Journal of Scientific Computing (2022) 91:14
https://doi.org/10.1007/s10915-022-01791-1

Optimal Error Estimates of a Time-Spectral Method for Fractional Diffusion Problems with Low Regularity Data

Hao Luo[1] · Xiaoping Xie[2]

Received: 6 June 2021 / Revised: 20 December 2021 / Accepted: 19 January 2022 /
Published online: 24 February 2022

Abstract
This paper is devoted to the error analysis of a time-spectral algorithm for fractional diffusion problems of order α $(0 < \alpha < 1)$. The solution regularity in the Sobolev space is revisited and new regularity results in the Besov space are established. A time-spectral algorithm is developed which adopts a standard spectral method and a conforming linear finite element method for temporal and spatial discretizations, respectively. Optimal error estimates are derived with nonsmooth data. Particularly, a sharp temporal convergence rate $1 + 2\alpha$ is shown theoretically and numerically.

Keywords Fractional diffusion problem · Finite element method · Spectral method · Jacobi polynomial · Low regularity · Besov space · Optimal error estimate

1 Introduction

Let $T > 0$ be a finite time. This paper considers the following time fractional diffusion equation:

$$\begin{cases} \mathrm{D}_{0+}^{\alpha}(u - u_0) - \Delta u = f & \text{in } \Omega \times (0, T), \\ u = 0 & \text{on } \partial\Omega \times (0, T), \\ u(0) = u_0 & \text{in } \Omega, \end{cases} \tag{1}$$

where $u_0,\ f$ are known data, $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ is a convex polygonal domain and D_{0+}^{α} is a Riemann–Liouville fractional differential operator with order $\alpha \in (0, 1)$; see Sect. 2.

The problem (1) is widely used in modeling of anomalous diffusion process [45,46] and anomalous transport [38,65], for its capability of accurately describing models with non-

This work was supported in part by National Natural Science Foundation of China (12171340, 11771312).

✉ Xiaoping Xie
xpxie@scu.edu.cn

1 School of Mathematical Sciences, Peking University, Beijing 100871, China

2 School of Mathematics, Sichuan University, Chengdu 610064, China

locality and historical memory [23,52]. For theoretical study to the problem, e.g. the weak solution and its regularity, we refer to [16,33,35,54].

Many numerical methods have been developed in the past a dozen years. Among existing works, three types of temporal discretization are most prevailing, i.e., the finite difference method (L-type schemes [2,24,36,41] and convolution quadrature methods [12,14,62,64]), the finite element method [25,27,29,30] and the spectral method [26,35,55,66]. Under certain circumstances, problem (1) has an equivalent form

$$\begin{cases} u_t - \mathrm{D}_{0+}^{1-\alpha}\,\Delta u = g & \text{in } \Omega \times (0,T), \\ u = 0 & \text{on } \partial\Omega \times (0,T), \\ u(0) = u_0 & \text{in } \Omega, \end{cases} \tag{2}$$

where $g = \mathrm{D}_{0+}^{1-\alpha} f$. In the literature, both (1) and (2) are called time fractional diffusion equations or time fractional subdiffusion equations. For the solution regularity and numerical analysis of problem (2), especially in the case of nonsmooth data, we refer the reader to [28,42,43,48,50,51].

It is well-known that the solution to problem (1) generally has boundary singularity (near $0+$) in temporal direction. If $f = 0$ and u_0 is (or not) smooth, or $u_0 = 0$ and f is smooth, then one can obtain growth estimates of the solution [19,20] or even find out the leading singular term t^α of the solution [33]. Due to the singularity, the accuracy $2-\alpha$ of the L1 scheme [36] deteriorates into 1 in the case of $f = 0$ and $u_0 \neq 0$, whether the initial data u_0 is smooth or not [21]. In the same situation, a piecewise constant discontinuous Galerkin (DG) semidiscretization was analyzed in [44]. Let us summarize the error estimate results of [21,44] as follows: for any temporal grid node $t_j = j\tau$ with $j = 1, 2, \ldots, J$ and $\tau = T/J$,

$$\left\| (u-U)(t_j) \right\|_{L^2(\Omega)} \leq C \begin{cases} t_j^{\alpha-1}\tau \left\| u_0 \right\|_{\dot H^2(\Omega)}, & \text{for L1 in [22]} \\ t_j^{-1}\tau \left\| u_0 \right\|_{L^2(\Omega)}, & \text{for L1 in [22] and DG in [45].} \end{cases} \tag{3}$$

Hence, if $u_0 \in L^2(\Omega)$, then the first order accuracy under $L^\infty(0,T;L^2(\Omega))$-norm is only achieved far away from the origin, and the global convergence rate degenerates as t_j approaches to zero; and if $u_0 \in \dot H^2(\Omega)$, then the global rate reduces to τ^α. These error estimates in (3) coincide with the solution regularity in Sobolev space (see Theorem 3.1):

$$\begin{cases} \epsilon^{3/2} \left| u \right|_{H^{1/2-\epsilon}(0,T;L^2(\Omega))} \leq C_{\alpha,T} \left\| u_0 \right\|_{L^2(\Omega)}, & 0 < \epsilon \leq 1/2, \\ \sqrt{2\gamma-\gamma^2} \left\| u \right\|_{H^{(1+\alpha\gamma)/2}(0,T;L^2(\Omega))} \leq C_{\alpha,T,\Omega} \left\| u_0 \right\|_{\dot H^\gamma(\Omega)}, & 0 < \gamma < 2. \end{cases}$$

To improve the temporal accuracy, graded meshes were used in [24,49,58] and some correction techniques were proposed in [13,22,31,63]. However, most of the existing works using graded meshes require some assumption of growth estimate on the true solution, and the analysis of correction schemes for (3) are mainly based on the Laplace transform, which is only applicable for uniform temporal grids, and the obtained convergence rates have the form $t_j^{-q}\tau^p$ with $0 < q \leq p$ (like (3)), which deteriorates near the origin. In [32], several technical stability results were developed to establish the optimal first order accuracy of a piecewise constant DG method on graded meshes. Also, spectral methods with singular basis functions were presented [7,55] but so far no rigorous convergence analysis is available with low regularity data. In [9], a multi-domain Petrov–Galerkin spectral method with a singular basis and geometrically graded meshes was proposed, and the exponential decay was verified numerically with nonsmooth initial data. We are also aware of the recent work [8], where an exponentially convergent rational approximation scheme has been proposed for the spatial semi-discretization of (1).

In the 1980s, Gui and Babuška [18] established the optimal approximation order $1 + 2\beta$ under the L^2-norm of the Legendre orthogonal expansion for the singular function $(x + 1)^\beta$ on $(-1, 1)$. Later Babuška and Suri [4] extended this to a p-version finite element method for solving two dimensional elliptic equations, and proved the sharp rate 2β under the energy norm, by assuming that the solution has the explicit singular expression r^β around the origin. Note that these functions have boundary singularities as well but the achieved convergence rates agree with their regularity in the Besov space.

In view of the boundary singularity of the solution to problem (1), one may wonder whether this happens to the convergence behavior of a time-spectral method. For simplicity, let us start with a fractional ordinary differential equation

$$\mathrm{D}_{0+}^{\alpha}(y - y_0) + \lambda y = 0 \quad \text{in} \quad (0, T], \tag{4}$$

where $y_0 \in \mathbb{R}$ and $\lambda > 0$. Invoking the Laplace transform gives the solution expression

$$y(t) = y_0 \sum_{k=0}^{\infty} \frac{(-\lambda t^\alpha)^k}{\Gamma(\alpha k + 1)}, \quad 0 \le t \le T. \tag{5}$$

Note that for a given fixed (small) $\lambda > 0$, we have $y \in H^{1/2+\alpha-\epsilon}(0, T)$ for any $\epsilon > 0$ (see Remark 3.1). We adopt a standard Legendre spectral method with polynomial degree $M \in \mathbb{N}$ to seek an approximation Y_M and use Y_{100} as a reference solution. Figure 1 plots the convergence order $1 + 2\alpha$ under L^2-norm in the case that $\lambda = y_0 = T = 1$. This agrees with the Besov regularity of (5): $\mathbb{B}_{-\alpha,0}^{1+2\alpha-\epsilon}(0, T)$ for any $\epsilon > 0$; see Lemma 3.4. However, if λ is extremely large or goes to infinity, then we see from Lemma 4.1 that the convergence rate will be ruined (we also refer the reader to [9, Section 1.2] for detailed numerical investigations in this case).

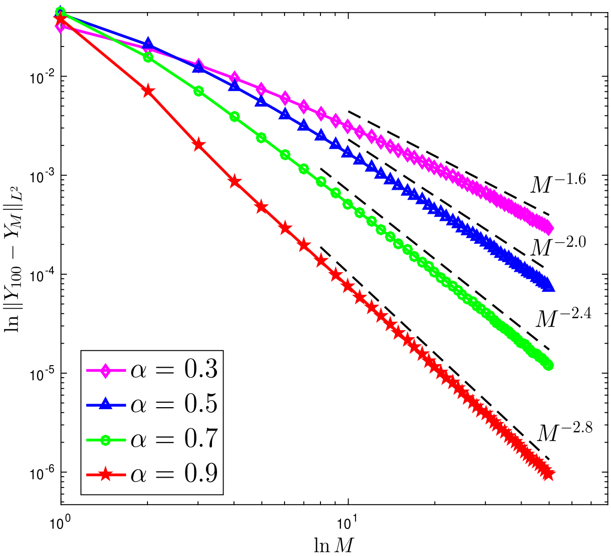

Fig. 1 Discretization errors of (4) with $\lambda = y_0 = T = 1$. The predicted accuracy is $M^{-1-2\alpha}$

As for the model problem (1) itself, although there exists a space-time spectral method that has been proposed in [35], to our best knowledge, no such rate $1+2\alpha$ has been mentioned numerically and established rigorously. It is nontrivial to obtain this since now the impact of large λ comes from the negative Laplacian operator $-\Delta$ (or its discrete version $-\Delta_h$). This motivates us to revisit the convergence analysis of the time-spectral method for time fractional diffusion problem (1). Can we prove the optimal approximation order in terms of Besov regularity with nonsmooth data? Especially, whether the accuracy $1 + 2\alpha$ can be established or improved?

In this work, we give partially positive answers to these questions mentioned above. Optimal error estimates with respect to the solution regularity in Besov space are established with low regularity data. Moreover, temporal convergence rates $1 + \alpha$ and $1 + 2\alpha$ under $H^{\alpha/2}(0, T; L^2(\Omega))$-norm and $L^2(0, T; \dot{H}^1(\Omega))$-norm are derived, respectively, which are sharp and cannot be improved even for smoother data.

The rest of this paper is organized as follows. In Sect. 2, we introduce some notations, including standard conventions, functional spaces and fractional calculus operators. Then in Sect. 3, we define the weak solution and establish its regularity results in Sobolev space and Besov space, and we present our main error estimates in Sect. 4. Finally, we conduct several numerical experiments in Sect. 5 and give some concluding remarks in Sect. 6.

2 Preliminary

For ease of notation, we make some standard conventions. For a Lebesgue measurable subset ω of $\mathbb{R}^l$ $(l = 1, 2, 3)$, we use $H^\gamma(\omega)$ $(\gamma \in \mathbb{R})$ and $H_0^\gamma(\omega)$ $(\gamma > 0)$ to denote two standard Sobolev spaces [59]. Given $1 \le p < \infty$, if ω is an interval and μ is a nonnegative measurable function on ω, then $L_\mu^p(\omega)$ denotes the weighted L^p-space, and the symbol $\langle a, b\rangle_\mu$ means $\int_\omega ab\mu$ whenever $ab \in L_\mu^1(\omega)$; if ω is a Lebesgue measurable set of $\mathbb{R}^l (l = 1, 2, 3, 4)$, then $\langle a, b\rangle_\omega$ stands for $\int_\omega ab$ whenever $ab \in L^1(\omega)$; if X is a Banach space, then $\langle \cdot, \cdot\rangle_X$ means the duality pairing between X^* (the dual space of X) and X. In particular, if X is a Hilbert space, then $(\cdot, \cdot)_X$ means its inner product. If X and Y are two Banach spaces, then $[X, Y]_{\theta,2}$ is the interpolation space constructed by the well-known K-method [5]. For $k \in \mathbb{N}$ and any d-polytope $\omega \subset \mathbb{R}^d (d = 1, 2, 3)$, $P_k(\omega)$ denotes the set of all polynomials defined on ω with degree no more than k.

As we all know, $L^2(\Omega)$ has an orthonormal basis $\{\phi_n\}_{n=0}^\infty$ such that [11]

$$\begin{cases} -\Delta\phi_n = \lambda_n\phi_n, & \text{in } \Omega,\\ \phi_n = 0, & \text{on } \partial\Omega, \end{cases}$$

where $\{\lambda_n\}_{n=0}^\infty$ is a nondecreasing real positive sequence and $\lambda_0 = \|\nabla\phi_0\|_{L^2(\Omega)}^2 > 0$ depends only on Ω. For any $\gamma \in \mathbb{R}$, define

$$\dot{H}^\gamma(\Omega) := \left\{ \sum_{n=0}^\infty c_n\phi_n : \sum_{n=0}^\infty \lambda_n^\gamma c_n^2 < \infty \right\}$$

and equip this space with the inner product

$$\left(\sum_{n=0}^\infty c_n\phi_n, \sum_{n=0}^\infty d_n\phi_n \right)_{\dot{H}^\gamma(\Omega)} := \sum_{n=0}^\infty \lambda_n^\gamma c_n d_n, \quad \text{for all } \sum_{n=0}^\infty c_n\phi_n, \sum_{n=0}^\infty d_n\phi_n \in \dot{H}^\gamma(\Omega).$$

The induced norm is denoted by $\|\cdot\|_{\dot H^\gamma(\Omega)} = \sqrt{(\cdot,\cdot)_{\dot H^\gamma(\Omega)}}$. Note that $\dot H^\gamma(\Omega)$ is a Hilbert space and has an orthonormal basis $\{\lambda_n^{-\gamma/2}\phi_n\}_{n=0}^\infty$.

Let ${}_0H^2(0,T) := \{v \in H^2(0,T) : v(0) = v'(0) = 0\}$ and equip this space with the norm $\|v\|_{{}_0H^2(0,T)} := \|v''\|_{L^2(0,T)}$. Given any $0 < \gamma < 2$, we introduce the space

$$ {}_0H^\gamma(0,T) := [L^2(0,T), {}_0H^2(0,T)]_{\gamma/2,2}. $$

Applying the interpolation theorem of bounded linear operators [39, Theorem 1.6] yields

$$ \|v\|_{[L^2(0,T),H^2(0,T)]_{\gamma/2,2}} \le \|v\|_{{}_0H^\gamma(0,T)} \quad \forall\, v \in {}_0H^\gamma(0,T). \tag{6} $$

In addition, if $0 < \gamma < 1/2$, then by [37, Chapter 1], the relation ${}_0H^\gamma(0,T) = H^\gamma(0,T)$ holds in the sense of equivalent norms, and in this case (i.e., $0 < \gamma < 1/2$) we have an alternative norm, which is defined by

$$ |w|_{H^\gamma(0,T)} := \left(\int_{\mathbb{R}} |\xi|^{2\gamma} \left|\mathcal{F}(w\chi_{(0,T)})(\xi)\right|^2 \,\mathrm{d}\xi \right)^{1/2} \quad \forall w \in H^\gamma(0,T), $$

where $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the Fourier transform and $\chi_{(0,T)}$ denotes the indicator function of $(0,T)$.

