Math. Comp. (accepted in November 2018)

SUBDIFFUSION WITH A TIME-DEPENDENT COEFFICIENT: ANALYSIS AND NUMERICAL SOLUTION

BANGTI JIN, BUYANG LI, AND ZHI ZHOU

Abstract. In this work, a complete error analysis is presented for fully discrete solutions of the subdiffusion equation with a time-dependent diffusion coefficient, obtained by the Galerkin finite element method with conforming piecewise linear finite elements in space and backward Euler convolution quadrature in time. The regularity of the solutions of the subdiffusion model is proved for both nonsmooth initial data and incompatible source term. Optimal-order convergence of the numerical solutions is established using the proven solution regularity and a novel perturbation argument via freezing the diffusion coefficient at a fixed time. The analysis is supported by numerical experiments.

1. Introduction

Let $\Omega \subset \mathbb{R}^d$ $(d \geq 1)$ be a convex polyhedral domain with a boundary $\partial\Omega$. Consider the following fractional-order parabolic problem for the function $u(x,t)$:

$$\begin{cases} \partial_t^\alpha u(x,t) - \nabla \cdot (a(x,t)\nabla u(x,t)) = f(x,t) & (x,t) \in \Omega \times (0,T], \\ u(x,t) = 0 & (x,t) \in \partial\Omega \times (0,T], \\ u(x,0) = u_0(x) & x \in \Omega, \end{cases} \tag{1.1}$$

where $T > 0$ is a fixed final time, $f \in L^\infty(0,T;L^2(\Omega))$ and $u_0 \in L^2(\Omega)$ are given source term and initial data, respectively, and $a(x,t) \in \mathbb{R}^{d\times d}$ is a symmetric matrix-valued diffusion coefficient such that for some constant $\lambda \geq 1$

$$\lambda^{-1}|\xi|^2 \leq a(x,t)\xi \cdot \xi \leq \lambda|\xi|^2, \qquad \forall\, \xi \in \mathbb{R}^d,\ \forall\,(x,t) \in \Omega \times (0,T], \tag{1.2}$$

$$|\partial_t a(x,t)| + |\nabla_x a(x,t)| + |\nabla_x \partial_t a(x,t)| \leq c, \qquad \forall\,(x,t) \in \Omega \times (0,T]. \tag{1.3}$$

The notation $\partial_t^\alpha u(t)$ denotes the Caputo derivative in time of order $\alpha \in (0,1)$, defined by [16, p. 70]

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

The literature on the numerical analysis of the subdiffusion problem is vast; see [21, 11, 9, 15] for a rather incomplete list and the overview [10] (and the references therein). The work [11] analyzed two spatially semidiscrete schemes, i.e., Galerkin finite element method (FEM) and lumped mass method, and derived nearly optimal order error estimates for the homogeneous problem. The inhomogeneous case was analyzed in [9]. See [15] for a finite volume element discretization, and [14] for a unified approach. There are a number of fully discrete schemes, e.g., convolution quadrature [36, 12], piecewise polynomial interpolation [33, 21, 2, 7, 35, 32], discontinuous Galerkin method [26, 27]; and some of them have an $O(\tau)$ rate for nonsmooth data, with τ being time step size. However, all these works analyzed only the case that the diffusion coefficient a is independent of the time t. These works mostly employ Laplace transform and its discrete analogue for analysis, which are not directly applicable to the case of a time-dependent coefficient. Recently, Mustapha [28] analyzed the spatially semidiscrete Galerkin FEM for (1.1) using a novel energy argument, and proved optimal-order convergence rates for both smooth and nonsmooth initial data (with a zero source term) based on certain assumptions on the regularity of the PDE's solution.

2010 *Mathematics Subject Classification.* Primary: 65M30, 65M15, 65M12.
Key words and phrases. subdiffusion, time-dependent coefficient, Galerkin finite element method, convolution quadrature, perturbation argument, error estimate.

In this article, using a novel perturbation argument, we present a new approach to analyze a fully discrete scheme for problem (1.1) based on the Galerkin FEM in space and backward Euler (BE) convolution quadrature in time, covering initial data and source term simultaneously. The main contributions of this paper are as follows. First, we give a complete existence, uniqueness and regularity theory for problem (1.1) in Theorems 2.1–2.3, which are crucial to the error analysis. Second, we derive sharp error estimates for the spatially semidiscrete Galerkin FEM. This is achieved by combining error estimates for a time-independent coefficient and a perturbation argument in time. Third, we derive nearly sharp error estimates for the fully discrete method. All error estimates are given directly in terms of the regularity of the initial data and source term, under mild regularity assumptions on the diffusion coefficient $a(x,t)$ that are weaker than the assumptions in [28]; see Remark 2.2 for the precise statement.

There are a few relevant works on standard parabolic problems with a time-dependent coefficient [24, 30, 31, 19]. For example, Luskin and Rannacher [24] proved optimal order error estimates for both spatially semidiscrete and fully discrete problems (by BE method) using a novel energy argument, and Sammon [30] analyzed fully discrete schemes with linear multistep methods. Our error analysis relies crucially on a perturbation argument, using basic estimates given in Lemmas 3.1 and 3.2, which are of independent interest. Generally, the idea of freezing coefficients and perturbation in time has been proved very useful in combination with energy estimates [31] and L^p estimates [1, 18, 20]. In this work, we have successfully adapted the idea to the subdiffuion model.

The rest of the paper is organized as follows. In Section 2, we discuss temporal and spatial regularity of the solution for nonsmooth problem data. Then in Section 3, we prove optimal-order convergence of the spatially semidiscrete Galerkin FEM for both homogeneous and inhomogeneous problems. In Section 4, we present the error analysis for the fully discrete FEM and prove first-order convergence in time. Last, in Section 5, we present numerical examples to support the theoretical analysis. Throughout, the notation c, with or without a subscript, denotes a generic positive constant, which may differ at each occurrence, but is always independent of the mesh size h and step size τ.

2. Regularity theory

In this section we investigate the regularity of the solutions of problem (1.1). For any function $f(x,t)$ defined on $\Omega\times(0,T)$, we denote by $f(t)$ the function $f(\cdot,t)$. Let $-\Delta : H_0^1(\Omega)\cap H^2(\Omega)\to L^2(\Omega)$ be the negative Laplacian operator with a zero Dirichlet boundary condition, and $\{(\lambda_j,\varphi_j)\}$ be its eigenvalues ordered nondecreasingly (with multiplicity counted) and the corresponding eigenfunctions normalized in the $L^2(\Omega)$ norm. For any $r\geq 0$, we denote the space $\dot H^r(\Omega)=\{v\in L^2(\Omega):(-\Delta)^{\frac{r}{2}}v\in L^2(\Omega)\}$, with the norm [34, Chapter 3]

$$\|v\|^2_{\dot H^r(\Omega)}=\sum_{j=1}^{\infty}\lambda_j^r(v,\varphi_j)^2.$$

Then we have $\dot H^0(\Omega)=L^2(\Omega)$, $\dot H^1(\Omega)=H_0^1(\Omega)$, and $\dot H^2(\Omega)=H^2(\Omega)\cap H_0^1(\Omega)$.

2.1. Subdiffusion with a time-independent coefficient.

First we recall basic properties of the subdiffusion model with a time-independent diffusion coefficient, i.e., $a(x,t)=a(x)$. Accordingly, we denote by $A: H_0^1(\Omega)\cap H^2(\Omega)\to L^2(\Omega)$ an elliptic operator, defined by

$$A\phi(x) := -\nabla\cdot(a(x)\nabla\phi(x)),$$

and consider the problem

$$\partial_t^\alpha u(t)+Au(t)=f(t)\quad t\in(0,T],\quad \text{with } u(0)=u_0. \tag{2.1}$$

This problem has been studied in [3, 4, 23, 25, 29, 13]. The following maximal L^p-regularity holds [3].

Lemma 2.1. *If $u_0=0$ and $f\in L^p(0,T;L^2(\Omega))$ with $1<p<\infty$, then problem (2.1) has a unique solution $u\in L^p(0,T;\dot H^2(\Omega))$ such that $\partial_t^\alpha u\in L^p(0,T;L^2(\Omega))$ and*

$$\|u\|_{L^p(0,T;\dot H^2(\Omega))}+\|\partial_t^\alpha u\|_{L^p(0,T;L^2(\Omega))}\leq c\|f\|_{L^p(0,T;L^2(\Omega))},$$

where the constant c does not depend on f and T.

By means of Laplace transform, the solution $u(t)$ can be represented by [13, Section 4]

$$u(t) = F(t)u_0 + \int_0^t E(t-s)f(s)\mathrm{d}s, \tag{2.2}$$

where the solution operators $F(t)$ and $E(t)$ are defined by

$$F(t) := \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}z^{\alpha-1}(z^\alpha + A)^{-1}\,\mathrm{d}z, \tag{2.3}$$

$$E(t) := \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}(z^\alpha + A)^{-1}\,\mathrm{d}z, \tag{2.4}$$

with integration over a contour $\Gamma_{\theta,\delta} \subset \mathbb{C}$ (oriented with an increasing imaginary part):

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

Throughout, we fix $\theta \in (\frac{\pi}{2}, \pi)$ so that $z^\alpha \in \Sigma_{\alpha\theta} \subset \Sigma_\theta := \{0 \ne z \in \mathbb{C} : \arg(z) \le \theta\}$, for all $z \in \Sigma_\theta$. The next lemma gives smoothing properties of $F(t)$ and $E(t)$, which follow from the resolvent estimate

$$\|(z + A)^{-1}\| \le c_\phi |z|^{-1}, \quad \forall z \in \Sigma_\phi,\ \forall\,\phi \in (0, \pi), \tag{2.5}$$

where $\|\cdot\|$ denotes the operator norm from $L^2(\Omega)$ to $L^2(\Omega)$.

Lemma 2.2. *The operators F and E defined in* (2.3) *and* (2.4) *satisfy the following properties.*

(i) $t^{-\alpha}\|A^{-1}(F(t) - I)\| + \|F(t) - I\| \le c, \quad \forall\, t \in (0, T]$;
(ii) $t^{1-\alpha}\|E(t)\| + t^{2-\alpha}\|E'(t)\| + t\|AE(t)\| \le c, \quad \forall\, t \in (0, T]$;
(iii) $t^\alpha\|AF(t)\| + t^{1-\beta\alpha}\|A^{-\beta}F'(t)\| \le c, \quad \forall\, t \in (0, T], \beta \in [0, 1]$.

Proof. Parts (i) and (ii) have been shown in [13, Lemma 3.4]. By letting $\delta = t^{-1}$ in $\Gamma_{\theta,\delta}$ and $z = s\cos\varphi + \mathrm{i}s\sin\varphi$, using (2.5), we have (with $|\mathrm{d}z|$ being the arc length of $\Gamma_{\theta,\delta}$)

$$\begin{aligned}\|AF(t)\| &= \left\|\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}z^{\alpha-1}A(z^\alpha + A)^{-1}\,\mathrm{d}z\right\| \le c\int_{\Gamma_{\theta,\delta}} e^{\Re(z)t}|z|^{\alpha-1}\,|\mathrm{d}z| \\ &\le c\int_\delta^\infty e^{st\cos\theta}s^{\alpha-1}\mathrm{d}s + c\int_{-\theta}^\theta e^{\cos\varphi}\delta^\alpha\mathrm{d}\varphi \le ct^{-\alpha}.\end{aligned}$$

Next, direct computation gives $F'(t) = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}z^\alpha(z^\alpha + A)\,\mathrm{d}z$, and thus by the estimate (2.5),

$$\|F'(t)\| \le c\int_{\Gamma_{\theta,\delta}} e^{\Re(z)t}|z|^\alpha|z|^{-\alpha}\,|\mathrm{d}z| \le ct^{-1},$$

which shows the assertion for $\beta = 0$. Meanwhile, by the identity $z^\alpha(z^\alpha + A)^{-1} = I - A(z^\alpha + A)^{-1}$, we have $F'(t) = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}z^\alpha(z^\alpha + A)\,\mathrm{d}z = -\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}A(z^\alpha + A)\,\mathrm{d}z$, and thus

$$\|A^{-1}F'(t)\| \le c\int_{\Gamma_{\theta,\delta}} e^{\Re(z)t}|z|^{-\alpha}\,|\mathrm{d}z| \le ct^{-1+\alpha}.$$

This shows the assertion for $\beta = 1$. Then the desired bound on $t^{1-\beta\alpha}\|A^{-\beta}F'(t)\|$ in part (iii) follows by interpolation. □

2.2. **Regularity theory for subdiffusion with a time-dependent coefficient.** Now we develop the regularity theory for the case of a time-dependent diffusion coefficient. The work [37] gave some interior Hölder estimates for bounded measurable coefficients. Recently, Kubica et al [17] showed the unique existence by approximating the coefficients by smooth functions, and derived several regularity estimates. We shall provide a different approach to derive regularity estimates in Sobolev spaces, which are essential for the error analysis in Sections 3 and 4.

