_t u \in L^p(0, T; L^\infty(\Omega)) \text{ is proved in Thm. 3.2})$$

$$\leq C\|(u_{n-1} - u(t_{n-1}))_{n=1}^{m}\|_{L^{p'}(L^2(\Omega))}$$

$$\leq C_\epsilon \|(u_{n-1} - u(t_{n-1}))_{n=1}^{m}\|_{L^1(L^2(\Omega))} + \epsilon \|(u_{n-1} - u(t_{n-1}))_{n=1}^{m}\|_{L^\infty(L^2(\Omega))}$$

$$= C_\epsilon \tau \sum_{n=1}^{m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)} + \epsilon \max_{1 \leq n \leq m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)},$$

where we have used Hölder's and Young's inequalities in deriving the last inequality (with a constant ϵ which can be arbitrarily small at the expense of enlarging the constant C_ϵ). The estimate above furthermore implies

$$
\begin{aligned}
\|\mathcal{E}_{m,4}\|_{L^2(\Omega)} &\leq C\tau \sum_{n=1}^{m} \|g_\tau(t_n) - \widetilde{g}_\tau(t_n)\|_{L^2(\Omega)} \\
&\leq C\tau \sum_{n=1}^{m} \|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} w(t_n) - \bar{\partial}_\tau^{\alpha(t_n, u(t_{n-1}))} w(t_n)\|_{L^2(\Omega)} \\
&\leq C_\epsilon \tau \sum_{n=1}^{m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)} + \epsilon \max_{1 \leq n \leq m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)}.
\end{aligned}
\tag{4.28}
$$

Last, we note that

$$
\|\mathcal{E}_{m,3}\|_{L^2(\Omega)} \leq C\tau \sum_{n=1}^{m} \|g_n - g_\tau(t_n)\|_{L^2(\Omega)} \leq C\tau \sum_{n=1}^{m} \left\|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})}(u_n - u(t_n))\right\|_{L^2(\Omega)}.
\tag{4.29}
$$

Collecting the above estimates of $\|(\mathcal{E}_{m,j})_{n=1}^m\|_{L^2(\Omega)}$, $j = 3, 4, 5, 6$, we have

$$
\begin{aligned}
\|\mathcal{E}_{m,1}\|_{L^2(\Omega)} &\leq \|\mathcal{E}_{m,3}\|_{L^2(\Omega)} + \|\mathcal{E}_{m,4}\|_{L^2(\Omega)} + \|\mathcal{E}_{m,5}\|_{L^2(\Omega)} + \|\mathcal{E}_{m,6}\|_{L^2(\Omega)} \\
&\leq C\tau + C\tau \sum_{n=1}^{m} \left\|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})}(u_n - u(t_n))\right\|_{L^2(\Omega)} \\
&\quad + C_\epsilon \tau \sum_{n=1}^{m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)} + \epsilon \max_{1 \leq n \leq m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)}.
\end{aligned}
\tag{4.30}
$$

Substituting (4.19) and (4.30) into (4.17), and then substituting the result into (4.15), we obtain

$$
\begin{aligned}
\max_{1 \leq n \leq m} \|u_n - u(t_n)\|_{L^2(\Omega)} &\leq C\tau + C\tau \sum_{n=1}^{m} \left\|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})}(u_n - u(t_n))\right\|_{L^2(\Omega)} \\
&\quad + C_\epsilon \tau \sum_{n=1}^{m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)} + \epsilon \max_{0 \leq n \leq m-1} \|u_n - u(t_n)\|_{L^2(\Omega)}.
\end{aligned}
$$

By choosing a sufficiently small ϵ, the last term on the right-hand side can be absorbed by the left-hand side. Then, using Gronwall's inequality, we obtain

$$
\begin{aligned}
\max_{1 \leq n \leq m} \|u_n - u(t_n)\|_{L^2(\Omega)} &\leq C\tau + C\tau \sum_{n=1}^{m} \left\|\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})}(u_n - u(t_n))\right\|_{L^2(\Omega)} \\
&\leq C\tau + C\tau \sum_{j=1}^{m} \|\bar{\partial}_\tau^{\beta}(u_j - u(t_j))\|_{L^2(\Omega)},
\end{aligned}
\tag{4.31}
$$

where we have used Lemma 4.3 (to be proved below) in deriving the last inequality. It remains to estimate $\tau \sum_{j=1}^{m} \|\bar{\partial}_\tau^{\beta}(u_j - u(t_j))\|_{L^2(\Omega)}$.

Part II: Estimation for $\tau \sum_{j=1}^{m} \|\bar{\partial}_\tau^{\beta}(u_j - u(t_j))\|_{L^2(\Omega)}$.

Let β be a number satisfying $\alpha_* < \beta < 1$. Then, applying the operator $\bar{\partial}_\tau^\beta$ on (4.14) and summing up the $L^2(\Omega)$-errors on different time levels, we obtain

$$
\tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(u_n - u(t_n))\|_{L^2(\Omega)}
$$

$$
\leq \tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(I_n - I(t_n))\|_{L^2(\Omega)} + \tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(J_n - J(t_n))\|_{L^2(\Omega)}
$$

$$
\leq \tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(I_n - I(t_n))\|_{L^2(\Omega)} + \tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta \mathcal{E}_{n,1}\|_{L^2(\Omega)} + \tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta \mathcal{E}_{n,2}\|_{L^2(\Omega)}, \tag{4.32}
$$

where we have used the decomposition $J_m - J(t_m) = \mathcal{E}_{n,1} + \mathcal{E}_{n,2}$ in (4.17).

Since $I_n - I(t_n)$ and $\mathcal{E}_{n,2}$ correspond to the error of discretizing a heat equation with the source term $\Delta u_0 + f$ and g, respectively, the following error estimates hold (*cf.* Appendix C, choosing $p = 1$ in Lem. C.1):

$$
\tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(I_n - I(t_n))\|_{L^2(\Omega)} \leq C\tau(\|u_0\|_{\dot{H}^2} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))}) \leq C\tau, \tag{4.33}
$$

$$
\tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(\mathcal{E}_{n,2})\|_{L^2(\Omega)} \leq C\tau\|g\|_{W^{1,1}(0,T;L^2(\Omega))} \leq C\tau, \tag{4.34}
$$

where we have used $g \in W^{1,1}(0,T;L^2(\Omega))$, as shown in (4.18).

