An Analysis of the L1 Scheme for the Subdiffusion Equation with Nonsmooth Data

BANGTI JIN†

Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK

RAYTCHO LAZAROV, AND ZHI ZHOU

Department of Mathematics, Texas A&M University, College Station, Texas 77843-3368 USA

The subdiffusion equation with a Caputo fractional derivative of order $\alpha \in (0,1)$ in time arises in a wide variety of practical applications, and it is often adopted to model anomalous subdiffusion processes in heterogeneous media. The L1 scheme is one of the most popular and successful numerical methods for discretizing the Caputo fractional derivative in time. The scheme was analyzed earlier independently by Lin and Xu (2007) and Sun and Wu (2006), and an $O(\tau^{2-\alpha})$ convergence rate was established, under the assumption that the solution is twice continuously differentiable in time. However, in view of the smoothing property of the subdiffusion equation, this regularity condition is restrictive, since it does not hold even for the homogeneous problem with a smooth initial data. In this work, we revisit the error analysis of the scheme, and establish an $O(\tau)$ convergence rate for both smooth and nonsmooth initial data. The analysis is valid for more general sectorial operators. In particular, the L1 scheme is applied to one-dimensional space-time fractional diffusion equations, which involves also a Riemann-Liouville derivative of order $\beta \in (3/2, 2)$ in space, and error estimates are provided for the fully discrete scheme. Numerical experiments are provided to verify the sharpness of the error estimates, and robustness of the scheme with respect to data regularity.

Keywords: fractional diffusion, L1 scheme, error estimates, space-time fractional diffusion.

1. Introduction

We consider the model initial–boundary value problem for the following fractional order parabolic differential equation for $u(x,t)$:

$$
\begin{aligned}
\partial_t^\alpha u - \Delta u &= f, && \text{in } \Omega & T \geqslant t > 0, \\
u &= 0, && \text{on } \partial\Omega & T \geqslant t > 0, \\
u(0) &= v, && \text{in } \Omega,
\end{aligned}
\tag{1.1}
$$

where Ω is a bounded convex polygonal domain in $\mathbb{R}^d$ $(d = 1,2,3)$ with a boundary $\partial\Omega$, v is a given function on Ω, and $T > 0$ is a fixed value. Here $\partial_t^\alpha u$ $(0 < \alpha < 1)$ denotes the left-sided Caputo fractional derivative of order α with respect to t and it is defined by (see, e.g. (Kilbas *et al.*, 2006, pp. 91))

$$
\partial_t^\alpha u(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} u'(s) \, ds,
\tag{1.2}
$$

---

†Corresponding author. Email: bangti.jin@gmail.com

where $\Gamma(\cdot)$ is Euler's Gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

The model (1.1) is known to capture well the dynamics of subdiffusion processes, in which the mean square variance grows at a rate slower than that in a Gaussian process (cf. Bouchaud & Georges (1990)), and has found a number of applications. For example, subdiffusion has been successfully used to describe thermal diffusion in media with fractal geometry (cf. Nigmatulin (1986)), highly heterogeneous aquifer (cf. Adams & Gelhar (1992)) and underground environmental problem (cf. Hatano & Hatano (1998)). At a microscopic level, the particle motion can be described by a continuous time random walk, in which the waiting time of the particle motion follows a heavy tailed distribution, as opposed to a Gaussian process, which is characteristic of the normal diffusion equation. The macroscopic counterpart is a diffusion equation with a Caputo fractional derivative in time, i.e., (1.1).

The derivation and study of accurate numerical methods for the model (1.1) with provable (possibly optimal-order) error bounds have attracted considerable interest in recent years, and a number of efficient numerical schemes, notably the finite difference method, have been developed. There are two predominant approximations: the L1 type approximation and Grünwald-Letnikov type approximation. The former is essentially of finite difference nature, whereas the latter is often based on convolution quadrature. We refer interested readers to (Jin *et al.*, 2014b, Section 1 and Table 1) for an updated overview of these numerical methods and relevant convergence rate results.

In this paper, for reasons to be explained below, we revisit the L1 scheme, which was independently developed and analyzed in Sun & Wu (2006) and Lin & Xu (2007). To this end, we divide the interval $[0,T]$ into a uniform grid with a time step size $\tau = T/N$, $N \in \mathbb{N}$, so that $0 = t_0 < t_1 < \ldots < t_N = T$, and $t_n = n\tau$, $n = 0, \ldots, N$. The L1 approximation of the Caputo fractional derivative $\partial_t^\alpha u(x, t_n)$ is given by (Lin & Xu, 2007, Section 3)

$$
\begin{aligned}
\partial_t^\alpha u(x, t_n) &= \frac{1}{\Gamma(1-\alpha)} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \frac{\partial u(x,s)}{\partial s} (t_n - s)^{-\alpha} ds \\
&\approx \frac{1}{\Gamma(1-\alpha)} \sum_{j=0}^{n-1} \frac{u(x, t_{j+1}) - u(x, t_j)}{\tau} \int_{t_j}^{t_{j+1}} (t_n - s)^{-\alpha} ds \\
&= \sum_{j=0}^{n-1} b_j \frac{u(x, t_{n-j}) - u(x, t_{n-j-1})}{\tau^\alpha} \\
&= \tau^{-\alpha} [b_0 u(x, t_n) - b_{n-1} u(x, t_0) + \sum_{j=1}^{n-1} (b_j - b_{j-1}) u(x, t_{n-j})] =: L_1^n(u).
\end{aligned}
\tag{1.3}
$$

where the weights b_j are given by

$$
b_j = ((j+1)^{1-\alpha} - j^{1-\alpha}) / \Gamma(2-\alpha), \ j = 0, 1, \ldots, N-1.
$$

It was shown in (Lin & Xu, 2007, equation (3.3)) (see also (Sun & Wu, 2006, Lemma 4.1)) that the local truncation error of the L1 approximation is bounded by $c\tau^{2-\alpha}$ for some constant c depending only on u, provided that the solution u is twice continuously differentiable. Further, upon discretizing the spatial derivative(s) using the finite difference method, finite element method or spectral method, we arrive at a fully discrete scheme. Since its first appearance, the L1 scheme has been extensively used in practice, and currently it is one of the most popular and successful numerical methods for solving the time fractional diffusion equation (1.1), including the case of nonsmooth data arising in inverse problems (see, e.g., Jin & Rundell (2012)).

Note that the fractional diffusion operator has only very limited smoothing property, especially for t close to zero. For example, for the homogeneous equation with an initial data $v \in L^2(\Omega)$, we have the following stability estimate (Sakamoto & Yamamoto, 2011, Theorem 2.1)

$$\|\partial_t^\alpha u\|_{L^2(\Omega)} \leqslant ct^{-\alpha}\|v\|_{L^2(\Omega)}.$$

That is, the α-th order Caputo derivative is already unbounded, not to mention the high-order derivatives. In view of the limited regularity of the solution of (1.1), the high regularity required in the convergence analysis in these useful works is restrictive.

Hence the C^2-regularity assumption generally does not hold for problem (1.1), and the case of nonsmooth data is not covered by the existing error analysis. Our numerical experiments indicate that the $O(\tau^{2-\alpha})$ convergence rate actually does not hold even for smooth initial data v. This is clearly seen from the numerical results in Table 1, where `rate` denotes the empirical convergence rate, and the numbers in the bracket denote the theoretical rates based on the local truncation error. These computational results show that the scheme is only first-order accurate even when the initial data v is smooth. This observation necessitates revisiting the convergence analysis of the L1 scheme, especially for the case of nonsmooth problem data.

The goal of this paper is to fill the gap between the existing convergence theory and the numerical experiments, namely, establishing of optimal error bounds that are expressed directly in terms of the regularity of the problem data. For the standard parabolic equation, such bounds are well known (see, e.g. (Thomée, 2006, Chapter 7)). To the best of our knowledge, there is no such error analysis of the L1 scheme in the case of nonsmooth data. Such theory and numerical experiments are provided in this work.

Table 1. The L^2-norm of the error $\|U_h^n - u(t_n)\|_{L^2(\Omega)}$ for problem (1.1) with $f \equiv 0$ and (a) $v = \chi_{(0,1/2)}$ and (b) $v = x(1-x)$, at $t = 0.1$, computed by the fully discrete scheme (2.8) with $h = 2^{-12}$.

α	N	10	20	40	80	160	320	rate
0.1	(a)	3.30e-4	1.62e-4	8.02e-5	3.99e-5	1.99e-5	9.93e-6	≈ 1.01 (1.90)
	(b)	4.97e-4	2.44e-4	1.21e-4	6.00e-5	2.99e-5	1.49e-5	≈ 1.00 (1.90)
0.5	(a)	3.04e-3	1.44e-3	6.96e-4	3.41e-4	1.68e-4	8.35e-5	≈ 1.03 (1.50)
	(b)	4.61e-3	2.18e-3	1.05e-3	5.16e-4	2.54e-4	1.34e-5	≈ 1.02 (1.50)
0.9	(a)	1.31e-2	6.46e-3	3.17e-3	1.56e-3	7.68e-4	3.78e-4	≈ 1.02 (1.10)
	(b)	1.94e-2	9.67e-3	4.78e-3	2.36e-3	1.16e-3	5.75e-4	≈ 1.02 (1.10)

In Theorem 3.4, we present an optimal $O(\tau)$ convergence rate for the fully discrete scheme based on the L1 scheme (1.3) in time and the Galerkin finite element method in space for both smooth and nonsmooth data, i.e., $v \in L^2(\Omega)$ and $Av \in L^2(\Omega)$ ($A = -\Delta$ with a homogeneous Dirichlet boundary condition), respectively. For example, for $v \in L^2(\Omega)$ and $U_h^0 = P_h v$, for the fully discrete solution U_h^n, there holds

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c(\tau t_n^{-1} + h^2 t_n^{-\alpha})\|v\|_{L^2(\Omega)}.$$

Surprisingly, for both $v \in L^2(\Omega)$ and $Av \in L^2(\Omega)$, the error estimate deteriorates as time t approaches zero, but for any fixed time $t_n > 0$, it can achieve a first-order convergence. Extensive numerical experiments confirm the optimality of the convergence rates. Our estimates are derived using the techniques developed by Lubich *et al.* (1996) for convolution quadrature and in the interesting recent work of McLean & Mustapha (2014) on a piecewise constant discontinuous Galerkin method. The proof essentially boils down to some delicate estimates of the kernel function, which involves the polylogarithmic function. Finally, we note that

our results are applicable to more general sectorial operators, including the very interesting case of the space time fractional differential problem involving a Riemann-Liouville derivative in space, cf. Jin *et al.* (2014a).