We define a time-dependent elliptic operator $A(t) : \dot{H}^2(\Omega) \to L^2(\Omega)$ by

$$A(t)\phi = -\nabla \cdot (a(x, t)\nabla\phi), \quad \forall \phi \in \dot{H}^2(\Omega).$$

Under condition (1.3), the following estimate holds:

$$\|(A(t) - A(s))v\|_{L^2(\Omega)} \le c|t - s|\|v\|_{H^2(\Omega)}. \tag{2.6}$$

First we give the existence, uniqueness and regularity of solutions to problem (1.1) with $u_0 = 0$.

Theorem 2.1. *Under conditions* (1.2)-(1.3)*, with* $u_0 = 0$ *and* $f \in L^p(0,T;L^2(\Omega))$*,* $1/\alpha < p < \infty$*, problem* (1.1) *has a unique solution* $u \in C([0,T];L^2(\Omega)) \cap L^p(0,T;\dot H^2(\Omega))$ *such that* $\partial_t^\alpha u \in L^p(0,T;L^2(\Omega))$*.*

Proof. For any $\theta \in [0,1]$, consider the following fractional-order parabolic problem

$$\partial_t^\alpha u(t) + A(\theta t)u(t) = f(t), \quad t \in (0,T], \quad \text{with } u(0) = 0, \tag{2.7}$$

and define a set

$$D = \{\theta \in [0,1] : \text{ (2.7) has a solution } u \in L^p(0,T;\dot H^2(\Omega)) \text{ such that } \partial_t^\alpha u \in L^p(0,T;L^2(\Omega))\}.$$

Lemma 2.1 implies $0 \in D$ and so $D \neq \emptyset$.

For any $\theta \in D$, by rewriting (2.7) as

$$\partial_t^\alpha u(t) + A(\theta t_0)u(t) = f(t) + (A(\theta t_0) - A(\theta t))u(t), \quad t \in (0,T], \quad \text{with } u(0) = 0, \tag{2.8}$$

and by applying Lemma 2.1 in the time interval $(0,t_0)$, we obtain

$$\begin{aligned}
&\|\partial_t^\alpha u\|_{L^p(0,t_0;L^2(\Omega))} + \|u\|_{L^p(0,t_0;H^2(\Omega))} \\
\leq& c\|f\|_{L^p(0,t_0;L^2(\Omega))} + c\|(A(\theta t_0) - A(\theta t))u(t)\|_{L^p(0,t_0;L^2(\Omega))} \\
\leq& c\|f\|_{L^p(0,t_0;L^2(\Omega))} + c\|(t_0 - t)u(t)\|_{L^p(0,t_0;H^2(\Omega))},
\end{aligned} \tag{2.9}$$

where the last line follows from (2.6). Let $g(t) = \|u\|^p_{L^p(0,t;H^2(\Omega))}$, which satisfies $g'(t) = \|u(t)\|^p_{H^2(\Omega)}$. Then (2.9) and integration by parts imply

$$\begin{aligned}
g(t_0) &\leq c\|f\|^p_{L^p(0,t_0;L^2(\Omega))} + c\int_0^{t_0} (t_0 - t)^p g'(t)\mathrm{d}t \\
&= c\|f\|^p_{L^p(0,t_0;L^2(\Omega))} + cp\int_0^{t_0} (t_0 - t)^{p-1} g(t)\mathrm{d}t \\
&\leq c\|f\|^p_{L^p(0,t_0;L^2(\Omega))} + c\int_0^{t_0} g(t)\mathrm{d}t,
\end{aligned}$$

which implies (via the standard Gronwall's inequality)

$$g(t_0) \leq c\|f\|^p_{L^p(0,t_0;L^2(\Omega))}, \quad \text{i.e., } \|u\|_{L^p(0,t_0;H^2(\Omega))} \leq c\|f\|_{L^p(0,t_0;L^2(\Omega))}.$$

Substituting the last inequality into (2.9) yields

$$\|\partial_t^\alpha u\|_{L^p(0,t_0;L^2(\Omega))} + \|u\|_{L^p(0,t_0;H^2(\Omega))} \leq c\|f\|_{L^p(0,t_0;L^2(\Omega))}. \tag{2.10}$$

Since the estimate (2.10) is independent of $\theta \in D$, D is a closed subset of $[0,1]$.

Now we show that D is also open with respect to the subset topology of $[0,1]$. In fact, if $\theta_0 \in D$, then problem (2.7) can be rewritten as

$$\partial_t^\alpha u(t) + A(\theta_0 t)u(t) + (A(\theta t) - A(\theta_0 t))u(t) = f(t), \quad t \in (0,T], \quad \text{with } u(0) = 0, \tag{2.11}$$

which is equivalent to

$$\Big[1 + (\partial_t^\alpha + A(\theta_0 t))^{-1}(A(\theta t) - A(\theta_0 t))\Big]u(t) = (\partial_t^\alpha + A(\theta_0 t))^{-1} f(t). \tag{2.12}$$

It follows from (2.10) that the operator $(\partial_t^\alpha + A(\theta_0 t))^{-1}(A(\theta t) - A(\theta_0 t))$ is small in the sense that

$$\|(\partial_t^\alpha + A(\theta_0 t))^{-1}(A(\theta t) - A(\theta_0 t))\|_{L^p(0,T;H^2(\Omega)) \to L^p(0,T;H^2(\Omega))} \leq c|\theta - \theta_0|.$$

Thus for θ sufficiently close to θ_0, the operator $1 + (\partial_t^\alpha + A(\theta_0 t))^{-1}(A(\theta t) - A(\theta_0 t))$ is invertible on $L^p(0,T;\dot H^2(\Omega))$, which implies $\theta \in D$. Thus D is open with respect to the subset topology of $[0,1]$. Since D is both closed and open respect to the subset topology of $[0,1]$, $D = [0,1]$. Further, note that for $1/\alpha < p < \infty$, the inequality (2.10) and the condition $u(0) = 0$ directly imply $u \in C([0,T];L^2(\Omega))$ [13, Theorem 2.1], which completes the proof of the theorem. ∎

The following generalized Gronwall's inequality is useful ([5, Lemma 6.3] and [8, Exercise 3, p. 190]).

Lemma 2.3. *Let the function $\varphi(t) \geq 0$ be continuous for $0 < t \leq T$. If*

$$\varphi(t) \leq at^{-1+\alpha} + b\int_0^t (t-s)^{-1+\beta}\varphi(s)\mathrm{d}s, \quad 0 < t \leq T,$$

for some constants $a, b \geq 0$, $\alpha, \beta > 0$, then there is a constant $c = c(b, T, \alpha, \beta)$ such that

$$\varphi(t) \leq cat^{-1+\alpha}, \quad 0 < t \leq T.$$

Next we give the spatial regularity of the solution u for the case $f = 0$.

Theorem 2.2. *Under conditions (1.2)-(1.3), with $u_0 \in \dot{H}^\beta(\Omega)$, $0 \leq \beta \leq 2$, and $f = 0$, problem (1.1) has a unique solution $u \in C([0,T]; L^2(\Omega)) \cap C((0,T]; \dot{H}^2(\Omega))$ such that $\partial_t^\alpha u \in C((0,T]; L^2(\Omega))$, and*

$$\|u(t)\|_{H^2(\Omega)} \leq ct^{-(1-\beta/2)\alpha}\|u_0\|_{\dot{H}^\beta(\Omega)}.$$

Proof. The existence and uniqueness of a solution can be proved in the same way as Theorem 2.1 based on the a priori estimate below. We rewrite problem (1.1) as

$$\partial_t^\alpha u(t) + A(t_0)u(t) = (A(t_0) - A(t))u(t) + f(t), \quad t \in (0,T], \quad \text{with } u(0) = u_0,$$

Then the solution $u(t)$ can be represented by

$$u(t) = F(t;t_0)u_0 + \int_0^t E(t-s;t_0)(A(t_0) - A(s))u(s)\mathrm{d}s + \int_0^t E(t-s;t_0)f(s)\mathrm{d}s, \tag{2.13}$$

where the operators $F(t;t_0)$ and $E(t;t_0)$ are defined respectively by

$$F(t;t_0) := \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}z^{\alpha-1}(z^\alpha + A(t_0))^{-1}\,\mathrm{d}z \quad \text{and} \quad E(t;t_0) := \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}(z^\alpha + A(t_0))^{-1}\,\mathrm{d}z.$$

In the case $f = 0$, applying $A(t_0)$ to both sides of (2.13) yields

$$A(t_0)u(t) = A(t_0)F(t;t_0)u_0 + \int_0^t A(t_0)E(t-s;t_0)(A(t_0) - A(s))u(s)\mathrm{d}s.$$

Then conditions (1.2)-(1.3) and Lemma 2.2(ii) imply

$$\begin{aligned}
\|u(t_0)\|_{H^2(\Omega)} &\leq c\|A(t_0)F(t_0;t_0)u_0\|_{L^2(\Omega)} + c\int_0^{t_0} \|A(t_0)E(t_0-s;t_0)\|\|(A(t_0) - A(s))u(s)\|_{L^2(\Omega)}\mathrm{d}s \\
&\leq ct_0^{-(1-\beta/2)\alpha}\|u_0\|_{\dot{H}^\beta(\Omega)} + c\int_0^{t_0} (t_0-s)^{-1}(t_0-s)\|u(s)\|_{H^2(\Omega)}\mathrm{d}s \\
&= ct_0^{-(1-\beta/2)\alpha}\|u_0\|_{\dot{H}^\beta(\Omega)} + c\int_0^{t_0} \|u(s)\|_{H^2(\Omega)}\mathrm{d}s, \quad \forall\, t_0 \in (0,T].
\end{aligned}$$

The desired estimate follows from the generalized Gronwall's inequality in Lemma 2.3. It remains to show $u \in C([0,T]; L^2(\Omega)) \cap C((0,T]; H^2(\Omega))$. Indeed, note that (by fixing $t_0 = 0$)

$$\|u(t) - u_0\|_{L^2(\Omega)} \leq \|F(t;0)u_0 - u_0\|_{L^2(\Omega)} + \int_0^t (t-s)^{-\alpha}s\|u(s)\|_{H^2(\Omega)}\mathrm{d}s,$$

which together with the bound on $\|u(s)\|_{H^2(\Omega)}$ implies

$$\begin{aligned}
&\lim_{t\to 0^+} \|u(t) - u_0\|_{L^2(\Omega)} \\
&\leq \lim_{t\to 0^+} \|F(t;0)u_0 - u_0\|_{L^2(\Omega)} + \lim_{t\to 0^+} c\int_0^t (t-s)^{\alpha-1}s^{1-(1-\beta/2)\alpha}\mathrm{d}s\|u_0\|_{\dot{H}^\beta(\Omega)} = 0.
\end{aligned}$$

i.e., $\lim_{t\to 0^+} u(t) = u_0$ in $L^2(\Omega)$. The rest of the assertion follows similarly. This completes the proof. □

To analyze the temporal regularity, we first give three technical lemmas.