By using decomposition (4.21) and the stability of the backward Euler method (*cf.* Lem. C.2 in Appendix C), we have

$$
\tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(\mathcal{E}_{n,1})_{n=1}^m\|_{L^2(\Omega)} \leq \sum_{j=3}^6 \tau \sum_{n=1}^m \|\bar{\partial}_\tau^\beta(\mathcal{E}_{n,j})\|_{L^2(\Omega)}
$$

$$
\leq C\tau \sum_{n=1}^m \left(\|g_n - g_\tau(t_n)\|_{L^2(\Omega)} + \|g_\tau(t_n) - \widetilde{g}_\tau(t_n)\|_{L^2(\Omega)} \right)
$$

$$
+ C\tau \sum_{n=1}^m \left(\|\widetilde{g}_\tau(t_n) - g_\tau^*(t_n)\|_{L^2(\Omega)} + \|g_\tau^*(t_n) - g(t_n)\|_{L^2(\Omega)} \right)
$$

$$
\leq C\tau + C_\epsilon \tau \sum_{n=1}^m \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)}
$$

$$
+ \epsilon \max_{1 \leq n \leq m} \|u_{n-1} - u(t_{n-1})\|_{L^2(\Omega)}
$$

$$
+ C\tau \sum_{n=1}^m \left\| \bar{\partial}_\tau^{\alpha(t_n, u_{n-1})}(u_n - u(t_n)) \right\|_{L^2(\Omega)}
$$

(here we have used (4.22), (4.26), (4.28) and (4.29))

$$
\leq C\tau + C\tau \sum_{n=1}^m \left\| \bar{\partial}_\tau^{\alpha(t_n, u_{n-1})}(u_n - u(t_n)) \right\|_{L^2(\Omega)} \quad \text{(here we have used (4.31))}
$$

$$
\leq C\tau + C\tau \sum_{n=1}^m \tau \sum_{j=1}^n t_{n+1-j}^{\beta-\alpha_*-1} \|\bar{\partial}_\tau^\beta(u_j - u(t_j))\|_{L^2(\Omega)}, \tag{4.35}
$$

where we have used Lemma 4.3 (to be shown below) in the last inequality. Then, substituting (4.35) into (4.32) and using (4.33) and (4.34), we obtain

$$\tau \sum_{j=1}^{m} \|\bar{\partial}_\tau^\beta (u_j - u(t_j))\|_{L^2(\Omega)} \le C\tau + C\tau \sum_{n=1}^{m} \tau \sum_{j=1}^{n} t_{n+1-j}^{\beta - \alpha_* - 1} \|\bar{\partial}_\tau^\beta (u_j - u(t_j))\|_{L^2(\Omega)}.$$

By applying a generalized Gronwall's inequality (*cf.* Lem. D.1 in Appendix D), we can remove the second term on the right-hand side and therefore obtain the following estimate (when τ is sufficiently small):

$$\tau \sum_{j=1}^{m} \|\bar{\partial}_\tau^\beta (u_j - u(t_j))\|_{L^2(\Omega)} \le C\tau.$$

Substituting this into (4.31) yields the desired error estimate in Theorem 4.2.   □

In the proof of Theorem 4.2, in particular (4.31) and (4.35), we have used the following result.

Lemma 4.3. *If $\beta > \alpha_* \ge \alpha(t_n, u_{n-1})$ and $v_n \in L^2(\Omega)$, $n = 1, \dots, N$, with $v_0 = 0$, then*

$$\tau \sum_{n=1}^{m} \left\| \bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} v_n \right\|_{L^2(\Omega)} \le C\tau \sum_{n=1}^{m} \tau \sum_{j=0}^{n} t_{n+1-j}^{\beta - \alpha_* - 1} \|\bar{\partial}_\tau^\beta v_j\|_{L^2(\Omega)}.$$

Proof. We use the identity $\bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} v_n = \bar{\partial}_\tau^{\alpha(t_n, u_{n-1}) - \beta} \bar{\partial}_\tau^\beta v_n = \sum_{j=0}^{n} d_{n-j} \bar{\partial}_\tau^\beta v_j$, where the coefficients d_j can be expressed by using Cauchy's integral formula:

$$
\begin{aligned}
|d_j| = \left| \frac{1}{2\pi i} \int_{|\zeta|=1} \left(\frac{1-\zeta}{\tau} \right)^{\alpha(t_n, u_{n-1}) - \beta} \zeta^{-j-1} d\zeta \right| &= \left| \frac{\tau}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1 - e^{-\tau z}}{\tau} \right)^{\alpha(t_n, u_{n-1}) - \beta} e^{t_j z} dz \right| \\
&\le C\tau \int_{\Gamma_\theta^\tau} |z|^{\alpha(t_n, u_{n-1}) - \beta} e^{-Ct_j |z|} |dz| \\
&\le C\tau t_{j+1}^{\beta - \alpha_* - 1},
\end{aligned}
$$

where we have used (4.5) in the last inequality, which furthermore implies the following estimate:

$$\tau \sum_{n=1}^{m} \left\| \bar{\partial}_\tau^{\alpha(t_n, u_{n-1})} v_n \right\|_{L^2(\Omega)} \le C\tau \sum_{n=1}^{m} \tau \sum_{j=0}^{n} t_{n+1-j}^{\beta - \alpha_* - 1} \|\bar{\partial}_\tau^\beta v_j\|_{L^2(\Omega)}.$$