Let X be a separable Hilbert space with an inner product $(\cdot,\cdot)_X$ and an orthonormal basis $\{e_n : n \in \mathbb{N}\}$. For any $\gamma \in \mathbb{R}$, let $H^\gamma(0,T;X)$ be a usual vector-valued Sobolev space defined by (cf. [37, Section 1.3])

$$ H^\gamma(0,T;X) := \left\{ \sum_{n=0}^\infty v_n e_n : \sum_{n=0}^\infty \|v_n\|^2_{H^\gamma(0,T)} < \infty \right\}, \tag{7} $$

with the norm

$$ \|v\|_{H^\gamma(0,T;X)} := \left(\sum_{n=0}^\infty \|(v,e_n)_X\|^2_{H^\gamma(0,T)} \right)^{1/2} \quad \forall\, v \in H^\gamma(0,T;X). $$

The spaces $L^2_{\mu^{a,b}}(0,T;X)$ and ${}_0H^\gamma(0,T;X)$ for $0 < \gamma < 2$ can be defined similarly as (7).

For $a, b > -1$, let $\{S_k^{a,b}\}_{k=0}^\infty$ be the family of shifted Jacobi polynomials on $(0,T)$ with respect to the weight $\mu^{a,b}(t) = (T-t)^a t^b$; see "Appendix A". Given $\gamma \ge 0$, we introduce the Besov space (also known as the weighted Sobolev space, cf. [3]) defined by

$$ \mathbb{B}^\gamma_{a,b}(0,T) := \left\{ \sum_{k=0}^\infty v_k S_k^{a,b} : \sum_{k=0}^\infty (1+k^{2\gamma})\xi_k^{a,b} v_k^2 < \infty \right\}, \tag{8} $$

where $\xi_k^{a,b}$ is given by (80), and endow this space with the norm

$$ \|v\|_{\mathbb{B}^\gamma_{a,b}(0,T)} := \left(\sum_{k=0}^\infty (1+k^{2\gamma})\xi_k^{a,b} v_k^2 \right)^{1/2} \quad \forall\, v = \sum_{k=0}^\infty v_k S_k^{a,b} \in \mathbb{B}^\gamma_{a,b}(0,T). $$

In addition, for any separable Hilbert space X, the vector-valued space $\mathbb{B}^\gamma_{a,b}(0,T;X)$ can be defined in a similar way as that of (7).

To the end, let us introduce the Riemann–Liouville fractional calculus operators and list some important lemmas. For any $\gamma > 0$ and $v \in L^1(0,T;X)$, define the fractional integrals of order γ as follows:

$$\left(\mathrm{D}_{0+}^{-\gamma} v\right)(t) := \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} v(s)\,\mathrm{d}s, \quad t \in (0,T),$$
$$\left(\mathrm{D}_{T-}^{-\gamma} v\right)(t) := \frac{1}{\Gamma(\gamma)} \int_t^T (s-t)^{\gamma-1} v(s)\,\mathrm{d}s, \quad t \in (0,T),$$

where $\Gamma(\cdot)$ denotes the Gamma function $\Gamma(z) := \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t$ for $z > 0$. For $k-1 < \gamma < k$ with positive integer $k \in \mathbb{N}_+$, define the left-sided and right-sided Riemann–Liouville fractional derivative operators of order γ respectively by

$$\mathrm{D}_{0+}^{\gamma} := \mathrm{D}^k\,\mathrm{D}_{0+}^{\gamma-k}, \quad \mathrm{D}_{T-}^{\gamma} := (-\,\mathrm{D})^k\,\mathrm{D}_{T-}^{\gamma-k},$$

where D is the first-order generalized derivative operator.

Lemma 2.1 [10] *If* $-1/2 < \gamma < 1/2$ *and* $v,\ w \in H^{\max\{0,\gamma\}}(0,T)$, *then*

$$\left\langle \mathrm{D}_{0+}^{\gamma} v, \mathrm{D}_{T-}^{\gamma} v\right\rangle_{(0,T)} = \cos(\gamma\pi)\,|v|_{H^\gamma(0,T)}^2,$$
$$\cos(\gamma\pi)\left\|\mathrm{D}_{0+}^{\gamma} v\right\|_{L^2(0,T)}^2 \le \left\langle \mathrm{D}_{0+}^{\gamma} v, \mathrm{D}_{T-}^{\gamma} v\right\rangle_{(0,T)} \le \sec(\gamma\pi)\left\|\mathrm{D}_{0+}^{\gamma} v\right\|_{L^2(0,T)}^2,$$
$$\left\langle \mathrm{D}_{0+}^{2\gamma} v, w\right\rangle_{H^\gamma(0,T)} = \left\langle \mathrm{D}_{0+}^{\gamma} v, \mathrm{D}_{T-}^{\gamma} w\right\rangle_{(0,T)} \le |v|_{H^\gamma(0,T)}\,|w|_{H^\gamma(0,T)}.$$

Lemma 2.2 [40] *If* $v \in {}_0H^\gamma(0,T)$ *with* $0 < \gamma < 2$, *then*

$$C_1\left\|\mathrm{D}_{0+}^{\gamma} v\right\|_{L^2(0,T)} \le \|v\|_{{}_0H^\gamma(0,T)} \le C_2\left\|\mathrm{D}_{0+}^{\gamma} v\right\|_{L^2(0,T)},$$

where C_1 *and* C_2 *depend only on* γ.

3 Weak Solution and Regularity

This section is to revisit the solution regularity of problem (1) in terms of proper Sobolev spaces and establish new regularity results in Besov spaces.

Following [27,35], we first introduce the weak solution to problem (1). To do so, set

$$\mathcal{X} := H^{\alpha/2}(0,T;L^2(\Omega)) \cap L^2(0,T;\dot{H}^1(\Omega)), \tag{9}$$

and endow this space with the norm

$$\|\cdot\|_{\mathcal{X}} := \left(|\cdot|_{H^{\alpha/2}(0,T;L^2(\Omega))}^2 + \|\cdot\|_{L^2(0,T;\dot{H}^1(\Omega))}^2\right)^{1/2}.$$

Assuming that $f + \mathrm{D}_{0+}^{\alpha} u_0 \in \mathcal{X}^*$, we call $u \in \mathcal{X}$ a weak solution to problem (1) if

$$\left\langle \mathrm{D}_{0+}^{\alpha} u, v\right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla u, \nabla v\rangle_{L^2(0,T;L^2(\Omega))} = \left\langle f + \mathrm{D}_{0+}^{\alpha} u_0, v\right\rangle_{\mathcal{X}} \quad \forall\, v \in \mathcal{X}. \tag{10}$$

As mentioned in [27, Remark 2.2], the well-posedness of the weak formulation (10) follows from the Lax–Milgram theorem and Lemma 2.1. More precisely, if $f + \mathrm{D}_{0+}^{\alpha} u_0 \in \mathcal{X}^*$, then problem (1) admits a unique weak solution in the sense of (10) such that

$$\|u\|_{\mathcal{X}} \le C_\alpha \left\|f + \mathrm{D}_{0+}^{\alpha} u_0\right\|_{\mathcal{X}^*}.$$

In the sequel, set $\gamma_0 := \min\{2, 1/\alpha\}$. To establish more elaborate regularity estimates, we apply the Galerkin method that reduces (10) to a family of ordinary differential equations, to which the solutions can be used to recover the weak solution to (10) through a series expression; see the lemma below.

Proposition 3.1 *Assume $f \in L^2(0,T;\dot H^{-1}(\Omega))$ and $u_0 \in \dot H^\gamma(\Omega)$ where $\gamma > 1-\gamma_0$. The solution to (10) is given by $u = \sum_{n=0}^\infty y_n\phi_n$, where $y_n \in H^{\alpha/2}(0,T)$ satisfies*

$$\left\langle \mathrm{D}_{0+}^\alpha \left(y_n - y_{n,0}\right), z\right\rangle_{H^{\alpha/2}(0,T)} + \lambda_n \langle y_n, z\rangle_{(0,T)} = \langle f_n, z\rangle_{(0,T)}, \tag{11}$$

for all $z \in H^{\alpha/2}(0,T)$, where $y_{n,0} = \langle u_0, \phi_n\rangle_{\dot H^\gamma(\Omega)}$ and $f_n = \langle f, \phi_n\rangle_{\dot H^1(\Omega)}$.

Proof The proof here is actually in line with that of [27, Theorem 3.1], where the case $0 < \alpha < 1/2$ has been considered. The case of $1/2 \le \alpha < 1$ follows similarly. □

3.1 Regularity in Sobolev Space

We now revisit the Sobolev regularity of the solution to (10). Thanks to Proposition 3.1, this can be done by investigating problem (11), which, in a general form, is equivalent to

$$\mathrm{D}_{0+}^\alpha (y - y_0) + \lambda y = g, \tag{12}$$

where $\lambda > 0$, $y_0 \in \mathbb{R}$ and $g \in L^2(0,T)$. Indeed, in [27, Lemmas 3.1–3.3] we have established corresponding regularity results via variational approach, and Sobolev regularities of the weak solution have been given in [27, Theorems 3.1–3.3].

For the case $u_0 = 0$, if $f \in L^2(0,T;\dot H^\gamma(\Omega))$ with $-1 \le \gamma \le 1$, then

$$\|u\|_{{}_0H^{\alpha(1+\gamma/2)}(0,T;L^2(\Omega))} + \|u\|_{L^2(0,T;\dot H^{2+\gamma}(\Omega))} \le C_{\alpha,T}\,\|f\|_{L^2(0,T;\dot H^\gamma(\Omega))}, \tag{13}$$

and based on the proof of [27, Theorem 3.3], if $f \in {}_0H^\beta(0,T;\dot H^\gamma(\Omega))$ with $-1 \le \gamma \le 1$ and $0 < \beta \le 1$, then we have

$$\|u\|_{{}_0H^{\alpha+\beta}(0,T;L^2(\Omega))} + \|u\|_{{}_0H^\beta(0,T;\dot H^{2+\gamma}(\Omega))} \le C_{\alpha,\beta,T}\,\|f\|_{{}_0H^\beta(0,T;\dot H^\gamma(\Omega))}. \tag{14}$$

However, the implicit constant in [27, Lemma 3.2] blows up as the corresponding parameter θ goes to $1/\alpha - 1$ and the estimate in [27, Theorem 3.1] for the homogeneous case $f = 0$ is not optimal. Therefore, in this section, we provide some improved results by using the Mittag-Leffler function [47]

$$E_{\alpha,\nu}(z) := \sum_{k=0}^\infty \frac{z^k}{\Gamma(\alpha k + \nu)}, \quad z \in \mathbb{C}, \quad \nu \in \mathbb{R}. \tag{15}$$

Given any $t > 0$, it is well-known that (cf. [23])

$$\left|E_{\alpha,\nu}(-t)\right| \le \frac{C_{\alpha,\nu}}{1+t}. \tag{16}$$

In addition, applying Laplace transform gives the solution to (12) with $g = 0$:

$$y(t) = y_0 E_{\alpha,1}(-\lambda t^\alpha), \quad 0 \le t \le T. \tag{17}$$

Lemma 3.1 *Assume $\lambda \ge \lambda_* > 0$, then the function $y(t)$ defined by (17) satisfies*

$$\begin{aligned}&\eta(\epsilon)\lambda^{\gamma_0/2-1-\epsilon}\,\|y\|_{H^\alpha(0,T)} + \lambda^{(\gamma_0-1)/2}\,\|y\|_{H^{\alpha/2}(0,T)}\\ &+\eta(\epsilon)\lambda^{\gamma_0/2-\epsilon}\,\|y\|_{L^2(0,T)} \le C_{\alpha,\lambda_*,T}\,|y_0|,\end{aligned} \tag{18}$$

where

$$\begin{cases} \eta(\epsilon) = 1, & \epsilon = 0, & \text{if } \alpha \neq 1/2, \\ \eta(\epsilon) = \sqrt{\epsilon}, & \epsilon \in (0, 1/2], & \text{if } \alpha = 1/2. \end{cases} \tag{19}$$

Moreover, we have

$$\begin{cases} \epsilon^{3/2} \left| y \right|_{H^{1/2-\epsilon}(0,T)} \leq C_{\alpha,T} \left| y_0 \right|, \quad 0 < \epsilon \leq 1/2, & (20) \\ \sqrt{2\gamma - \gamma^2} \left\| y \right\|_{H^{(1+\alpha\gamma)/2}(0,T)} \leq C_{\alpha,\lambda_*} \lambda^{\gamma/2} \left| y_0 \right|, \quad 0 < \gamma < 2. & (21) \end{cases}$$

Proof We first prove (18). Since the case $0 < \alpha < 1/2$ has been given by [27, Lemma 3.1], we only consider the case $1/2 \leq \alpha < 1$, which says that

$$\eta(\epsilon)\lambda^{-1-\epsilon} \left\| y \right\|_{H^\alpha(0,T)} + \lambda^{-1/2} \left\| y \right\|_{H^{\alpha/2}(0,T)} + \eta(\epsilon)\lambda^{-\epsilon} \left\| y \right\|_{L^2(0,T)} \leq C_{\alpha,\lambda_*,T} \lambda^{-\frac{1}{2\alpha}} \left| y_0 \right|. \tag{22}$$

By (16) and direct calculations, we have

$$\begin{aligned} &\left| \eta(\epsilon) \right|^2 \lambda^{-2-2\epsilon} \left| \left(\mathrm{D}_{0+}^\alpha (y - y_0) \right)(t) \right|^2 + \lambda^{-1} \left| \left(\mathrm{D}_{0+}^{\alpha/2} y \right)(t) \right|^2 + \left| \eta(\epsilon) \right|^2 \lambda^{-2\epsilon} \left| y(t) \right|^2 \\ &\quad \leq C_\alpha \left| y_0 \right|^2 \frac{\left| \eta(\epsilon) \right|^2 \lambda^{-2\epsilon} + \lambda^{-1} t^{-\alpha}}{(1 + \lambda t^\alpha)^2}, \end{aligned} \tag{23}$$

for all $0 < t \leq T$. It is evident that

$$\begin{aligned} \int_0^T \frac{\lambda^{-1} t^{-\alpha}}{(1 + \lambda t^\alpha)^2} \, \mathrm{d}t &\leq \int_0^\infty \frac{\lambda^{-1} t^{-\alpha}}{(1 + \lambda t^\alpha)^2} \, \mathrm{d}t \leq \int_0^{\lambda^{-\frac{1}{\alpha}}} \lambda^{-1} t^{-\alpha} \, \mathrm{d}t + \int_{\lambda^{-\frac{1}{\alpha}}}^\infty \lambda^{-3} t^{-3\alpha} \, \mathrm{d}t \\ &= C_\alpha \lambda^{-1/\alpha}. \end{aligned} \tag{24}$$

If $\alpha = 1/2$, then for $0 < \epsilon \leq 1/2$ it holds

$$\int_0^T \frac{\lambda^{-2\epsilon}}{(1 + \lambda t^\alpha)^2} \, \mathrm{d}t = \lambda^{-\frac{1}{\alpha}} \int_0^T \frac{(\lambda t^\alpha)^{1/\alpha - 2\epsilon}}{(1 + \lambda t^\alpha)^2} t^{2\alpha\epsilon - 1} \, \mathrm{d}t \leq \lambda^{-\frac{1}{\alpha}} \int_0^T t^{2\alpha\epsilon - 1} \, \mathrm{d}t = \frac{C_{\alpha,T}}{\epsilon} \lambda^{-1/\alpha};$$

and if $1/2 < \alpha < 1$, then using a similar manner for estimating (24) gives

$$\int_0^T \frac{1}{(1 + \lambda t^\alpha)^2} \, \mathrm{d}t \leq C_\alpha \lambda^{-1/\alpha}.$$

Hence, by Lemma 2.2, plugging the above estimates into (23) implies

$$\eta(\epsilon)\lambda^{-1-\epsilon} \left\| y - y_0 \right\|_{{}_0H^\alpha(0,T)} + \lambda^{-1/2} \left\| y \right\|_{{}_0H^{\alpha/2}(0,T)} + \eta(\epsilon)\lambda^{-\epsilon} \left\| y \right\|_{L^2(0,T)} \leq C_{\alpha,T} \lambda^{-\frac{1}{2\alpha}} \left| y_0 \right|,$$

which, together with the fact (6) and the assumption $\lambda \geq \lambda_* > 0$, yields (22) immediately.