Lemma 2.4. *Let conditions (1.2) and (1.3) be fulfilled, and u be the solution to problem (1.1) with $u_0 \in L^2(\Omega)$ and $f = 0$. Then there holds*

$$\|\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(t-s;t_0)(A(t_0) - A(s))u(s)\mathrm{d}s\big|_{t=t_0}\|_{L^2(\Omega)} \leq c\|u_0\|_{L^2(\Omega)}.$$

Proof. Let $\mathrm{I} = \frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(t-s;t_0)(A(t_0)-A(s))u(s)\mathrm{d}s|_{t=t_0}$. Then

$$\begin{aligned}\mathrm{I} = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big(&\int_0^{t_0+\varepsilon} E(t_0+\varepsilon-s;t_0)(A(t_0)-A(s))u(s)\mathrm{d}s\\ &-\int_0^{t_0} E(t_0-s;t_0)(A(t_0)-A(s))u(s)\mathrm{d}s\Big) =: \lim_{\varepsilon\to 0}\Lambda(\varepsilon).\end{aligned}\tag{2.14}$$

If $\varepsilon > 0$, then

$$\begin{aligned}\Lambda(\varepsilon) = &\frac{1}{\varepsilon}\int_{t_0}^{t_0+\varepsilon} E(t_0+\varepsilon-s;t_0)(A(t_0)-A(s))u(s)\mathrm{d}s\\ &+\int_0^{t_0}\frac{E(t_0+\varepsilon-s;t_0)-E(t_0-s;t_0)}{\varepsilon}(A(t_0)-A(s))u(s)\mathrm{d}s =: \mathrm{I}_+ + \mathrm{II}_+.\end{aligned}\tag{2.15}$$

By applying Lemma 2.2(ii), (2.6) and Theorem 2.2, we deduce

$$\begin{aligned}\|\mathrm{I}_+\|_{L^2(\Omega)} &\le c\varepsilon^{-1}\int_{t_0}^{t_0+\varepsilon}\|E(t_0+\varepsilon-s;t_0)\|\|(A(t_0)-A(s))u(s)\|_{L^2(\Omega)}\mathrm{d}s\\ &\le c\varepsilon^{-1}\int_{t_0}^{t_0+\varepsilon}|t_0+\varepsilon-s|^{\alpha-1}|t_0-s|\|u(s)\|_{H^2(\Omega)}\mathrm{d}s\\ &\le c\varepsilon^{-1}\int_{t_0}^{t_0+\varepsilon}|t_0+\varepsilon-s|^{\alpha-1}|t_0-s|s^{-\alpha}\|u_0\|_{L^2(\Omega)}\mathrm{d}s\\ &\le c\varepsilon^{-1}\int_{t_0}^{t_0+\varepsilon}|t_0+\varepsilon-s|^{\alpha-1}(\varepsilon)t_0^{-\alpha}\|u_0\|_{L^2(\Omega)}\mathrm{d}s \le c\varepsilon^{\alpha}t_0^{-\alpha}\|u_0\|_{L^2(\Omega)},\end{aligned}$$

and similarly,

$$\begin{aligned}\|\mathrm{II}_+\|_{L^2(\Omega)} &= \left\|\int_0^{t_0}\int_0^1 E'(t_0+\theta\varepsilon-s;t_0)(A(t_0)-A(s))u(s)\,\mathrm{d}\theta\mathrm{d}s\right\|_{L^2(\Omega)}\\ &\le c\int_0^1\int_0^{t_0}(t_0+\theta\varepsilon-s)^{\alpha-2}(t_0-s)\|u(s)\|_{H^2(\Omega)}\,\mathrm{d}s\mathrm{d}\theta\\ &\le c\int_0^{t_0}(t_0-s)^{\alpha-1}\|u(s)\|_{H^2(\Omega)}\,\mathrm{d}s\\ &\le c\int_0^{t_0}(t_0-s)^{\alpha-1}s^{-\alpha}\|u_0\|_{L^2(\Omega)}\,\mathrm{d}s \le c\|u_0\|_{L^2(\Omega)}.\end{aligned}$$

If $-t_0 < \varepsilon < 0$, then

$$\begin{aligned}\Lambda(\varepsilon) = &\varepsilon^{-1}\int_{t_0-|\varepsilon|}^{t_0} E(t_0-s;t_0)(A(t_0)-A(s))u(s)\mathrm{d}s\\ &+\int_0^{t_0-|\varepsilon|}\frac{E(t_0-s;t_0)-E(t_0-|\varepsilon|-s;t_0)}{\varepsilon}(A(t_0)-A(s))u(s)\mathrm{d}s =: \mathrm{I}_- + \mathrm{II}_-,\end{aligned}$$

and similarly, we obtain

$$\|\mathrm{I}_-\|_{L^2(\Omega)} \le c\varepsilon^{\alpha}(t_0+\varepsilon)^{-\alpha}\|u_0\|_{L^2(\Omega)} \quad\text{and}\quad \|\mathrm{II}_-\|_{L^2(\Omega)} \le c\|u_0\|_{L^2(\Omega)}.$$

Combining the preceding estimates yields the assertion. □

Lemma 2.5. *Let conditions* (1.2) *and* (1.3) *be fulfilled, and u be the solution to problem* (1.1) *with $f \in C([0,T];L^2(\Omega))$, $\int_0^t (t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s < \infty$ and $u_0 = 0$. Then there holds*

$$\int_0^t (t-s)^{\alpha-1}\|u(s)\|_{H^2(\Omega)}\mathrm{d}s \le c\|f(0)\|_{L^2(\Omega)} + c\int_0^t (t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

Proof. By the solution representation (2.13) with $u_0 = 0$, we have

$$A(t_0)u(t_0) = \int_0^{t_0} A(t_0)E(t_0 - s; t_0)f(s)\mathrm{d}s + \int_0^{t_0} A(t_0)E(t_0 - s; t_0)(A(s) - A(t_0))u(s)\mathrm{d}s := \mathrm{I} + \mathrm{II}.$$

It follows directly from the definition of the operators $E(s;t_0)$ and $F(s;t_0)$ that the identity $A(t_0)E(s;t_0) = -\frac{\mathrm{d}}{\mathrm{d}s}F(s;t_0) = \frac{\mathrm{d}}{\mathrm{d}s}(I - F(s;t_0))$ holds. So upon changing variables and integration by parts, we obtain

$$\begin{aligned}\mathrm{I} &= \int_0^{t_0} A(t_0)E(s;t_0)f(t_0 - s)\mathrm{d}s = \int_0^{t_0} \frac{\mathrm{d}}{\mathrm{d}s}(I - F(s;t_0))f(t_0 - s)\mathrm{d}s \\ &= (F(t_0;t_0) - I)f(0) - \int_0^{t_0} (I - F(s;t_0))\frac{\mathrm{d}}{\mathrm{d}s}f(t_0 - s)\mathrm{d}s,\end{aligned}$$

where we have used the identity $F(0;t_0) = I$. Thus, by Lemma 2.2(i), we obtain

$$\|\mathrm{I}\|_{L^2(\Omega)} \le c\|f(0)\|_{L^2(\Omega)} + c\int_0^{t_0} \|f'(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

Similarly, by Lemma 2.2(ii) and (2.6), for the term II, we have

$$\|\mathrm{II}\|_{L^2(\Omega)} \le c\int_0^{t_0} (t_0 - s)^{-1}|t_0 - s|\|A(t_0)u(s)\|_{L^2(\Omega)}\mathrm{d}s = c\int_0^{t_0} \|A(t_0)u(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

Let $g(t) = \int_0^t (t-s)^{\alpha-1}\|u(s)\|_{H^2(\Omega)}\mathrm{d}s$. Then the last two estimates together give

$$\begin{aligned}g(t) &\le c\int_0^t (t-s)^{\alpha-1}(\|\mathrm{I}\|_{L^2(\Omega)} + \|\mathrm{II}\|_{L^2(\Omega)})\,\mathrm{d}s \\ &\le c\int_0^t (t-s)^{\alpha-1}\big(\|f(0)\|_{L^2(\Omega)} + \int_0^s \|f'(\xi)\|_{L^2(\Omega)}\mathrm{d}\xi + \int_0^\xi \|u(\xi)\|_{H^2(\Omega)}\mathrm{d}\xi\big)\,\mathrm{d}s \\ &\le ct^\alpha\|f(0)\|_{L^2(\Omega)} + c\int_0^t (t-s)^\alpha\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s + c\int_0^t g(s)\,\mathrm{d}s,\end{aligned}$$

where the last line follows directly from the semigroup property of Riemann-Liouville integral and change of integration orders. Now Gronwall's inequality gives

$$g(t) \le ct^\alpha\|f(0)\|_{L^2(\Omega)} + c\int_0^t (t-s)^\alpha\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s,$$

from which the desired assertion follows directly. □

Lemma 2.6. *Let conditions* (1.2) *and* (1.3) *be fulfilled, and* u *be the solution to problem* (1.1) *with* $f \in C([0,T];L^2(\Omega))$, $\int_0^t (t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s < \infty$ *and* $u_0 = 0$. *Then there holds*

$$\|\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(t-s;t_0)(A(t_0) - A(s))u(s)\mathrm{d}s\big|_{t=t_0}\|_{L^2(\Omega)} \le c\|f(0)\|_{L^2(\Omega)} + c\int_0^{t_0} (t_0 - s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

Proof. For any small $\varepsilon > 0$, we employ the splitting (2.14). By Lemma 2.2(ii) and (2.6), we bound the term I_+ by

$$\begin{aligned}\|\mathrm{I}_+\|_{L^2(\Omega)} &\le \varepsilon^{-1}\int_{t_0}^{t_0+\varepsilon} \|E(t_0 + \varepsilon - s;t_0)\|\|(A(t_0) - A(s))u(s)\|_{L^2(\Omega)}\mathrm{d}s \\ &\le c\varepsilon^{-1}\int_{t_0}^{t_0+\varepsilon} |t_0 + \varepsilon - s|^{\alpha-1}|t_0 - s|\|u(s)\|_{H^2(\Omega)}\mathrm{d}s \\ &\le c\int_{t_0}^{t_0+\varepsilon} |t_0 + \varepsilon - s|^{\alpha-1}\|u(s)\|_{H^2(\Omega)}\mathrm{d}s.\end{aligned}$$

This estimate, Hölder's inequality and Theorem 2.1 directly imply $\lim_{\varepsilon\to0^+}\|\mathrm{I}_+\|_{L^2(\Omega)} = 0$. For the term II_+, Lemma 2.2(ii) gives

$$\|\mathrm{II}_+\|_{L^2(\Omega)} = \Big\|\int_0^{t_0}\int_0^1 E'(t_0 + \theta\varepsilon - s;t_0)(A(t_0) - A(s))u(s)\,\mathrm{d}\theta\mathrm{d}s\Big\|_{L^2(\Omega)}$$

$$
\begin{aligned}
&\le c\int_0^1\int_0^{t_0}(t_0+\theta\varepsilon-s)^{\alpha-2}(t_0-s)\|u(s)\|_{H^2(\Omega)}\,\mathrm{d}s\mathrm{d}\theta\\
&\le c\int_0^{t_0}(t_0-s)^{\alpha-1}\|u(s)\|_{H^2(\Omega)}\,\mathrm{d}s,
\end{aligned}
$$

which together with Lemma 2.5 yields the assertion for $\varepsilon>0$. Similar estimates hold for the case $\varepsilon<0$, and this completes the proof of the lemma. □

Remark 2.1. *Note that the bound on* II_+ *in Lemma 2.4 blows up for* $\alpha\to 1^-$*:*

$$
\int_0^{t_0}(t_0-s)^{\alpha-1}s^{-\alpha}\,\mathrm{d}s = B(\alpha,1-\alpha),
$$

and in view of the asymptotics $B(\alpha,1-\alpha)=O((1-\alpha)^{-1})$ *as* $\alpha\to1^-$*, it blows up at a rate* $1/(1-\alpha)$*. Actually, this can be avoided by the following alternative argument:*

$$
\begin{aligned}
\|\mathrm{II}_+\|_{L^2(\Omega)} &= \Big\|\int_0^{t_0}\int_0^1 E'(t_0+\theta\varepsilon-s;t_0)A(t_0)^{1/2}A(t_0)^{1/2}(1-A(t_0)^{-1}A(s))u(s)\,\mathrm{d}\theta\mathrm{d}s\Big\|_{L^2(\Omega)}\\
&\le c\int_0^1\int_0^{t_0}(t_0+\theta\varepsilon-s)^{\alpha/2-2}(t_0-s)\|A(t_0)^{1/2}u(s)\|_{L^2(\Omega)}\,\mathrm{d}s\mathrm{d}\theta\\
&\le c\int_0^{t_0}(t_0-s)^{\alpha/2-1}\|u(s)\|_{H^1(\Omega)}\,\mathrm{d}s\\
&\le c\int_0^{t_0}(t_0-s)^{\alpha/2-1}s^{-\alpha/2}\|u_0\|_{L^2(\Omega)}\,\mathrm{d}s\le c\|u_0\|_{L^2(\Omega)},
\end{aligned}
$$

where the first inequality is due to (2.6)*, Corollary 3.1 below and interpolation. The same argument can be applied to the term* II_+ *in Lemma 2.6. Thus, the involved constants are bounded for* $\alpha\to1^-$*.*

Now we can give the temporal regularity of the solution u.