The rest of the paper is organized as follows. In Section 2, we recall preliminaries on the fully discrete scheme, and derive the solution representation for the semidiscrete and fully discrete schemes, which play an important role in the error analysis. The full technical details of the convergence analysis are presented in Section 3. In Section 4 we consider the adaptation of the scheme to a one-dimensional time-space fractional differential equation, and derive optimal convergence rate. It appears to be the first error estimates expressed directly in terms of the data regularity for such equation with nonsmooth data. Numerical results are presented in Section 5 to confirm the convergence theory and the robustness of the scheme. Throughout, the notation c, with or without a subscript, denotes a generic constant, which may differ at different occurrences, but it is always independent of the mesh size h and the time step size τ.

2. Preliminary

In this part, we give the semidiscrete and fully discrete schemes, based on a standard Galerkin method in space and the L1 approximation in time.

2.1 *Semidiscrete scheme*

Since the solution $u : (0, T] \to L^2(\Omega)$ can be analytically extended to the sector $\{z \in \mathbb{C}; z \neq 0, |\arg z| < \pi/2\}$ (Sakamoto & Yamamoto, 2011, Theorem 2.1), when $f \equiv 0$, we may apply the Laplace transform to equation (1.1) to deduce

$$z^{\alpha}\widehat{u}(z) + A\widehat{u}(z) = z^{\alpha-1}v, \qquad (2.1)$$

with the operator $A = -\Delta$ with a homogeneous Dirichlet boundary condition. Hence the solution $u(t)$ can be represented by

$$u(t) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt}(z^{\alpha}I + A)^{-1}z^{\alpha-1}v \, dz, \qquad (2.2)$$

where the contour $\Gamma_{\theta,\delta}$ is given by

$$\Gamma_{\theta,\delta} = \{z \in \mathbb{C} : |z| = \delta, |\arg z| \leqslant \theta\} \cup \{z \in \mathbb{C} : z = \rho e^{\pm i\theta}, \rho \geqslant \delta\}.$$

Throughout, we choose the angle $\theta \in (\pi/2, \pi)$. Then $z^{\alpha} \in \Sigma_{\theta'}$ with $\theta' = \alpha\theta < \pi$ for all $z \in \Sigma_{\theta} := \{z \in \mathbb{C} : |\arg z| \leqslant \theta\}$. Then there exists a constant c which depends only on θ and α such that

$$\|(z^{\alpha}I + A)^{-1}\| \leqslant cz^{-\alpha}, \quad \forall z \in \Sigma_{\theta}. \qquad (2.3)$$

Now we introduce the spatial semidiscrete scheme based on the Galerkin finite element method. Let $\mathscr{T}_h$ be a shape regular and quasi-uniform triangulation of the domain Ω into d-simplexes, denoted by T. Then over the triangulation $\mathscr{T}_h$ we define a continuous piecewise linear finite element space X_h by

$$X_h = \left\{v_h \in H_0^1(\Omega) : v_h|_T \text{ is a linear function}, \forall T \in \mathscr{T}_h\right\}.$$

On the space X_h, we define the $L^2(\Omega)$-orthogonal projection $P_h : L^2(\Omega) \to X_h$ and the Ritz projection $R_h : H_0^1(\Omega) \to X_h$, respectively, by

$$(P_h\varphi, \chi) = (\varphi, \chi) \quad \forall \chi \in X_h,$$
$$(\nabla R_h\varphi, \nabla\chi) = (\nabla\varphi, \nabla\chi) \quad \forall \chi \in X_h,$$

where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$-inner product. Then the semidiscrete Galerkin scheme for problem (1.1) reads: find $u_h(t) \in X_h$ such that

$$(\partial_t^\alpha u_h, \chi) + (\nabla u_h, \nabla \chi) = (f, \chi) \quad \forall \chi \in X_h, \tag{2.4}$$

with $u_h(0) = v_h \in X_h$. Upon introducing the discrete Laplacian $\Delta_h : X_h \to X_h$ defined by

$$-(\Delta_h \varphi, \chi) = (\nabla \varphi, \nabla \chi) \quad \forall \varphi, \chi \in X_h,$$

the spatial semidiscrete scheme (2.4) can be rewritten into

$$\partial_t^\alpha u_h(t) + A_h u_h(t) = f_h(t), \ t > 0 \tag{2.5}$$

with $u_h(0) = v_h \in X_h$, $f_h = P_h f$ and $A_h = -\Delta_h$. Like before, the solution u_h to (2.5) with $f_h \equiv 0$ can be represented by

$$u_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha + A_h)^{-1} z^{\alpha-1} v_h \, dz. \tag{2.6}$$

Further, for later analysis, we let $w_h = u_h - v_h$. Then w_h satisfies the problem:

$$\partial_t^\alpha w_h + A_h w_h = -A_h v_h,$$

with $w_h(0) = 0$. The Laplace transform gives

$$z^\alpha \widehat{w}_h(z) + A_h \widehat{w}_h(z) = -z^{-1} A_h v_h.$$

Hence, $\widehat{w}_h(z) = K_1(z) v_h$, with

$$K_1(z) = -z^{-1}(z^\alpha I + A_h)^{-1} A_h,$$

and the desired representation for $w_h(t)$ follows from the inverse Laplace transform

$$w_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} K_1(z) v_h \, dz. \tag{2.7}$$

The semidiscrete solution u_h satisfies the following estimates from Jin *et al.* (2013a). The log factor in the estimates in (Jin *et al.*, 2013a, Section 3) can be removed using the operator trick in (Bazhlekova *et al.*, 2014, Section 3).

THEOREM 2.1 Let u and u_h be the solutions of problems (1.1) and (2.5) with $f \equiv 0$ and $f_h \equiv 0$, respectively. Then the following error estimates hold.

(a) If $Av \in L^2(\Omega)$ and $v_h = R_h v$, then

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} \leqslant ch^2 \|Av\|_{L^2(\Omega)}.$$

(b) If $v \in L^2(\Omega)$ and $v_h = P_h v$, then

$$\|u(t_n) - u_h(t)\|_{L^2(\Omega)} \leqslant ch^2 t^{-\alpha} \|v\|_{L^2(\Omega)}.$$

2.2 *Fully discrete scheme*

Now we describe the fully discrete scheme based on the L1 approximation (1.3): find $U_h^n \in X_h$ for $n = 1, 2, \ldots, N$

$$(b_0 I + \tau^\alpha A_h) U_h^n = b_{n-1} U_h^0 + \sum_{j=1}^{n-1} (b_{j-1} - b_j) U_h^{n-j} + \tau^\alpha F_h^n, \tag{2.8}$$

with $U_h^0 = v_h$ and $F_h^n = P_h f(t_n)$. We focus on the homogeneous case, i.e., $f \equiv 0$, and use the technique in Lubich *et al.* (1996) and McLean & Mustapha (2014) for convergence analysis. Throughout, we denote by

$$\widetilde{\omega}(\xi) = \sum_{j=0}^{\infty} \omega_j \xi^j$$

the generating function of a sequence $\{\omega_j\}_{j=0}^{\infty}$. To analyze the fully discrete scheme (2.8), we first derive a discrete analogue of the solution representation (2.7). The fully discrete solution $W_h^n := U_h^n - U_h^0$ satisfies the following time-stepping scheme for $n = 1, 2, \ldots, N$

$$L_1^n(W_h) + A_h W_h^n = -A_h v_h,$$

with $W_h^0 = 0$. Next multiplying both sides of the equation by ξ^n and summing from 1 to ∞ yields

$$\sum_{n=1}^{\infty} L_1^n(W_h) \xi^n + A_h \widetilde{W}_h(\xi) = -\frac{\xi}{1-\xi} A_h v_h.$$

Now we focus on the term $\sum_{n=1}^{\infty} L_1^n(W_h) \xi^n$. By the definition of the difference operator L_1^n, we have

$$\sum_{n=1}^{\infty} L_1^n(W_h) \xi^n = \tau^{-\alpha} \sum_{n=1}^{\infty} \left(b_0 W_h^n + \sum_{j=1}^{n-1} (b_j - b_{j-1}) W_h^{n-j} \right) \xi^n$$

$$= \tau^{-\alpha} \sum_{n=1}^{\infty} \left(\sum_{j=0}^{n-1} b_j W_h^{n-j} \right) \xi^n - \tau^{-\alpha} \sum_{n=1}^{\infty} \left(\sum_{j=1}^{n-1} b_{j-1} W_h^{n-j} \right) \xi^n$$

$$:= I - II.$$

Using the fact $W_h^0 = 0$ and the convolution rule of generating functions (discrete Laplace transform), the first term I can be written as

$$I = \tau^{-\alpha} \sum_{n=1}^{\infty} (\sum_{j=0}^{n} b_j W_h^{n-j}) \xi^n = \tau^{-\alpha} \widetilde{b}(\xi) \widetilde{W}_h(\xi).$$

Similarly, the second term II can be written as

$$II = \tau^{-\alpha} \sum_{n=1}^{\infty} (\sum_{j=1}^{n} b_{j-1} W_h^{n-j}) \xi^n = \tau^{-\alpha} \xi \sum_{n=1}^{\infty} (\sum_{j=0}^{n-1} b_j W_h^{n-1-j}) \xi^{n-1} = \tau^{-\alpha} \xi \widetilde{b}(\xi) \widetilde{W}_h(\xi).$$

Hence, we arrive at

$$\sum_{n=1}^{\infty} L_1^n(W_h) \xi^n = \tau^{-\alpha} (1 - \xi) \widetilde{b}(\xi) \widetilde{W}_h(\xi).$$

Next we derive a proper representation for $\widetilde{b}(\xi)$:

$$\widetilde{b}(\xi) = \frac{1}{\Gamma(2-\alpha)} \sum_{j=0}^{\infty} ((j+1)^{1-\alpha} - j^{1-\alpha})\xi^j$$

$$= \frac{1-\xi}{\xi\Gamma(2-\alpha)} \sum_{j=1}^{\infty} j^{1-\alpha}\xi^j = \frac{(1-\xi)\mathrm{Li}_{\alpha-1}(\xi)}{\xi\Gamma(2-\alpha)},$$

where $\mathrm{Li}_p(z)$ denotes the polylogarithm function defined by (see Lewin (1981))

$$\mathrm{Li}_p(z) = \sum_{j=1}^{\infty} \frac{z^j}{j^p}.$$

The polylogarithm function $\mathrm{Li}_p(z)$ is well defined for $|z| < 1$, and it can be analytically continued to the split complex plane $\mathbb{C} \setminus [1,\infty)$; see Flajolet (1999). With $z = 1$, it recovers the Riemann zeta function $\zeta(p) = \mathrm{Li}_p(1)$. Therefore, the fully discrete solution $\widetilde{W}_h(\xi)$ can be represented by

$$\widetilde{W}_h(\xi) = -\frac{\xi}{1-\xi}\left(\frac{(1-\xi)^2}{\xi\tau^\alpha\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(\xi) + A_h\right)^{-1} A_h v_h.$$