This proves the desired result.   □

5. Numerical results

Example 5.1. We consider (1.3) in a two-dimensional rectangular domain $\Omega = (0, 1) \times (0, 1)$ with both linear and nonlinear fractional orders:

$$
\begin{aligned}
&(i) \ \alpha(x, t) = 0.5 \pm 0.1 * \cos(t), \\
&(ii) \ \alpha(x, u) = 0.5 \pm 0.2 * \cos(u),
\end{aligned}
$$

with the following source term and initial value:

$$f(x, t) = \cos(t) \cos(\pi x) \cos(\pi y) \quad \text{and} \quad u_0(x) = x(1 - x)y(1 - y). \tag{5.1}$$
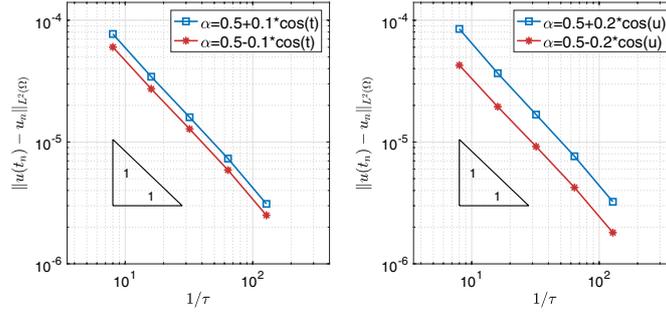
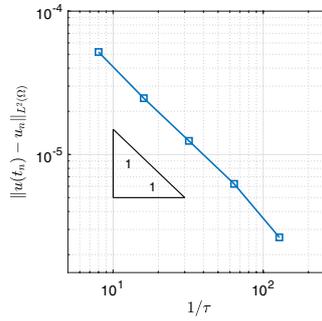FIGURE 2. Errors of numerical solutions in Example 5.1.

FIGURE 3. Errors of numerical solutions in Example 5.2.

We solve the problem by the proposed time-stepping method (4.1), with piecewise linear finite element method in space using a sufficiently small mesh size $h = 2^{-6}$ so that the spatial discretization error is negligible in observing the order of convergence in time. The reference solution for computing the errors of numerical solutions is obtained by using a much smaller time stepsize $\tau = 2^{-9}$. The errors of numerical solutions at time $T = 1$ for the linear and nonlinear fractional orders are presented in Figure 2, which shows that the numerical solutions have first-order convergence, consistent with the theoretical result proved in Theorem 4.2.

Example 5.2. Convergence of numerical solutions in Theorem 4.2 is based on condition (3.1), which allows the fractional order $\alpha(x, t, u)$ to be discontinuous in x. Here we consider the fractional PDE problem (1.3) in a two-dimensional rectangular domain $\Omega = (0, 1) \times (0, 1)$ with a discontinuous nonlinear fractional order:

$$\alpha(x, u) = \begin{cases} 0.5 - 0.1 * \cos(u) & \text{for} \quad x + y > 1, \\ 0.5 - 0.2 * \cos(u) & \text{for} \quad x + y \leq 1, \end{cases} \tag{5.2}$$

with the same source term and initial value in (5.1).

The problem is discretized by using the proposed time-stepping method (4.1) and the linear finite element method in space with a sufficiently small mesh size $h = 2^{-8}$ so that the spatial discretization error is relatively negligible. Similarly, a surrogate for the exact solution $u(x, t_n)$ is determined by using a much smaller time step $\tau = 2^{-8}$. The errors of numerical solutions at time $T = 1$ are presented in Figure 3, which shows that the numerical solutions have first-order convergence, consistent with the theoretical analysis.

6. Conclusion

We have proved well-posedness of the nonlinear problem (1.3) under condition (2.1), which requires the fractional order $\alpha(x, t, u)$ and proportion $b(x, t)$ to be measurable in x and t, and Lipschitz continuous in the unknown solution u. We have proposed a linearly implicit method for discretizing problem (1.3). The proposed method only requires solving a linear equation at every time level, using the convolution quadrature for approximating the variable-order fractional derivative. By requiring the functions $\alpha(x, t, u)$ and $b(x, t)$ to be Lipschitz continuous in both t and u, and measurable in x, we managed to prove first-order convergence of the proposed method under proved regularity results. The numerical results in Examples 5.1 and 5.2 illustrate the convergence rates of numerical solutions with several different choices of $\alpha(x, t, u)$, with continuous and discontinuous dependence on x, respectively.

Extension of the error analysis to full discretizations still remains challenging. For full discretization with finite element method in space, the error can be decomposed into two parts: the error between the continuous solution and the semidiscrete finite element solution (spatial error), and the error between the semidiscrete finite element solution and the fully discrete solution (temporal error). The first part of the error requires more delicate estimates on the spatial regularity of the solution, while the current paper mainly focuses on proving temporal regularity estimates. The second part of the error may be estimated similarly as the current paper, but needs to be estimated in the framework of finite element spaces in which the discrete H^2 norm should replace the H^2 norm used in the current paper.

Appendix A. An inhomogeneous Sobolev embedding result

In this appendix, we prove the following inhomogeneous Sobolev embedding result, which is used in Part I of the proof for Theorem 3.2.

Lemma A.1. *For $p_0 > \frac{4}{4-d}$ (or equivalently $1 - \frac{d}{4} - \frac{1}{p_0} > 0$), there holds*

$$W^{1,p_0}(0, T; L^2(\Omega)) \cap L^{p_0}(0, T; \dot{H}^2(\Omega)) \hookrightarrow C^\gamma([0, T]; L^\infty(\Omega))$$

where γ can be any number satisfying $0 < \gamma < 1 - \frac{d}{4} - \frac{1}{p_0}$.

Proof. First, the Sobolev embedding in time implies $W^{1,p_0}(0, T; L^2(\Omega)) \hookrightarrow C^{1-\frac{1}{p_0}}([0, T]; L^2(\Omega))$, which implies

$$\|u(t) - u(s)\|_{L^2(\Omega)} \le C|t - s|^{1-\frac{1}{p_0}} \|u\|_{W^{1,p_0}(0,T;L^2(\Omega))}. \tag{A.1}$$

Second, the inhomogeneous Sobolev embedding in Proposition 1.2.10 of [33] says that

$$W^{1,p_0}(0, T; L^2(\Omega)) \cap L^{p_0}(0, T; \dot{H}^2(\Omega)) \hookrightarrow C\left([0, T]; (L^2(\Omega), \dot{H}^2(\Omega))_{1-\frac{1}{p_0}, p_0}\right),$$

which implies

$$\|u(t) - u(s)\|_{(L^2(\Omega), \dot{H}^2(\Omega))_{1-\frac{1}{p_0}, p_0}} \le C\left(\|u\|_{W^{1,p_0}(0,T;L^2(\Omega))} + \|u\|_{L^{p_0}(0,T;\dot{H}^2(\Omega))}\right). \tag{A.2}$$