Theorem 2.3. *Let conditions* (1.2)-(1.3) *be fulfilled, and* u *be the solution to problem* (1.1)*.*

(i) *For* $u_0\in\dot H^\beta(\Omega)$*,* $0\le\beta\le2$*, and* $f=0$*, then*

$$
\|u'(t)\|_{L^2(\Omega)}\le ct^{-(1-\alpha\beta/2)}\|u_0\|_{\dot H^\beta(\Omega)}.
$$

(ii) *For* $u_0=0$*,* $f\in C([0,T];L^2(\Omega))$ *and* $\int_0^t(t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s<\infty$*, then*

$$
\|u'(t)\|_{L^2(\Omega)}\le ct^{-(1-\alpha)}\|f(0)\|_{L^2(\Omega)}+c\int_0^t(t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s.
$$

(iii) *For* $u_0=0$ *and* $f\in L^p(0,T;L^2(\Omega))$ *with* $2/\alpha<p<\infty$*, then*

$$
\|u(t)\|_{H^1(\Omega)}\le c\|f\|_{L^p(0,t;L^2(\Omega))}.
$$

Proof. The proof employs the solution representation (2.13). By Lemma 2.2(iii), we have

$$
\Big\|\frac{\mathrm{d}}{\mathrm{d}t}F(t;t_0)u_0\big|_{t=t_0}\Big\|_{L^2(\Omega)}\le ct_0^{-(1-\alpha\beta/2)}\|u_0\|_{\dot H^\beta(\Omega)}.
$$

This and Lemma 2.4 yield the assertion in part (i).

To show part (ii), differentiating (2.13) with respect to t yields

$$
u'(t_0)=\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(t-s;t_0)(A(t_0)-A(s))u(s)\mathrm{d}s\Big|_{t=t_0}+\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(s;t_0)f(t-s)\mathrm{d}s\Big|_{t=t_0}. \tag{2.16}
$$

In view of the identity

$$
\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(s;t_0)f(t-s)\mathrm{d}s=E(t;t_0)f(0)+\int_0^t E(s;t_0)f'(t-s)\mathrm{d}s,
$$

by Lemma 2.2(ii), we have

$$\begin{aligned}\left\|\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t E(s;t_0)f(t-s)\mathrm{d}s\right\|_{L^2(\Omega)} &\le \|E(t;t_0)f(0)\|_{L^2(\Omega)} + \int_0^t \|E(s;t_0)f'(t-s)\|_{L^2(\Omega)}\mathrm{d}s,\\ &\le ct^{-(1-\alpha)}\|f(0)\|_{L^2(\Omega)} + c\int_0^t s^{\alpha-1}\|f'(t-s)\|_{L^2(\Omega)}\,\mathrm{d}s.\end{aligned} \tag{2.17}$$

This and Lemma 2.6 complete the proof of part (ii).

Last, for the choice $2/\alpha < p < \infty$, Lemma 2.1 implies

$$\begin{aligned}u \in L^p(0,T;H^2(\Omega)) \cap W^{\alpha,p}(0,T;L^2(\Omega)) &\hookrightarrow W^{\alpha/2,p}(0,T;(L^2(\Omega),H^2(\Omega))_{1/2})\\ &= W^{\alpha/2,p}(0,T;H^1(\Omega)) \hookrightarrow C([0,T];H^1(\Omega)),\end{aligned}$$

where $(L^2(\Omega),H^2(\Omega))_{1/2}$ denotes the complex interpolation space between $L^2(\Omega)$ and $H^2(\Omega)$, and the last embedding is a consequence of [13, equation (2.3)]. Then the proof of Theorem 2.3 is complete. □

Remark 2.2. *In the error analysis, the work* [28] *requires the following conditions on the coefficient* $a(x,t)$*:* $a(x,t), \partial_t a(x,t) \in L^\infty(0,T;W^{1,\infty}(\Omega))$ *and* $\partial_{tt}^2 a(x,t) \in L^\infty(0,T;L^\infty(\Omega))$*, which are more stringent than* (1.3)*. Further, the work* [28] *has assumed the following regularity on the solution* u *to the homogeneous problem: for* $0 \le p \le q \le 2$*,*

$$\|u(t)\|_{\dot H^q(\Omega)} + t\|u'(t)\|_{\dot H^q(\Omega)} \le ct^{-(q-p)\alpha/2}\|u_0\|_{\dot H^p(\Omega)}.$$

In contrast, for the homogeneous problem, we proved the following estimates under assumption (1.3)*:*

$$t^{(1-\beta/2)\alpha}\|u(t)\|_{H^2(\Omega)} + t^{1-\alpha\beta/2}\|u'(t)\|_{L^2(\Omega)} \le c\|u_0\|_{\dot H^\beta(\Omega)}$$

and similar estimates for the inhomogeneous problem. It is worth noting that unlike the argument in [28]*, the error analysis below does not need the regularity* $\|u'(t)\|_{\dot H^2(\Omega)}$*, which allows us to relax the regularity assumption on the coefficient* $a(x,t)$*.*

Remark 2.3. *Our discussions focus on the low regularity in space, i.e.,* $u(t) \in H^2(\Omega)$*, which is sufficient for the error analysis of the piecewise linear FEM in Section 3. These results cannot be further improved for* $u_0 \in L^2(\Omega)$ *or* $f \in L^p(0,T;L^2(\Omega))$*, due to the limited smoothing properties of the solution operators* (*at most of order two in space*)*. For smoother problem data, one may expect higher spatial regularity of the solution. For example, for the homogeneous problem with a time-independent elliptic operator, there holds for any* $\beta \ge 0$ [29]

$$\|u(t)\|_{\dot H^{2+\beta}(\Omega)} \le ct^{-\alpha}\|u_0\|_{\dot H^\beta(\Omega)}, \quad t > 0.$$

Naturally, one may expect similar estimates for the case of a time-dependent elliptic operator, provided both the domain Ω *and the coefficient* $a(x,t)$ *are sufficiently smooth. Further, note that the regularity analysis extends straightforwardly to the slightly more general elliptic operators with the potential and convective terms, provided that the coefficients in the lower-order terms have suitable regularity.*

3. Semi-discrete Galerkin finite element method

In this part we investigate the semidiscrete Galerkin FEM. Let $\mathcal{T}_h$ be a shape regular quasi-uniform triangulation of the domain Ω into simplicial elements, and h be the maximal diameter of the elements. Let $S_h \subset H_0^1(D)$ be the space of continuous piecewise linear functions over the triangulation $\mathcal{T}_h$. Then we define the $L^2(\Omega)$ orthogonal projection $P_h : L^2(\Omega) \to S_h$ by

$$(P_h\varphi,\chi) = (\varphi,\chi) \quad \forall\,\varphi \in L^2(\Omega),\ \forall\,\chi \in S_h.$$

The operator P_h satisfies the following error estimate

$$\|P_h\varphi - \varphi\|_{L^2(\Omega)} + h\|\nabla(P_h\varphi - \varphi)\|_{L^2(\Omega)} \le ch^q\|\varphi\|_{H^q(\Omega)}, \quad \varphi \in \dot H^q(\Omega),\ q = 1,2.$$

The spatially semidiscrete FEM for problem (1.1) reads: find $u_h(t) \in S_h$ such that

$$(\partial_t^\alpha u_h(t),\chi) + (a(\cdot,t)\nabla u_h(t),\nabla\chi) = (f(\cdot,t),\chi), \quad \forall\chi \in S_h,\ t \in (0,T], \quad \text{with } u_h(0) = P_h u_0. \tag{3.1}$$

Then we define a time-dependent operator $A_h(t) : S_h \to S_h$ by

$$(A_h(t)v_h, \chi) = (a(\cdot, t)\nabla v_h, \nabla\chi), \quad \forall\, v_h, \chi \in S_h.$$

Under condition (1.2), $A_h(t) : S_h \to S_h$ is bounded and invertible on S_h, and problem (3.1) can be rewritten as

$$\partial_t^\alpha u_h(t) + A_h(t)u_h(t) = P_h f(t), \quad \forall\, t \in (0, T], \quad \text{with } u_h(0) = P_h u_0. \tag{3.2}$$

3.1. **Perturbation lemmas.** In this part we give two crucial perturbation results. We need a time-dependent Ritz projection operator $R_h(t) : H_0^1(\Omega) \to S_h$ defined by

$$(a(\cdot, t)\nabla R_h(t)\varphi, \nabla\chi) = (a(\cdot, t)\nabla\varphi, \nabla\chi), \quad \forall\varphi \in H_0^1(\Omega), \chi \in S_h. \tag{3.3}$$

The operator $R_h(t)$ satisfies the following approximation property [24, p. 99]:

$$\|R_h(t)\varphi - \varphi\|_{L^2(\Omega)} + h\|\nabla(R_h(t)\varphi - \varphi)\|_{L^2(\Omega)} \le ch^q\|\varphi\|_{H^q(\Omega)}, \quad \varphi \in \dot{H}^q(\Omega), q = 1, 2. \tag{3.4}$$

Lemma 3.1. *Under conditions* (1.2)-(1.3)*, the following estimate holds:*

$$\|(I - A_h(t)^{-1}A_h(s))v_h\|_{L^2(\Omega)} \le c|t - s|\|v_h\|_{L^2(\Omega)}, \quad \forall\, v_h \in S_h.$$

Proof. For any given $v_h \in S_h$, let $\varphi_h = A_h(s)v_h$ and $w_h = A_h(t)^{-1}\varphi_h$. Then

$$(A_h(t)w_h, \chi) = (\varphi_h, \chi) = (A_h(s)v_h, \chi), \quad \forall\, \chi \in S_h,$$

which implies

$$(a(\cdot, t)\nabla w_h, \nabla\chi) = (a(\cdot, s)\nabla v_h, \nabla\chi), \quad \forall\, \chi \in S_h.$$

Consequently,

$$(a(\cdot, t)\nabla(w_h - v_h), \nabla\chi) = ((a(\cdot, s) - a(\cdot, t))\nabla v_h, \nabla\chi), \quad \forall\, \chi \in S_h.$$

Let $\phi \in H_0^1(\Omega)$ be the weak solution of the elliptic problem

$$(a(\cdot, t)\nabla\phi, \nabla\xi) = ((a(\cdot, s) - a(\cdot, t))\nabla v_h, \nabla\xi), \quad \forall\, \xi \in H_0^1(\Omega). \tag{3.5}$$

By Lax-Milgram theorem, ϕ satisfies the following *a priori* estimate:

$$\|\phi\|_{H^1(\Omega)} \le c\|(a(\cdot, s) - a(\cdot, t))\nabla v_h\|_{L^2(\Omega)} \le c|t - s|\|v_h\|_{H^1(\Omega)}.$$

Thus, with the Ritz projection $R_h(t)$, cf. (3.3), we have $w_h - v_h = R_h(t)\phi$.

By the error estimate (3.4) and the inverse inequality,

$$\begin{aligned}\|w_h - v_h - \phi\|_{L^2(\Omega)} &\le ch\|\phi\|_{H^1(\Omega)} \le ch|t - s|\|v_h\|_{H^1(\Omega)}\\ &\le c|t - s|\|v_h\|_{L^2(\Omega)}.\end{aligned}$$

Thus, the triangle inequality implies

$$\|w_h - v_h\|_{L^2(\Omega)} \le c|t - s|\|v_h\|_{L^2(\Omega)} + \|\phi\|_{L^2(\Omega)}. \tag{3.6}$$

For any $\varphi \in L^2(\Omega)$, let $\xi \in \dot{H}^2(\Omega)$ be the solution of the elliptic problem

$$-\nabla \cdot (a(\cdot, t)\nabla\xi) = \varphi.$$

Then $\|\xi\|_{H^2(\Omega)} \le c\|\varphi\|_{L^2(\Omega)}$. By substituting ξ into (3.5), we obtain

$$\begin{aligned}|(\phi, \varphi)| &= |(a(\cdot, t)\nabla\phi, \nabla\xi)|\\ &= |((a(\cdot, s) - a(\cdot, t))\nabla v_h, \nabla\xi)|\\ &= |(v_h, \nabla \cdot (a(\cdot, s) - a(\cdot, t))\nabla\xi)|\\ &\le c|t - s|\|v_h\|_{L^2(\Omega)}\|\xi\|_{H^2(\Omega)}\\ &\le c|t - s|\|v_h\|_{L^2(\Omega)}\|\varphi\|_{L^2(\Omega)}.\end{aligned}$$

This implies (via duality)

$$\|\phi\|_{L^2(\Omega)} = \sup_{\varphi \in L^2(\Omega)} \frac{|(\phi, \varphi)|}{\|\varphi\|_{L^2(\Omega)}} \le c|t - s|\|v_h\|_{L^2(\Omega)}.$$

Substituting the last inequality back into (3.6), we deduce

$$\|w_h - v_h\|_{L^2(\Omega)} \le c|t-s|\|v_h\|_{L^2(\Omega)}.$$

This completes the proof of Lemma 3.1. □

Remark 3.1. *Note that the semidiscrete operator $A_h(t)$ is self-adjoint. Then Lemma 3.1 together with a duality argument yields*

$$\|(I - A_h(s)A_h(t)^{-1})v_h\|_{L^2(\Omega)} \le c|t-s|\|v_h\|_{L^2(\Omega)}, \quad \forall v_h \in S_h.$$

Consequently,

$$\begin{aligned}\|(A_h(t) - A_h(s))v_h\|_{L^2(\Omega)} &\le \|(I - A_h(s)A_h(t)^{-1})A_h(t)v_h\|_{L^2(\Omega)} \\ &\le c|t-s|\|A_h(t)v_h\|_{L^2(\Omega)}.\end{aligned}$$

Further, the interpolation between $\beta = 0, 1$ yields

$$\|A_h^\beta(t)(I - A_h(t)^{-1}A_h(s))v_h\|_{L^2(\Omega)} \le c|t-s|\|A_h^\beta(t)v_h\|_{L^2(\Omega)}.$$

The following result is the continuous analogue of Lemma 3.1, and it is independent interest.