If $0 < \epsilon \leq 1/2$, then using Lemma 2.1 and 16 gives

$$\begin{aligned} \left| y \right|^2_{H^{1/2-\epsilon}(0,T)} \leq \csc^2 \epsilon\pi \left\| \mathrm{D}_{0+}^{1/2-\epsilon} y \right\|^2_{L^2(0,T)} &\leq \frac{C_\alpha \left| y_0 \right|^2}{\epsilon^2} \int_0^T \frac{t^{2\epsilon - 1}}{(1 + \lambda t^\alpha)^2} \, \mathrm{d}t \\ &\leq \frac{C_\alpha \left| y_0 \right|^2}{\epsilon^2} \int_0^T t^{2\epsilon - 1} \, \mathrm{d}t = \frac{C_{\alpha,T} \left| y_0 \right|^2}{\epsilon^3}, \end{aligned}$$

which proves (20).

To the end, we consider (21). Since $0 < \gamma < 2$, we have $1/2 < (1+\alpha\gamma)/2 < \alpha + 1/2$. By (16), Lemma 2.2 and straightforward calculations, we get

$$\begin{aligned}\|y - y_0\|^2_{{}_0H^{(1+\alpha\gamma)/2}(0,T)} &\le C_\alpha \int_0^T \left|\left(\mathrm{D}_{0+}^{(1+\alpha\gamma)/2}(y-y_0)\right)(t)\right|^2 \mathrm{d}t\\ &\le C_\alpha |y_0|^2 \int_0^T \frac{\lambda^2 t^{\alpha(2-\gamma)-1}}{(1+\lambda t^\alpha)^2}\,\mathrm{d}t,\end{aligned}$$

and we estimate the integral in a similar way for (24) to obtain

$$\|y - y_0\|_{{}_0H^{(1+\alpha\gamma)/2}(0,T)} \le \frac{C_\alpha \lambda^{\gamma/2} |y_0|}{\sqrt{2\gamma - \gamma^2}}.$$

Therefore, from (6) and the assumption $\lambda \ge \lambda_* > 0$ it follows (21). □

Remark 3.1 For any fixed $\lambda > 0$, from Lemma 3.1 we see that the highest regularity of $y = y_0 E_{\alpha,1}(-\lambda t^\alpha)$ is no more than $H^{1/2+\alpha}(0,T)$. Although we can establish higher regularity (cf. [34]) for the smooth part $Y = y - y_0 S_\lambda$, where

$$S_\lambda(t) := 1 - \frac{\lambda t^\alpha}{\Gamma(\alpha+1)}, \quad 0 \le t \le T,$$

the final regularity for $y = Y + y_0 S_\lambda$ is dominated by S_λ, which belongs to $H^{1/2+\alpha-\epsilon}(0,T)$ due to the singular term t^α.

Combining Proposition 3.1 and Lemma 3.1 gives the following conclusion.

Theorem 3.1 *If $f = 0$ and $u_0 \in \dot H^\gamma(\Omega)$ with $\gamma > 1 - \gamma_0$, then*

$$\begin{aligned}&\eta(\epsilon)\,\|u\|_{H^\alpha(0,T;\dot H^{\gamma_0+\gamma-2-\epsilon}(\Omega))} + \|u\|_{H^{\alpha/2}(0,T;\dot H^{\gamma_0+\gamma-1}(\Omega))} + \eta(\epsilon)\,\|u\|_{L^2(0,T;\dot H^{\gamma_0+\gamma-\epsilon}(\Omega))}\\ &\quad \le C_{\alpha,T,\Omega}\,\|u_0\|_{\dot H^\gamma(\Omega)},\end{aligned}$$

where ϵ, $\eta(\epsilon)$ are defined by (19). Moreover,

$$\begin{cases} \epsilon^{3/2}\,|u|_{H^{1/2-\epsilon}(0,T;L^2(\Omega))} \le C_{\alpha,T}\,\|u_0\|_{L^2(\Omega)}, & 0 < \epsilon \le 1/2,\\ \sqrt{2\gamma-\gamma^2}\,\|u\|_{H^{(1+\alpha\gamma)/2}(0,T;L^2(\Omega))} \le C_{\alpha,T,\Omega}\,\|u_0\|_{\dot H^\gamma(\Omega)}, & 0 < \gamma < 2,\\ \sqrt{2\gamma-\gamma^2}\,\|u\|_{H^{(1+\alpha\gamma)/2}(0,T;\dot H^1(\Omega))} \le C_{\alpha,T,\Omega}\,\|u_0\|_{\dot H^{\gamma+1}(\Omega)}, & 0 < \gamma < 2.\end{cases}$$

3.2 Regularity in Besov Space

We then consider the regularity of the solution to (10) in Besov spaces. As before, we start from the auxiliary problem (12) and split it into two cases: $y_0 \ne 0,\ g = 0$; and $y_0 = 0,\ g \ne 0$.

To move on, let us present a useful expression of the Mittag-Leffler function (15); see [15, Theorem 2.1].

Lemma 3.2 [15] *If $\nu < 1 + \alpha$, then for all $0 < t < \infty$, it holds that*

$$E_{\alpha,\nu}(-t) = \frac{1}{\pi\alpha}\int_0^\infty \frac{r\sin\nu\pi - t\sin(\alpha-\nu)\pi}{r^2 + t^2 + 2rt\cos\alpha\pi}\, r^{(1-\nu)/\alpha} e^{-r^{1/\alpha}}\,\mathrm{d}r.$$

Clearly, by definition (15), we have $E_{\alpha,\nu}(0) = 1/\Gamma(\nu)$ for any $\nu \in \mathbb{R}\backslash\{-1,-2,\ldots\}$ and thus $E_{\alpha,\nu}(-t)$ is bounded near $t = 0+$. Below, we give a refined estimate that implies the asymptotic behavior of $E_{\alpha,\nu}(-t)$. Note that this has no contraction with the boundness around $t = 0+$ and what we are interested in is $t \to \infty$.

Lemma 3.3 *If* $\nu < 1$, *then for all* $0 < t < \infty$,

$$\left|E_{\alpha,\nu}(-t)\right| \le C_\alpha \Gamma(1-\nu+\theta\alpha) t^{-\theta}, \tag{25}$$

where $0 \le \theta \le 1$. *Moreover, if* $\alpha - \nu \in \mathbb{Z}$, *then* (25) *holds with* $0 \le \theta \le 2$.

Proof By Lemma 3.2, we have

$$E_{\alpha,\nu}(-t) = \frac{1}{\pi\alpha}\int_0^\infty r^{(1-\nu)/\alpha} e^{-r^{1/\alpha}} \varphi_{\alpha,\nu}(r,t)\,\mathrm{d}r, \tag{26}$$

where

$$\varphi_{\alpha,\nu}(r,t) := \frac{r\sin\nu\pi - t\sin(\alpha-\nu)\pi}{r^2+2rt\cos\alpha\pi + t^2}.$$

In light of the estimate

$$r^2 + 2rt\cos\alpha\pi + t^2 \ge \frac{1+\cos\alpha\pi}{2}(r+t)^2 \tag{27}$$

and the arithmetic-geometric mean inequality (cf. [57, page 4])

$$r^{1-\theta}t^\theta \le (1-\theta)r + \theta t, \quad 0 \le \theta \le 1, \tag{28}$$

we obtain that

$$\left|\varphi_{\alpha,\nu}(r,t)\right| \le \frac{C_\alpha}{r+t} \le C_\alpha r^{\theta-1}t^{-\theta}. \tag{29}$$

Therefore, inserting (29) into (26) gives

$$\begin{aligned}\left|E_{\alpha,\nu}(-t)\right| &\le C_\alpha t^{-\theta}\int_0^\infty e^{-r^{1/\alpha}}(r^{1/\alpha})^{(\theta-1)\alpha+1-\nu}\,\mathrm{d}r \\ &= C_\alpha t^{-\theta}\int_0^\infty e^{-s}s^{\theta\alpha-\nu}\,\mathrm{d}s = C_\alpha\Gamma(1+\theta\alpha-\nu)t^{-\theta},\end{aligned}$$

which establishes (25). Moreover, if $\alpha - \nu \in \mathbb{Z}$, then

$$\varphi_{\alpha,\nu}(r,t) = \frac{r\sin\nu\pi}{r^2+2rt\cos\alpha\pi+t^2},$$

and from (27) and (28) it follows

$$\left|\varphi_{\alpha,\nu}(r,t)\right| \le C_\alpha r^{\theta-1}t^{-\theta}, \quad 0 \le \theta \le 2,$$

which maintains the estimate (25) and enlarges the range of θ. This completes the proof. □

Based on Lemma 3.3, we are able to establish the following lemma.

Lemma 3.4 *Assume* $-1 \le \theta \le 1$ *with* $1 + 2\alpha\theta > 0$, *then the function* y *defined by* (17) *satisfies*

$$\|y\|_{\mathbb{B}^{1+2\alpha\theta-\epsilon}_{-\alpha,0}(0,T)} \le \frac{C_{\alpha,T}}{\sqrt{\epsilon}}\lambda^\theta\,|y_0|\,, \tag{30}$$

for any $0 < \epsilon \le 1 + 2\alpha\theta$.

Proof By (16), it is evident that $y \in L^2_{\mu^{-\alpha,0}}(0,T)$ and we have the orthogonal decomposition $y = \sum_{k=0}^{\infty} y_k S_k^{-\alpha,0}$, where $\{S_k^{-\alpha,0}\}_{k=0}^{\infty}$ denotes the shifted Jacobi polynomials on $(0,T)$ with respect to the weight $\mu^{-\alpha,0}(t) = (T-t)^{-\alpha}$ (see "Appendix A"), and

$$y_k = \frac{1}{\xi_k^{-\alpha,0}} \left\langle y, S_k^{-\alpha,0} \right\rangle_{\mu^{-\alpha,0}}, \quad \xi_k^{-\alpha,0} = \frac{T^{1-\alpha}}{2k+1-\alpha}. \tag{31}$$

According to the definition (8), it suffices to investigate the asymptotic behavior of the coefficient y_k. Let us fix $k \in \mathbb{N}_+$. By Rodrigues' formula (79), we have

$$\left\langle y, S_k^{-\alpha,0} \right\rangle_{\mu^{-\alpha,0}} = \frac{(-1)^k}{T^k k!} \left\langle y, \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mu^{k-\alpha,k} \right\rangle_{(0,T)}. \tag{32}$$

Using integration by parts gives

$$\left\langle y, \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mu^{k-\alpha,k} \right\rangle_{(0,T)} = \sum_{i=0}^{k-1} (-1)^i \big(\zeta_i(T) - \zeta_i(0)\big) + (-1)^k \left\langle y^{(k)}, \mu^{k-\alpha,k} \right\rangle_{(0,T)},$$

where

$$\zeta_i(t) = y^{(i)}(t) \left(\frac{\mathrm{d}^{k-i-1}}{\mathrm{d}t^{k-i-1}} \mu^{k-\alpha,k} \right)(t) = y^{(i)}(t) \sum_{j=0}^{k-i-1} C_{\alpha,i,j,k} \mu^{k-\alpha-j,i+1+j}(t).$$

From (17) we have the identity

$$y^{(i)}(t) = -\lambda y_0 t^{\alpha-i} E_{\alpha,\alpha+1-i}(-\lambda t^\alpha), \quad 0 \le i < k, \tag{33}$$

and it follows that $\zeta_i(0) = \zeta_i(T) = 0$ for all $0 \le i < k$. Thus, we obtain the relation

$$\left\langle y, \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mu^{k-\alpha,k} \right\rangle_{(0,T)} = (-1)^k \left\langle y^{(k)}, \mu^{k-\alpha,k} \right\rangle_{(0,T)}. \tag{34}$$

Invoking (33) and Lemma 3.3 yields the inequality

$$\left| y^{(k)}(t) \right| \le C_\alpha \Gamma(k-\theta\alpha) \left| y_0 \right| \lambda^\theta t^{\theta\alpha-k}, \tag{35}$$

where $-1 \le \theta \le 1$. In addition, we have a useful formula (cf. [56, Appendix, (A.6)])

$$\left\| \mu^{a,b} \right\|_{L^1(0,T)} = T^{a+b+1} \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)}, \quad a, b > -1, \tag{36}$$

which, together with (35), indicates that

$$\begin{aligned} \left| \left\langle y^{(k)}, \mu^{k-\alpha,k} \right\rangle_{(0,T)} \right| &\le C_\alpha \left| y_0 \right| \lambda^\theta \Gamma(k-\theta\alpha) \left\| \mu^{k-\alpha,\theta\alpha} \right\|_{L^1(0,T)} \\ &= C_{\alpha,T} \left| y_0 \right| \lambda^\theta T^k \cdot \frac{\Gamma(k+1-\alpha)\Gamma(k-\theta\alpha)}{\Gamma(k+2+(\theta-1)\alpha)}. \end{aligned} \tag{37}$$

Observe Stirling's formula [1, Eq. (6.1.38)]

$$\Gamma(z+1) = \sqrt{2\pi} z^{z+1/2} e^{\theta(z)/(12z)-z}, \quad 0 < \theta(z) < 1, \ z > 0. \tag{38}$$

Therefore, collecting (32),(34) and (37) gives

$$\left| \left\langle y, S_k^{-\alpha,0} \right\rangle_{\mu^{-\alpha,0}} \right| \le C_{\alpha,T} \left| y_0 \right| \lambda^\theta \frac{\Gamma(k+1-\alpha)\Gamma(k-\theta\alpha)}{k!\Gamma(k+2+(\theta-1)\alpha)} \le C_{\alpha,T} \left| y_0 \right| \lambda^\theta k^{-2-2\theta\alpha}. \tag{39}$$

As a result, for any $0 < \epsilon \le 1 + 2\alpha\theta$, we have

$$\begin{aligned}
&\|y\|^2_{\mathbb{B}^{1+2\alpha\theta-\epsilon}_{-\alpha,0}(0,T)} \\
&= \sum_{k=0}^{\infty} \frac{1 + k^{2+4\alpha\theta-2\epsilon}}{\xi_k^{-\alpha,0}} \left| \left\langle y, S_k^{-\alpha,0} \right\rangle_{\mu^{-\alpha,0}} \right|^2 \le C_{\alpha,T} \, |y_0|^2 \, \lambda^{2\theta} \left(1 + \sum_{k=1}^{\infty} k^{-1-2\epsilon} \right) \\
&\le C_{\alpha,T} \, |y_0|^2 \, \lambda^{2\theta} \left(1 + \int_1^{\infty} r^{-1-2\epsilon} \, \mathrm{d}r \right) \le \frac{C_{\alpha,T}}{\epsilon} \, |y_0|^2 \, \lambda^{2\theta},
\end{aligned} \tag{40}$$

which shows (30) and finishes the proof of this lemma. □

Remark 3.2 As mentioned in Remark 3.1, the best regularity of $y(t) = y_0 E_{\alpha,1}(-\lambda t^\alpha)$ is no more than $H^{1/2+\alpha}(0, T)$. However, from Lemma 3.4 we know that $y \in \mathbb{B}^{1+2\alpha-\epsilon}_{-\alpha,0}(0, T)$, which can not be improved due to the singular term t^α, and the optimal rate $1 + 2\alpha$ of the standard Legendre spectral method under L^2-norm has been validated numerically in Fig. 1.