The real interpolation between (A.1) and (A.2) yields

$$\|u(t) - u(s)\|_{(L^2(\Omega), \dot{H}^2(\Omega))_{(1-\frac{1}{p_0})\theta, 2}} \le C|t - s|^{(1-\frac{1}{p_0})(1-\theta)} \left(\|u\|_{W^{1,p_0}(0,T;L^2(\Omega))} + \|u\|_{L^{p_0}(0,T;\dot{H}^2(\Omega))}\right) \tag{A.3}$$

which holds for all $\theta \in (0, 1)$. Note that

$$(L^2(\Omega), \dot{H}^2(\Omega))_{(1-\frac{1}{p_0})\theta, 2} = \dot{H}^{2(1-\frac{1}{p_0})\theta}(\Omega) \hookrightarrow L^\infty(\Omega) \quad \text{if} \quad \left(1 - \frac{1}{p_0}\right)\theta > \frac{d}{4}. \tag{A.4}$$

If $p_0 > \frac{4}{4-d}$ then $1 - \frac{1}{p_0} > \frac{d}{4}$ and therefore, there exists $\theta \in (0,1)$ satisfying $\left(1 - \frac{1}{p_0}\right)\theta > \frac{d}{4}$. For any such θ, (A.3) and (A.4) imply

$$\|u(t) - u(s)\|_{L^\infty(\Omega)} \le C|t - s|^{(1 - \frac{1}{p_0})(1-\theta)} \left(\|u\|_{W^{1,p_0}(0,T;L^2(\Omega))} + \|u\|_{L^{p_0}(0,T;\dot{H}^2(\Omega))} \right). \tag{A.5}$$

This proves that $u \in C^\gamma([0,T]; L^\infty(\Omega))$ with

$$\gamma = \left(1 - \frac{1}{p_0}\right)(1 - \theta) = 1 - \frac{1}{p_0} - \left(1 - \frac{1}{p_0}\right)\theta.$$

Since $\left(1 - \frac{1}{p_0}\right)\theta$ can be arbitrarily close to (but larger than) $\frac{d}{4}$, it follows that γ can arbitrarily close to (but smaller than) $1 - \frac{1}{p_0} - \frac{d}{4}$. $\qquad \square$

APPENDIX B. AN INTERPOLATION INEQUALITY

In this section we prove the following lemma, which is used in the proof of Theorem 2.7.

Lemma B.1. *For $p > \frac{4}{4-d}$ and $w \in Y$, the following estimate holds for $t_* \in (0,T]$:*

$$\|w\|_{L^\infty(0,t_*;L^\infty(\Omega))} \le C_\varepsilon \|w\|_{L^1(0,t_*;L^2(\Omega))} + \varepsilon \left(\|\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} + \|w\|_{L^p(0,t_*;\dot{H}^2(\Omega))} \right),$$

where ε can be arbitrarily small at expense of enlarging the constant C_ε, which is independent of $t_* \in (0,T]$ (but may depend on T).

Proof. For any function $w \in W^{1,p}(0,t_*;L^2(\Omega)) \cap L^p(0,t_*;\dot{H}^2(\Omega))$ with $w(0) = 0$, we can extend w to $W^{1,p}(0,2T;L^2(\Omega)) \cap L^p(0,2T;\dot{H}^2(\Omega))$ by setting

$$w(t) = \begin{cases} w(t) & \text{for} \quad t \in [0, t_*], \\ w(2t_* - t) & \text{for} \quad t \in [t_*, 2t_*], \\ 0 & \text{for} \quad t \in [2t_*, 2T]. \end{cases}$$

Then

$$\|\partial_t w\|_{L^p(0,2T;L^2(\Omega))} + \|w\|_{L^p(0,2T;\dot{H}^2(\Omega))} \le C \left(\|\partial_t w\|_{L^p(0,t_*;L^2(\Omega))} + \|w\|_{L^p(0,t_*;\dot{H}^2(\Omega))} \right).$$

Since $W^{1,p}(0,2T;L^2(\Omega)) \cap L^p(0,2T;\dot{H}^2(\Omega))$ is compactly embedded into $L^\infty(0,2T;L^\infty(\Omega))$ when $p > \frac{4}{4-d}$ (see Appendix A), it follows from Lemma B.2 that

$$\|w\|_{L^\infty(0,2T;L^\infty(\Omega))} \le C_\varepsilon \|w\|_{L^1(0,2T;L^2(\Omega))} + C^{-1}\varepsilon \left(\|\partial_t w\|_{L^p(0,2T;L^2(\Omega))} + \|w\|_{L^p(0,2T;\dot{H}^2(\Omega))} \right),$$

where the constant C depends only on T (independent of $t_* \in (0,T]$). The above two estimates together imply the desired result of Lemma B.1. $\qquad \square$

Lemma B.2 ([1], Lem. 9.1). *Let Y, Z and W be Banach spaces such that Y is compactly embedded into Z and Z is continuously embedded into W. Then*

$$\|y\|_Z \le C_\varepsilon \|y\|_W + \varepsilon \|y\|_Y,$$

where ε can be arbitrarily small at the expense of enlarging the constant C_ε.

APPENDIX C. ERROR AND STABILITY ESTIMATES FOR THE HEAT EQUATION

In this appendix, we present an error estimate and a stability estimate of the backward Euler method for the heat equation. These estimates are used in the proof of Theorem 4.2.

Let

$$v_n = \tau \sum_{j=1}^{n} E_{n-j}(\Delta u_0 + f(t_j)) \quad \text{and} \quad v(t) = \int_0^t E(t-s)(\Delta u_0 + f(s))\mathrm{d}s, \tag{C.1}$$

where the discrete and continuous heat semigroup E_n and $E(t)$ are defined in (4.12) and (3.7), respectively. The functions v_n and $v(t)$ defined above are the numerical solution (with backward Euler method) and exact solution of the heat equation, respectively, with source term f and initial value u_0.

It is well known that the backward Euler method has first-order convergence in time. In the following lemma, we show that the discrete fractional-order derivative of the error also has first-order convergence in a temporal L^p norm.