Simple calculation shows that the function $\widetilde{W}_h(\xi)$ is analytic at $\xi = 0$. Hence the Cauchy theorem implies that for ρ small enough, there holds

$$W_h^n = -\frac{1}{2\pi\mathrm{i}} \int_{|\xi|=\rho} \frac{1}{(1-\xi)\xi^n}\left(\frac{(1-\xi)^2}{\xi\tau^\alpha\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(\xi) + A_h\right)^{-1} A_h v_h \, d\xi.$$

Upon changing variable $\xi = e^{-z\tau}$, we obtain

$$W_h^n = -\frac{1}{2\pi\mathrm{i}} \int_{\Gamma^0} e^{zt_{n-1}} \frac{\tau}{1-e^{-z\tau}}\left(\frac{(1-e^{-z\tau})^2}{e^{-z\tau}\tau^\alpha\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(e^{-z\tau}) + A_h\right)^{-1} A_h v_h \, dz,$$

where the contour $\Gamma^0 := \{z = -\ln(\rho)/\tau + \mathrm{i}y : |y| \leqslant \pi/\tau\}$ is oriented counterclockwise. By deforming the contour Γ^0 to $\Gamma_\tau := \{z \in \Gamma_{\theta,\delta} : |\Im(z)| \leqslant \pi/\tau\}$ and using the periodicity of the exponential function, we obtain the following alternative representation for W_h^n

$$W_h^n = -\frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\tau} e^{zt_{n-1}} \frac{\tau}{1-e^{-z\tau}}\left(\frac{(1-e^{-z\tau})^2}{e^{-z\tau}\tau^\alpha\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(e^{-z\tau}) + A_h\right)^{-1} A_h v_h \, dz. \tag{2.9}$$

This representation forms the basis of the error analysis.

3. Error analysis of the fully discrete scheme

In this part, we derive optimal error estimates for the fully discrete scheme (2.8). The analysis is based on the representations of the semidiscrete and fully discrete solutions, i.e., (2.7) and (2.9). Upon subtracting them, we may write the difference between W_h^n and $w_h(t_n)$ as

$$w_h(t_n) - W_h^n = I + II,$$

where the terms *I* and *II* are given by

$$I = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta} \setminus \Gamma_\tau} e^{z t_n} K_1(z) v_h dz$$

and

$$II = \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{z t_n} (K_1(z) - e^{-z\tau} K_2(z)) v_h dz.$$

Here the kernel functions $K_1(z)$ and $K_2(z)$ are given by

$$K_1(z) := -z^{-1} (z^\alpha + A_h)^{-1} A_h =: z^{-1} B_1(z) \tag{3.1}$$

and

$$K_2(z) := -\frac{\tau}{1 - e^{-z\tau}} \left(\frac{1 - e^{-z\tau}}{\tau^\alpha} \psi(z\tau) + A_h \right)^{-1} A_h =: \frac{\tau}{1 - e^{-z\tau}} B_2(z), \tag{3.2}$$

respectively, and the auxiliary function ψ is define by

$$\psi(z\tau) = \frac{e^{z\tau} - 1}{\Gamma(2 - \alpha)} \mathrm{Li}_{\alpha-1}(e^{-z\tau}).$$

Since the function $|e^{-z\tau}|$ is uniformly bounded on the contour Γ_τ, we deduce

$$\begin{aligned}
\|K_1(z) - e^{-z\tau} K_2(z)\| &\leqslant |e^{-z\tau}| \|K_1(z) - K_2(z)\| + |1 - e^{-z\tau}| \|K_1(z)\| \\
&\leqslant c \|K_1(z) - K_2(z)\| + c|z|\tau \|K_1(z)\| \\
&\leqslant c \|K_1(z) - K_2(z)\| + c\tau,
\end{aligned} \tag{3.3}$$

where the last line follows from the inequality, in view of (2.3):

$$\|K_1(z)\| = |z|^{-1} \| - I + z^\alpha (z^\alpha + A_h)^{-1} \| \leqslant c|z|^{-1}.$$

Thus it suffices to establish a bound on $\|K_1(z) - K_2(z)\|$, which will be carried out below. First we give bounds on the function $\chi(z) = \tau^{-1}(1 - e^{-z\tau})$.

LEMMA 3.1 Let $\chi(z) = \tau^{-1}(1 - e^{-z\tau})$. Then for all $z \in \Gamma_\tau$, there hold for some $c_1, c_2 > 0$

$$|\chi(z) - z| \leqslant c|z|^2 \tau \quad \text{and} \quad c_1|z| \leqslant |\chi(z)| \leqslant c_2|z|.$$

Proof. We note that $|z|\tau \leqslant \pi / \sin\theta$ for $z \in \Gamma_\tau$. Then the first assertion follows by

$$|\chi(z) - z| \leqslant |z|^2 \tau \left| \sum_{j=0}^\infty \frac{|z|^j \tau^j}{(j+2)!} \right| \leqslant c|z|^2 \tau \quad \text{for } z \in \Gamma_\tau.$$

Next we consider the second claim. The upper bound on $\chi(z)$ is trivial. Thus it suffices to verify the lower bound. To this end, we split the contour Γ_τ into three disjoint parts $\Gamma_\tau = \Gamma_\tau^+ \cup \Gamma_\tau^c \cup \Gamma_\tau^-$, with Γ_τ^+ and Γ_τ^- being the rays in the upper and lower half planes, respectively, and Γ_τ^c being the circular arc. Here we set $\xi = -z\tau$ with $\rho \equiv |\xi| \in (0, \pi / \sin\theta)$. We first consider the case of $z \in \Gamma_\tau^+$, for which $\xi = \rho e^{-i(\pi-\theta)}$, $\rho \in (1, \pi / \sin\theta)$. Using the trivial inequality $|\cos(\rho \sin\theta)| \leqslant 1$, we obtain

$$\begin{aligned}
\left| \frac{1 - e^{-z\tau}}{z\tau} \right| = \left| \frac{e^\xi - 1}{\xi} \right| &= \frac{|e^{-\rho\cos\theta} \cos(\rho\sin\theta) - 1 - i e^{-\rho\cos\theta} \sin(\rho\sin\theta)|}{\rho} \\
&\geqslant \frac{(e^{-2\rho\cos\theta} + 1 - 2e^{-\rho\cos\theta})^{1/2}}{\rho} = \frac{e^{-\rho\cos\theta} - 1}{\rho} \geqslant -\cos\theta > 0
\end{aligned} \tag{3.4}$$

due to positivity and monotonicity of $(e^{-\rho\cos\theta} - 1)/\rho$ as a function of ρ over the interval $(0,\infty)$. The case of $z \in \Gamma_\tau^-$ follows analogously. Last, we consider $z \in \Gamma_\tau^c$, the circular arc. In this case, by means of Taylor expansion, we have

$$\chi(z) = z\left(1 + \sum_{j=1}^{\infty}(-1)^j\frac{z^j\tau^j}{(j+1)!}\right).$$

From this and the fact that $\rho = |z\tau| < 1$, it follows directly that $|\chi(z)| \geqslant c|z|$ for $z \in \Gamma_\tau^c$. This completes the proof of the lemma.   □

Then by the trivial inequality $\|B_1(z)\| \leqslant c$ and Lemma 3.1, we deduce that

$$\begin{aligned}\|K_1(z) - K_2(z)\| &\leqslant |z^{-1} - \chi(z)^{-1}|\|B_1(z)\| + |\chi(z)|^{-1}\|B_1(z) - B_2(z)\|\\ &\leqslant \frac{|z - \chi(z)|}{|z\chi(z)|} + c|z|^{-1}\|B_1(z) - B_2(z)\|\\ &\leqslant c\tau + c|z|^{-1}\|B_1(z) - B_2(z)\|.\end{aligned} \tag{3.5}$$

Thus it suffices to establish a bound on $\|B_1(z) - B_2(z)\|$. This will be done using a series of lemmas. To this end, first we recall an important singular expansion of the function $\mathrm{Li}_p(e^{-z})$ (Flajolet, 1999, Theorem 1).

LEMMA 3.2 For $p \neq 1, 2, \ldots$, the function $\mathrm{Li}_p(e^{-z})$ satisfies the singular expansion

$$\mathrm{Li}_p(e^{-z}) \sim \Gamma(1-p)z^{p-1} + \sum_{k=0}^{\infty}(-1)^k\zeta(p-k)\frac{z^k}{k!} \quad \text{as } z \to 0, \tag{3.6}$$

where ζ is the Riemann zeta function.

REMARK 3.1 The singular expansion in Lemma 3.2 is stated only for $z \to 0$. However, we note that the expansion is valid in the sector $\Sigma_{\theta,\delta}$ (Flajolet, 1999, pp. 377, proof of Lemma 2).

We shall need the following results. The first result gives the absolute convergence of a special series involving the Riemann zeta function ζ.

LEMMA 3.3 Let $|z| \leqslant \pi/\sin\theta$ with $\theta \in (\pi/2, 5\pi/6)$, and $p = \alpha - 1$. Then the series (3.6) converges absolutely.

Proof. Using the following well-known functional equation for the Riemann zeta function (see e.g., Knopp & Robins (2001) for a short proof): for $z \notin \mathbb{Z}$, there holds

$$\zeta(1-z) = \frac{2}{(2\pi)^z}\cos\left(\frac{z\pi}{2}\right)\Gamma(z)\zeta(z),$$

we obtain for $p = \alpha - 1 \in (-1, 0)$

$$\begin{aligned}\zeta(p-k) &= \zeta(1 - (1-p+k))\\ &= \frac{2}{(2\pi)^{1-p+k}}\cos\left(\frac{(1-p+k)\pi}{2}\right)\Gamma(1-p+k)\zeta(1-p+k).\end{aligned}$$

By Stirling's formula for the Gamma function $\Gamma(x)$, $x \to \infty$ (Abramowitz & Stegun, 1964, pp. 257)

$$\Gamma(x+1) = x^{x+1}e^{-x}\sqrt{\frac{2\pi}{x}}\left(1 + O(x^{-1})\right)$$

and that $\zeta(1-p+k) \to 1$ as $k \to \infty$, we have

$$\lim_{k\to\infty} \sqrt[k]{\frac{|\zeta(p-k)||z|^k}{k!}} \leqslant \frac{1}{2\sin\theta} \quad \forall |z| \leqslant \pi/\sin\theta.$$

Since for $\theta \in (\pi/2, 5\pi/6)$, $2\sin\theta > 1$, the series converges absolutely. □

Next we state an error estimate for the function $\frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau)$ with respect to z^α.