Corollary 3.1. *Under conditions* (1.2)-(1.3)*, the following estimate holds:*

$$\|(I - A(t)^{-1}A(s))v\|_{L^2(\Omega)} \le c|t-s|\|v\|_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega).$$

Proof. For any $v \in \dot H^2(\Omega)$, there holds $A_h(s)R_h(s)v = P_hA(s)v$ [34, equation (1.34), p. 11]. By the standard error estimates for Galerkin FEM,

$$\begin{aligned}&\|A_h(t)^{-1}A_h(s)R_h(s)v - A(t)^{-1}A(s)v\|_{L^2(\Omega)} \\ =&\|(A_h(t)^{-1}P_h - A(t)^{-1})A(s)v\|_{L^2(\Omega)} \\ \le& ch^2\|A(s)v\|_{L^2(\Omega)} \le ch^2\|v\|_{H^2(\Omega)}.\end{aligned}$$

Then by Lemma 3.1 and the triangle inequality, we deduce

$$\begin{aligned}\|(I - A(t)^{-1}A(s))v\|_{L^2(\Omega)} &\le \|(I - A_h(t)^{-1}A_h(s))R_h(s)v\|_{L^2(\Omega)} + ch^2\|v\|_{H^2(\Omega)} \\ &\le c|t-s|\|R_h(s)v\|_{L^2(\Omega)} + ch^2\|v\|_{H^2(\Omega)} \\ &\le c|t-s|\|v\|_{L^2(\Omega)} + c|t-s|\|R_h(s)v - v\|_{L^2(\Omega)} + ch^2\|v\|_{H^2(\Omega)} \\ &\le c|t-s|\|v\|_{L^2(\Omega)} + (c|t-s| + c)h^2\|v\|_{H^2(\Omega)}, \quad \forall v \in \dot H^2(\Omega).\end{aligned}$$

Then the assertion follows by letting $h \to 0$ and noting that the space $\dot H^2(\Omega)$ is dense in $H_0^1(\Omega)$. □

Lemma 3.2. *Under conditions* (1.2)-(1.3)*, the following estimate holds:*

$$\|(R_h(t) - R_h(s))v\|_{L^2(\Omega)} \le ch^2|t-s|\|v\|_{H^2(\Omega)}, \quad \forall v \in \dot H^2(\Omega). \tag{3.7}$$

Proof. By the definition of Ritz projection, cf. (3.3), the difference $\eta_h = R_h(t)v - R_h(s)v \in S_h$ satisfies

$$(a(\cdot,s)\nabla\eta_h, \nabla\chi) = ((a(\cdot,t) - a(\cdot,s))\nabla(v - R_h(t)v), \nabla\chi), \quad \forall \chi \in S_h.$$

Let $\eta \in H_0^1(\Omega)$ be the weak solution of the elliptic problem

$$(a(\cdot,s)\nabla\eta, \nabla\xi) = ((a(\cdot,t) - a(\cdot,s))\nabla(v - R_h(t)v), \nabla\xi), \quad \forall \xi \in H_0^1(\Omega). \tag{3.8}$$

By the definition of $R_h(s)$, cf. (3.3), $\eta_h = R_h(s)\eta$ and by the error estimate (3.4), there holds

$$\begin{aligned}\|\eta_h - \eta\|_{L^2(\Omega)} &\le ch\|\eta\|_{H^1(\Omega)} \le ch\|(a(\cdot,t) - a(\cdot,s))\nabla(v - R_h(t)v)\|_{L^2(\Omega)} \\ &\le ch^2|t-s|\|v\|_{H^2(\Omega)}.\end{aligned}$$

The triangle inequality implies

$$\|\eta_h\|_{L^2(\Omega)} \le ch^2|t-s|\|v\|_{H^2(\Omega)} + \|\eta\|_{L^2(\Omega)}. \tag{3.9}$$

Next we use a duality argument to bound $\|\eta\|_{L^2(\Omega)}$. For any $\varphi \in L^2(\Omega)$, let $\xi \in \dot H^2(\Omega)$ be the solution of the elliptic problem

$$-\nabla \cdot (a(\cdot,s)\nabla\xi) = \varphi.$$

Upon substituting ξ into (3.8), we obtain

$$\begin{aligned}|(\eta,\varphi)| &= |(a(\cdot,s)\nabla\eta,\nabla\xi)| = |((a(\cdot,t)-a(\cdot,s))\nabla(v-R_h(t)v),\nabla\xi)|\\ &= |(v-R_h(t)v,\nabla\cdot((a(\cdot,t)-a(\cdot,s))\nabla\xi)|\\ &\le c\|v-R_h(t)v\|_{L^2(\Omega)}|t-s|\|\xi\|_{H^2(\Omega)}\\ &\le ch^2\|v\|_{H^2(\Omega)}|t-s|\|\varphi\|_{L^2(\Omega)},\end{aligned}$$

which implies (via duality)

$$\|\eta\|_{L^2(\Omega)} \le ch^2\|v\|_{H^2(\Omega)}|t-s|.$$

Substituting the above inequality into (3.9) yields Lemma 3.2. □

3.2. **Semidiscrete scheme and error estimates.** By the discrete maximal L^p-regularity, one can show the existence and uniqueness of a FEM solution $u_h(t)$. We also have the following stability estimates. The proof is identical with that for Theorems 2.1–2.3, using the estimates in Section 3.1, and hence it is omitted.

Theorem 3.1. *Let conditions* (1.2)-(1.3) *be fulfilled, and u_h be the solution to problem* (3.1)*.*

(i) *For $u_0 \in \dot H^\beta(\Omega)$, $0\le\beta\le 2$, and $f=0$, then*

$$\|A_hu_h(t)\|_{L^2(\Omega)} \le ct^{-(1-\beta/2)\alpha}\|u_0\|_{\dot H^\beta(\Omega)} \quad \text{and} \quad \|u_h'(t)\|_{L^2(\Omega)} \le ct^{-(1-\alpha\beta/2)}\|u_0\|_{\dot H^\beta(\Omega)}.$$

(ii) *For $u_0=0$, $f\in C([0,T];L^2(\Omega))$ and $\int_0^t (t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s<\infty$, then*

$$\|A_hu_h\|_{L^p(0,T;L^2(\Omega))} + \|\partial_t^\alpha u_h\|_{L^p(0,T;L^2(\Omega))} \le c\|f\|_{L^p(0,T;L^2(\Omega))},$$

$$\|u_h'(t)\|_{L^2(\Omega)} \le ct^{-(1-\alpha)}\|f(0)\|_{L^2(\Omega)} + c\int_0^t (t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s,$$

$$\int_0^t (t-s)^{\alpha-1}\|A_hu_h(s)\|_{L^2(\Omega)}\mathrm{d}s \le c\|f(0)\|_{L^2(\Omega)} + c\int_0^t (t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

(iii) *For $u_0=0$ and $f\in L^p(0,T;L^2(\Omega))$ with $p\in(2/\alpha,\infty)$, then*

$$\|u_h(t)\|_{H^1(\Omega)} \le c\|f\|_{L^p(0,t;L^2(\Omega))}.$$

Now we derive error estimates for the semidiscrete solution u_h. Problem (3.1) can be rewritten as

$$\partial_t^\alpha u_h(t) + A_h(t_0)u_h(t) = P_hf(t) + (A_h(t_0)-A_h(t))u_h(t), \quad t\in(0,T], \quad \text{with } u_h(0)=P_hu_0,$$

whose solution is given by

$$u_h(t) = F_h(t;t_0)P_hu_0 + \int_0^t E_h(t-s;t_0)\Big(P_hf(s) + (A_h(t_0)-A_h(s))u_h(s)\Big)\mathrm{d}s, \tag{3.10}$$

where the semidiscrete solution operators $F_h(t;t_0)$ and $E_h(t;t_0)$ are defined respectively by

$$F_h(t;t_0) := \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}z^{\alpha-1}(z^\alpha+A_h(t_0))^{-1}\,\mathrm{d}z \quad \text{and} \quad E_h(t;t_0) := \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt}(z^\alpha+A_h(t_0))^{-1}\,\mathrm{d}z.$$

Let $e_h = P_hu - u_h$. Then by (2.13) and (3.10), e_h can be represented by

$$\begin{aligned}e_h(t) =&(P_hF(t;t_0)u_0 - F_h(t;t_0)P_hu_0) + \int_0^t (P_hE(t-s;t_0) - E_h(t-s;t_0)P_h)f(s)\mathrm{d}s\\ &+ \int_0^t (P_hE(t-s;t_0) - E_h(t-s;t_0)P_h)(A(t_0)-A(s))u(s)\mathrm{d}s\\ &+ \int_0^t E_h(t-s;t_0)\Big(P_h(A(t_0)-A(s))u(s) - (A_h(t_0)-A_h(s))u_h(s)\Big)\mathrm{d}s\\ =:& \sum_{i=1}^4 \mathrm{I}_i(t).\end{aligned} \tag{3.11}$$

The terms $\mathrm{I}_1(t)$ and $\mathrm{I}_2(t)$ represent the errors for the homogeneous and inhomogeneous problems with a time-independent operator $A(t_0)$, respectively, which have been analyzed: [11, Theorem 3.7] implies

$$\|\mathrm{I}_1(t_0)\|_{L^2(\Omega)} \le c t_0^{-(1-\beta/2)\alpha} h^2 \|u_0\|_{\dot H^\beta(\Omega)}, \quad \beta \in [0,2], \tag{3.12}$$

and by the argument in [9], there holds (with $\ell_h = \log(1+1/h)$)

$$\|\mathrm{I}_2(t_0)\|_{L^2(\Omega)} \le c h^2 \ell_h^2 \|f\|_{L^\infty(0,T;L^2(\Omega))}. \tag{3.13}$$

It remains to bound the two terms $\mathrm{I}_3(t)$ and $\mathrm{I}_4(t)$, which are given below. We shall discuss the homogeneous and inhomogeneous problems separately.