Lemma C.1. *If $u_0 \in \dot{H}^2(\Omega)$ and $f \in W^{1,1}(0,T;L^2(\Omega))$, then the following error estimate holds for $0 \le \beta < 1$:*

$$\|\bar{\partial}_\tau^\beta v_n - \partial_t^\beta v(t_n)\|_{L^2(\Omega)} \le C\tau\, t_n^{-\beta}(\|u_0\|_{\dot{H}^2(\Omega)} + \|f(0)\|_{L^2(\Omega)})$$

$$+ C\tau \int_0^{t_n} (t_n - s)^{-\beta}\|\partial_t f(s)\|_{L^2(\Omega)}\mathrm{d}s. \tag{C.2}$$

In particular, if $0 \le \beta < 1 \le p < 1/\beta$ then

$$\left(\tau \sum_{n=1}^{N} \|\bar{\partial}_\tau^\beta v_n - \partial_t^\beta v(t_n)\|_{L^2(\Omega)}\right)^{\frac{1}{p}} \le C\tau \left(\|u_0\|_{\dot{H}^2(\Omega)} + \|f\|_{W^{1,1}(0,T;L^2(\Omega))}\right). \tag{C.3}$$

Proof. We decompose the error in the following way:

$$\bar{\partial}_\tau^\beta(v_n - v(t_n)) = (\bar{\partial}_\tau^\beta v_n - \partial_t^\beta v(t_n)) + (\partial_t^\beta v(t_n) - \bar{\partial}_\tau^\beta v(t_n)). \tag{C.4}$$

The second part on the right-hand side has already been estimated in Lemma 4.1, which implies that

$$\left\|\left(\partial_t^\beta v(t_n) - \bar{\partial}_\tau^\beta v(t_n)\right)_{n=1}^{N}\right\|_{L^p(L^2(\Omega))} \le C\tau\|^R\partial_t^\beta \partial_t v\|_{L^p(0,T;L^2(\Omega))} \le C\tau, \tag{C.5}$$

where we have used the regularity result $^R\partial_t^\beta \partial_t v \in L^p(0,T;L^2(\Omega))$, as shown in Theorem 3.2 (the case $b \equiv 0$ reduces to the heat equation).

The first part on the right-hand side of (C.4) can be estimated in the following way. We split the source term into two parts:

$$\Delta u_0 + f(t) = [\Delta u_0 + f(0)] + 1 * \partial_t f(t),$$

and estimate the error corresponding to each part, separately.

In the first case, when the source term is $w = \Delta u_0 + f(0)$ (independent of time), we have

$$\|\bar{\partial}_\tau^\beta v_n - \partial_t^\beta v(t_n)\|_{L^2(\Omega)}$$

$$= \left\| \frac{1}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1 - e^{-\tau z}}{\tau}\right)^{\beta - 1} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta\right)^{-1} w e^{t_n z}\mathrm{d}z \right.$$

$$\left. - \frac{1}{2\pi i} \int_{\Gamma_\theta} z^{\beta - 1}(z - \Delta)^{-1} w e^{t_n z}\mathrm{d}z \right\|_{L^2(\Omega)}$$

$$\leq C \int_{\Gamma_\theta^\tau} \left\| \left(\frac{1 - e^{-\tau z}}{\tau} \right)^{\beta-1} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta \right)^{-1} - z^{\beta-1}(z - \Delta)^{-1} \right\| \|w\|_{L^2(\Omega)} e^{-Ct_n|z|} |\mathrm{d}z|$$

$$+ C \int_{\Gamma_\theta \setminus \Gamma_\theta^\tau} |z|^{\beta-1} \|(z - \Delta)^{-1}\| \|w\|_{L^2(\Omega)} e^{-Ct_n|z|} |\mathrm{d}z|$$

$$\leq \int_{\Gamma_\theta^\tau} C\tau |z|^{\beta-1} e^{-Ct_n|z|} \|w\|_{L^2(\Omega)} |\mathrm{d}z| + C \int_{\Gamma_\theta \setminus \Gamma_\theta^\tau} |z|^{\beta-1} e^{-Ct_n|z|} \|w\|_{L^2(\Omega)} |\mathrm{d}z|$$

$$\leq C\tau t_{n+1}^{-\beta} \|w\|_{L^2(\Omega)} = C\tau t_{n+1}^{-\beta} \|\Delta u_0 + f(0)\|_{L^2(\Omega)}, \tag{C.6}$$

where the second to last inequality follows from (4.5).

In the second case, when the source term is $1 * \partial_t f(t)$, we denote

$$J_{n,\epsilon}(t) = \sum_{j=0}^{\infty} J_j \delta_{t_j + \epsilon}(t), \quad \text{with} \quad J_j = \frac{\tau}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1 - e^{-\tau z}}{\tau} \right)^{\beta} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta \right)^{-1} e^{t_j z} \mathrm{d}z,$$

$$J(t) = \frac{1}{2\pi i} \int_{\Gamma_\theta} z^\beta (z - \Delta)^{-1} e^{tz} \mathrm{d}z,$$

and express the error as

$$\bar{\partial}_\tau^\beta v_n - \partial_t^\beta v(t_n) = \sum_{j=1}^{n} J_{n-j}[1 * \partial_t f(t_j)] - \int_0^{t_n} J(t_n - s)[1 * \partial_t f](s) \mathrm{d}s$$

$$= \left[\lim_{\epsilon \to 0} (J_{n,\epsilon} * 1) - (J * 1) \right] * \partial_t f(t_n).$$