LEMMA 3.4 Let $\psi(z) = \frac{e^z-1}{\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(e^{-z})$. Then for the choice $\theta \in (\pi/2, 5\pi/6)$, there holds

$$|\frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau) - z^\alpha| \leqslant c|z|^2\tau^{2-\alpha} \quad \forall z \in \Gamma_\tau.$$

Proof. Upon noting the fact $0 \leqslant |z\tau| \leqslant \pi/\sin\theta$ and using Taylor expansion and (3.6), we deduce that for $z \in \Gamma_\tau$, there holds

$$e^{z\tau} - 1 = \sum_{j=1}^\infty \frac{(z\tau)^j}{j!},$$

$$\mathrm{Li}_{\alpha-1}(e^{-z\tau}) = \Gamma(2-\alpha)(z\tau)^{\alpha-2} + \sum_{k=0}^\infty (-1)^k \zeta(1-\alpha-k)\frac{(z\tau)^k}{k!}.$$

Hence, the function $\psi(z)$ can be represented by

$$\psi(z\tau) = \sum_{j=1}^\infty \frac{(z\tau)^j}{j!}\left[(z\tau)^{\alpha-2} + \sum_{k=0}^\infty \frac{(-1)^k\zeta(-\alpha-k)}{\Gamma(2-\alpha)}\frac{(z\tau)^k}{k!}\right]$$

$$= \sum_{j=1}^\infty \frac{(z\tau)^{\alpha+j-2}}{j!} + \sum_{j=1}^\infty \frac{(z\tau)^j}{j!}\sum_{k=0}^\infty \frac{(-1)^k\zeta(-\alpha-k)}{\Gamma(2-\alpha)}\frac{(z\tau)^k}{k!},$$

and

$$\frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau) = \tau^{-\alpha}\sum_{l=0}^\infty (-1)^l\frac{(z\tau)^{l+1}}{(l+1)!}\left[\sum_{j=1}^\infty \frac{(z\tau)^{\alpha+j-2}}{j!} + \sum_{j=1}^\infty \frac{(z\tau)^j}{j!}\sum_{k=0}^\infty \frac{(-1)^k\zeta(-\alpha-k)}{\Gamma(2-\alpha)}\frac{(z\tau)^k}{k!}\right]$$

$$= z^\alpha + \frac{\zeta(\alpha-1)}{\Gamma(2-\alpha)}z^2\tau^{2-\alpha} + O(z^{2+\alpha}\tau^2).$$

In view of the choice $\theta \in (\pi/2, 5\pi/6)$ and Lemma 3.3, the bound is uniform, since the series converges uniformly for $z \in \Gamma_\tau$. Consequently,

$$|\frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau) - z^\alpha| \leqslant |z|^2\tau^{2-\alpha}\left(-\frac{\zeta(\alpha-1)}{\Gamma(2-\alpha)} + O((z\tau)^\alpha)\right) \leqslant c|z|^2\tau^{2-\alpha},$$

from which the desired assertion follows. □

The next result gives a uniform lower bound on the function $\psi(z)$ on the contour Γ_τ.

LEMMA 3.5 Let $\psi(z) = \frac{e^z-1}{\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(e^{-z})$. Then for any θ close to $\pi/2$, there holds for any $\delta < \pi/2\tau$

$$|\psi(z\tau)| \geqslant c > 0 \quad \forall z \in \Gamma_\tau.$$

Proof. Since for $z \in \Gamma_\tau$, $|\Im z| \leqslant \pi/\tau$ and $z \notin (-\infty, 0]$, by (McLean & Mustapha, 2014, Lemma 1), there holds

$$\psi(z) = c_\alpha \int_0^\infty \frac{s^{\alpha-1}}{1 - e^{-z-s}} \frac{1 - e^{-s}}{s} \, ds,$$

with the constant $c_\alpha = \sin(\pi(1-\alpha))/\pi$. To prove the assertion, we again split the contour Γ_τ into $\Gamma = \Gamma_\tau^+ \cup \Gamma_\tau^c \cup \Gamma_\tau^-$, where Γ_τ^+ and Γ_τ^- are the rays in the upper and lower half planes, respectively, and Γ_τ^c is the circular arc of the contour Γ_τ, and discuss the three cases separately. We first consider the case $z \in \Gamma_\tau^+$ and set $z\tau = \rho e^{i\theta} = \rho \cos\theta + i\rho \sin\theta$ with $\delta < \rho < \pi/\sin\theta$. Upon letting $r = \rho \cos\theta$ and $\phi = \rho \sin\theta$ then

$$\begin{aligned}
\psi(z\tau) &= c_\alpha \int_0^\infty \frac{s^{\alpha-1}}{1 - e^{-(r+i\phi)-s}} \frac{1 - e^{-s}}{s} \, ds \\
&= c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1 - e^{-s})}{1 - e^{-r-s}\cos\phi + ie^{-r-s}\sin\phi} \, ds \\
&= c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1 - e^{-s})(1 - e^{-r-s}\cos\phi - ie^{-r-s}\sin\phi)}{(1 - e^{-r-s}\cos\phi)^2 + e^{-2r-2s}\sin^2\phi} \, ds.
\end{aligned}$$

It suffices to show that the real part

$$\Re\psi(z\tau) = c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1 - e^{-s})(1 - e^{-r-s}\cos\phi)}{1 - 2e^{-r-s}\cos\phi + e^{-2r-2s}} \, ds$$

is bounded from below by some positive constant c. First we consider the case $\phi = \rho \sin\theta \in [\pi/2, \pi]$, for which $\cos\phi \leqslant 0$ and thus

$$\begin{aligned}
0 &< 1 - e^{-r-s}\cos\phi \leqslant 1 - 2e^{-r-s}\cos\phi \\
&\leqslant 1 - 2e^{-r-s}\cos\phi + e^{-2r-2s}.
\end{aligned}$$

Consequently,

$$\Re\psi(z\tau) \geqslant c_\alpha \int_0^\infty s^{\alpha-2}(1 - e^{-s}) \, ds = c_0.$$

Next we consider the case $\phi \in (0, \pi/2)$, for which $\cos\phi > 0$. Further we fix $\theta = \pi/2$, and thus $r = \rho \cos\theta = 0$ and $e^{-r}\cos\phi = \cos(\rho \sin\theta) = \cos\rho > 0$. Then

$$1 - e^{-r-s}\cos\phi > 1 - e^{-s} \quad \text{and} \quad 0 \leqslant 1 - 2e^{-r-s}\cos\phi + e^{-2r-2s} \leqslant 2,$$

and accordingly the real part $\Re\psi(z\tau)$ simplifies to

$$\begin{aligned}
\Re\psi(z\tau) &= c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1 - e^{-s})(1 - e^{-s}\cos\rho)}{1 - 2e^{-s}\cos\rho + e^{-2s}} \, ds \\
&\geqslant \frac{c_\alpha}{2} \int_0^\infty s^{\alpha-2}(1 - e^{-s})^2 \, ds \geqslant c_1.
\end{aligned}$$

Then by continuity of $\Re\psi(z\tau)$, we may choose an angle $\theta \in (\pi/2, 5\pi/6)$ such that for any $z \in \Gamma_\tau^+$, there holds $\Re\psi(z\tau) \geqslant c_2$. Repeating the above argument shows also the assertion for the case $z \in \Gamma_\tau^-$. It remains to show the case $z \in \Gamma_\tau^c$. For any fixed $\rho \in (0, \pi/2)$ and $\theta \in [-\pi/2, \pi/2]$, $\cos\phi = \cos(\rho \sin\theta) \geqslant 0$, $r = \rho \cos\theta \geqslant 0$. Consequently

$$1 - e^{-r-s}\cos\phi \geqslant 1 - e^{-s}\cos\phi \geqslant 1 - e^{-s},$$

and

$$1 - 2e^{-r-s}\cos\phi + e^{-2r-2s} \leqslant 1 + e^{-2r-2s} \leqslant 2.$$

These two inequalities directly imply

$$\Re\psi(z\tau) = c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1-e^{-s})(1-e^{-r-s}\cos\phi)}{1-2e^{-r-s}\cos\phi + e^{-2r-2s}}\,ds$$

$$\geqslant \frac{c_\alpha}{2}\int_0^\infty s^{\alpha-2}(1-e^{-s})^2\,ds \geqslant c_3.$$

Then by continuity, we may choose an angle $\theta > \pi/2$ such that for $z \in \Gamma_\tau^c$, there holds $\Re\psi(z\tau) \geqslant c_4 > 0$. $\square$

The next result shows a "sector-preserving" property of the mapping $\chi_1(z)$: there exists some $\theta_0 < \pi$, such that $\chi_1(z) \in \Sigma_{\theta_0}$ for all $z \in \Sigma_\theta$. This property plays a fundamental role in the error analysis below.

LEMMA 3.6 Let $\psi(z) = \frac{e^z-1}{\Gamma(2-\alpha)}\mathrm{Li}_{\alpha-1}(e^{-z})$ and $\chi_1(z) = \frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau)$. Then there exists some $\theta_0 \in (\pi/2, \pi)$ such that $\chi_1(z) \in \Sigma_{\theta_0}$ for all $z \in \Sigma_\theta$.

Proof. Like before, for $z\tau = \rho e^{i\theta}$, we denote by $r = \rho\cos\theta$, $\phi = \rho\sin\theta$ and $c_\alpha = \sin(\pi(1-\alpha))/\pi$. Then the real part $\Re\chi_1(z)$ and the imaginary part $\Im\chi_1(z)$ of the kernel $\chi_1(z)$ are given by

$$\Re\chi_1(z) = \frac{c_\alpha}{\tau^\alpha}\int_0^\infty \frac{s^{\alpha-2}(1-e^{-s})(1+e^{-2r-s}-e^{-r-s}\cos\phi - e^{-r}\cos\phi)}{1-2e^{-r-s}\cos\phi+e^{-2r-2s}}\,ds \tag{3.7}$$

and

$$\Im\chi_1(z) = \frac{c_\alpha}{\tau^\alpha}\int_0^\infty \frac{s^{\alpha-2}(1-e^{-s})^2 e^{-r}\sin\phi}{1-2e^{-r-s}\cos\phi+e^{-2r-2s}}\,ds, \tag{3.8}$$

respectively. Obviously, for $\theta \in [-\pi/2, \pi/2]$ then $r = \rho\cos\theta \geqslant 0$, $0 < e^{-r} \leqslant 1$ and thus

$$1 + e^{-2r-s} - e^{-r-s}\cos\phi - e^{-r}\cos\phi \geqslant 1 + e^{-2r-s} - e^{-r-s} - e^{-r}$$

$$= (1-e^{-r-s})(1-e^{-r}) \geqslant (1-e^{-s})(1-e^{-r}).$$

Meanwhile, $r \geqslant 0$ implies $0 \leqslant 1 - 2e^{-r-s}\cos\phi + e^{-2r-2s} \leqslant 4$, and consequently

$$\Re\chi_1(z) \geqslant \frac{c_\alpha(1-e^{-r})}{4\tau^\alpha}\int_0^\infty s^{\alpha-2}(1-e^{-s})^2\,ds = \frac{c'_\alpha}{\tau^\alpha}(1-e^{-r}),$$

with the constant $c'_\alpha = \frac{c_\alpha}{4}\int_0^\infty s^{\alpha-2}(1-e^{-s})^2\,ds$.