For another case: $y_0 = 0,\ g \ne 0$, according to [23, Proposition 5.10], the solution y to (12) can be represented as follows

$$y(t) = \int_0^t (t - s)^{\alpha-1} E_{\alpha,\alpha}(-\lambda(t - s)^\alpha) g(s) \, \mathrm{d}s, \quad 0 < t \le T, \tag{41}$$

and by the proof of [27, Lemma 3.3], if $g \in {}_0H^\beta(0, T)$ with $\beta \ge 0$, then we have

$$\|y\|_{{}_0H^{\alpha+\beta}(0,T)} + \lambda \, \|y\|_{L^2(0,T)} \le C_{\alpha,\beta,T} \, \|g\|_{{}_0H^\beta(0,T)} \, .$$

In particular, for $g(t) = g_0 t^\sigma$ with $\sigma > -1/2$, a direct computation gives

$$y(t) = \Gamma(\sigma + 1) g_0 t^{\alpha+\sigma} E_{\alpha,\alpha+\sigma+1}(-\lambda t^\alpha), \quad 0 < t \le T. \tag{42}$$

Analogous to Lemma 3.4, we have the following regularity estimate.

Lemma 3.5 *Assume $\sigma > -1/2$ and $0 \le \theta \le 1$, then the function y defined by (42) satisfies*

$$\|y\|_{\mathbb{B}^{1+2\sigma+2\alpha\theta-\epsilon}_{-\alpha,0}(0,T)} \le \frac{C_{\alpha,\sigma,T}}{\sqrt{\epsilon}} \, |g_0| \, \lambda^{\theta-1}, \tag{43}$$

for all $0 < \epsilon \le 1 + 2\sigma$. Moreover, if $\alpha + \sigma \in \mathbb{N}$, then we can take $\theta \in [0, 2]$.

Proof Since $\sigma > -1/2$, again by (16), we have $y \in L^2_{\mu^{-\alpha,0}}(0, T)$, and we have the orthogonal decomposition $y = \sum_{k=0}^{\infty} y_k S_k^{-\alpha,0}$, where y_k and $\xi_k^{-\alpha,0}$ are defined in (31).

If $\alpha + \sigma \notin \mathbb{N}$, then

$$y^{(k)}(t) = \Gamma(\sigma + 1) g_0 t^{\alpha+\sigma-k} E_{\alpha,\alpha+\sigma+1-k}(-\lambda t^\alpha) \quad \forall k \in \mathbb{N}. \tag{44}$$

For $k > \alpha + \sigma$, invoking Lemma 3.3 yields the estimate:

$$\left| y^{(k)}(t) \right| \le C_\alpha \Gamma(\sigma + 1) \, |g_0| \, \lambda^{\theta-1} t^{\theta\alpha+\sigma-k} \Gamma(k - \sigma - \theta\alpha), \quad 0 < t \le T, \tag{45}$$

where $0 \le \theta \le 1$. Hence, by using Rodrigues' formula (79) and integration by parts, a similar argument as that for (34) gives

$$\left\langle y, S_k^{-\alpha,0} \right\rangle_{\mu^{-\alpha,0}} = \frac{1}{T^k k!} \left\langle y^{(k)}, \mu^{k-\alpha,k} \right\rangle_{(0,T)}, \tag{46}$$

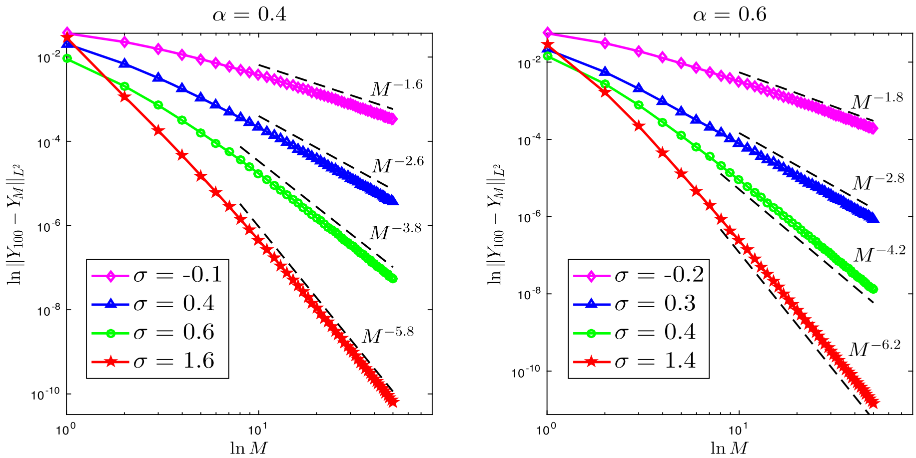

Fig. 2 Discretization errors of (12) with $\lambda = T = 1$, $y_0 = 0$ and $g(t) = t^\sigma$. The predicted accuracies are $M^{-1-2\sigma-2\alpha}$ and $M^{-1-2\sigma-4\alpha}$, respectively for $\sigma + \alpha \notin \mathbb{N}$ and $\sigma + \alpha \in \mathbb{N}$

and from (36) and (45) we obtain

$$\left|\frac{1}{T^k k!}\left\langle y^{(k)}, \mu^{k-\alpha,k}\right\rangle_{(0,T)}\right| \le C_{\alpha,\sigma}\,|g_0|\,\lambda^{\theta-1}\frac{\Gamma(k-\sigma-\theta\alpha)}{T^k\Gamma(k+1)}\left\|\mu^{k-\alpha,\sigma+\theta\alpha}\right\|_{L^1(0,T)}$$
$$\le C_{\alpha,\sigma,T}\,|g_0|\,\lambda^{\theta-1}\cdot\frac{\Gamma(k-\alpha+1)}{\Gamma(k+1)}\cdot\frac{\Gamma(k-\sigma-\theta\alpha)}{\Gamma(k-\alpha+\sigma+\theta\alpha+2)}.$$

This together with (38) implies that

$$\left|\left\langle y, S_k^{-\alpha,0}\right\rangle_{\mu^{-\alpha,0}}\right| \le C_{\alpha,\sigma,T}\,|g_0|\,\lambda^{\theta-1}k^{-2-2\sigma-2\theta\alpha}. \tag{47}$$

Note that for $k < \alpha + \sigma$, using (16),(28) and (44) gives

$$\left|y^{(k)}(t)\right| \le C_\alpha\Gamma(\sigma+1)\,|g_0|\,\lambda^{\theta-1}t^{\theta\alpha+\sigma-k}, \quad 0 < t \le T,$$

where $0 \le \theta \le 1$. Thus the estimate (47) holds true for all $k \in \mathbb{N}_+$ and $\theta \in [0, 1]$.

On the other hand, if $\alpha + \sigma \in \mathbb{N}$, then we find that for all $k \in \mathbb{N}$,

$$y^{(k)}(t) = \begin{cases} \Gamma(\sigma+1)g_0t^{\alpha+\sigma-k}E_{\alpha,\alpha+\sigma+1-k}(-\lambda t^\alpha), & \text{if } k \le \alpha+\sigma,\\ -\lambda\Gamma(\sigma+1)g_0t^{2\alpha+\sigma-k}E_{\alpha,2\alpha+\sigma+1-k}(-\lambda t^\alpha), & \text{if } k > \alpha+\sigma.\end{cases}$$

Thus, we still have the identity (46) and by Lemma 3.3, the estimate (47) holds true for all $k \in \mathbb{N}_+$ and $\theta \in [0, 2]$. Hence, using the proof of (40) leads to the desired estimate (43). □

To illustrate the maximal regularity in Lemma 3.5, we adopt a standard Legendre spectral method to solve (12) with $\lambda = T = 1$, $y_0 = 0$ and $g(t) = t^\sigma$. In Fig. 2, we report the numerical results under L^2-norm and observe the convergence rates $M^{-1-2\sigma-2\alpha}$ and $M^{-1-2\sigma-4\alpha}$, respectively for $\alpha + \sigma \notin \mathbb{N}$ and $\alpha + \sigma \in \mathbb{N}$.

Finally, recall Proposition 3.1, which says $u = \sum_{n=0}^{\infty} y_n\phi_n$ with y_n solving the ordinary differential equation (11), and gathering Lemmas 3.4 and 3.5 implies the following regularity results.

Theorem 3.2 *If $f = 0$ and $u_0 \in \dot H^\gamma(\Omega)$ with $\gamma > 1 - \gamma_0$, then*

$$\|u\|_{\mathbb{B}^{1+\alpha(\gamma-\beta)-\epsilon}_{-\alpha,0}(0,T;\dot H^\beta(\Omega))} \leq \frac{C_{\alpha,T}}{\sqrt{\epsilon}} \|u_0\|_{\dot H^\gamma(\Omega)}, \tag{48}$$

where $\gamma - 2 \leq \beta \leq \gamma + 2$ and $0 < \epsilon \leq 1 + \alpha(\gamma - \beta)$.

Proof By definition, we have

$$\|u\|^2_{\mathbb{B}^{1+\alpha(\gamma-\beta)-\epsilon}_{-\alpha,0}(0,T;\dot H^\beta(\Omega))} = \sum_{n=0}^{\infty} \lambda_n^\beta \|y_n\|^2_{\mathbb{B}^{1+\alpha(\gamma-\beta)-\epsilon}_{-\alpha,0}(0,T)}.$$

Since $u_0 = \sum_{n=0}^\infty y_{n,0}\phi_n$ and $y_n = y_{n,0}E_{\alpha,1}(-\lambda_n t^\alpha)$, it follows from Lemma 3.4 that

$$\|u\|^2_{\mathbb{B}^{1+\alpha(\gamma-\beta)-\epsilon}_{-\alpha,0}(0,T;\dot H^\beta(\Omega))} \leq \frac{C_{\alpha,T}}{\epsilon} \sum_{n=0}^{\infty} \lambda_n^\gamma \left|y_{n,0}\right|^2 = \frac{C_{\alpha,T}}{\epsilon} \|u_0\|^2_{\dot H^\gamma(\Omega)},$$

provided that $-1 \leq (\gamma - \beta)/2 \leq 1$ and $0 < \epsilon \leq 1 + \alpha(\gamma - \beta)$. □

Theorem 3.3 *Assume $\sigma > -1/2$ and $\gamma \geq -1$. If $u_0 = 0$ and $f = t^\sigma v$ with $v \in \dot H^\gamma(\Omega)$, then*

$$\|u\|_{\mathbb{B}^{1+2\sigma+\alpha(2+\gamma-\beta)-\epsilon}_{-\alpha,0}(0,T;\dot H^\beta(\Omega))} \leq \frac{C_{\alpha,\sigma,T}}{\sqrt{\epsilon}} \|v\|_{\dot H^\gamma(\Omega)},$$

where $\gamma \leq \beta \leq 2+\gamma$ and $0 < \epsilon \leq 1+2\sigma$. Particularly, if $\alpha+\sigma \in \mathbb{N}$, then $\gamma-2 \leq \beta \leq \gamma+2$.

4 Discretization and Error Analysis

Let $\mathcal{K}_h$ be a conventional conforming and shape regular simplical triangulation of Ω that consists of d-simplexes, and we use h to denote the maximum diameter of the elements in $\mathcal{K}_h$. Define

$$X_h := \left\{ v_h \in \dot H^1(\Omega) : \ v_h|_K \in P_1(K) \quad \forall K \in \mathcal{K}_h \right\}.$$

In the sequel, let $\Omega_T := \Omega \times (0, T)$ and $M \in \mathbb{N}$. The time-spectral method for problem (1) reads as follows: find $U \in P_M(0, T) \otimes X_h$ such that

$$\left\langle \mathrm{D}^\alpha_{0+} U, V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla U, \nabla V \rangle_{\Omega_T} = \left\langle f + \mathrm{D}^\alpha_{0+} u_0, V \right\rangle_{P_M(0,T)\otimes X_h} \tag{49}$$

for all $V \in P_M(0, T) \otimes X_h$. It is easy to see that (49) admits a unique solution $U \in P_M(0, T) \otimes X_h$ such that

$$\|U\|_{\mathcal{X}} \leq C_\alpha \left\| f + \mathrm{D}^\alpha_{0+} u_0 \right\|_{\mathcal{X}^*},$$

where the space $\mathcal{X}$ and its norm $\|\cdot\|_{\mathcal{X}}$ are defined in Sect. 3.

We now present our main error estimates. Note that all of the following results are optimal. Also, we emphasis that, in Theorem 4.1, the best temporal convergence orders $1 + \alpha$ and $1 + 2\alpha$ under $H^{\alpha/2}(0, T; L^2(\Omega))$-norm and $L^2(0, T; \dot H^1(\Omega))$-norm are sharp and cannot be improved even for smoother initial data. As for Theorem 4.2, we restrict ourselves to a special case $f(t, x) = t^\sigma v(x)$, and the obtained temporal convergence rates $1 + 2\sigma + \alpha$ and $1 + 2\sigma + 2\alpha$ are also sharp even for smoother v. Those predictions are verified by our numerical results in Sect. 5.

Theorem 4.1 *If $f = 0$ and $u_0 \in \dot{H}^\gamma(\Omega)$ with $1 - \gamma_0 < \gamma \leq 3$, then*

$$\|u - U\|_{L^2(0,T;\dot{H}^1(\Omega))} \lesssim \left(\epsilon_h h^{\min\{1,\gamma_0+\gamma-1\}} + M^{-1-\alpha(\gamma-1)}\right) \|u_0\|_{\dot{H}^\gamma(\Omega)}, \tag{50}$$

$$\|u - U\|_{H^{\alpha/2}(0,T;L^2(\Omega))} \lesssim \left(h^{\min\{2,\gamma_0+\gamma-1\}} + M^{-1-\alpha\min\{1,\gamma-1\}}\right) \|u_0\|_{\dot{H}^\gamma(\Omega)}, \tag{51}$$

where $\epsilon_h = 1$ if $\alpha \neq 1/2$, and $\epsilon_h = \sqrt{|\ln h|}$ if $\alpha = 1/2$.

Theorem 4.2 *Assume $-1 \leq \gamma \leq 1$ and $\sigma > (\alpha - 1)/2$. If $u_0 = 0$ and $f(t,x) = t^\sigma v(x)$ with $v \in \dot{H}^\gamma(\Omega)$, then*

$$\|u - U\|_{H^{\alpha/2}(0,T;L^2(\Omega))} \lesssim \left(h^{\min\{2,\gamma+2\}} + M^{-1-2\sigma-\alpha\min\{1,\gamma+1\}}\right) \|v\|_{\dot{H}^\gamma(\Omega)}, \tag{52}$$

$$\|u - U\|_{L^2(0,T;\dot{H}^1(\Omega))} \lesssim \left(h^{\min\{1,\gamma+1\}} + M^{-1-2\sigma-\alpha(\gamma+1)}\right) \|v\|_{\dot{H}^\gamma(\Omega)}, \tag{53}$$

In addition, if $\alpha + \sigma \in \mathbb{N}$ and $v \in \dot{H}^\gamma(\Omega)$ with $-1 \leq \gamma \leq 3$, then (53) still holds true and the temporal rate in (52) shall be $M^{-1-2\sigma-\alpha\min\{3,\gamma+1\}}$.

Remark 4.1 By Remark 4.4 and the proof of Theorem 4.1, if the true solution is $u(x,t) = t^\sigma\phi(x)$ with $\sigma > (\alpha-1)/2$ and $\phi \in \dot{H}^2(\Omega)$, then we have

$$\begin{aligned} \|u - U\|_{L^2(0,T;\dot{H}^1(\Omega))} &\lesssim \left(h + M^{-1-2\sigma}\right) \|\phi\|_{\dot{H}^2(\Omega)}, \\ \|u - U\|_{H^{\alpha/2}(0,T;L^2(\Omega))} &\lesssim \left(h^2 + M^{\alpha-1-2\sigma}\right) \|\phi\|_{\dot{H}^2(\Omega)}. \end{aligned} \tag{54}$$

Remark 4.2 Although in Theorem 4.2 we only give the rigorous error estimates for a special singular right hand side $f = t^\sigma v(x)$, we hope this is helpful for understanding the convergence behavior for general f. Indeed, regardless of the spatial regularity, we conclude that: (i) if f is smooth in time but not vanishing at $t = 0$, then the best temporal convergence rate is $1+2\alpha$; if f is smooth (away from the original point) and behaves like t^σ near $t = 0+$, then the best rate is $1+2\sigma+2\alpha$; in addition, if $\alpha+\sigma \in \mathbb{N}$, then we have a faster rate $1+2\sigma+4\alpha$.