Then, the same as the estimate in (C.6) (without w this time), we can obtain

$$\| \lim_{\epsilon \to 0} (J_{n,\epsilon} * 1)(t_n) - (J * 1)(t_n) \|$$

$$= \left\| \frac{1}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1 - e^{-\tau z}}{\tau} \right)^{\beta-1} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta \right)^{-1} e^{t_n z} \mathrm{d}z - \frac{1}{2\pi i} \int_{\Gamma_\theta} z^{\beta-1}(z - \Delta)^{-1} e^{t_n z} \mathrm{d}z \right\|$$

$$\leq C\tau t_{n+1}^{-\beta}.$$

Meanwhile, the estimates

$$\| \lim_{\epsilon \to 0} (J_{n,\epsilon} * 1)(t_n) - \lim_{\epsilon \to 0} (J_{n,\epsilon} * 1)(t) \| \leq C\tau t_{n+1}^{-\beta} \quad \text{and} \quad \|(J * 1)(t_n) - (J * 1)(t)\| \leq C\tau t_{n+1}^{-\beta}$$

can be done similarly to page 536 of [19] (using the Taylor expansion of the functions $J_{n,\epsilon} * 1$ and $J * 1$). To summarize, we have

$$\| \lim_{\epsilon \to 0} (J_{n,\epsilon} * 1)(t) - (J * 1)(t) \| \leq C\tau t^{-\beta} \quad \forall t \in (0, T],$$

and therefore

$$\|\bar{\partial}_\tau^\beta v_n - \partial_t^\beta v(t_n)\|_{L^2(\Omega)} = \|[\lim_{\epsilon \to 0} (J_{n,\epsilon} * 1) - (J * 1)] * \partial_t f(t_n)\|$$

$$\leq C\tau \int_0^{t_n} (t_n - s)^{-\beta} \|\partial_t f(s)\|_{L^2(\Omega)} \mathrm{d}s. \tag{C.7}$$

Combining the two inequalities (C.6) and (C.7) yields the desired error estimate (C.3). Since $1 \leq p < 1/\beta$, it follows that $t^{-p\beta}$ is integrable in time. As a result, taking the discrete temporal L^p norm of (C.2) yields (C.3). This completes the proof of Lemma C.1. $\qquad \square$

In Part II of the proof for Theorem 4.2, we have used the following stability result of the backward Euler method.

Lemma C.2. *If $f_n \in L^2(\Omega)$, $n = 1, 2, \ldots$, and $0 \le \beta < 1$, then the function*

$$v_n = \tau \sum_{j=1}^{n} E_{n-j} f_j$$

satisfies

$$\tau \sum_{j=1}^{n} \|\bar{\partial}_\tau^\beta v_j\|_{L^2(\Omega)} \le C\tau \sum_{j=1}^{n} \|f_j\|_{L^2(\Omega)} \quad \forall n = 1, \ldots, N, \tag{C.8}$$

where the constant C is independent of τ, n and N, but may depend on T.

Proof. Applying $\bar{\partial}_\tau^\beta$ to the expression of v_n yields $\bar{\partial}_\tau^\beta v_n = \tau \sum_{j=1}^{n} \bar{\partial}_\tau^\beta E_{n-j} f_j$, with

$$\|\bar{\partial}_\tau^\beta E_n\| = \left\| \frac{\tau}{2\pi i} \int_{\Gamma_\theta^\tau} \left(\frac{1 - e^{-\tau z}}{\tau} \right)^\beta \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta \right)^{-1} e^{t_n z} \mathrm{d}z \right\|$$

$$\le C\tau \int_{\Gamma_\theta^\tau} |z|^{\beta-1} e^{-Ct_n|z|} |\mathrm{d}z| \le C\tau t_{n+1}^{-\beta},$$

where the last inequality follows from (4.5). Thus

$$\tau \sum_{n=1}^{m} \|\bar{\partial}_\tau^\beta v_n\|_{L^2(\Omega)} \le C\tau \sum_{n=1}^{m} C\tau \sum_{j=1}^{n} t_{n+1-j}^{-\beta} \|f_j\|_{L^2(\Omega)} \le C\tau \sum_{j=1}^{m} \|f_j\|_{L^2(\Omega)},$$

where we have interchanged the order of summation and used the property

$$\tau \sum_{n=j}^{m} t_{n+1-j}^{-\beta} \le C \int_0^{t_{m+1}} t^{-\beta} \mathrm{d}t \le CT^{1-\beta}.$$

This proves the stability estimate (C.8). □

APPENDIX D. A GENERALIZED GRONWALL'S INEQUALITY

We have used the following version of Gronwall's inequality in the proof of Theorem 4.2.

Lemma D.1. *Let A, B and $0 \le \gamma < 1$ be positive constants, and let $t_n = n\tau$, $n = 0, 1, \ldots, N$, be a partition of the time interval $[0, T]$ with a stepsize $\tau \le \frac{1}{2} \left(\frac{1-\gamma}{2B} \right)^{\frac{1}{1-\gamma}}$. If $y_n \in \mathbb{R}$, $n = 1, \ldots, N$, satisfy*

$$\tau \sum_{n=1}^{m} y_n \le A + B\tau \sum_{n=1}^{m} \tau \sum_{j=1}^{n} t_{n+1-j}^{-\gamma} y_j, \tag{D.1}$$

then

$$\tau \sum_{n=1}^{N} y_n \le C_B A, \tag{D.2}$$

where the constant C_B is independent of τ and N, but may depend on B, T and γ.

Proof. Let $0 = T_0 < T_1 < \cdots < T_L = T$, with $T_l = N_l \tau$, be a partition of the interval $[0, T]$ such that

$$\frac{T_*}{2} \leq T_l - T_{l-1} \leq T_*, \quad \text{with} \quad T_* := \left(\frac{1-\gamma}{2B}\right)^{\frac{1}{1-\gamma}}.$$

Let

$$\mathcal{N}_l = \{n : N_{l-1} < n \leq N_l\} \quad \text{for} \quad l = 1, \ldots, L,$$

and denote $Y_l = \tau \sum_{n \in \mathcal{N}_l} y_n$. Then setting $m = N_k$ in (D.1) yields the following inequality for $k = 1, \ldots, L$:

$$\sum_{l=1}^{k} Y_l \leq A + B\tau \sum_{n=1}^{N_k} \tau \sum_{j=1}^{n} t_{n+1-j}^{-\gamma} y_j$$

$$= A + B\sum_{l=1}^{k-1} \tau \sum_{n \in \mathcal{N}_l} \tau \sum_{j=1}^{n} t_{n+1-j}^{-\gamma} y_j + B\tau \sum_{n \in \mathcal{N}_k} \tau \sum_{j=1}^{n} t_{n+1-j}^{-\gamma} y_j$$

$$= A + B\sum_{l=1}^{k-1} \tau \sum_{n \in \mathcal{N}_l} \tau \sum_{j=1}^{n} t_{n+1-j}^{-\gamma} y_j + B\tau \sum_{n \in \mathcal{N}_k} \tau \sum_{j=1}^{N_{k-1}} t_{n+1-j}^{-\gamma} y_j$$