Next we consider the case $|\theta| > \pi/2$. It suffices to consider the case $\theta > \pi/2$, and the other case $\theta < -\pi/2$ can be treated analogously. Let $z\tau = \rho e^{i\theta}$ with $\rho \in (0,\infty)$. First, clearly, for $\phi = \rho\sin\theta \in [\pi/2, \pi]$, $\cos\phi \leqslant 0$, and there holds

$$0 < 1 - 2e^{-r-s}\cos\phi + e^{-2r-2s} \leqslant (1+e^{-r-s})^2,$$

and thus

$$\Re\chi_1(z) \geqslant \frac{c_\alpha}{\tau^\alpha(1+e^{-r})^2}\int_0^\infty s^{\alpha-2}(1-e^{-s})\,ds > 0.$$

Second, we consider the case $\phi = \rho \sin\theta \in (0, \pi/2)$. There are two possible situations: (a) $1 - e^{-r}\cos\phi \geqslant 0$ and (b) $1 - e^{-r}\cos\phi < 0$. In case (a), we have

$$1 + e^{-2r-s} - e^{-r-s}\cos\phi - e^{-r}\cos\phi$$
$$\geqslant 1 + e^{-2r-s}\cos^2\phi - e^{-r-s}\cos\phi - e^{-r}\cos\phi$$
$$= (1 - e^{-r-s}\cos\phi)(1 - e^{-r}\cos\phi) \geqslant (1 - e^{-s})(1 - e^{-r}\cos\phi).$$

Consequently,

$$\Re\chi_1(z) \geqslant \frac{c_\alpha(1 - e^{-r}\cos\phi)}{\tau^\alpha} \int_0^\infty \frac{s^{\alpha-2}(1 - e^{-s})^2}{1 - 2e^{-r-s}\cos\phi + e^{-2r-2s}} \, ds \geqslant 0.$$

In case (b), we may further assume $\Re\chi_1(z) < 0$, otherwise the statement follows directly. Then appealing to (3.7) and using the trivial inequality $|\cos\phi| \leqslant 1$, we deduce that $e^{-r} > 1$ and

$$1 - e^{-r}\cos\phi < 0 \quad \text{and} \quad e^{-2r} - e^{-r}\cos\phi > 0.$$

With the help of these two inequalities, and the assumption $\Re\chi_1(z) < 0$, we arrive at

$$0 > 1 + e^{-2r-s} - e^{-r-s}\cos\phi - e^{-r}\cos\phi$$
$$= (1 - e^{-r}\cos\phi) + e^{-s}(e^{-2r} - e^{-r}\cos\phi)$$
$$\geqslant (\cos\phi - e^{-r}) + e^{-s}(e^{-2r} - e^{-r}\cos\phi)$$
$$\geqslant e^{-r}(\cos\phi - e^{-r}) + e^{-s}(e^{-2r} - e^{-r}\cos\phi)$$
$$= (1 - e^{-s})(e^{-r}\cos\phi - e^{-2r}),$$

where the first and fourth inequalities follow from $\Re\chi_1(z) < 0$ and $e^{-r} > 1$, respectively. Consequently,

$$|\Re\chi_1(z)| \leqslant \frac{c_1}{\tau^\alpha}(e^{-2r} - e^{-r}\cos\phi),$$

with the constant

$$c_1 = c_1(r, \phi) = c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1 - e^{-s})^2}{1 - 2e^{-r-s}\cos\phi + e^{-2r-2s}} \, ds.$$

Meanwhile, it follows directly from (3.8) that

$$|\Im\chi_1(z)| = \frac{c_1}{\tau^\alpha}e^{-r}\sin\phi.$$

Therefore,

$$\frac{|\Im\chi_1(z)|}{|\Re\chi_1(z)|} \geqslant \frac{\sin(\rho\sin\theta)}{e^{-\rho\cos\theta} - \cos(\rho\sin\theta)} =: g(\rho).$$

Now set $g_1(\rho) = \sin(\rho\sin\theta)$ and $g_2(\rho) = e^{-\rho\cos\theta} - \cos(\rho\sin\theta)$. Since for $\rho \in (0, \pi/(2\sin\theta))$ and $\theta > \pi/2$,

$$\lim_{\rho\to 0} g(\rho) = -\tan\theta, \quad g_1(\rho), g_2(\rho) \geqslant 0, \quad g_1'(\rho) \leqslant 0 \quad \text{and} \quad g_2'(\rho) \geqslant 0,$$

i.e., the function $g(\rho)$ is monotonically decreasing on the interval $[0, \pi/2\sin\theta]$, we deduce

$$\inf_{\rho\in(0,\pi/(2\sin\theta))} g(\rho) = g(\pi/(2\sin\theta)) = e^{\pi\cot\theta/2} > 0.$$

This completes the proof of the lemma. □

REMARK 3.2 By the proof of Lemma 3.6, the sector Σ_{θ_0} depends on the choice of the angle θ. For $\theta \to \pi/2$, it is contained in the sector $\Sigma_{3\pi/4-\varepsilon}$, for any $\varepsilon > 0$.

LEMMA 3.7 Let θ be close to $\pi/2$, and $\delta < \pi/2\tau$. Then for $K_1(z)$ and $K_2(z)$ defined in (3.1) and (3.2), respectively, there holds

$$\|K_1(z) - K_2(z)\| \leqslant c\tau \quad \forall z \in \Gamma_\tau.$$

Proof. For the operators $B_1(z)$ and $B_2(z)$ defined in (3.1) and (3.2), respectively, we have

$$B_1(z) = -I + z^\alpha (z^\alpha + A_h)^{-1} \quad \text{and} \quad B_2(z) = -I + \chi_1(z)(\chi_1(z) + A_h)^{-1},$$

where $\chi_1(z) = \frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau)$. Then by Lemma 3.4

$$
\begin{aligned}
\|B_1(z) - B_2(z)\| &\leqslant |\chi_1(z) - z^\alpha| \|(\chi_1(z) + A_h)^{-1}\| + |z|^\alpha \|(\chi_1(z) + A_h)^{-1} - (z^\alpha + A_h)^{-1}\| \\
&\leqslant c|z|^2 \tau^{2-\alpha} \|(\chi_1(z) + A_h)^{-1}\| + |z|^\alpha |\chi_1(z) - z^\alpha| \|(\chi_1(z) + A_h)^{-1}(z^\alpha + A_h)^{-1}\| \\
&\leqslant c|z|^2 \tau^{2-\alpha} \|(\chi_1(z) + A_h)^{-1}\|.
\end{aligned}
$$

Now we note that

$$|\chi_1(z)| = |\chi(z)| \tau^{1-\alpha} |\psi(z\tau)| \geqslant c|z| \tau^{1-\alpha}.$$

By Lemma 3.6, $\chi(z) \in \Sigma_{\theta_0}$ for some $\theta_0 \in (\pi/2, \pi)$. Thus by Lemma 3.5 and the resolvent estimate (2.3), we have

$$\|K_1(z) - K_2(z)\| \leqslant c|z|^{-1} |z|^2 \tau^{2-\alpha} |\chi_1(z)|^{-1} + c\tau \leqslant c\tau,$$

and the desired estimate follows immediately. □

Now we can state an error estimate for the discretization error in time for nonsmooth initial data, i.e., $v \in L^2(\Omega)$.

THEOREM 3.1 Let u_h and U_h^n be the solutions of problems (2.5) and (2.8) with $v \in L^2(\Omega)$, $U_h^0 = v_h = P_h v$ and $f \equiv 0$, respectively. Then there holds

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c\tau t_n^{-1} \|v\|_{L^2(\Omega)}.$$

Proof. It suffices to bound the terms *I* and *II*. With the choice $\delta = t_n^{-1}$ and (2.3), we arrive at the following bound for the term *I*

$$
\begin{aligned}
\|I\|_{L^2(\Omega)} &\leqslant c\tau \|v_h\|_{L^2(\Omega)} \left(\int_{1/t_n}^{\pi/(\tau \sin \theta)} e^{rt_n \cos \theta}\, dr + \int_{-\theta}^{\theta} e^{\cos \psi} t_n^{-1}\, d\psi \right) \\
&\leqslant ct_n^{-1} \tau \|v_h\|_{L^2(\Omega)}.
\end{aligned}
\tag{3.9}
$$

By Lemma 3.7 and direct calculation, we bound the term *II* by

$$
\begin{aligned}
\|II\|_{L^2(\Omega)} &\leqslant c \int_{\pi/(\tau \sin \theta)}^{\infty} e^{rt_n \cos \theta} r^{-1}\, dr \|v_h\|_{L^2(\Omega)} \\
&\leqslant c\tau \|v_h\|_{L^2(\Omega)} \int_0^{\infty} e^{rt_n \cos \theta}\, dr \leqslant c\tau t_n^{-1} \|v_h\|_{L^2(\Omega)}.
\end{aligned}
\tag{3.10}
$$

Combining estimates (3.9) and (3.10) yields $\|w_h(t_n) - W_h^n\|_{L^2(\Omega)} \leqslant c\tau t_n^{-1} \|v_h\|_{L^2(\Omega)}$ and the desired result follows from the identity $U_h^n - u_h(t_n) = W_h^n - w_h(t_n)$ and the stability of the projection P_h in $L^2(\Omega)$. □

REMARK 3.3 The $L^2(\Omega)$ stability of the L1 scheme follows directly from Theorem 3.1.

Next we turn to smooth initial data, i.e., $v \in D(A) = H^2(\Omega) \cap H_0^1(\Omega)$. To this end, we first state an alternative estimate on the solution kernels.