Lemma 3.3. *Under conditions* (1.2) *and* (1.3)*, for $u_0 \in L^2(\Omega)$ and $f = 0$, for the term $\mathrm{I}_3(t)$, there holds*

$$\|\mathrm{I}_3(t_0)\|_{L^2(\Omega)} \le c h^2 \|u_0\|_{L^2(\Omega)}.$$

Proof. By the definitions of the operators $E(t;t_0)$ and $E_h(t;t_0)$, we have

$$\|\mathrm{I}_3(t_0)\|_{L^2(\Omega)} \le c\int_0^{t_0}\int_{\Gamma_{\theta,\delta}} |e^{z(t_0-s)}| \big\|(z^\alpha + A(t_0))^{-1} - (z^\alpha + A_h(t_0))^{-1}P_h\big\|_{L^2(\Omega)} \|(A(t_0)-A(s))u(s)\|\,|\mathrm{d}z|\mathrm{d}s.$$

By condition (1.2), for any $z \in \Gamma_{\theta,\delta}$, we have [6, p. 820]

$$\big\|(z^\alpha + A(t_0))^{-1} - (z^\alpha + A_h(t_0))^{-1}P_h\big\| \le c h^2,$$

where the constant c is independent of z. Meanwhile, condition (1.3) implies

$$\|(A(t_0)-A(s))u(s)\|_{L^2(\Omega)} \le c|t_0 - s|\|u(s)\|_{H^2(\Omega)}.$$

Thus by Theorem 2.2,

$$\begin{aligned}\|\mathrm{I}_3(t_0)\|_{L^2(\Omega)} &\le c h^2 \int_0^{t_0} (t_0-s)^{-1}(t_0-s)\|u(s)\|_{H^2(\Omega)}\mathrm{d}s \le c h^2 \int_0^{t_0} \|u(s)\|_{H^2(\Omega)}\mathrm{d}s \\ &= c h^2 \int_0^{t_0} s^{-\alpha}\|u_0\|_{L^2(\Omega)}\mathrm{d}s \le c h^2 \|u_0\|_{L^2(\Omega)}. \end{aligned}\tag{3.14}$$

This completes the proof of the lemma. □

Lemma 3.4. *Under conditions* (1.2) *and* (1.3)*, for $u_0 \in L^2(\Omega)$ and $f = 0$, for the term $\mathrm{I}_4(t)$, there holds*

$$\|\mathrm{I}_4(t_0)\|_{L^2(\Omega)} \le c h^2 \|u_0\|_{L^2(\Omega)} + c\int_0^{t_0} \|e_h(s)\|_{L^2(\Omega)}\mathrm{d}s.$$

Proof. Let $e_h = P_h u - u_h$. Using the identity $P_h A(s) = A_h(s) R_h(s)$ [34, (1.34), p. 11] and the triangle inequality, we derive

$$\begin{aligned}\|\mathrm{I}_4(t_0)\|_{L^2(\Omega)} &= \Big\| \int_0^{t_0} E_h(t_0-s;t_0)\Big((A_h(t_0)R_h(t_0) - A_h(s)R_h(s))u(s) - (A_h(t_0)-A_h(s))u_h(s)\Big)\mathrm{d}s\Big\|_{L^2(\Omega)} \\ &\le \Big\|\int_0^{t_0} E_h(t_0-s;t_0)(A_h(t_0)-A_h(s))e_h(s)\mathrm{d}s\Big\|_{L^2(\Omega)} \\ &\quad + \Big\|\int_0^{t_0} E_h(t_0-s;t_0)\Big(A_h(t_0)(R_h(t_0)-P_h)u(s) - A_h(s)(R_h(s)-P_h)u(s)\Big)\mathrm{d}s\Big\|_{L^2(\Omega)} \\ &=: \mathrm{I}_{4,1}(t_0) + \mathrm{I}_{4,2}(t_0).\end{aligned}$$

For the term $\mathrm{I}_{4,1}(t_0)$, by Lemmas 2.2(ii) and 3.1, we have

$$\begin{aligned}\mathrm{I}_{4,1}(t_0) &= \Big\|\int_0^{t_0} A_h(t_0)E_h(t_0-s;t_0)(I - A_h(t_0)^{-1}A_h(s))e_h(s)\mathrm{d}s\Big\|_{L^2(\Omega)} \\ &\le \int_0^{t_0} \|A_h(t_0)E_h(t_0-s;t_0)\| \|(I - A_h(t_0)^{-1}A_h(s))e_h(s)\|_{L^2(\Omega)}\mathrm{d}s \\ &\le c\int_0^{t_0} (t_0-s)^{-1}(t_0-s)\|e_h(s)\|_{L^2(\Omega)}\mathrm{d}s = c\int_0^{t_0}\|e_h(s)\|_{L^2(\Omega)}\mathrm{d}s.\end{aligned}$$

For the term $\mathrm{I}_{4,2}(t_0)$, by the triangle inequality, we further split it into

$$\mathrm{I}_{4,2}(t_0) \leq \Big\| \int_0^{t_0} E_h(t_0-s;t_0)A_h(t_0)(R_h(t_0)-R_h(s))u(s)\mathrm{d}s \Big\|_{L^2(\Omega)} + \Big\| \int_0^{t_0} E_h(t_0-s;t_0)(A_h(t_0)-A_h(s))(R_h(s)-P_h)u(s)\Big)\mathrm{d}s \Big\|_{L^2(\Omega)} =: \mathrm{I}'_{4,2}(t_0) + \mathrm{I}''_{4,2}(t_0).$$

Now by Lemmas 2.2(ii) and 3.2 and Theorem 2.2, we bound $\mathrm{I}'_{4,2}(t_0)$ by

$$\begin{aligned}
\mathrm{I}'_{4,2}(t_0) &\leq \int_0^{t_0} \|E_h(t_0-s;t_0)A_h(t_0)\| \|(R_h(t_0)-R_h(s))u(s)\|_{L^2(\Omega)}\mathrm{d}s \\
&\leq c\int_0^{t_0} (t_0-s)^{-1}(t_0-s)h^2\|u(s)\|_{H^2(\Omega)}\mathrm{d}s \\
&\leq ch^2\int_0^{t_0} s^{-\alpha}\|u_0\|_{L^2(\Omega)}\mathrm{d}s \leq ch^2\|u_0\|_{L^2(\Omega)}.
\end{aligned}$$

Likewise, by Lemma 3.1 and Theorem 2.2, we bound $I''_{4,2}(t_0)$ by

$$\begin{aligned}
\mathrm{I}''_{4,2}(t_0) &= \Big\| \int_0^{t_0} A_h(t_0)E_h(t_0-s;t_0)(I-A_h(t_0)^{-1}A_h(s))(R_h(s)-P_h)u(s)\Big)\mathrm{d}s \Big\|_{L^2(\Omega)} \\
&\leq \int_0^{t_0} \|A_h(t_0)E_h(t_0-s;t_0)\| \|(I-A_h(t_0)^{-1}A_h(s))(R_h(s)-P_h)u(s)\|_{L^2(\Omega)}\mathrm{d}s \\
&\leq c\int_0^{t_0} (t_0-s)^{-1}(t_0-s)\|(R_h(s)-P_h)u(s)\|_{L^2(\Omega)}\mathrm{d}s \\
&\leq ch^2\int_0^{t_0} \|u(s)\|_{H^2(\Omega)}\mathrm{d}s \leq ch^2\int_0^{t_0} s^{-\alpha}\|u_0\|_{L^2(\Omega)}\mathrm{d}s \leq ch^2\|u_0\|_{L^2(\Omega)}.
\end{aligned}$$

The desired assertion follows by combining the preceding estimates. □

Now we can state the main result of this part, i.e., error estimate on the semidiscrete solution u_h.

Theorem 3.2. *Under conditions* (1.2) *and* (1.3), *for* $u_0 \in L^2(\Omega)$ *and* $f = 0$, *there holds*

$$\|u(t)-u_h(t)\|_{L^2(\Omega)} \leq ch^2t^{-\alpha}\|u_0\|_{L^2(\Omega)}.$$

Proof. Substituting (3.12) and Lemmas 3.3 and 3.4 into (3.11) yields

$$\|P_hu(t_0)-u_h(t_0)\|_{L^2(\Omega)} \leq ct_0^{-\alpha}h^2\|u_0\|_{L^2(\Omega)} + c\int_0^{t_0} \|P_hu(s)-u_h(s)\|_{L^2(\Omega)}\mathrm{d}s, \quad \forall\, t_0 \in (0,T].$$

By Gronwall's inequality from Lemma 2.3, we obtain

$$\|P_hu(t)-u_h(t)\|_{L^2(\Omega)} \leq ct^{-\alpha}h^2\|u_0\|_{L^2(\Omega)}, \quad \forall\, t \in (0,T].$$

By the approximation property of P_h and Theorem 2.2, we have

$$\|u(t_0)-P_hu(t_0)\|_{L^2(\Omega)} \leq ch^2\|u(t_0)\|_{H^2(\Omega)} \leq ct_0^{-\alpha}h^2\|u_0\|_{L^2(\Omega)}.$$

The last two estimates together imply the desired result. □

A similar error estimate holds for the inhomogeneous problem.

Theorem 3.3. *Under conditions* (1.2) *and* (1.3), *for* $u_0 = 0$ *and* $f \in L^\infty(0,T;L^2(\Omega))$, *there holds*

$$\|u(t)-u_h(t)\|_{L^2(\Omega)} \leq ch^2\ell_h^2\|f\|_{L^\infty(0,t;L^2(\Omega))}, \quad \text{with } \ell_h = \log(1+1/h).$$

Proof. The proof is similar to Theorem 3.2, in view of (3.13), and the following estimates:

$$\begin{aligned}
&\|\mathrm{I}_3(t_0)\|_{L^2(\Omega)} \leq ch^2\|f\|_{L^\infty(0,t_0;L^2(\Omega))}, \\
&\|\mathrm{I}_4(t_0)\|_{L^2(\Omega)} \leq ch^2\|f\|_{L^\infty(0,t_0;L^2(\Omega))} + c\int_0^{t_0} \|e_h(s)\|_{L^2(\Omega)}\mathrm{d}s,
\end{aligned}$$

which follow similarly as Lemmas 3.3 and 3.4. Actually, the first follows from (3.14) and Theorem 2.1 by

$$\|\mathrm{I}_3(t_0)\|_{L^2(\Omega)} \le ch^2 \int_0^{t_0} \|u(s)\|_{H^2(\Omega)}\mathrm{d}s \le ch^2\|f\|_{L^\infty(0,t_0;L^2(\Omega))}.$$

Similarly, the second follows from the expressions of $\mathrm{I}_{4,1}'$ and $\mathrm{I}_{4,2}''$ in Lemma 3.4, and Theorem 2.1. □

Remark 3.2. *We have only discussed discretization by piecewise linear finite elements. It is of much interest to extend the analysis to high-order finite elements. This seems missing even for the case of a time-independent diffusion coefficient when problem data are nonsmooth, partly due to the limited smoothing property of the solution operators* [10].

4. Time discretization

Now we study the time discretization of problem (1.1). We divide the time interval $[0,T]$ into a uniform grid, with $t_n = n\tau$, $n = 0,\ldots,N$, and $\tau = T/N$ being the time step size. Then we approximate the Riemann-Liouville fractional derivative

$${}^R\partial_t^\alpha\varphi(t) = \frac{1}{\Gamma(1-\alpha)}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t (t-s)^{-\alpha}\varphi(s)\mathrm{d}s$$

by the backward Euler (BE) convolution quadrature (with $\varphi^j = \varphi(t_j)$) [22, 12]:

$${}^R\partial_t^\alpha\varphi(t_n) \approx \tau^{-\alpha}\sum_{j=0}^n b_j\varphi^{n-j} := \bar\partial_\tau^\alpha\varphi^n, \quad \text{with } \sum_{j=0}^\infty b_j\xi^j = (1-\xi)^\alpha.$$

The fully discrete scheme for problem (1.1) reads: find $u_h^n \in S_h$ such that

$$\bar\partial_\tau^\alpha(u_h^n - u_h^0) + A_h(t_n)u_h^n = P_hf(t_n), \quad n = 1,2,\ldots,N, \tag{4.1}$$

with the initial condition $u_h^0 = P_hu_0 \in S_h$. Similar to the semidiscrete case, for a given $m \in \mathbb{N}$ with $1 \le m \le N$, we rewrite (4.1) as

$$\bar\partial_\tau^\alpha(u_h^n - u_h^0) + A_h(t_m)u_h^n = P_hf(t_n) + (A_h(t_m) - A_h(t_n))u_h^n. \tag{4.2}$$

By means of discrete Laplace transform, the fully discrete solution $u_h^m \in S_h$ is given by

$$u_h^m = F_{\tau,m}^m u_h^0 + \tau\sum_{k=1}^m E_{\tau,m}^{m-k}[P_hf(t_k) + (A_h(t_m) - A_h(t_k))u_h^k], \tag{4.3}$$

where the fully discrete operators $F_{\tau,m}^n$ and $E_{\tau,m}^n$ are respectively defined by (with $\delta_\tau(\xi) = (1-\xi)/\tau$)

$$F_{\tau,m}^n = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}^\tau} e^{zn\tau}\delta_\tau(e^{-z\tau})^{\alpha-1}(\delta_\tau(e^{-z\tau})^\alpha + A_h(t_m))^{-1}\,\mathrm{d}z, \tag{4.4}$$

$$E_{\tau,m}^n = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}^\tau} e^{zn\tau}(\delta_\tau(e^{-z\tau})^\alpha + A_h(t_m))^{-1}\,\mathrm{d}z, \tag{4.5}$$

with the contour $\Gamma_{\theta,\delta}^\tau := \{z \in \Gamma_{\theta,\delta} : |\Im(z)| \le \pi/\tau\}$ (oriented with an increasing imaginary part).

The next lemma gives elementary properties of the kernel $\delta_\tau(e^{-z\tau})$.