4.1 Technical Lemmas

Let $R_h : \dot{H}^1(\Omega) \longrightarrow X_h$ be the well-known Ritz projection operator

$$\langle \nabla(I - R_h)v, \nabla v_h \rangle = 0 \quad \forall\, v_h \in X_h, \tag{55}$$

for which we have the standard estimate [53]

$$\|(I - R_h)v\|_{L^2(\Omega)} + h\,\|(I - R_h)v\|_{\dot{H}^1(\Omega)} \lesssim h^\gamma\,\|v\|_{\dot{H}^\gamma(\Omega)} \quad \forall\, v \in \dot{H}^\gamma(\Omega), \quad 1 \leq \gamma \leq 2.$$

For $a, b > -1$, denote by $\Phi_M^{a,b} : L^2_{\mu^{a,b}}(0,T) \to P_M(0,T)$ the $L^2_{\mu^{a,b}}$-orthogonal projection onto $P_M(0,T)$ (cf. "Appendix A"). Assume X is a separable Hilbert space with an orthonormal basis $\{e_n\}_{n=0}^\infty$ and $P_M([0,T];X)$ is the space of all polynomials (degree $\leq M$) over $(0,T)$ with coefficients in X (see [60]). Following [26, Section 4], we can extend $\Phi_M^{a,b}$ to the vector-valued case $\Phi_{M,X}^{a,b} : L^2_{\mu^{a,b}}(0,T;X) \to P_M([0,T];X)$ by that: for all $v = \sum_{n=0}^\infty v_n e_n$ with $v_n \in L^2_{\mu^{a,b}}(0,T)$, define $\Phi_{M,X}^{a,b} v := \sum_{n=0}^\infty e_n \Phi_M^{a,b} v_n$. Clearly, the case $X = \mathbb{R}$ boils down to $\Phi_{M,\mathbb{R}}^{a,b} = \Phi_M^{a,b}$.

Recall that $\{\lambda_n^{-\gamma/2}\phi_n\}_{n=0}^\infty$ is an orthonormal basis of $\dot{H}^\gamma(\Omega)$, and by definition, we claim that

$$\Phi_{M,\dot{H}^\gamma(\Omega)}^{a,b} v = \Phi_{M,L^2(\Omega)}^{a,b} v \quad \forall\, v \in \dot{H}^\gamma(\Omega), \tag{56}$$

for all $\gamma > 0$. Suppose X and Y are two separable Hilbert spaces and $A : X \to Y$ is a bounded linear operator. Analogous to [26, Lemma 5.6], we have the commutativity

$$\Phi^{a,b}_{M,Y} A v = A \Phi^{a,b}_{M,X} v \quad \forall v \in L^2_{\mu^{a,b}}(0, T; X). \tag{57}$$

In particular, there holds that

$$\Phi^{-\alpha,0}_{M,X_h} R_h v = R_h \Phi^{-\alpha,0}_{M,\dot{H}^1(\Omega)} v = R_h \Phi^{-\alpha,0}_{M,L^2(\Omega)} v \quad \forall v \in \mathcal{X}. \tag{58}$$

Moreover, by Lemma A.1, for all $v \in \mathcal{X}$,

$$\left\langle \mathrm{D}^\alpha_{0+}(I - \Phi^{-\alpha,0}_{M,X_h}) R_h v, V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} = 0 \quad \forall V \in P_M(0, T) \otimes X_h. \tag{59}$$

This can be verified directly by definition; we also refer to the discussion in [26, Remark 4.2]. For clarity, we provide the detailed proof of (57) in "Appendix B".

When no confuse arises, we simply write $\Phi^{-\alpha,0}_{M,L^2(\Omega)} u = \Phi^{-\alpha,0}_M u$, where $u = \sum_{n=0}^\infty y_n \phi_n$ is the unique weak solution to (10) and $y_n \in H^{\alpha/2}(0, T)$. By the stability estimate (85), we have that

$$\left\| \Phi^{-\alpha,0}_M u \right\|_{H^{\alpha/2}(0,T;\dot{H}^\gamma(\Omega))} \lesssim \|u\|_{H^{\alpha/2}(0,T;\dot{H}^\gamma(\Omega))}, \quad 0 \le \gamma \le 2, \tag{60}$$

and moreover,

$$u - \Phi^{-\alpha,0}_M u = \sum_{n=0}^\infty \phi_n (I - \Phi^{-\alpha,0}_M) y_n, \tag{61}$$

which means the projection error of u is related to the scalar case $(I - \Phi^{-\alpha,0}_M) y_n$, where y_n solves the ordinary differential equation (11).

In what follows, we give some nontrivial projection error bounds in terms of the solution to (12). The decay estimates of y_k established in Lemmas 3.4 and 3.5 are enough to provide optimal rate in the $L^2_{\mu^{-\alpha,0}}(0, T)$-norm. Moreover, we shall utilize it to prove the error estimate under the fractional norm $|\cdot|_{H^{\alpha/2}(0,T)}$.

Lemma 4.1 *Let y be given by (17) with $\lambda > 0$ and $y_0 \in \mathbb{R}$, then*

$$\left\| (I - \Phi^{-\alpha,0}_M) y \right\|_{L^2_{\mu^{-\alpha,0}}(0,T)} \le C_{\alpha,T} |y_0| \lambda^\theta M^{-1-2\alpha\theta}, \tag{62}$$

where $-1 \le \theta \le 1$ and $1 + 2\alpha\theta > 0$.

Proof According to the proof of Lemma 3.4,

$$\left\| (I - \Phi^{-\alpha,0}_M) y \right\|^2_{L^2_{\mu^{-\alpha,0}}(0,T)} = \sum_{k=M+1}^\infty |y_k|^2 \xi_k^{-\alpha,0} \le C_{\alpha,T} \sum_{k=M+1}^\infty \frac{|y_k|^2}{k},$$

where y_k and $\xi_k^{-\alpha,0}$ are defined by (31). Thanks to (39), we have

$$\begin{aligned} \left\| (I - \Phi^{-\alpha,0}_M) y \right\|^2_{L^2_{\mu^{-\alpha,0}}(0,T)} &\le C_{\alpha,T} |y_0|^2 \lambda^{2\theta} \sum_{k=M+1}^\infty k^{-3-4\theta\alpha} \\ &\le C_{\alpha,T} |y_0|^2 \lambda^{2\theta} \int_M^\infty r^{-3-4\theta\alpha} \, \mathrm{d}r = C_{\alpha,T} |y_0|^2 \lambda^{2\theta} M^{-2-4\theta\alpha}. \end{aligned} \tag{63}$$

This establishes (62) and finishes the proof of this lemma. □

Remark 4.3 It is clear that $L^2_{\mu^{-\alpha,0}}(0,T) \hookrightarrow L^2(0,T)$. Hence, we conclude from Lemma 4.1 that

$$\left\| (I - \Phi_M^{-\alpha,0}) y \right\|_{L^2(0,T)} \le C_{\alpha,T} \, |y_0| \, \lambda^\theta M^{-1-2\alpha\theta},$$

where $-1 \le \theta \le 1$ and $1 + 2\alpha\theta > 0$.

Lemma 4.2 *Let y be given by (17) with $\lambda > 0$ and $y_0 \in \mathbb{R}$, then*

$$\left| (I - \Phi_M^{-\alpha,0}) y \right|_{H^{\alpha/2}(0,T)} \le C_{\alpha,T} \, |y_0| \, \lambda^\theta M^{\alpha-1-2\alpha\theta}, \tag{64}$$

where $-1 \le \theta \le 1$ and $1 + 2\alpha\theta > \alpha$.

Proof By (18), we see $y \in H^{\alpha/2}(0,T)$ and from [17, Theorem 1.4.4.3], we know that $v \in L^2_{\mu^{-\alpha}}(0,T)$, where $\mu^{-\alpha}(t) = (T-t)^{-\alpha} t^{-\alpha}$. Recall the coefficient y_k defined by (31) and set $z := (I - \Phi_M^{-\alpha,0}) y = \sum_{k=M+1}^{\infty} y_k S_k^{-\alpha,0}$. By (82), we have

$$\mathrm{D}_{0+}^{\alpha/2} z = \sum_{k=M+1}^{\infty} \frac{y_k \Gamma(k+1)}{\Gamma(k+1-\alpha/2)} t^{-\alpha/2} S_k^{-\alpha/2}. \tag{65}$$

Here, we changed the order of the summation and the fractional derivative operator and the identity (65) holds true in $L^2(0,T)$. To verify this, it is enough to check that $\mathrm{D}_{0+}^{\alpha/2} y = \sum_{k=0}^{\infty} y_k \, \mathrm{D}_{0+}^{\alpha/2} S_k^{-\alpha,0}$ in $L^2(0,T)$. Indeed, this follows from Lemma 2.1 and the fact $\lim_{M\to\infty} |(I - \Phi_M^{-\alpha,0}) y|_{H^{\alpha/2}(0,T)} = 0$; see Lemma A.1.

On the other hand, since $z \in L^2_{\mu^{0,-\alpha}}(0,T)$, we have the orthogonal expansion (see (81)):

$$z = \sum_{k=0}^{\infty} z_k S_k^{0,-\alpha} \quad \text{with } z_k = \frac{2k+1-\alpha}{T^{1-\alpha}} \left\langle z, S_k^{0,-\alpha} \right\rangle_{\mu^{0,-\alpha}},$$

and it follows from (83) that

$$\mathrm{D}_{T-}^{\alpha/2} z = \sum_{k=0}^{\infty} \frac{z_k \Gamma(k+1)}{\Gamma(k+1-\alpha/2)} (T-t)^{-\alpha/2} S_k^{-\alpha/2}.$$

Similar with (65), this holds true in $L^2(0,T)$. Hence, from Lemma 2.1 and the orthogonality of $\{S_k^{-\alpha/2}\}_{k=0}^{\infty}$ with respect to the weight $\mu^{-\alpha/2}(t) = (T-t)^{-\alpha/2} t^{-\alpha/2}$, we obtain

$$\cos(\alpha\pi/2) |z|^2_{H^{\alpha/2}(0,T)} = \left\langle \mathrm{D}_{0+}^{\alpha/2} z, \mathrm{D}_{T-}^{\alpha/2} z \right\rangle_{(0,T)} = \sum_{k=M+1}^{\infty} \frac{y_k z_k \Gamma(k+1) T^{1-\alpha}}{(2k+1-\alpha)\Gamma(k+1-\alpha)}. \tag{66}$$

By Rodrigues' formula (79), we have

$$y_k = (-1)^k \frac{2k+1-\alpha}{T^{k+1-\alpha} k!} \left\langle y, \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mu^{k-\alpha,k} \right\rangle_{(0,T)},$$

$$z_k = (-1)^k \frac{2k+1-\alpha}{T^{k+1-\alpha} k!} \left\langle z, \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mu^{k,k-\alpha} \right\rangle_{(0,T)}.$$

It follows from (32) and (34) that

$$\left\langle y, \frac{\mathrm{d}^k}{\mathrm{d}t^k}\mu^{k-\alpha,k}\right\rangle_{(0,T)} = (-1)^k \left\langle y^{(k)}, \mu^{k-\alpha,k}\right\rangle_{(0,T)}.$$

When $k \geq M+1$, we have $z^{(k)}(t) = y^{(k)}(t)$. Therefore, applying the proof of (34) gives

$$\left\langle z, \frac{\mathrm{d}^k}{\mathrm{d}t^k}\mu^{k,k-\alpha}\right\rangle_{(0,T)} = (-1)^k \left\langle y^{(k)}, \mu^{k,k-\alpha}\right\rangle_{(0,T)}.$$

Substituting the above two identities into (66) implies

$$\cos(\alpha\pi/2)|z|^2_{H^{\alpha/2}(0,T)} = \sum_{k=M+1}^{\infty} \frac{2k+1-\alpha}{T^{1-\alpha}} \cdot \frac{\left\langle y^{(k)}, \mu^{k-\alpha,k}\right\rangle_{(0,T)}}{T^k\Gamma(k+1-\alpha)} \cdot \frac{\left\langle y^{(k)}, \mu^{k,k-\alpha}\right\rangle_{(0,T)}}{T^k\Gamma(k+1)}. \tag{67}$$

In view of (37), it follows that

$$\frac{\left\langle y^{(k)}, \mu^{k-\alpha,k}\right\rangle_{(0,T)}}{T^k\Gamma(k+1-\alpha)} \leq C_{\alpha,T} \cdot \frac{|y_0|\,\lambda^{\beta}\Gamma(k-\beta\alpha)}{\Gamma(k+2+(\beta-1)\alpha)}, \tag{68}$$

for all $-1 \leq \beta \leq 1$. Analogously, one has

$$\frac{\left\langle y^{(k)}, \mu^{k,k-\alpha}\right\rangle_{(0,T)}}{T^k\Gamma(k+1)} \leq C_{\alpha,T} \cdot \frac{|y_0|\,\lambda^{\gamma}\Gamma(k-\gamma\alpha)}{\Gamma(k+2+(\gamma-1)\alpha)}, \tag{69}$$

for arbitrary $-1 \leq \gamma \leq 1$.

Consequently, plugging the above two estimates into (67) and applying Stirling's formula (38) yield that

$$\begin{aligned} |z|^2_{H^{\alpha/2}(0,T)} &\leq C_{\alpha,T}\,|y_0|^2\,\lambda^{\beta+\gamma} \sum_{k=M+1}^{\infty} \frac{k\Gamma(k-\beta\alpha)}{\Gamma(k+2+(\beta-1)\alpha)} \cdot \frac{\Gamma(k-\gamma\alpha)}{\Gamma(k+2+(\gamma-1)\alpha)} \\ &\leq C_{\alpha,T}\,|y_0|^2\,\lambda^{\beta+\gamma} \sum_{k=M+1}^{\infty} k^{2\alpha-3-2\alpha(\beta+\gamma)}. \end{aligned}$$

Let $\theta = (\beta+\gamma)/2 \in [-1, 1]$ be such that $1+2\alpha\theta > \alpha$. Then using the proof of (63) implies the desired estimate (64) and concludes the proof of this lemma. □

Remark 4.4 To get (64), the key is to establish (68) and (69), which are easy to obtain for the singular function $y(t) = t^{\sigma}$. Hence, according to the proofs of Lemmas 4.1 and 4.2, it is not hard to conclude the following estimate:

$$\left\|(I-\Phi_M^{-\alpha,0})y\right\|_{L^2_{\mu^{-\alpha,0}}(0,T)} + M^{-\alpha}\left|(I-\Phi_M^{-\alpha,0})y\right|_{H^{\alpha/2}(0,T)} \leq C_{\alpha,\sigma,T}M^{-1-2\sigma},$$

where $\sigma > (\alpha-1)/2$. We mention that the optimal rate $1+2\sigma$ of the Legendre projection under L^2-norm has already been proved in [18, Theorem 5]. See also some similar results in [61, Lemma 3.3 and Theorem 3.5] for the estimate of $(I-\Phi_M^{\alpha,0})y$ under the weighted $L^2_{\mu^{\alpha,0}}$-norm.

For the particular case: $y_0 = 0,\ g(t) = g_0 t^{\sigma}$, we can also prove optimal projection error bounds for the solution to the auxiliary problem (12). Indeed, by the previous two estimates

(45) and (47), we are able to establish the corresponding key results like (68) and (69). Since the proof techniques are almost the same as that of Lemmas 4.1 and 4.2, we omit the details and only list the main results as follows.