$$+ B\tau \sum_{n \in \mathcal{N}_k} \tau \sum_{j=N_{k-1}+1}^{n} t_{n+1-j}^{-\gamma} y_j$$

$$=: A + G_{1,k} + G_{2,k} + G_{3,k},$$

with

$$G_{3,k} = B\tau \sum_{j=N_{k-1}+1}^{N_k} \tau \sum_{n=j}^{N_k} t_{n+1-j}^{-\gamma} y_j \leq \frac{BT_*^{1-\gamma}}{1-\gamma} Y_k,$$

where we have changed the order of summation and used the following property: if $N_{k-1} + 1 \leq j \leq N_k$ then

$$\tau \sum_{n=j}^{N_k} t_{n+1-j}^{-\gamma} \leq \tau \sum_{n=0}^{N_k - N_{k-1}} t_{n+1}^{-\gamma} \leq \int_0^{T_k - T_{k-1}} t^{-\gamma} \mathrm{d}t \leq \frac{T_*^{1-\gamma}}{1-\gamma}.$$

We also note that

$$G_{2,k} = B\tau \sum_{j=1}^{N_{k-1}} \tau \sum_{n=N_{k-1}+1}^{N_k} t_{n+1-j}^{-\gamma} y_j \leq \frac{BT_*^{1-\gamma}}{1-\gamma} \sum_{l=1}^{k-1} Y_l,$$

$$G_{1,k} = B\tau \sum_{j=1}^{N_{k-1}} \tau \sum_{n=j}^{N_{k-1}} t_{n+1-j}^{-\gamma} y_j \leq \frac{BT^{1-\gamma}}{1-\gamma} \sum_{l=1}^{k-1} Y_l.$$

Therefore,

$$\sum_{l=1}^{k} Y_l \leq A + \frac{2BT^{1-\gamma}}{1-\gamma} \sum_{l=1}^{k-1} Y_l + \frac{BT_*^{1-\gamma}}{1-\gamma} Y_k.$$

Our choice of $T_* = \left(\frac{1-\gamma}{2B}\right)^{\frac{1}{1-\gamma}}$ satisfies that $\frac{BT_*^{1-\gamma}}{1-\gamma} = \frac{1}{2}$, then we obtain

$$\sum_{l=1}^{k} Y_l \leq 2A + \frac{4BT^{1-\gamma}}{1-\gamma} \sum_{l=1}^{k-1} Y_l.$$

By iterating this inequality for $k = 1, \ldots, L$ (a bounded number of times, with $L \leq 2T/T_*$), we obtain the desired estimate $\sum_{l=1}^{L} Y_l \leq CA$. $\square$

References

[1] G. Akrivis, B. Li and C. Lubich, Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. *Math. Comput.* **86** (2017) 1527–1552.

[2] L. Banjai and M. López-Fernández, Efficient high order algorithms for fractional integrals and fractional differential equations. *Numer. Math.* **141** (2019) 289–317.

[3] P. Baveye, P. Vandevivere, B.L. Hoyle, P.C. DeLeo and D.S. de Lozada, Environmental impact and mechanisms of the biological clogging of saturated soils and aquifer materials. *Crit. Rev. Environ. Sci. Technol.* **28** (2006) 123–191.

[4] J. Bear, Dynamics of Fluids in Porous Media. Elsevier, New York (1972).

[5] L.C. Becker, Resolvents and solutions of weakly singular linear Volterra integral equations. *Nonlinear Anal.* **74** (2011) 1892–1912.

[6] D. Benson, R. Schumer, M.M. Meerschaert and S.W. Wheatcraft, Fractional dispersion, Lévy motions, and the MADE tracer tests. *Transp. Porous Media* **42** (2001) 211–240.

[7] J. Bergh and J. Löfström, Interpolation Spaces. An Introduction. Springer, Berlin-Heidleberg (1976).

[8] V. Bogachev, Measure Theory I. Springer, Berlin-Heidelberg (1978).

[9] T.A. Burton and D.P. Dwiggins, Resolvents of integral equations with continuous kernels. *Nonlinear Stud.* **18** (2011) 293–305.

[10] C. Cheng, V. Thomée and L. Wahlbin, Finite element approximation of a parabolic integro-differential equation with a weakly singular kernel. *Math. Comput.* **58** (1992) 587–602.

[11] E. Cuesta, C. Lubich and C. Palencia, Convolution quadrature time discretization of fractional diffusion-wave equations. *Math. Comput.* **75** (2006) 673–696.

[12] W. Deng, Finite element method for the space and time fractional Fokker–Planck equation. *SIAM J. Numer. Anal.* **47** (2008) 204–226.

[13] P. Embrechts and M. Maejima, Selfsimilar Processes. In: *Princeton Series in Applied Mathematics.* Princeton University Press, Princeton, NJ (2002).

[14] L.C. Evans, Partial Differential Equations, 2nd edition. In: Vol. 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence (2010).

[15] L. Gandossi and U.V. Estorff, *An overview of hydraulic fracturing and other formation stimulation technologies for shale gas production*, Scientific and Technical Research Reports. Joint Research Centre of the European Commission; Publications Office of the European Union (2015). DOI:10.2790/379646.

[16] L. Grafakos, Classical Fourier Analysis, 3rd edition. In: Vol. 249 of *Graduate Texts in Mathematics.* Springer, New York (2014).

[17] B. Jin, R. Lazarov and Z. Zhou, Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. *SIAM J. Sci. Comput.* **38** (2016) A146–A170.

[18] B. Jin, B. Li and Z. Zhou, Correction of high-order BDF convolution quadrature for fractional evolution equations. *SIAM J. Sci. Comput.* **39** (2017) A3129–A3152.

[19] B. Jin, B. Li and Z. Zhou, An analysis of the Crank–Nicolson method for subdiffusion. *IMA J. Numer. Anal.* **38** (2018) 518–541.

[20] B. Jin, B. Li and Z. Zhou, Subdiffusion with a time-dependent coefficient: analysis and numerical solution. *Math. Comput.* **88** (2019) 2157–2186.