Lemma 4.1. *For any $\theta \in (\pi/2,\pi)$, there exists $\theta' \in (\pi/2,\pi)$ and positive constants c, c_1, c_2 (independent of τ) such that for all $z \in \Gamma_{\theta,\delta}^\tau$*

$$c_1|z| \le |\delta_\tau(e^{-z\tau})| \le c_2|z|, \quad \delta_\tau(e^{-z\tau}) \in \Sigma_{\theta'},$$
$$|\delta_\tau(e^{-z\tau}) - z| \le c\tau|z|^2, \qquad |\delta_\tau(e^{-z\tau})^\alpha - z^\alpha| \le c\tau|z|^{1+\alpha}.$$

By the solution representations (3.10) and (4.3), the temporal error $e_h^m = u_h^m - u_h(t_m)$ satisfies

$$e_h^m = (F_h(t_m;t_m)P_hu_0 - F_{\tau,m}^m u_h^0) + \big(\tau\sum_{k=1}^m E_{\tau,m}^{m-k}P_hf(t_k) - \sum_{k=1}^m\int_{t_{k-1}}^{t_k} E_h(t_m-s)P_hf(s)\mathrm{d}s\big)$$
$$+ \big(\tau\sum_{k=1}^m E_{\tau,m}^{m-k}(A_h(t_m) - A_h(t_k))u_h^k - \sum_{k=1}^m\int_{t_{k-1}}^{t_k} E_h(t_m-s)(A_h(t_m) - A_h(s))u_h(s)ds\big)$$

$$(4.6) \qquad = \sum_{i=1}^{3} \mathrm{I}_i^m.$$

For the first two terms, there hold [12, Theorem 3.5]

$$\|\mathrm{I}_1^m\|_{L^2(\Omega)} \le c\tau t_m^{-(1-\alpha\beta/2)}\|u_0\|_{\dot H^\beta(\Omega)}, \quad \beta\in[0,2],$$
$$\|\mathrm{I}_2^m\|_{L^2(\Omega)} \le c\tau t_m^{-(1-\alpha)}\|f(0)\|_{L^2(\Omega)} + c\tau\int_0^{t_m}(t_m-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s.$$

To estimate I_3^m, we need two preliminary bounds on the operator $E_{\tau,m}^n$.

Lemma 4.2. *For the operator $E_{\tau,m}^{m-k}$ defined in (4.5), there holds for any $\beta\in[0,1]$*

$$\Big\|[\tau A_h^\beta(t_m)E_{\tau,m}^{m-k} - \int_{t_{k-1}}^{t_k} A_h^\beta(t_m)E_h(t_m-s;t_m)\,\mathrm{d}s]\Big\| \le c\tau^2(t_m-t_k+\tau)^{-(2-(1-\beta)\alpha)}.$$

Proof. First we consider the case $\beta=0$. By the definition of the operator $E_h(t;t_m)$, we have

$$\begin{aligned}\int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\,\mathrm{d}s &= \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}}(z^\alpha+A_h(t_m))^{-1}\int_{t_{k-1}}^{t_k} e^{z(t_m-s)}\,\mathrm{d}s\,\mathrm{d}z\\ &= \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{z(t_m-t_k)}z^{-1}(e^{z\tau}-1)(z^\alpha+A_h(t_m))^{-1}\,\mathrm{d}z.\end{aligned}$$

This and the defining relation (4.5) yield

$$\begin{aligned}&\tau E_{\tau,m}^{m-k} - \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\,\mathrm{d}s\\ =&\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}^\tau} e^{z(t_m-t_k)}\left[\tau(\delta_\tau(e^{-z\tau})^\alpha+A_h(t_m))^{-1} - z^{-1}(e^{z\tau}-1)(z^\alpha+A_h(t_m))^{-1}\right]\,\mathrm{d}z\\ &-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}\setminus\Gamma_{\theta,\delta}^\tau} e^{z(t_m-t_k)}z^{-1}(e^{z\tau}-1)(z^\alpha+A_h(t_m))^{-1}\,\mathrm{d}z := \mathrm{I}+\mathrm{II}.\end{aligned}$$

For $k<m$, let $\delta=(t_m-t_k+\tau)^{-1}$ and $z=s\cos\varphi+\mathrm{i}s\sin\varphi$. By Lemma 4.1 and (2.5), we obtain

$$\big\|\tau(\delta_\tau(e^{-z\tau})^\alpha+A_h(t_m))^{-1} - z^{-1}(e^{z\tau}-1)(z^\alpha+A_h(t_m))^{-1}\big\| \le c\tau^2|z|^{-\alpha+1}, \quad \forall z\in\Gamma_{\theta,\delta}^\tau.$$

Then the bound on the term I follows by

$$\|\mathrm{I}\| \le c\tau^2\int_\delta^{\frac{\pi}{\tau\sin\theta}} e^{s(t_m-t_k)\cos\theta}s^{-\alpha+1}\mathrm{d}s + c\tau^2\int_{-\theta}^{\theta} e^{\cos\varphi}\delta^{-\alpha+2}\mathrm{d}\varphi \le c\tau^2(t_m-t_k+\tau)^{\alpha-2}.$$

Similarly, Taylor expansion of $e^{z\tau}$, (2.5) and Lemma 4.1 bound the term II by

$$\begin{aligned}\|\mathrm{II}\| &\le c\tau\int_{\Gamma_{\theta,\delta}\setminus\Gamma_{\theta,\delta}^\tau}|e^{z(t_m-t_k)}||z|^{-\alpha}|\mathrm{d}z| \le c\tau\int_{\frac{\pi}{\tau\sin\theta}}^\infty e^{s(t_m-t_k)\cos\theta}s^{-\alpha}\mathrm{d}s\\ &\le c\tau^2\int_{\frac{\pi}{\tau\sin\theta}}^\infty e^{s(t_m-t_k)\cos\theta}s^{1-\alpha}\mathrm{d}s \le c\tau^2(t_m-t_k)^{-(2-\alpha)}.\end{aligned}$$

For $k=m$, there hold

$$\|\mathrm{I}\| \le c\tau^2\int_\delta^{\frac{\pi}{\tau\sin\theta}} s^{-\alpha+1}\mathrm{d}s + c\tau^2\int_{-\theta}^{\theta}\delta^{-\alpha+2}\mathrm{d}\varphi \le c\tau^\alpha,$$
$$\|\mathrm{II}\| \le c\int_{\Gamma_{\theta,\delta}\setminus\Gamma_{\theta,\delta}^\tau}|z|^{-\alpha-1}|\mathrm{d}z| \le c\int_{\frac{\pi}{\tau\sin\theta}}^\infty s^{-\alpha-1}\mathrm{d}s \le c\tau^\alpha.$$

The proof for the case $\beta=1$ is analogous, and the intermediate case $\beta\in(0,1)$ follows by interpolation. ☐

The next result gives the smoothing property of the operator $E_{\tau,m}^n$.

Lemma 4.3. *For the operator* $E^n_{\tau,m}$ *defined in* (4.5), *there holds*

$$\|A_h(t_m)E^n_{\tau,m}\| \le c(t_n+\tau)^{-1}, \quad n = 0,1,\ldots,N.$$

Proof. Upon letting $\delta = (t_n+\tau)^{-1}$ in $\Gamma^\tau_{\theta,\delta}$ and $z = s\cos\varphi + \mathrm{i}s\sin\varphi$, by (2.5) and Lemma 4.1, we have

$$\begin{aligned}\|A_h(t_m)E^n_{\tau,m}\| &= \left\| \frac{1}{2\pi\mathrm{i}} \int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n} A_h(t_m)(\delta_\tau(e^{-z\tau})^\alpha + A_h(t_m))^{-1}\,\mathrm{d}z \right\| \\ &\le c\int_{(t_n+\tau)^{-1}}^{\frac{\pi}{\tau\sin\theta}} e^{st_n\cos\theta}\mathrm{d}s + c\int_{-\theta}^{\theta} e^{\cos\varphi}(t_n+\tau)^{-1}\mathrm{d}\varphi \le c(t_n+\tau)^{-1}.\end{aligned}$$

This completes the proof of the lemma. □

Below we analyze the scheme (4.1) for the homogeneous and inhomogeneous problems separately.

4.1. Error estimate for the homogeneous problem.

First we analyze the homogeneous problem. It suffices to bound the term I_3^m in the splitting (4.6).

Lemma 4.4. *Under conditions* (1.2)-(1.3), *for* $u_0 \in L^2(\Omega)$ *and* $f = 0$, *there holds*

$$\|\mathrm{I}_3^m\|_{L^2(\Omega)} \le c\tau\log(1+t_m/\tau)t_m^{-1}\|u_0\|_{L^2(\Omega)} + c\tau\sum_{k=1}^m \|e_h^k\|_{L^2(\Omega)}.$$

Proof. Let $e_h^k = u_h^k - u_h(t_k)$, and $Q(t) = (A_h(t_m) - A_h(t))u_h(t)$. Then we split the summand of I_3^m into

$$\begin{aligned}&\tau E^{m-k}_{\tau,m}(A_h(t_m)-A_h(t_k))u_h^k - \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)Q(s)\,\mathrm{d}s \\ =&\big(\tau E^{m-k}_{\tau,m}(A_h(t_m)-A_h(t_k))e_h^k\big) + \Big(\tau E^{m-k}_{\tau,m} - \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\,\mathrm{d}s\Big)Q(t_k) \\ &+ \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)(Q(t_k)-Q(s))\,\mathrm{d}s =: \mathrm{I}_k + \mathrm{II}_k + \mathrm{III}_k.\end{aligned}$$

It remains to bound the terms I_k, II_k and III_k. First, Lemmas 4.3 and 3.1 bound the term $\|\mathrm{I}_k\|_{L^2(\Omega)}$ by:

$$\begin{aligned}\|\mathrm{I}_k\|_{L^2(\Omega)} &= \tau\|A_h(t_m)E^{m-k}_{\tau,m}(I - A_h^{-1}(t_m)A_h(t_k))e_h^k\|_{L^2(\Omega)} \\ &\le c\tau(t_m-t_k+\tau)^{-1}\|(I - A_h^{-1}(t_m)A_h(t_k))e_h^k\|_{L^2(\Omega)} \le c\tau\|e_h^k\|_{L^2(\Omega)}.\end{aligned}$$

Second, by Lemma 4.2 (with $\beta = 0$) and Remark 3.1, we bound the term II_k by

$$\begin{aligned}\|\mathrm{II}_k\|_{L^2(\Omega)} &\le \Big\|\tau E^{m-k}_{\tau,m} - \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\,\mathrm{d}s\Big\|\|Q(t_k)\|_{L^2(\Omega)} \\ &\le c\tau^2(t_m-t_k+\tau)^{\alpha-1}\|A_h(t_m)u_h(t_k)\|_{L^2(\Omega)},\end{aligned}$$

and consequently, by Theorem 3.1(i), we deduce

$$\begin{aligned}\sum_{k=1}^m \|\mathrm{II}_k\|_{L^2(\Omega)} &\le c\tau^2\sum_{k=1}^m (t_m-t_k+\tau)^{\alpha-1}\|A_h(t_m)u_h(t_k)\|_{L^2(\Omega)} \\ &\le c\tau^2\|u_0\|_{L^2(\Omega)}\sum_{k=1}^m (t_m-t_k+\tau)^{\alpha-1}t_k^{-\alpha} \le c\tau\|u_0\|_{L^2(\Omega)},\end{aligned}$$

where the last line follows from the inequality

$$\tau\sum_{k=1}^m (t_m-t_k+\tau)^{\alpha-1}t_k^{-\alpha} \le c.$$

Last, for the third term III_k, with $k=1$, by Lemma 3.1, we have

$$\|\mathrm{III}_1\|_{L^2(\Omega)} \le \int_0^\tau \|E_h(t_m-s;t_m)Q(\tau)\|_{L^2(\Omega)}\,\mathrm{d}s + \int_0^\tau \|E_h(t_m-s;t_m)Q(s)\|_{L^2(\Omega)}\,\mathrm{d}s$$

$$
\begin{aligned}
&= \int_0^\tau \|A_h(t_m)E_h(t_m-s;t_m)A_h(t_m)^{-1}Q(\tau)\|_{L^2(\Omega)}\,\mathrm{d}s \\
&\quad + \int_0^\tau \|A_h(t_m)E_h(t_m-s;t_m)A_h(t_m)^{-1}Q(s)\|_{L^2(\Omega)}\,\mathrm{d}s \\
&\leq c\int_0^\tau (t_m-s)^{-1}((t_m-\tau)+(t_m-s))\,\mathrm{d}s\|u_0\|_{L^2(\Omega)} \leq c\tau\|u_0\|_{L^2(\Omega)}.
\end{aligned}
$$