Lemma 4.3 *Let y be given by (42) with $g_0 \in \mathbb{R}$, $\lambda > 0$ and $\sigma > -1/2$. Then for any $0 \le \theta \le 1$, we have*

$$\left\| \left(I - \Phi_M^{-\alpha,0}\right) y \right\|_{L^2_{\mu^{-\alpha,0}}(0,T)} \le C_{\alpha,\sigma,T} \, |g_0| \, \lambda^{\theta-1} M^{-1-2\sigma-2\theta\alpha}, \tag{70}$$

and if $\sigma > (\alpha-1)/2$, then

$$\left| (I - \Phi_M^{-\alpha,0}) y \right|_{H^{\alpha/2}(0,T)} \le C_{\alpha,\sigma,T} \, |g_0| \, \lambda^{\theta-1} M^{\alpha-1-2\sigma-2\alpha\theta}. \tag{71}$$

Moreover, if $\alpha + \sigma \in \mathbb{N}$, then we can take $\theta \in [0,2]$ for both (70) and (71).

Below, we present a lemma that connects the above projection errors with our desired estimates. Recall that the space $\mathcal{X}$ is given in (9) and $\mathcal{X}^*$ denotes its dual space.

Lemma 4.4 *If $f + \mathrm{D}_{0+}^{\alpha} u_0 \in \mathcal{X}^*$, then*

$$\begin{aligned} \|u - U\|_{H^{\alpha/2}(0,T;L^2(\Omega))} \lesssim{} & \left\| (I - \Phi_M^{-\alpha,0}) u \right\|_{\mathcal{X}} + \|(I - R_h) u\|_{H^{\alpha/2}(0,T;L^2(\Omega))} \\ & + \left\| (I - R_h) \Phi_M^{-\alpha,0} u \right\|_{H^{\alpha/2}(0,T;L^2(\Omega))}, \end{aligned} \tag{72}$$

and moreover,

$$\|u - U\|_{L^2(0,T;\dot{H}^1(\Omega))} \lesssim \|(I - R_h) u\|_{\mathcal{X}} + \left\| (I - \Phi_M^{-\alpha,0}) u \right\|_{L^2(0,T;\dot{H}^1(\Omega))}. \tag{73}$$

Proof By (10), for any $V \in P_M(0,T) \otimes X_h$, we have

$$\left\langle \mathrm{D}_{0+}^{\alpha} u, V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla u, \nabla V \rangle_{\Omega_T} = \left\langle f + \mathrm{D}_{0+}^{\alpha} u_0, V \right\rangle_{\mathcal{X}},$$

which, together with (49), gives the error equation

$$\left\langle \mathrm{D}_{0+}^{\alpha} (u - U), V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla (u - U), \nabla V \rangle_{\Omega_T} = 0 \quad \forall\, V \in P_M(0,T) \otimes X_h. \tag{74}$$

Hence it follows that

$$\begin{aligned} & \left\langle \mathrm{D}_{0+}^{\alpha} (U - W), V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla (U - W), \nabla V \rangle_{\Omega_T} \\ & \quad = \left\langle \mathrm{D}_{0+}^{\alpha} (u - W), V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla (u - W), \nabla V \rangle_{\Omega_T}, \end{aligned}$$

where $W = \Phi_{M,X_h}^{-\alpha,0} R_h u$. Applying (55) and (59) and the fact $\Phi_{M,X_h}^{-\alpha,0} R_h u = R_h \Phi_M^{-\alpha,0} u$ (cf. (58)) yields the identity

$$\begin{aligned} & \left\langle \mathrm{D}_{0+}^{\alpha} (U - W), V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \langle \nabla (U - W), \nabla V \rangle_{\Omega_T} \\ & \quad = \left\langle \mathrm{D}_{0+}^{\alpha} (u - R_h u), V \right\rangle_{H^{\alpha/2}(0,T;L^2(\Omega))} + \left\langle \nabla (I - \Phi_M^{-\alpha,0}) u, \nabla V \right\rangle_{\Omega_T}, \end{aligned}$$

and thus taking $V = U - W$ implies

$$\|U - W\|_{\mathcal{X}} \lesssim \|(I - R_h) u\|_{H^{\alpha/2}(0,T;L^2(\Omega))} + \left\| (I - \Phi_M^{-\alpha,0}) u \right\|_{L^2(0,T;\dot{H}^1(\Omega))}.$$

Now using the triangle inequality and the stability result (85) gives

$$\begin{aligned}\|u-U\|_{H^{\alpha/2}(0,T;L^2(\Omega))} &\lesssim \|u-W\|_{H^{\alpha/2}(0,T;L^2(\Omega))}+\|U-W\|_{H^{\alpha/2}(0,T;L^2(\Omega))}\\ &\lesssim \left\|(I-\Phi_M^{-\alpha,0})u\right\|_{H^{\alpha/2}(0,T;L^2(\Omega))}\\ &\quad+\left\|(I-R_h)\Phi_M^{-\alpha,0}u\right\|_{H^{\alpha/2}(0,T;L^2(\Omega))}+\|U-W\|_{\mathcal{X}}\\ &\lesssim \left\|(I-\Phi_M^{-\alpha,0})u\right\|_{\mathcal{X}}+\left\|(I-R_h)\Phi_M^{-\alpha,0}u\right\|_{H^{\alpha/2}(0,T;L^2(\Omega))}\\ &\quad+\|(I-R_h)u\|_{H^{\alpha/2}(0,T;L^2(\Omega))}.\end{aligned}$$

This establishes (72). Since (73) can be proved similarly, we conclude the poof. □

4.2 Proofs of Theorems 4.1 and 4.2

As the proof of Theorem 4.2 is parallel to that of Theorem 4.1, we only consider the latter.

Proof of Theorem 4.1 According to Theorem 3.1, we have

$$\|(I-R_h)u\|_{L^2(0,T;\dot H^1(\Omega))}\lesssim\begin{cases}\dfrac{1}{\sqrt{\epsilon}}h^{\min\{1,\gamma+1\}-\epsilon}\|u_0\|_{\dot H^\gamma(\Omega)}, & \alpha=1/2,\\ h^{\min\{1,\gamma_0+\gamma-1\}}\|u_0\|_{\dot H^\gamma(\Omega)}, & \alpha\neq 1/2,\end{cases}$$
$$\|(I-R_h)u\|_{H^{\alpha/2}(0,T;L^2(\Omega))}\lesssim h^{\min\{2,\gamma_0+\gamma-1\}}\|u_0\|_{\dot H^\gamma(\Omega)}.$$

For $\alpha=1/2$, we choose $\epsilon=1/(2+|\ln h|)$ to get

$$\|(I-R_h)u\|_{L^2(0,T;\dot H^1(\Omega))}\lesssim\sqrt{|\ln h|}\,h^{\min\{1,\gamma+1\}}\|u_0\|_{\dot H^\gamma(\Omega)}.$$

Besides, there holds that

$$\left\|(I-R_h)\Phi_M^{-\alpha,0}u\right\|_{H^{\alpha/2}(0,T;L^2(\Omega))}\lesssim h^{\min\{2,\gamma_0+\gamma-1\}}\left\|\Phi_M^{-\alpha,0}u\right\|_{H^{\alpha/2}(0,T;\dot H^{\gamma_0+\gamma-1}(\Omega))}.$$

Thanks to Theorem 3.1 and (60), we obtain

$$\left\|(I-R_h)\Phi_M^{-\alpha,0}u\right\|_{H^{\alpha/2}(0,T;L^2(\Omega))}\lesssim h^{\min\{2,\gamma_0+\gamma-1\}}\|u_0\|_{\dot H^\gamma(\Omega)}.$$

Invoking (61),Proposition 3.1,Lemma 4.1, and Remark 4.3 gives the estimate

$$\left\|(I-\Phi_M^{-\alpha,0})u\right\|_{L^2(0,T;\dot H^1(\Omega))}\lesssim M^{-1-\alpha(\gamma-1)}\|u_0\|_{\dot H^\gamma(\Omega)},$$

and similarly, by Lemma 4.2 we conclude that

$$\left\|(I-\Phi_M^{-\alpha,0})u\right\|_{\mathcal{X}}\lesssim M^{-1-\alpha\min\{1,\gamma-1\}}\|u_0\|_{\dot H^\gamma(\Omega)}.$$

Combining these estimates with Lemma 4.4 leads to (51) and (50). This completes the proof of Theorem 4.1. □

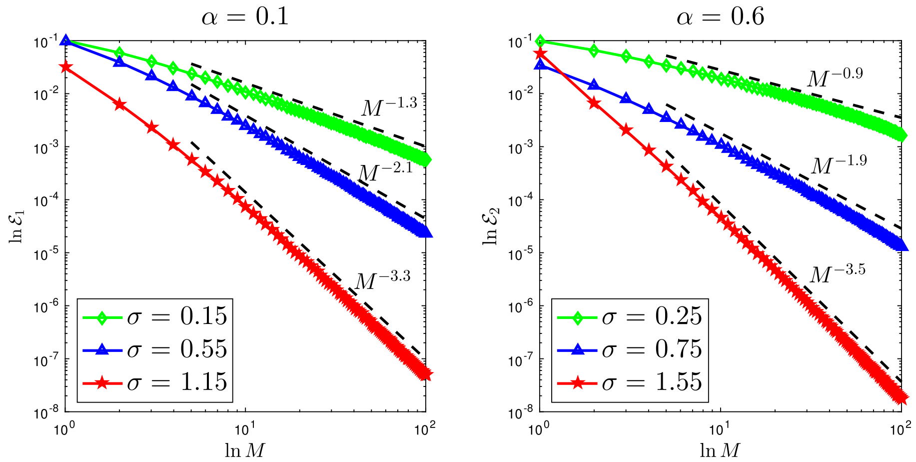

Fig. 3 Temporal errors of Example 1 with $h = 2^{-10}$

5 Numerical Tests

This section presents several numerical experiments to validate our theoretical predictions. For simplicity, we take $T = 1$, $\Omega = (0, 1)$ or $(0, 1)^2$. As spatial discretization errors have been investigated in [27], we are concerned with the temporal convergence behaviors of

$$\mathcal{E}_1 := \|\widehat{u} - U\|_{L^2(0,T;\dot{H}^1(\Omega))} \quad \text{and} \quad \mathcal{E}_2 := \|\widehat{u} - U\|_{H^{\alpha/2}(0,T;L^2(\Omega))},$$

where $\widehat{u}$ is the reference solution.

5.1 One-Dimensional Tests

We first consider three experiments in one spatial dimension: $\Omega = (0, 1)$. The reference solution $\widehat{u}$ corresponds to $M = 200$ and $h = 2^{-10}$.

Example 1 This example is to verify Remark 4.1 with a priorly known solution

$$u(x, t) = t^{\sigma} \sin \pi x, \quad (x, t) \in \Omega_T,$$

where $\sigma > (\alpha - 1)/2$. Temporal discretization errors are plotted in Fig. 3, which shows $\mathcal{E}_1 = O(M^{-1-2\sigma})$ and $\mathcal{E}_2 = O(M^{\alpha-1-2\sigma})$. This agrees well with the rates given in (54).

Example 2 To verify Theorem 4.1, we consider $f = 0$ and

$$u_0(x) = \theta x(1 - x)^{\gamma-1/2} + (1 - \theta) \sin \pi x, \quad x \in \Omega,$$

where $1 - \gamma_0 < \gamma \leq 1.5$ and $\theta \in \{0, 1\}$. For $\theta = 1$, a direct calculation yields that $u_0 \in \dot{H}^{\gamma-\epsilon}(\Omega)$; and for $\theta = 0$, we have $u_0 \in \dot{H}^{\beta}(\Omega)$ with any $\beta > 0$, since $\sin \pi x$ is an eigenfunction of $-\Delta$ on $\Omega = (0, 1)$ with the homogeneous Dirichlet boundary condition. Numerical outputs are plotted in Fig. 4, which implies that $\mathcal{E}_1 = O(M^{-1-\alpha \min\{2,\gamma-1\}})$ and $\mathcal{E}_2 = O(M^{-1-\alpha \min\{1,\gamma-1\}})$. These coincide with the sharp estimates established in Theorem 4.1.

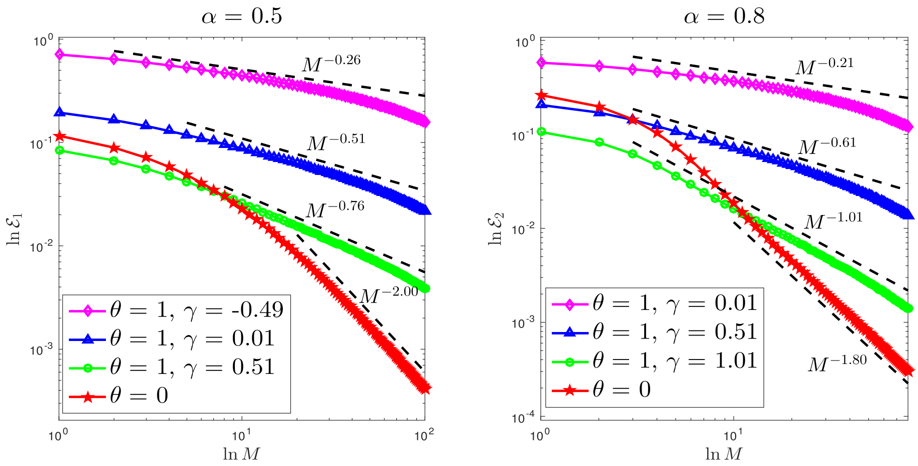

Fig. 4 Temporal errors of Example 2 with $h = 2^{-10}$

Example 3 This example is to verify Theorem 4.2 with $u_0 = 0$ and $f(x,t) = t^\sigma v(x)$, where $\sigma > (\alpha - 1)/2$ and

$$v(x) = \theta x^{\gamma - 1/2}(1 - x) + (1 - \theta)\sin \pi x, \quad x \in \Omega,$$

with $-0.5 < \gamma \le 1.5$ and $\theta \in \{0, 1\}$.

We first consider $\alpha + \sigma \notin \mathbb{N}$ and $\theta = 1$. Note that $v \in \dot{H}^{\gamma - \epsilon}(\Omega)$ and from Fig. 5 we conclude that

$$\mathcal{E}_1 = O\left(M^{-1-2\sigma-\alpha \min\{2,\gamma+1\}}\right) \quad \text{and} \quad \mathcal{E}_2 = O\left(M^{-1-2\sigma-\alpha \min\{1,\gamma+1\}}\right). \tag{75}$$

Then we take $\alpha + \sigma \in \mathbb{N}$ and $\theta = 0$. In this situation, we have $v \in \dot{H}^\beta(\Omega)$ with any $\beta > 0$ and according to Fig. 6, we observe faster convergence rates

$$\mathcal{E}_1 = O(M^{-1-2\sigma-4\alpha}) \quad \text{and} \quad \mathcal{E}_2 = O(M^{-1-2\sigma-3\alpha}). \tag{76}$$

Both this and the previous case are conformable to the sharp error bounds in Theorem 4.2.

5.2 Two-Dimensional Tests

We provide two more examples in two spatial dimensions: $\Omega = (0, 1)^2$. The reference solution $\widehat{u}$ corresponds to $M = 120$ and $h = 2^{-5}$.