LEMMA 3.8 Let θ be close to $\pi/2$, and $\delta < \pi/2\tau$. Further, let $K_1^s(z) = -z^{-1}(z^\alpha + A_h)^{-1}$ and $K_2^s(z) = -\chi(z)^{-1}(\chi_1(z) + A_h)^{-1}$. Then for any $z \in \Sigma_{\delta,\theta}$, there hold

$$\|K_1^s(z) - K_2^s(z)\| \leqslant c|z|^{-\alpha}\tau.$$

Proof. Like before, we set $B_1^s(z) = -(z^\alpha + A_h)^{-1}$ and $B_2^s(z) = -(\chi_1(z) + A_h)^{-1}$. Then by Lemmas 3.4 and 3.5, we have

$$\|B_1^s(z) - B_2^s(z)\| \leqslant |\chi_1(z) - z^\alpha|\|(z^\alpha + A_h)^{-1}\|\|(\chi_1(z) + A_h)^{-1}\|$$
$$\leqslant c|z|^2\tau^{2-\alpha}|z|^{-\alpha}|\chi_1(z)|^{-1} \leqslant c|z|^{1-\alpha}\tau.$$

The rest follows analogously to the derivation (3.3). □

Now we can state an error estimate for $Av \in L^2(\Omega)$.

THEOREM 3.2 Let u_h and U_h^n be the solutions of problems (2.5) and (2.8) with $v \in \dot{H}^2(\Omega)$, $U_h^0 = v_h = R_h v$ and $f \equiv 0$, respectively. Then there holds

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c\tau t_n^{\alpha-1}\|Av\|_{L^2(\Omega)}.$$

Proof. Let $K_1^s(z) = -z^{-1}(z^\alpha + A_h)^{-1}$ and $K_2^s(z) = -\chi(z)^{-1}(\chi_1(z) + A_h)^{-1}$. Then we can rewrite the error as

$$w_h(t_n) - W_h^n = \frac{1}{2\pi i}\int_{\Gamma_{\theta,\delta}\backslash\Gamma_\tau} e^{zt_n}K_1^s(z)A_h v_h dz$$
$$+ \frac{1}{2\pi i}\int_{\Gamma_\tau} e^{zt_n}(K_1^s(z) - e^{-z\tau}K_2^s(z))A_h v_h dz = I + II. \tag{3.11}$$

By Lemma 3.8 we have for $z \in \Gamma_\tau$

$$\|K_1^s(z) - e^{-z\tau}K_2^s(z)\| \leqslant c|z|^{-\alpha}\tau.$$

By setting $\delta = 1/t_n$ and for all $z \in \Gamma_{\delta,\theta}$, we derive the following bound for the term II

$$\|II\|_{L^2(\Omega)} \leqslant c\tau\|A_h v_h\|_{L^2(\Omega)}\left(\int_{1/t_n}^{\pi/(\tau\sin\theta)} e^{rt_n\cos\theta}r^{-\alpha}dr + \int_{-\theta}^{\theta} e^{\cos\psi}t_n^{\alpha-1}d\psi\right)$$
$$\leqslant ct_n^{\alpha-1}\tau\|A_h v_h\|_{L^2(\Omega)}. \tag{3.12}$$

Now (2.3) implies that for all $z \in \Gamma_{\delta,\theta}$

$$\|I\|_{L^2(\Omega)} \leqslant c\|A_h v_h\|_{L^2(\Omega)}\int_{\pi/(\tau\sin\theta)}^{\infty} e^{rt_n\cos\theta}r^{-\alpha-1}dr$$
$$\leqslant c\tau\|A_h v_h\|_{L^2(\Omega)}\int_0^\infty e^{rt_n\cos\theta}r^{-\alpha}dr \leqslant c\tau t_n^{\alpha-1}\|A_h v_h\|_{L^2(\Omega)}. \tag{3.13}$$

Then the desired result follows directly from (3.12), (3.13) and the identities $U_h^n - u_h(t_n) = W^n - w_h(t_n)$ and $A_h R_h = P_h A$. □

REMARK 3.4 The convergence behavior of the L1 scheme is identical with that for the convolution quadrature generated by the backward Euler method, which also converges at an $O(\tau)$ rate, cf. Jin *et al.* (2014b). In particular for smooth initial data $v \in D(A)$, the time discretization error by both schemes contains a singularity $t_n^{\alpha-1}$. This singularity reflects the limited smoothing property of the solution u (Sakamoto & Yamamoto, 2011, Theorem 2.1)

$$\|\partial_t^\alpha u(t)\|_{L^2(\Omega)} \leqslant c\|Av\|_{L^2(\Omega)},$$

whereas the first order derivative $u'(t)$ is unbounded at $t = 0$.

EXAMPLE 3.3 To illustrate the convergence rate in Theorem 3.2, we give a trivial example. Consider the following initial value problem for the fractional ordinary differential equation:

$$\partial_t^\alpha u + u = 0, \quad \forall t > 0, \quad \text{with } u(0) = 1.$$

The exact solution u at $t = \tau$ is given by $u(\tau) = E_{\alpha,1}(-\tau^\alpha)$, where $E_{\alpha,1}(z) = \sum_{k=0}^\infty z^k/\Gamma(\alpha k + 1)$ is the Mittag-Leffler function. For small τ, the L1 scheme at the first step is given by

$$U^1 = (1 + \Gamma(2-\alpha)\tau^\alpha)^{-1} = 1 + \sum_{n=1}^\infty (-1)^n (\Gamma(2-\alpha)\tau^\alpha)^n.$$

Then the difference between U^1 and $u(\tau)$ is given by

$$u(\tau) - U^1 = (\Gamma(2-\alpha) - \Gamma(\alpha+1)^{-1})\tau^\alpha + c_\tau \tau^{2\alpha},$$

with $c_\tau = \sum_{n=2}^\infty (-1)^n (\Gamma(n\alpha+1)^{-1} - \Gamma(2-\alpha)^n)\tau^{(n-2)\alpha}$. Since $|c_\tau| \leqslant c_0$ for small τ, we deduce that

$$|u(\tau) - U^1| \sim \tau^\alpha = t_1^{\alpha-1}\tau.$$

This confirms the convergence order in Theorem 3.2.

Last, the error estimates for the fully discrete scheme (2.8) follow from Theorems 2.1, 3.1 and 3.2 and the triangle inequality.

THEOREM 3.4 Let u and U_h^n be the solutions of problems (1.1) and (2.8) with $U_h^0 = v_h$ and $f \equiv 0$, respectively. Then the following estimates hold.

(a) If $v \in D(A)$ and $v_h = R_h v$, then for $n \geqslant 1$

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c(\tau t_n^{\alpha-1} + h^2)\|Av\|_{L^2(\Omega)}.$$

(b) If $v \in L^2(\Omega)$ and $v_h = P_h v$, then for $n \geqslant 1$

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c(\tau t_n^{-1} + h^2 t_n^{-\alpha})\|v\|_{L^2(\Omega)}.$$

REMARK 3.5 For $v \in D(A)$, we can also choose $v_h = P_h v$ by the stability of the L1 scheme. Hence, by interpolation we deduce

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c(\tau t_n^{-1+\alpha\sigma} + h^2 t_n^{-\alpha(1-\sigma)})\|A^\sigma v\|_{L^2(\Omega)}, \quad 0 \leqslant \sigma \leqslant 1.$$

4. Time-space fractional differential problem

The convergence theory developed in Section 3 may be extended to more general sectorial operators A, i.e., (a) The resolvent set $\rho(A)$ contains the sector $\{z : \theta \leqslant |\arg z| \leqslant \pi\}$ for some $\theta \in (0, \pi/4)$; (b) $\|(zI+A)^{-1}\| \leqslant M/|z|$ for $z \in \Sigma_{\pi-\theta}$ and some constant M. The technical restriction $\theta \in (0, \pi/4)$ stems from Remark 3.2. This in particular covers the Riemann-Liouville fractional derivative of order $\beta \in (3/2, 2)$; see Lemma 4.1 below. Specifically, we consider the following one-dimensional space-time fractional differential equation

$$
\begin{aligned}
\partial_t^\alpha u - {}_0^R\!D_x^\beta u &= f, \quad \text{in } \Omega = (0,1) \quad T \geqslant t > 0, \\
u &= 0, \quad \text{on } \partial\Omega \qquad\quad T \geqslant t > 0, \\
u(0) &= v, \quad \text{in } \Omega,
\end{aligned}
\tag{4.1}
$$

with $\alpha \in (0,1)$ and $\beta \in (3/2, 2)$. Here ${}_0^R\!D_x^\beta$ with $n-1 < \beta < n$, $n \in \mathbb{N}$, denotes the left-sided Riemann-Liouville fractional derivative ${}_0^R\!D_x^\beta u$ of order β defined by (Kilbas $et\ al.$, 2006, pp. 70):

$$
{}_0^R\!D_x^\beta u = \frac{1}{\Gamma(n-\beta)} \frac{d^n}{dx^n} \int_0^x (x-s)^{n-\beta-1} u(s) ds.
\tag{4.2}
$$

The right-sided version of Riemann-Liouville fractional derivative is defined analogously

$$
{}_x^R\!D_1^\beta u = \frac{(-1)^n}{\Gamma(n-\beta)} \frac{d^n}{dx^n} \int_x^1 (s-x)^{n-\beta-1} u(s) ds.
$$

The model (4.1) is often adopted to describe anomalous diffusion process involving both long range interactions and history mechanism.

The variational formulation of (4.1) is to find $u \in \widetilde{H}^{\beta/2}(\Omega) \equiv H_0^{\beta/2}(\Omega)$ such that (see Jin $et\ al.$ (2013b))

$$
(\partial_t^\alpha u, \varphi) + A(u, \varphi) = (f, \varphi) \quad \forall \varphi \in \widetilde{H}^{\beta/2}(\Omega),
\tag{4.3}
$$

with $u(0) = v$, where the sesquilinear form $A(\cdot, \cdot)$ is given by

$$
A(\varphi, \psi) = - \left({}_0^R\!D_x^{\beta/2}\varphi, \, {}_x^R\!D_1^{\beta/2}\psi \right).
$$

It is known (see (Ervin & Roop, 2006, Lemma 3.1) and (Jin $et\ al.$, 2013b, Lemma 4.2)) that the sesquilinear form $A(\cdot, \cdot)$ is coercive and bounded on the space $\widetilde{H}^{\beta/2}(\Omega)$. Then Riesz representation theorem implies that there exists a unique bounded linear operator $\widetilde{A} : \widetilde{H}^{\beta/2}(\Omega) \to H^{-\beta/2}(\Omega)$ such that

$$
A(\varphi, \psi) = \langle \widetilde{A}\varphi, \psi \rangle, \quad \forall \varphi, \psi \in \widetilde{H}^{\beta/2}(\Omega).
$$

Define $D(A) = \{\psi \in \widetilde{H}^{\beta/2}(\Omega) : \widetilde{A}\psi \in L^2(\Omega)\}$ and an operator $A : D(A) \to L^2(\Omega)$ by

$$
A(\varphi, \psi) = (A\varphi, \psi), \ \varphi \in D(A), \ \psi \in \widetilde{H}^{\beta/2}(\Omega).
\tag{4.4}
$$

We recall that the domain $D(A)$ consists of functions of the form ${}_0I_x^\beta f - ({}_0I_x^\beta f)(1) x^{\beta-1}$, where $f \in L^2(\Omega)$, cf. Jin $et\ al.$ (2013b). The term $x^{\beta-1} \in \widetilde{H}_L^{\beta-1+\delta}(\Omega)$, $\delta \in [1-\beta/2, 1/2)$, appears because it is in the kernel of the operator ${}_0^R\!D_x^\beta$. The presence of the term $x^{\beta-1}$ indicates that the solution u usually can only have limited regularity.