[21] Y. Kian, E. Soccorsi and M. Yamamoto, Finite difference/spectral approximations for the time-fractional diffusion equation. *Ann. Henri Poincaré* **19** (2018) 3955–3881.

[22] G.E. King, *Hydraulic fracturing 101: What every representative, environmentalist, regulator, reporter, investor, university researcher, neighbor and engineer should know about estimating frac risk and improving frac performance in unconventional gas and oil wells.* Society of Petroleum Engineers, SPE 152596 (2012).

[23] N. Kopteva, Error analysis of the L1 method on graded and uniform meshes for a fractional-derivative problem in two and three dimensions. *Math. Comput.* **88** (2019) 2135–2155.

[24] A. Kubica and M. Yamamoto, Initial-boundary value problems for fractional diffusion equations with time-dependent coefficients. *Fract. Calc. Appl. Anal.* **21** (2018) 276–311.

[25] T. Lee, L. Bocquet and B. Coasne, The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *Nat. Commun.* **29** (2016) 11890.

[26] L. Li and J.-G. Liu, A discretization of caputo derivatives with application to time fractional SDEs and gradient flows. *SIAM J. Numer. Anal.* **57** (2019) 2095–2120.

[27] H. Liao, W. McLean and J. Zhang, A discrete Grönwall inequality with applications to numerical schemes for subdiffusion problems. *SIAM J. Numer. Anal.* **57** (2019) 218–237.

[28] Y. Lin and C. Xu, Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.* **225** (2007) 1533–1552.

[29] C.F. Lorenzo and T.T. Hartley, Variable order and distributed order fractional operators. *Nonlinear Dyn.* **29** (2002) 57–98.

[30] C. Lubich, Convolution quadrature and discretized operational calculus. I and II. *Numer. Math.* **52** (1988) 129–145 and 413–425.

[31] C. Lubich, Convolution quadrature revisited. *BIT* **44** (2004) 503–514.

[32] C. Lubich, I.H. Sloan and V. Thomée, Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term. *Math. Comput.* **65** (1996) 1–17.

[33] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, 3rd edition. Birkhäuser Verlag, Basel (1995).

[34] M. Luskin and R. Rannacher, On the smoothing property of the Galerkin method for parabolic equations. *SIAM J. Numer. Anal.* **19** (1982) 93–113.

[35] C. Lv and C. Xu, Error analysis of a high order method for time-fractional diffusion equations. *SIAM J. Sci. Comput.* **38** (2016) A2699–A2724.

[36] W. McLean and K. Mustapha, Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. *Numer. Algorithms* **52** (2009) 69–88.

[37] M.M. Meerschaert and A. Sikorskii, Stochastic Models for Fractional Calculus. In: *De Gruyter Studies in Mathematics* (2011).

[38] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* **339** (2000) 1–77.

[39] R. Metzler and J. Klafter, The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A: Math. Gen.* **37** (2004) R161–R208.

[40] K. Mustapha, Time-stepping discontinuous Galerkin methods for fractional diffusion problems. *Numer. Math.* **130** (2015) 497–516.

[41] K. Mustapha, FEM for time-fractional diffusion equations, novel optimal error analyses. *Math. Comput.* **87** (2018) 2259–2272.

[42] K. Mustapha and W. McLean, Piecewise-linear, discontinuous Galerkin method for a fractional diffusion equation. *Numer. Algorithms* **56** (2011) 159–184.

[43] K. Mustapha and W. McLean, Uniform convergence for a discontinuous Galerkin, time-stepping method applied to a fractional diffusion equation. *IMA J. Numer. Anal.* **32** (2011) 906–925.

[44] K. Mustapha and D. Schötzau, Well-posedness of hp-version discontinuous Galerkin methods for fractional diffusion wave equation. *IMA J. Numer. Anal.* **34** (2013) 1426–1446.

[45] K. Mustapha, B. Abdallah and K.M. Furati, A discontinuous Petrov–Galerkin method for time-fractional diffusion equations. *SIAM J. Numer. Anal.* **52** (2014) 2512–2529.

[46] E.M. Ouhabaz, Gaussian estimates and holomorphy of semigroups. *Proc. Amer. Math. Soc.* **123** (1995) 1465–1474.

[47] R. Schumer, D.A. Benson, M.M. Meerschaert and B. Baeumer, Fractal mobile/immobile solute transport. *Water Resour. Res.* **39** (2003) 1–12.

[48] M. Stynes, E. O'Riordan and J.L. Gracia, Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. *SIAM J. Numer. Anal.* **55** (2017) 1057–1079.

[49] Z.-Z. Sun and X. Wu, A fully discrete difference scheme for a diffusion-wave system. *Appl. Numer. Math.* **56** (2006) 193–209.

[50] H. Sun, W. Chen and Y. Chen, Variable-order fractional differential operators in anomalous diffusion modeling. *Phys. A: Stat. Mech. Appl.* **388** (2009) 4586–4592.

[51] H. Sun, A. Chang, Y. Zhang and W. Chen, A review on variable-order fractional differential equations: mathematical foundations, physical models, numerical methods and applications. *Fract. Calc. Appl. Anal.* **22** (2019) 27–59.

[52] J. Sun, D. Nie and W. Deng, Error estimates for backward fractional Feynman–Kac equation with non-smooth initial data. *J. Sci. Comput.* **84** (2020) 1–23.

[53] H. Wang and X. Zheng, Wellposedness and regularity of the variable-order time-fractional diffusion equations. *J. Math. Anal. App.* **475** (2019) 1778–1802.

[54] L. Weis, A new approach to maximal L_p-regularity, edited by B. Herrenalb. In: Vol. 215 of *Lecture Notes in Pure and Applied Mathematics. Evolution Equations and their Applications in Physical and Life Sciences.* Dekker, New York (2001) 195–214.

[55] Y. Xian, M. Jin, H. Zhan and Y. Liu, Reactive transport of nutrients and bioclogging during dynamic disconnection process of stream and groundwater. *Water Resour. Res.* **55** (2019) 3882–3903.

[56] Y. Zhang, C. Green and B. Baeumer, Linking aquifer spatial properties and non-Fickian transport in mobile–immobile like alluvial settings. *J. Hydrol.* **512** (2014) 315–331.