For $\theta \in \{1, 2, 3\}$ and any $x = (x_1, x_2) \in \Omega$, define

$$v(x) := \begin{cases} \delta_z(x) & \text{if } \theta = 1, \\ \chi_{\{0<x_1<1/2\}}(x) & \text{if } \theta = 2, \\ \sin(\pi x_1)\sin(\pi x_2) & \text{if } \theta = 3, \end{cases} \tag{77}$$

where δ_z denotes the Dirac distribution centered at $z = (1/2, 1/2) \in \Omega$ and χ_ω is the indicator function of the region ω, i.e., $\chi(x) = 1$ for $x \in \omega$ and $\chi(x) = 0$ for $x \notin \omega$. It is not hard to

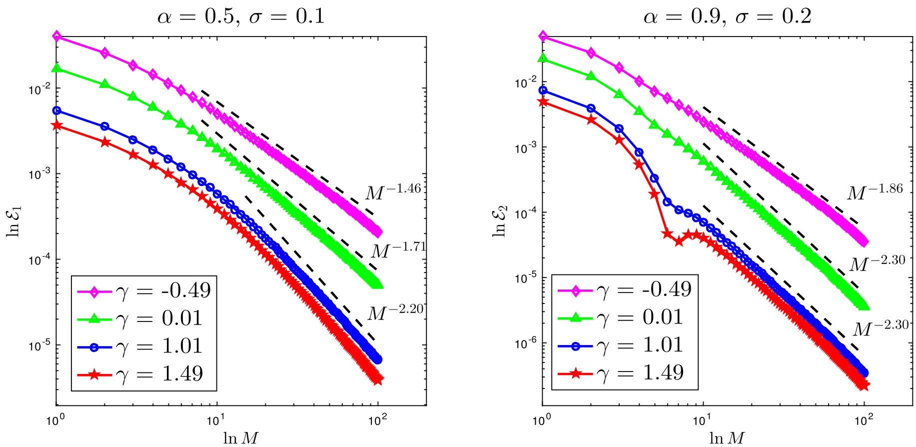

Fig. 5 Temporal errors of Example 3 with $\theta = 1$ and $h = 2^{-10}$

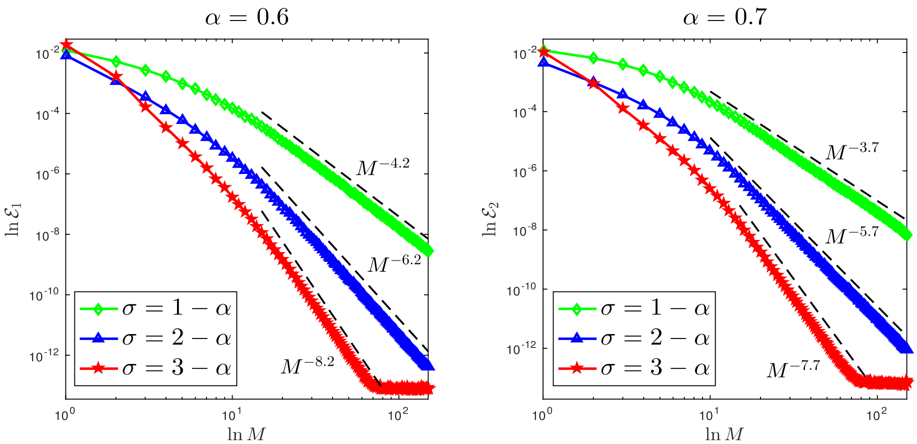

Fig. 6 Temporal errors of Example 3 with $\theta = 0$ and $h = 2^{-10}$

find that $v \in \dot{H}^{\gamma(\theta)-\epsilon}(\Omega)$, where

$$\gamma(\theta) := \begin{cases} -1 & \text{if } \theta = 1, \\ 1/2 & \text{if } \theta = 2, \\ \gamma \in (3, \infty) & \text{if } \theta = 3. \end{cases} \tag{78}$$

Example 4 In this test, we take $f = 0$ and $u_0 = v$ with v being defined in (77). From Fig. 7, we see that $\mathcal{E}_1 = O(M^{-1-\alpha \min\{2,\gamma(\theta)-1\}})$ and $\mathcal{E}_2 = O(M^{-1-\alpha \min\{1,\gamma(\theta)-1\}})$ for $\theta \in \{2, 3\}$, where $\gamma(\theta)$ is defined by (78). This verifies the estimates in Theorem 4.1.

Example 5 To the end, we consider $u_0 = 0$ and $f(x, t) = t^\sigma v(x)$ where $\sigma > (\alpha - 1)/2$ and v is defined in (77). In Fig. 8, we report the numerical results in (75) for $\alpha + \sigma \notin \mathbb{N}$ and observe the same rates, with γ being $\gamma(\theta)$ (cf. (78)). The case $\theta = 3$, $\alpha + \sigma \in \mathbb{N}$ has also been displayed in Fig. 9, which yields the same faster rates as in (76). These agree well with the theoretical predictions in Theorem 4.2.

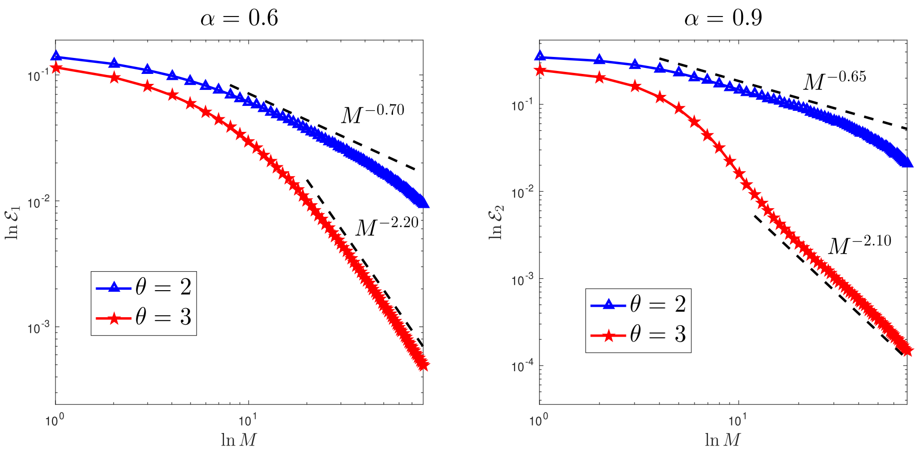

Fig. 7 Temporal errors of Example 4 with $h = 2^{-5}$

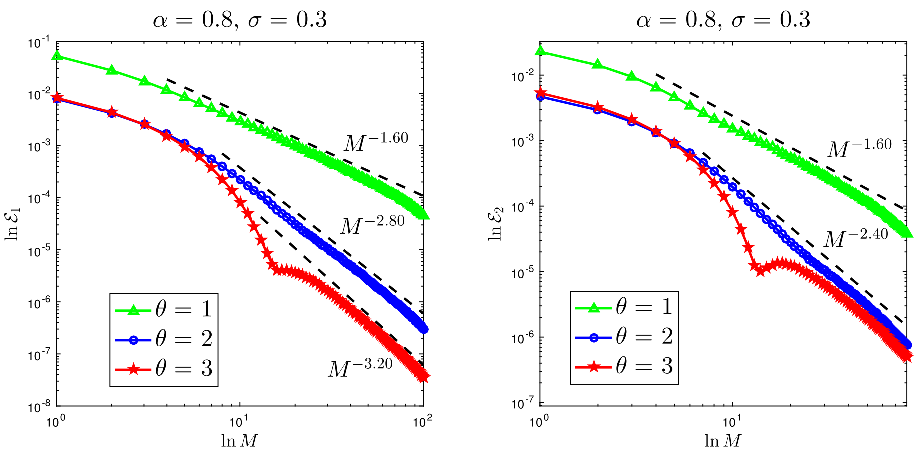

Fig. 8 Temporal errors of Example 5 with $h = 2^{-5}$

6 Conclusion

This paper is devoted to sharp error estimates of a time-spectral algorithm for time fractional diffusion problems of order α $(0 < \alpha < 1)$. Based on new regularity results in the Besov space, optimal convergence rates have been derived with low regularity data. Particularly, for the homogenuous case $f = 0$, optimal temporal convergence orders $1 + 2\alpha$ and $1 + \alpha$ under $L^2(0, T; \dot{H}^1(\Omega))$-norm and $H^{\alpha/2}(0, T; L^2(\Omega))$-norm have been shown theoretically and numerically.

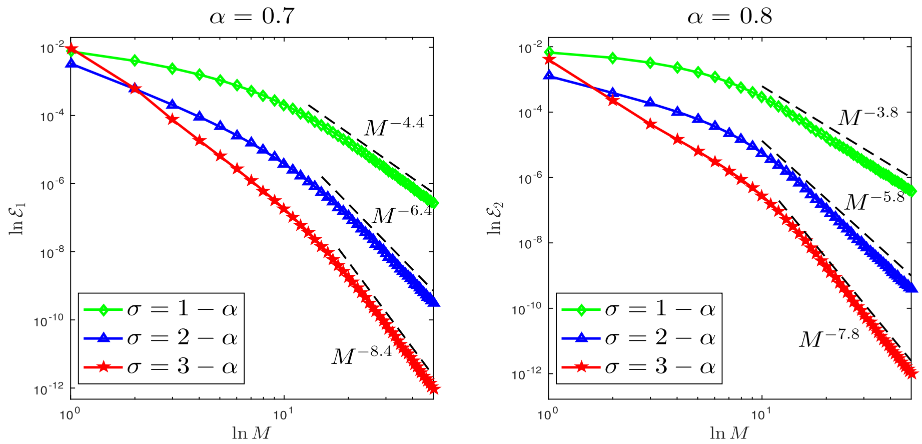

Fig. 9 Temporal errors of Example 5 with $\theta = 3$ and $h = 2^{-5}$

Funding Funding was provided by National Natural Science Foundation of China (Grant No. 11771312).

Data Availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors have not disclosed any conflict of interest.

A The Shifted Jacobi Polynomial

Given $a, b > -1$, the family of shifted Jacobi polynomial $\{S_k^{a,b}\}_{k=0}^{\infty}$ on $(0, T)$ are defined as follows:

$$\mu^{a,b}(t) S_k^{a,b}(t) = \frac{(-1)^k}{T^k k!} \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mu^{k+a,k+b}(t), \quad 0 < t < T, \tag{79}$$

where $\mu^{\nu,\theta}(t) = (T-t)^{\nu} t^{\theta}$ for all $-1 < \nu, \theta < \infty$. Note that (79) is also called Rodrigues' formula [56], which implies $\{S_k^{a,b}\}_{k=0}^{\infty}$ is orthogonal with respect to the weight $\mu^{a,b}$ on $(0, T)$, i.e.,

$$\left\langle S_k^{a,b}, S_l^{a,b} \right\rangle_{\mu^{a,b}} = \xi_k^{a,b} \delta_{kl},$$

where δ_{kl} denotes the Kronecker product and

$$\xi_k^{a,b} := \frac{T^{a+b+1} \Gamma(k+a+1) \Gamma(k+b+1)}{(2k+a+b+1) k! \Gamma(k+a+b+1)}. \tag{80}$$

As $\{S_k^{a,b}\}_{k=0}^{\infty}$ forms a complete orthogonal basis of $L^2_{\mu^{a,b}}(0, T)$, any $v \in L^2_{\mu^{a,b}}(0, T)$ admits a unique decomposition

$$v = \sum_{k=0}^{\infty} v_k S_k^{a,b} \quad \text{with } v_k = \frac{1}{\xi_k^{a,b}} \left\langle v, S_k^{a,b} \right\rangle_{\mu^{a,b}}, \tag{81}$$

and the $L^2_{\mu^{a,b}}$-orthogonal projection of v onto $P_M(0,T)$ is defined as $\Phi_M^{a,b} v := \sum_{k=0}^M v_k S_k^{a,b}$. For ease of notation, we shall set $S_k^a = S_k^{a,a},\ \mu^a = \mu^{a,a},\ \Phi_M^a = \Phi_M^{a,a}$, and all the superscripts are omitted when $a = 0$.

Thanks to [7, Lemma 2.5], a standard calculation gives

$$\mathrm{D}_{0+}^{\theta} S_k^{\beta-\theta,0} = \frac{\Gamma(k+1)}{\Gamma(k+1-\theta)} t^{-\theta} S_k^{\beta,-\theta}(t), \tag{82}$$

$$\mathrm{D}_{T-}^{\theta} S_k^{0,\beta-\theta} = \frac{\Gamma(k+1)}{\Gamma(k+1-\theta)} (T-t)^{-\theta} S_k^{-\theta,\beta}(t), \tag{83}$$

where $0 < \theta < 1$ and $-1 < \beta < \infty$.

Lemma A.1 *For any $v \in H^{\alpha/2}(0,T)$, it holds that*

$$\left\langle \mathrm{D}_{0+}^{\alpha}(I - \Phi_M^{-\alpha,0})v, q \right\rangle_{H^{\alpha/2}(0,T)} = 0 \quad \forall q \in P_M(0,T). \tag{84}$$

Consequently, we have the stability:

$$\left|\Phi_M^{-\alpha,0} v\right|_{H^{\alpha/2}(0,T)} \le C_\alpha \left|v\right|_{H^{\alpha/2}(0,T)}, \tag{85}$$

and the convergence: $\lim\limits_{M\to\infty} |(I - \Phi_M^{-\alpha,0})v|_{H^{\alpha/2}(0,T)} = 0.$

Proof Given any fixed $v \in H^{\alpha/2}(0,T)$, by [17, Theorem 1.4.4.3], we know that $v \in L^2_{\mu^{-\alpha,0}}(0,T)$. To prove (84), it is enough to consider $q = S_k^{0,-\alpha}$ for any $0 \le k \le M$. Thanks to (83), we have

$$\mathrm{D}_{T-}^{\alpha} q = \frac{\Gamma(k+1)}{\Gamma(k+1-\alpha)} (T-t)^{-\alpha} S_k^{-\alpha,0}.$$

Again, it follows from [17, Theorem 1.4.4.3] that $\mathrm{D}_{T-}^{\alpha} q \in (H^{\alpha/2}(0,T))^*$. Thus using the definition of $\Phi_M^{-\alpha,0}$ and Lemma 2.1 gives

$$\begin{aligned}
\left\langle \mathrm{D}_{0+}^{\alpha}(I - \Phi_M^{-\alpha,0})v, q \right\rangle_{H^{\alpha/2}(0,T)} &= \left\langle (I - \Phi_M^{-\alpha,0})v, \mathrm{D}_{T-}^{\alpha} q \right\rangle_{(H^{\alpha/2}(0,T),(H^{\alpha/2}(0,T))^*)} \\
&= \frac{\Gamma(k+1)}{\Gamma(k+1-\alpha)} \left\langle (I - \Phi_M^{-\alpha,0})v, (T-t)^{-\alpha} S_k^{-\alpha,0} \right\rangle_{(0,T)} \\
&= 0.
\end{aligned}$$

This establishes (84) and by Lemma 2.1, we have

$$\begin{aligned}
\cos(\alpha\pi/2) \left|\Phi_M^{-\alpha,0} v\right|^2_{H^{\alpha/2}(0,T)} &= \left\langle \mathrm{D}_{0+}^{\alpha} \Phi_M^{-\alpha,0} v, \Phi_M^{-\alpha,0} v \right\rangle_{H^{\alpha/2}(0,T)} \\
&= \left\langle \mathrm{D}_{0+}^{\alpha} v, \Phi_M^{-\alpha,0} v \right\rangle_{H^{\alpha/2}(0,T)} \\
&\le \left|\Phi_M^{-\alpha,0} v\right|_{H^{\alpha/2}(0,T)} \left|v\right|_{H^{\alpha/2}(0,T)},
\end{aligned}$$

which implies (85).

By (84) and the proof of (85), we find that

$$|(I - \Phi_M^{-\alpha,0})v|_{H^{\alpha/2}(0,T)} \le \sec(\alpha\pi/2) |v - q|_{H^{\alpha/2}(0,T)} \quad \forall q \in P_M(0,T).$$

Therefore, a standard density argument leads to

$$\lim_{M\to\infty} |(I - \Phi_M^{-\alpha,0})v|_{H^{\alpha/2}(0,T)} = 0.$$

This finishes the proof of this lemma. □

B Proof of the Commutativity (57)

Let $\{x_i\}_{i\in\mathbb{N}}$ and $\{y_j\}_{j\in\mathbb{N}}$ be the orthonormal basis of X and Y, respectively. Assume $Ax_i = \sum_{j=0}^{\infty} a_{ij} y_j$ with $a_{ij} \in \mathbb{R}$ for all $i \in \mathbb{N}$. It is clear that

$$\sum_{j=0}^{\infty} \left|a_{ij}\right|^2 = \|Ax_i\|_Y^2 \le \|A\|_{X\to Y}^2 \, \|x_i\|_X^2 \quad \text{for all } i \in \mathbb{N}, \tag{86}$$

where $\|A\|_{X\to Y}$ denotes the operator norm of A.