LEMMA 4.1 For $\beta \in (3/2, 2)$, the resolvent set $\rho(A)$ contains the sector $\{z : \theta \leqslant |\arg(z)| \leqslant \pi\}$ for any $\theta \in ((2 - \beta)\pi/2, \pi/4)$.

Proof. Let $\widetilde{u}$ be the zero extension of u. Recall that for $s \in (1/2, 1)$ and $\Omega = (0, 1)$,

$$\|u\|_{\widetilde{H}^s(\Omega)} = \||\xi|^s \mathscr{F}(\widetilde{u})(\xi)\|_{L^2(\mathbb{R})},$$

where $\mathscr{F}(\widetilde{u})$ is the Fourier transform of $\widetilde{u}$, is a consistent and well-defined norm on $\widetilde{H}^s(\Omega)$. Further, we note that for $u \in \widetilde{H}^{\beta/2}(\Omega)$ (see Ervin & Roop (2006) and Jin *et al.* (2013b))

$$\Re(A(u, u)) \geqslant c_0 \|u\|^2_{\widetilde{H}^{\beta/2}(\Omega)} \quad \text{with } c_0 = \cos((2 - \beta/2)\pi).$$

Further, we recall the fact that for $u \in C_0^\infty(\Omega)$ and $s \in (1/2, 1)$

$$\begin{aligned}
\|{}_0^R D_x^s u\|_{L^2(\Omega)} = \|{}_{-\infty}^R D_x^s \widetilde{u}\|_{L^2(\Omega)} &\leqslant \|{}_{-\infty}^R D_x^s \widetilde{u}\|_{L^2(\mathbb{R})} \\
&= \|\mathscr{F}({}_{-\infty}^R D_x^s \widetilde{u})\|_{L^2(\mathbb{R})} = \|\mathscr{F}({}_{-\infty}^R D_x^s \widetilde{u})\|_{L^2(\mathbb{R})} \\
&= \||\xi|^s \mathscr{F}(\widetilde{u})(\xi)\|_{L^2(\mathbb{R})} = \|u\|_{\widetilde{H}^s(\Omega)}.
\end{aligned}$$

By the density of $C_0^\infty(\Omega)$ in $\widetilde{H}^s(\Omega)$ for $s \in (1/2, 1)$ we obtain for all $u \in \widetilde{H}^{\beta/2}(\Omega)$

$$\|{}_0^R D_x^{\beta/2} u\|_{L^2(\Omega)} \leqslant \|u\|_{\widetilde{H}^{\beta/2}(\Omega)}.$$

Likewise, the right sided case follows:

$$\|{}_x^R D_1^{\beta/2} u\|_{L^2(\Omega)} \leqslant \|u\|_{\widetilde{H}^{\beta/2}(\Omega)}.$$

Thus for $u, v \in \widetilde{H}^{\beta/2}(\Omega)$, there holds

$$|A(u, v)| \leqslant c_1 \|u\|_{\widetilde{H}^{\beta/2}(\Omega)} \|v\|_{\widetilde{H}^{\beta/2}(\Omega)} \quad \text{with } c_1 = 1.$$

Then by Lemma 2.1 of Jin *et al.* (2014a), we conclude that the resolvent set $\rho(A)$ contains the sector $\{z : \theta \leqslant |\arg z| \leqslant \pi\}$ for all $\theta \in ((2 - \beta/2)\pi, \pi/2)$. In particular, for $\beta > 3/2$, we may choose $\theta \in ((2 - \beta/2)\pi, \pi/4)$. $\qquad\square$

By Lemma 4.1 and Remark 3.2, we can apply the theory in Sections 2 and 3 to derive a fully discrete scheme based on the L1 scheme in time and the Galerkin finite element approximation in space. First we partition the unit interval Ω into a uniform mesh with a mesh size $h = 1/M$. We then define V_h to be the set of continuous functions in V which are linear when restricted to the subintervals $[x_i, x_{i+1}]$, $i = 0, \ldots, M - 1$. Further, we define the discrete operator $A_h : V_h \to V_h$ by

$$(A_h \varphi, \chi) = A(\varphi, \chi) \quad \forall \varphi, \chi \in V_h.$$

The corresponding Ritz projection $R_h : \widetilde{H}^{\beta/2}(\Omega) \to V_h$ is defined by

$$A(R_h \psi, \chi) = A(\psi, \chi) \quad \forall \psi \in \widetilde{H}^{\beta/2}(\Omega), \ \chi \in V_h. \tag{4.5}$$

Then the fully discrete scheme for problem (4.1) based on the L1 approximation (1.3) reads: find U_h^n for $n = 1, 2, \ldots, N$

$$(I + \tau^\alpha A_h) U_h^n = b_n U_h^0 + \sum_{j=1}^n (b_{j-1} - b_j) U_h^{n-j} + \tau^\alpha F_h^n, \tag{4.6}$$

with $U_h^0 = v_h$ and $F_h^n = P_h f(t_n)$.

Last, we state the error estimates for the fully discrete scheme (4.6). These estimates follow from Theorems 3.1 and 3.2 and the error estimates for the semidiscrete Galerkin scheme, which can be proved using the operator trick in Jin *et al.* (2014a), and thus the proof is omitted.

THEOREM 4.1 Let u and U_h^n be the solutions of problems (4.1) and (4.6) with $U_h^0 = v_h$ and $f \equiv 0$, respectively. Then for $\delta \in [1 - \beta/2, 1/2)$ the following estimates hold.

(a) If $v \in D(A)$ and $v_h = R_h v$, then for $n \geqslant 1$

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c(\tau t_n^{\alpha-1} + h^{\beta-1+\delta}) \|Av\|_{L^2(\Omega)}.$$

(b) If $v \in L^2(\Omega)$ and $v_h = P_h v$, then for $n \geqslant 1$

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant c(\tau t_n^{-1} + h^{\beta-1+\delta} t_n^{-\alpha}) \|v\|_{L^2(\Omega)}.$$

5. Numerical experiments and discussions

Now we present numerical results to verify the convergence theory in Sections 3 and 4, i.e., the $O(\tau)$ convergence rate. We consider the subdiffusion case (1.1) and time-space fractional case (4.1) separately. For each model, we consider the following two initial data:

(a) $\Omega = (0,1)$, and $v = \sin(2\pi x) \in H^2(\Omega) \cap H_0^1(\Omega)$.

(b) $\Omega = (0,1)$, and $v = x^{-1/4} \in H^{1/4-\varepsilon}(\Omega)$, with $\varepsilon \in (0, 1/4)$;

We measure the error $e^n = u(t_n) - U_h^n$ by the normalized errors $\|e^n\|_{L^2(\Omega)}/\|v\|_{L^2(\Omega)}$. In the computations, we divide the unit interval $\Omega = (0,1)$ into M equally spaced subintervals with a mesh size $h = 1/M$. Likewise, we fix the time step size τ at $\tau = t/N$, where t is the time of interest. In this work, we only examine the temporal convergence rate, since the space convergence rate has been examined earlier in Jin *et al.* (2013a, 2014a). To this end, we take a small mesh size $h = 2^{-13}$, so that the spatial discretization error is negligible.

5.1 *Subdiffusion.*

The exact solution can be written explicitly as an infinite series involving the Mittag-Leffler function $E_{\alpha,\beta}(z)$ defined by

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)};$$

see Jin *et al.* (2013a) for the details. The Mittag-Leffler function $E_{\alpha,\beta}(z)$ can be evaluated efficiently by the algorithm developed in Seybold & Hilfer (2009). The numerical results for cases (a) and (b) are shown in Table 2, where `rate` refers to the empirical convergence rate, and the number in the bracket refers to the theoretical rate from Theorem 3.4. The results fully confirm the theoretical prediction: for both smooth and nonsmooth data, the fully discrete solution U_h^n converges at a rate $O(\tau)$, and the rate is independent of the fractional order α. Further, for fixed t, the error increases with the fractional order α. This might be attributed to the local decay behavior of the solution u: the larger is the fractional order α, the slower is the solution decay around $t = 0$. According to Remark 3.5, we have

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leqslant C(N^{-1} t_n^{\alpha\sigma} + h^2 t_n^{-\alpha(1-\sigma)}) \|A^\sigma v\|_{L^2(\Omega)}, \quad \sigma \in [0,1].$$

Thus the temporal error deteriorates as the time $t_n \to 0$ at a rate like t_n^α and $t_n^{\alpha/8-\alpha\varepsilon}$ for $Av \in L^2(\Omega)$ and $v \in D(A^{1/8-\varepsilon})$, with $\varepsilon \in (0, 1/8)$, respectively. In particular, for fixed N, the error behaves like $t_n^{1/2}$ and $t_n^{1/16}$ for cases (a) and (b) with $\alpha = 0.5$, respectively. This is clearly observed in Table 3, thereby showing the sharpness of the error estimates. Further, we observe that the L1 scheme fails to converge uniformly (in time) at a first order even though the initial data v in case (a) is very smooth, i.e., $v \in D(A^p)$ for any $p \geqslant 0$, cf. Table 3. This numerically confirms the observation in Remark 3.4.

Table 2. The L^2-norm of the error for the subdiffusion equation with initial data (a) and (b) at $t = 0.1$ with $h = 2^{-13}$, $\tau = 1/N$.