We claim first that by definition, $\Phi_{M,Z}^{a,b} v_n \overset{n\to\infty}{\to} \Phi_{M,Z}^{a,b} v$ in $L^2_{\mu^{a,b}}(0,T;Z)$ whenever $v_n \overset{n\to\infty}{\to} v$ in $L^2_{\mu^{a,b}}(0,T;Z)$ for $Z = X$ or Y. Besides, we have the identity

$$\begin{aligned} \|w\|^2_{L^2_{\mu^{a,b}}(0,T;Z)} &= \sum_{i=0}^{\infty} \|w_i\|^2_{L^2_{\mu^{a,b}}(0,T)} = \int_0^T \sum_{i=0}^{\infty} |w_i(t)|^2 \, \mu^{a,b}(t)\,\mathrm{d}t \\ &= \int_0^T \|w(t)\|_Z^2 \, \mu^{a,b}(t)\,\mathrm{d}t, \end{aligned}$$

for all $w = \sum_{i=0}^{\infty} w_i z_i \in L^2_{\mu^{a,b}}(0,T;Z)$, where $z_i = x_i$ or y_i and we used the monotone convergence theorem (see [6, Theorem 4.1, pp.90]). Based on this, let us verify $Av_n \to Av$ in $L^2_{\mu^{a,b}}(0,T;Y)$ provided that $v_n \overset{n\to\infty}{\to} v$ in $L^2_{\mu^{a,b}}(0,T;X)$. Indeed,

$$\begin{aligned} \|Av_n - Av\|^2_{L^2_{\mu^{a,b}}(0,T;Y)} &= \int_0^T \|(Av_n - Av)(t)\|_Y^2 \, \mu^{a,b}(t)\,\mathrm{d}t \\ &\le \|A\|^2_{X\to Y} \int_0^T \|(v_n - v)(t)\|_X^2 \, \mu^{a,b}(t)\,\mathrm{d}t \\ &= \|A\|^2_{X\to Y} \, \|v - v_n\|^2_{L^2_{\mu^{a,b}}(0,T;X)}. \end{aligned}$$

Now, we take $v_n = \sum_{i=0}^{n} v_i x_i$, which converges to v in $L^2_{\mu^{a,b}}(0,T;X)$. According to the above discussions, $\Phi_{M,X}^{a,b} v_n \overset{n\to\infty}{\to} \Phi_{M,X}^{a,b} v$ in $L^2_{\mu^{a,b}}(0,T;X)$ and $Av_n \overset{n\to\infty}{\to} Av$ in $L^2_{\mu^{a,b}}(0,T;Y)$. To prove (57), it is sufficient to establish

$$A\Phi_{M,X}^{a,b} v_n = \Phi_{M,Y}^{a,b} A v_n. \tag{87}$$

Consider $\xi_i^m = \sum_{j=0}^{m} a_{ij} y_j$, which converges to Ax_i by (86) and further implies that

$$\sum_{i=0}^{n} \xi_i^m \Phi_M^{a,b} v_i \overset{m\to\infty}{\to} \sum_{i=0}^{n} Ax_i \Phi_M^{a,b} v_i = A\Phi_{M,X}^{a,b} v_n \quad \text{in} \quad L^2_{\mu^{a,b}}(0,T;Y). \tag{88}$$

On the other hand, we find

$$\sum_{i=0}^{n} \xi_i^m \Phi_M^{a,b} v_i = \sum_{j=0}^{m} \left(\sum_{i=0}^{n} a_{ij} \Phi_M^{a,b} v_i \right) y_j = \Phi_{M,Y}^{a,b} \sum_{j=0}^{m} \left(\sum_{i=0}^{n} a_{ij} v_i \right) y_j = \Phi_{M,Y}^{a,b} \sum_{i=0}^{n} v_i \xi_i^m .$$

Since

$$\sum_{i=0}^{n} v_i \xi_i^m \overset{m\to\infty}{\to} \sum_{i=0}^{n} v_i A x_i = A v_n \quad \text{in} \quad L^2_{\mu^{a,b}}(0, T; Y),$$

we conclude that

$$\sum_{i=0}^{n} \xi_i^m \Phi_M^{a,b} v_i \overset{m\to\infty}{\to} \Phi_{M,Y}^{a,b} A v_n \quad \text{in} \quad L^2_{\mu^{a,b}}(0, T; Y).$$

This together with (88) proves (87) and completes the proof.

References

1. Abramovitz, A., Stegun, I.: Handbook of Mathematical Functions. Dover, New York (1972)
2. Alikhanov, A.: A new difference scheme for the time fractional diffusion equation. J. Comput. Phys. **280**, 424–438 (2015)
3. Babuška, I., Guo, B.: Direct and inverse approximation theorems for the p-version of the finite element method in the framework of weighted Besov spaces. Part I: approximability of functions in the weighted Besov spaces. SIAM J. Numer. Anal. **39**(5), 1512–1538 (2002)
4. Babuška, I., Suri, M.: The optimal convergence rate of the p-version of the finite element method. SIAM J. Numer. Anal. **24**(4), 750–776 (1987)
5. Bergh, J., Löfström, J.: Interpolation Spaces: An Introduction. Number 223 in Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, Berlin (1976)
6. Brézis, H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext, Springer, New York (2011)
7. Chen, S., Shen, J., Wang, L.: Generalized Jacobi functions and their applications to fractional differential equations. Math. Comput. **85**(300), 1603–1638 (2016)
8. Duan, B., Zhang, Z.: A rational approximation scheme for computing Mittag–Leffler function with discrete elliptic operator as input. J. Sci. Comput. **87**(3), 75 (2021)
9. Duan, B., Zheng, Z.: An exponentially convergent scheme in time for time fractional diffusion equations with non-smooth initial data. J. Sci. Comput. **80**(2), 717–742 (2019)
10. Ervin, V., Roop, J.: Variational formulation for the stationary fractional advection dispersion equation. Numer. Methods Partial Differ. Equ. **22**(3), 558–576 (2006)
11. Evans, L.: Partial Differential Equations, 2nd. Number 19 in Graduate Studies in Mathematics. American Mathematical Society (2010)
12. Ford, N., Xiao, J., Yan, Y.: A finite element method for time fractional partial differential equations. Fract. Calc. Appl. Anal. **14**(3), 454–474 (2011)
13. Ford, N., Yan, Y.: An approach to construct higher order time discretisation schemes for time fractional partial differential equations with nonsmooth data. Fract. Calc. Appl. Anal. **20**(5), 1076–1105 (2017)
14. Gao, G., Sun, H., Sun, Z.: Stability and convergence of finite difference schemes for a class of time-fractional sub-diffusion equations based on certain superconvergence. J. Comput. Phys. **280**, 510–528 (2015)
15. Gorenflo, R., Loutchko, J., Luchko, Y.: Computation of the Mittag–Leffler function $E_{\alpha,\beta}(z)$ and its derivative. Fract. Calc. Appl. Anal. **5**(4), 12–15 (2002)
16. Gorenflo, R., Luchko, Y., Yamamoto, M.: Time-fractional diffusion equation in the fractional Sobolev spaces. Fract. Calc. Appl. Anal. **18**(3), 799–820 (2015)
17. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Pitman, London (1985)
18. Gui, W., Babuška, I.: The h, p and h-p versions of the finite element method in 1 dimension. Part I. The error analysis of the p-version. Numer. Math. **49**(6), 577–612 (1986)
19. Jin, B., Lazarov, R., Pasciak, J., Zhou, Z.: Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion. IMA J. Numer. Anal. **35**(2), 561–582 (2015)

20. Jin, B., Lazarov, R., Zhou, Z.: Error estimates for a semidiscrete finite element method for fractional order parabolic equations. SIAM J. Numer. Anal. **51**(1), 445–466 (2013)
21. Jin, B., Lazarov, R., Zhou, Z.: An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data. IMA J. Numer. Anal. **36**, 197–221 (2015)
22. Jin, B., Li, B., Zhou, Z.: Correction of high-order BDF convolution quadrature for fractional evolution equations. SIAM J. Sci. Comput. **39**(6), A3129–A3152 (2017)
23. Kilbas, A., Srivastava, H., Trujillo, J.: Theory and Applications of Fractional Differential Equations, 1st. Number 204 in North-Holland Mathematics Studies. Elsevier, Amsterdam (2006)
24. Kopteva,N.: Error analysis of an L2-type method on graded meshes for a fractional-order parabolic problem. arXiv:1905.05070 (2019)
25. Li, B., Luo, H., Xie, X.: A space-time finite element method for fractional wave problems. Numer. Algorithms **85**(3), 1095–1121 (2020)
26. Li, B., Luo, H., Xie, X.: A time-spectral algorithm for fractional wave problems. J. Sci. Comput. **77**(2), 1164–1184 (2018)
27. Li, B., Luo, H., Xie, X.: Analysis of a time-stepping scheme for time fractional diffusion problems with nonsmooth data. SIAM J. Numer. Anal. **57**(2), 779–798 (2019)
28. Li, B., Luo, H., Xie, X.: Error estimates of a discontinuous Galerkin method for time fractional diffusion problems with nonsmooth data. In: Fractional Calculus and Applied Analysis, accepted. arXiv:1809.02015 (2018)
29. Li, B., Wang, T., Xie, X.: Analysis of a temporal discretization for a semilinear fractional diffusion equation. Comput. Math. Appl. **80**(10), 2115–2134 (2020)
30. Li, B., Wang, T., Xie, X.: Analysis of a time-stepping discontinuous Galerkin method for fractional diffusion-wave equation with nonsmooth data. J. Sci. Comput. **82**, 4 (2020)
31. Li, B., Wang, T., Xie, X.: Analysis of the L1 scheme for fractional wave equations with nonsmooth data. Comput. Math. Appl. **90**, 1–12 (2021)
32. Li, B., Wang, T., Xie, X.: Numerical analysis of two Galerkin discretizations with graded temporal grids for fractional evolution equations. J. Sci. Comput. **85**, 3 (2020)
33. Li, B., Xie, X.: Regularity of solutions to time fractional diffusion equations. Discrete Contin. Dyn. Syst. B **24**, 3195–3210 (2019)
34. Li, B., Xie, X., Zhang, S.: A new smoothness result for Caputo-type fractional ordinary differential equations. Appl. Math. Comput. **349**, 408–420 (2019)
35. Li, X., Xu, C.: A space-time spectral method for the time fractional diffusion equation. SIAM J. Numer. Anal. **47**(3), 2108–2131 (2009)
36. Lin, Y., Xu, C.: Finite difference/spectral approximations for the time-fractional diffusion equation. J. Comput. Phys. **225**(2), 1533–1552 (2007)
37. Lions, J., Magenes, E.: Non-homogeneous Boundary Value Problems and Applications, vol. 1. Springer, Berlin (1973)
38. Luchko, Y., Punzi, A.: Modeling anomalous heat transport in geothermal reservoirs via fractional diffusion equations. GEM Int. J. Geomath. **1**(2), 257–276 (2011)
39. Lunardi, A.: Interpolation Theory. Springer, Basel (1995)
40. Luo, H., Li, B., Xie, X.: Convergence analysis of a Petrov–Galerkin method for fractional wave problems with nonsmooth data. J. Sci. Comput. **80**(2), 957–992 (2019)
41. Lv, C., Xu, C.: Error analysis of a high order method for time-fractional diffusion equations. SIAM J. Sci. Comput. **38**(5), A2699–A2724 (2016)
42. McLean, W.: Regularity of solutions to a time-fractional diffusion equation. ANZIAM **52**(2), 123–138 (2010)
43. McLean, W., Mustapha, K.: Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. Numer. Algorithms **52**(1), 69–88 (2009)
44. McLean, W., Mustapha, K.: Time-stepping error bounds for fractional diffusion problems with non-smooth initial data. J. Comput. Phys. **293**(C), 201–217 (2015)
45. Metzler, R., Glöckle, W., Nonnenmacher, T.: Fractional model equation for anomalous diffusion. Physica A Stat. Mech. Appl. **211**(1), 13–24 (1994)
46. Metzler, R., Klafter, J.: The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Rep. **339**(1), 1–77 (2000)
47. Mittag-Leffler, G.: Sur la nouvelle function E_α. C. R. Acad. Sci. Paris **137**, 554–558 (1903)
48. Mustapha, K.: Time-stepping discontinuous Galerkin methods for fractional diffusion problems. Numer. Math. **130**(3), 497–516 (2015)
49. Mustapha, K., Abdallah, B., Furati, K.: A discontinuous Petrov–Galerkin method for time-fractional diffusion equations. SIAM J. Numer. Anal. **52**(5), 2512–2529 (2014)

50. Mustapha, K., McLean, W.: Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations. SIAM J. Numer. Anal. **51**(1), 491–515 (2013)
51. Mustapha, K., McLean, W.: Uniform convergence for a discontinuous Galerkin, time-stepping method applied to a fractional diffusion equation. IMA J. Numer. Anal. **32**(3), 906–925 (2012)
52. Podlubny, I.: Fractional Differential Equations, Volume 198 of Mathematics in Science and Engineering. Academic Press, Berlin (1999)
53. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations, 1st. Number 23 in Springer Series in Computational Mathematics. Springer, Berlin (2008)
54. Sakamoto, K., Yamamoto, M.: Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. J. Math. Anal. Appl. **382**(1), 426–447 (2011)
55. Shen, J., Sheng, C.: An efficient space-time method for time fractional diffusion equation. J. Sci. Comput. **81**(2), 1088–1110 (2019)
56. Shen, J., Tang, T., Wang, L.: Spectral Methods: Algorithms, Analysis and Applications. Number 41 in Springer Series in Computational Mathematics. Springer, Heidelberg (2011)
57. Stein, E., Shakarchi, R.: Functional Analysis-Introduction to Further Topics in Analysis. Number 4 in Princeton Lectures in Analysis. Princeton University Press, Princeton (2011)
58. Stynes, M., O'Riordan, E., Gracia, J.: Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. SIAM J. Numer. Anal. **55**(2), 1057–1079 (2017)
59. Tartar, L.: An Introduction to Sobolev Spaces and Interpolation Spaces. Number 3 in Lecture Notes of The Unione Matematica Italiana. Springer, Berlin (2007)
60. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems. Springer, Berlin (2006)
61. Wang, Z.-Q., Guo, Y.-L., Yi, L.-J.: An *hp*-version Legendre–Jacobi spectral collocation method for Volterra integro-differential equations with smooth and weakly singular kernels. Math. Comput. **86**(307), 2285–2324 (2017)
62. Xing, Y., Yan, Y.: A higher order numerical method for time fractional partial differential equations with nonsmooth data. J. Comput. Phys. **357**, 305–323 (2018)
63. Yan, Y., Khan, M., Ford, N.: An analysis of the modified L1 scheme for time-fractional partial differential equations with nonsmooth data. SIAM J. Numer. Anal. **56**(1), 210–227 (2018)
64. Yang, Y., Yan, Y., Ford, N.: Some time stepping methods for fractional diffusion problems with nonsmooth data. Comput. Methods Appl. Mat. **18**(1), 129–146 (2018)
65. Zaslavsky, G.: Chaos, fractional kinetics, and anomalous transport. Phys. Rep. **371**(6), 461–580 (2002)
66. Zheng, M., Liu, F., Turner, I., Anh, V.: A novel high order space-time spectral method for the time fractional Fokker–Planck equation. SIAM J. Sci. Comput. **37**(2), A701–A724 (2015)