α	N	10	20	40	80	160	320	rate
0.1	(a)	1.46e-4	7.18e-5	3.55e-5	1.77e-5	8.82e-6	4.40e-6	≈ 1.01 (1.00)
	(b)	3.95e-4	1.93e-4	9.57e-5	4.76e-5	2.38e-5	1.19e-5	≈ 1.00 (1.00)
0.5	(a)	1.22e-3	5.89e-4	2.88e-4	1.43e-4	7.08e-5	3.52e-5	≈ 1.01 (1.00)
	(b)	3.65e-3	1.73e-3	8.36e-4	4.09e-4	2.02e-4	1.00e-4	≈ 1.02 (1.00)
0.9	(a)	7.01e-3	3.05e-3	1.39e-3	6.53e-4	3.12e-4	1.50e-4	≈ 1.07 (1.00)
	(b)	1.54e-2	7.67e-3	3.79e-3	1.87e-3	9.23e-4	4.55e-4	≈ 1.02 (1.00)

Table 3. The L^2-norm of the error for the subdiffusion equation with initial data (a) and (b) with $\alpha = 0.5$, $h = 2^{-13}$ and $N = 10$.

t	1e-5	1e-6	1e-7	1e-8	1e-9	1e-10	rate
(a)	2.94e-3	1.05e-3	3.45e-4	1.11e-4	3.51e-5	1.11e-5	≈ 0.50 (0.50)
(b)	3.02e-3	2.56e-3	2.18e-3	1.86e-3	1.58e-3	1.35e-3	≈ 0.07 (0.06)

5.2 *Time-space fractional problem.*

Now we present numerical results for the space time fractional problem. Since the exact solution is not available in closed form, we compute the reference solution by the second-order backward difference scheme from Jin *et al.* (2014b) on a much finer mesh, i.e., $N = 1000$. The numerical results for case (a) with different α and β values are presented in Table 4. A first-order convergence is observed, and the convergence rate is independent of the time- and space-fractional orders. Interestingly, the observation is valid also for the case $\beta = 5/4$, for which the theory in Section 4 does not cover, awaiting further study. For a fixed α value, the error decreases with the increase of the fractional order β, which indicates that the solution decays faster as β approaches two. This is also consistent with the fact that the closer is the β value to unity, the more singular is the solution, and thus more challenging to approximate numerically, cf. Jin *et al.* (2013b). For the nonsmooth case (b), we are particularly interested in the case of small t. Thus we present the numerical results for $t = 0.1$, $t = 0.01$ and $t = 0.001$ in Table 5. The experiment shows the first order convergence is robust for nonsmooth data even if t is close to zero. Like before, for fixed n, let $t_n \to 0$, the error behaves like t_n^α for case (a), which is fully confirmed by the numerical results, cf. Table 6, indicating the sharpness of the estimate in Theorem 4.1.

6. Conclusions

In this paper we have revisited the popular L1 time-stepping scheme for discretizing the Caputo fractional derivative of order $\alpha \in (0, 1)$, arising in the modeling of subdiffusion, and rigorously established the first

Table 4. The L^2-norm of the error for the time-space fractional problem with initial data (a) at $t = 0.1$ with $h = 2^{-13}$, $\tau \to 0$.

α	$\beta \backslash N$	5	10	20	40	80	rate
	1.25	1.37e-3	6.55e-4	3.21e-4	1.59e-4	7.90e-5	$\approx$ 1.01 (1.00)
0.1	1.5	8.41e-4	4.03e-4	1.98e-4	9.78e-5	4.87e-5	$\approx$ 1.01 (1.00)
	1.75	5.08e-4	2.44e-4	1.19e-4	5.92e-5	2.94e-5	$\approx$ 1.01 (1.00)
	1.25	1.52e-2	6.69e-3	3.12e-3	1.50e-3	7.32e-4	$\approx$ 1.03 (1.00)
0.5	1.5	8.22e-3	3.70e-3	1.75e-3	8.49e-4	4.17e-4	$\approx$ 1.05 (1.00)
	1.75	4.69e-3	2.14e-3	1.02e-3	4.97e-4	2.45e-4	$\approx$ 1.03 (1.00)
	1.25	6.01e-2	3.19e-2	1.62e-2	8.07e-3	3.99e-4	$\approx$ 1.01 (1.00)
0.9	1.5	5.61e-2	2.86e-2	1.42e-2	6.95e-3	3.39e-3	$\approx$ 1.03 (1.00)
	1.75	3.58e-2	1.66e-2	7.74e-3	3.66e-3	1.75e-3	$\approx$ 1.07 (1.00)

Table 5. The L^2-norm of the error for the time-space fractional problem with initial data (b) with $\beta = 1.5$, $h = 2^{-13}$, $t = 0.1$, 0.01 and 0.001.

α	$t \backslash N$	5	10	20	40	80	rate
	0.1	1.53e-3	7.32e-4	3.59e-4	1.77e-4	8.83e-5	$\approx$ 1.01 (1.00)
0.1	0.01	1.74e-3	8.34e-4	4.08e-4	2.02e-4	1.01e-4	$\approx$ 1.01 (1.00)
	0.001	1.94e-3	9.31e-4	4.56e-4	2.25e-4	1.12e-4	$\approx$ 1.01 (1.00)
	0.1	1.39e-2	6.36e-3	3.01e-3	1.45e-3	7.11e-4	$\approx$ 1.04 (1.00)
0.5	0.01	1.22e-2	5.86e-3	2.85e-3	1.40e-3	6.89e-4	$\approx$ 1.03 (1.00)
	0.001	8.02e-3	3.83e-3	1.86e-3	9.12e-4	4.50e-4	$\approx$ 1.02 (1.00)
	0.1	1.99e-2	1.02e-2	5.15e-3	2.60e-3	1.31e-3	$\approx$ 1.00 (1.00)
0.9	0.01	1.21e-2	6.10e-3	3.05e-3	1.52e-3	7.53e-4	$\approx$ 1.00 (1.00)
	0.001	7.63e-3	3.83e-3	1.91e-3	9.51e-4	4.73e-4	$\approx$ 1.00 (1.00)

Table 6. The L^2-norm of the error for the time-space fractional problem with initial data (a) and (b), $\alpha = 0.5$ and $\beta = 1.5$, as $t \to 0$ with $h = 2^{-13}$ and $N = 5$.

t	1e-5	1e-6	1e-7	1e-8	1e-9	1e-10	rate
(a)	2.60e-3	8.90e-4	2.96e-4	9.71e-5	3.17e-5	1.03e-5	$\approx$ 0.48 (0.50)
(b)	4.74e-3	3.76e-3	3.02e-3	2.44e-3	1.99e-3	1.63e-3	$\approx$ 0.09 (—)

order convergence for both smooth and nonsmooth initial data. This result complements existing convergence analysis, which assumes a C^2 regularity in time. The extensive numerical experiments fully verify the sharpness of the estimates and robustness of the scheme. The convergence analysis is valid for more general sectorial operators, and in particular, it covers also the space-time fractional differential equations, for which we have also provided error estimates.

In view of the solution singularity for time t close to zero, it is natural to consider the L1 scheme on a nonuniform time mesh in order to arrive at a uniform first-order convergence. However, the generating function approach used in our analysis does not work directly in this case, and it is an interesting open question to rigorously establish error estimates directly in terms of the data regularity. A second interesting future research problem is the convergence rate analysis of the scheme for the diffusion wave equation,

which involves a Caputo fractional derivative of order $\alpha \in (1,2)$ in time.

Acknowledgment

References

ABRAMOWITZ, M. & STEGUN, I. A. (1964) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

ADAMS, E. E. & GELHAR, L. W. (1992) Field study of dispersion in a heterogeneous aquifer: 2. spatial moments analysis. *Water Res. Research*, **28**, 3293–3307.

BAZHLEKOVA, E., JIN, B., LAZAROV, R. & ZHOU, Z. (2014) An analysis of the Rayleigh-Stokes problem for the generalized second grade fluid. *Numer. Math.*, in press. arXiv:1401.8049.

BOUCHAUD, J.-P. & GEORGES, A. (1990) Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications. *Phys. Rep.*, **195**, 127–293.

ERVIN, V. J. & ROOP, J. P. (2006) Variational formulation for the stationary fractional advection dispersion equation. *Numer. Methods Partial Diff. Eq.*, **22**, 558–576.

FLAJOLET, P. (1999) Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.*, **215**, 371–381.

HATANO, Y. & HATANO, N. (1998) Dispersive transport of ions in column experiments: An explanation of long-tailed profiles. *Water Res. Research*, **34**, 1027–1033.

JIN, B., LAZAROV, R. & ZHOU, Z. (2013a) Error estimates for a semidiscrete finite element method for fractional order parabolic equations. *SIAM J. Numer. Anal.*, **51**, 445–466.

JIN, B., LAZAROV, R., PASCIAK, J. & RUNDELL, W. (2013b) Variational formulation of problems involving fractional order differential operators. *Math. Comp.*, in press. Available as arXiv:1307.4795.

JIN, B., LAZAROV, R., PASCIAK, J. & ZHOU, Z. (2014a) Error analysis of a finite element method for a space-fractional parabolic equation. *SIAM J. Numer. Anal.*, **52**, 2272–2294.

JIN, B., LAZAROV, R. & ZHOU, Z. (2014b) On two schemes for fractional diffusion and diffusion-wave equations. preprint, arXiv:1404.3800.

JIN, B. & RUNDELL, W. (2012) An inverse problem for a one-dimensional time-fractional diffusion problem. *Inverse Problems*, **28**, 075010, 19.

KILBAS, A., SRIVASTAVA, H. & TRUJILLO, J. (2006) *Theory and Applications of Fractional Differential Equations*. Amsterdam: Elsevier.

KNOPP, M. & ROBINS, S. (2001) Easy proofs of Riemann's functional equation for $\zeta(s)$ and of Lipschitz summation. *Proc. Amer. Math. Soc.*, **129**, 1915–1922.

LEWIN, L. (1981) *Polylogarithms and Associated Functions*. New York, Amsterdam: North-Holland.

LIN, Y. & XU, C. (2007) Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.*, **225**, 1533–1552.

LUBICH, C., SLOAN, I. H. & THOMÉE, V. (1996) Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term. *Math. Comp.*, **65**, 1–17.

MCLEAN, W. & MUSTAPHA, K. (2014) Time-stepping error bounds for fractional diffusion problems with non-smooth initial data. *J. Comput. Phys.*, in press. arXiv:1405.2140.

NIGMATULIN, R. (1986) The realization of the generalized transfer equation in a medium with fractal geometry. *Phys. Stat. Sol. B*, **133**, 425–430.

SAKAMOTO, K. & YAMAMOTO, M. (2011) Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *J. Math. Anal. Appl.*, **382**, 426–447.

SEYBOLD, H. & HILFER, R. (2009) Numerical algorithm for calculating the generalized Mittag-Leffler function. *SIAM J. Numer. Anal.*, **47**, 69–88.

SUN, Z.-Z. & WU, X. (2006) A fully discrete scheme for a diffusion wave system. *Appl. Numer. Math.*, **56**, 193–209.

THOMÉE, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*. Berlin: Springer.