Meanwhile, for $k>1$, we further split the term III_k into

$$
\begin{aligned}
\mathrm{III}_k &= \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\int_{t_k}^s Q'(\xi)\,\mathrm{d}\xi\,\mathrm{d}s \\
&= \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\int_{t_k}^s (A_h(t_m)-A_h(\xi))u_h'(\xi)\,\mathrm{d}\xi\,\mathrm{d}s \\
&\quad - \int_{t_{k-1}}^{t_k} E_h(t_m-s;t_m)\int_{t_k}^s A_h'(\xi)u_h(\xi)\,\mathrm{d}\xi\,\mathrm{d}s =: \mathrm{III}_{k,1}+\mathrm{III}_{k,2}.
\end{aligned}
$$

By Lemmas 3.1 and 2.2(ii), the term $\mathrm{III}_{k,1}$ for any $k>1$ can be bounded by

$$
\begin{aligned}
\|\mathrm{III}_{k,1}\|_{L^2(\Omega)} &\leq \int_{t_{k-1}}^{t_k} \|A_h(t_m)E_h(t_m-s;t_m)\|\int_s^{t_k} \|(I-A_h(t_m)^{-1}A_h(\xi))u_h'(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s \\
&\leq c\int_{t_{k-1}}^{t_k} (t_m-s)^{-1}\int_s^{t_k} (t_m-\xi)\|u_h'(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s \\
&\leq c\int_{t_{k-1}}^{t_k} (t_m-s)^{-1}\int_s^{t_k} (t_m-\xi)\xi^{-1}\|u_0\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s,
\end{aligned}
$$

where the last step is due to Theorem 3.1(i). Now we note the elementary inequality

$$
\begin{aligned}
\int_{t_{k-1}}^{t_k} (t_m-s)^{-1}\int_s^{t_k} (t_m-\xi)\xi^{-1}\mathrm{d}\xi\mathrm{d}s &= \int_{t_{k-1}}^{t_k} (t_m-\xi)\xi^{-1}\int_{t_{k-1}}^{\xi} (t_m-s)^{-1}\mathrm{d}s\mathrm{d}\xi \\
&\leq \int_{t_{k-1}}^{t_k} (t_m-\xi)(t_m-\xi)^{-1}\tau\xi^{-1}\mathrm{d}\xi = \tau\int_{t_{k-1}}^{t_k} \xi^{-1}\mathrm{d}\xi.
\end{aligned}
$$

Consequently,

$$
\|\mathrm{III}_{k,1}\|_{L^2(\Omega)} \leq c\tau\int_{t_{k-1}}^{t_k} \xi^{-1}\,\mathrm{d}\xi\|u_0\|_{L^2(\Omega)},
$$

and

$$
\sum_{k=2}^m \|\mathrm{III}_{k,1}\|_{L^2(\Omega)} \leq c\tau\int_\tau^{t_m} \xi^{-1}\,\mathrm{d}\xi\|u_0\|_{L^2(\Omega)} = c\tau\log(t_m/\tau).
$$

Similarly, by Theorem 3.1(i), the term $\mathrm{III}_{k,2}$ for any $k>1$ is bounded by

$$
\begin{aligned}
\|\mathrm{III}_{k,2}\|_{L^2(\Omega)} &= \int_{t_{k-1}}^{t_k} \|E_h(t_m-s;t_m)\|\int_s^{t_k} \|A_h'(\xi)u_h(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s \\
&\leq c\int_{t_{k-1}}^{t_k} (t_m-s)^{\alpha-1}\int_s^{t_k} \xi^{-\alpha}\,\mathrm{d}\xi\,\mathrm{d}s\|u_0\|_{L^2(\Omega)} \\
&\leq c\tau\int_{t_{k-1}}^{t_k} (t_m-s)^{\alpha-1}s^{-\alpha}\,\mathrm{d}s\|u_0\|_{L^2(\Omega)},
\end{aligned}
$$

where the last line follows from the inequality $\int_s^{t_k}\xi^{-\alpha}\mathrm{d}\xi \leq s^{-\alpha}\int_s^{t_k}\mathrm{d}\xi \leq s^{-\alpha}\tau$. Thus,

$$
\sum_{k=1}^m \|\mathrm{III}_{k,2}\|_{L^2(\Omega)} \leq c\tau\int_0^{t_m} (t_m-s)^{\alpha-1}s^{-\alpha}\,\mathrm{d}s\|u_0\|_{L^2(\Omega)} \leq c\|u_0\|_{L^2(\Omega)}.
$$

Hence, there holds

$$\sum_{k=1}^{m} \|\mathrm{III}_k\|_{L^2(\Omega)} \le c\tau(1+\log(t_m/\tau))\|u_0\|_{L^2(\Omega)}.$$

Combining the preceding estimates completes the proof of the lemma. □

Now we can state an error estimate for the homogeneous problem.

Theorem 4.1. *Under conditions* (1.2)-(1.3), $u_0 \in L^2(\Omega)$ *and* $f=0$, *there holds*

$$\|u_h^m - u_h(t_m)\|_{L^2(\Omega)} \le c\tau t_m^{-1}\log(1+t_m/\tau))\|u_0\|_{L^2(\Omega)}.$$

Proof. It follows from Lemma 4.4 that

$$\|e_h^m\|_{L^2(\Omega)} \le c\tau(t_m^{-1}+\log(1+t_m/\tau))\|u_0\|_{L^2(\Omega)} + c\tau\sum_{k=1}^{m}\|e_h^k\|_{L^2(\Omega)}.$$

The desired estimate follows from a variant of the discrete Gronwall's inequality [34, p. 258]. □

Remark 4.1. *The logarithmic factor* $\log(1+t_m/\tau)$ *is also present for the BE method for standard parabolic problems with a time-dependent diffusion coefficient* [24, Theorem 2, p. 95]. *For* $u_0 \in \dot H^\beta(\Omega)$, $\beta \in (0,2]$, *it may be improved:*

$$\|u_h^m - u_h(t_m)\|_{L^2(\Omega)} \le c\tau t_m^{-(1-\beta\alpha/2)}\|u_0\|_{\dot H^\beta(\Omega)}.$$

In fact, the argument of Lemma 4.4 (*together with Theorem 3.1(i)*) *implies*

$$\begin{aligned}
\|\mathrm{III}_{k,1}\|_{L^2(\Omega)} &\le c\int_{t_{k-1}}^{t_k}(t_m-s)^{-1}\int_s^{t_k}(t_m-\xi)\|u_h'(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s\\
&\le c\int_{t_{k-1}}^{t_k}(t_m-s)^{-1}\int_s^{t_k}(t_m-\xi)\xi^{-(1-\alpha\beta/2)}\|u_0\|_{\dot H^\beta(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s\\
&\le c\tau\int_{t_{k-1}}^{t_k}\xi^{-(1-\beta\alpha/2)}\,\mathrm{d}\xi\|u_0\|_{\dot H^\beta(\Omega)},
\end{aligned}$$

and

$$\begin{aligned}
\|\mathrm{III}_{k,2}\|_{L^2(\Omega)} &= \int_{t_{k-1}}^{t_k}\|E_h(t_m-s;t_m)\|\int_s^{t_k}\|A_h'(\xi)u_h(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s\\
&\le c\int_{t_{k-1}}^{t_k}(t_m-s)^{\alpha-1}\int_s^{t_k}\xi^{-(1-\beta/2)\alpha}\,\mathrm{d}\xi\,\mathrm{d}s\|u_0\|_{\dot H^\beta(\Omega)}\\
&\le c\tau\int_{t_{k-1}}^{t_k}(t_m-s)^{\alpha-1}s^{-(1-\beta/2)\alpha}\,\mathrm{d}s\|u_0\|_{\dot H^\beta(\Omega)}.
\end{aligned}$$

Thus, for any $\beta\in(0,2]$, *there holds*

$$\sum_{k=1}^{m}\left\|\int_{t_{k-1}}^{t_k}E(t_m-s;t_m)(Q(t_k)-Q(s))\,\mathrm{d}s\right\| \le c\tau\|u_0\|_{\dot H^\beta(\Omega)}.$$

Then the desired estimate follows by repeating the argument for Theorem 4.1.

4.2. **Error estimate for the inhomogeneous problem.** Now we give the temporal discretization error for the inhomogeneous problem.

Theorem 4.2. *Under conditions* (1.2)-(1.3), $u_0=0$, $f\in C([0,T];L^2(\Omega))$ *and* $\int_0^t(t-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s<\infty$ *for any* $0<t\le T$, *there holds*

$$\|u_h^m-u_h(t_m)\|_{L^2(\Omega)} \le c\tau\Big(t_m^{-(1-\alpha)}\|f(0)\|_{L^2(\Omega)} + \int_0^{t_m}(t_m-s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s\Big).$$

Proof. The argument is similar to Theorem 4.1, and thus we only sketch the main steps. It suffices to bound the terms II_k and III_k in Lemma 4.4. By Theorem 3.1(iii), Remark 3.1 and Lemma 4.2 (with $\beta = 1/2$), the following estimate holds:

$$\begin{aligned}\sum_{k=1}^m \|\mathrm{II}_k\|_{L^2(\Omega)} &\le c\tau^2 \sum_{k=1}^m (t_m - t_k + \tau)^{\alpha/2-1} \|A_h^{1/2}(t_m) u_h(t_k)\|_{L^2(\Omega)} \\ &\le c\tau^2 \sum_{k=1}^m (t_m - t_k + \tau)^{\alpha/2-1} \|f\|_{L^\infty(0,t_m;L^2(\Omega))} \le c\tau \|f\|_{L^\infty(0,t_m;L^2(\Omega))}.\end{aligned}$$

Meanwhile, upon noting $t_m \le T$, we have

$$\begin{aligned}\|f\|_{L^\infty(0,t_m;L^2(\Omega))} &\le \|f(0)\|_{L^2(\Omega)} + \int_0^{t_m} \|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s \\ &\le \|f(0)\|_{L^2(\Omega)} + c_T \int_0^{t_m} (t_m - s)^{\alpha-1} \|f'(s)\|_{L^2(\Omega)}\,\mathrm{d}s.\end{aligned}$$

Next, the two terms $\mathrm{III}_{k,1}$ and $\mathrm{III}_{k,2}$ can be bounded respectively by

$$\begin{aligned}\|\mathrm{III}_{k,1}\|_{L^2(\Omega)} &\le c\int_{t_{k-1}}^{t_k} (t_m - s)^{-1} \int_s^{t_k} (t_m - \xi) \|u_h'(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s \\ &= c\int_{t_{k-1}}^{t_k} (t_m - \xi)\|u_h'(\xi)\|_{L^2(\Omega)} \int_{t_{k-1}}^{\xi} (t_m - s)^{-1}\,\mathrm{d}s\,\mathrm{d}\xi \\ &\le c\tau \int_{t_{k-1}}^{t_k} \|u_h'(\xi)\|_{L^2(\Omega)}\mathrm{d}\xi,\end{aligned}$$

and

$$\begin{aligned}\|\mathrm{III}_{k,2}\|_{L^2(\Omega)} &\le \int_{t_{k-1}}^{t_k} \|E_h(t_m - s; t_m)\| \int_s^{t_k} \|A_h'(\xi) u_h(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s \\ &\le c\int_{t_{k-1}}^{t_k} (t_m - s)^{\alpha-1} \int_s^{t_k} \|A_h'(\xi) u_h(\xi)\|_{L^2(\Omega)}\,\mathrm{d}\xi\,\mathrm{d}s \\ &= c\int_{t_{k-1}}^{t_k} \|A_h'(\xi) u_h(\xi)\|_{L^2(\Omega)} \int_{t_{k-1}}^{\xi} (t_m - s)^{\alpha-1}\,\mathrm{d}s\mathrm{d}\xi \\ &\le c\tau \int_{t_{k-1}}^{t_k} (t_m - \xi)^{\alpha-1} \|A_h'(\xi) u_h(\xi)\|_{L^2(\Omega)}\mathrm{d}\xi.\end{aligned}$$

Then by Theorem 3.1(ii), we have

$$\begin{aligned}\sum_{k=1}^m \|\mathrm{III}_{k,1}\|_{L^2(\Omega)} &\le c\tau \int_0^{t_m} \|u_h'(s)\|_{L^2(\Omega)}\mathrm{d}s \\ &\le c\tau\Big(\int_0^{t_m} \|f(0)\|_{L^2(\Omega)}\mathrm{d}s + \int_0^{t_m}\int_0^s (s-\xi)^{\alpha-1}\|f'(\xi)\|_{L^2(\Omega)}\mathrm{d}\xi\mathrm{d}s\Big) \\ &= c\tau\Big(\int_0^{t_m} \|f(0)\|_{L^2(\Omega)}\mathrm{d}s + \int_0^{t_m} \|f'(\xi)\|_{L^2(\Omega)} \int_\xi^{t_m} (s-\xi)^{\alpha-1}\mathrm{d}s\mathrm{d}\xi\Big) \\ &\le c\tau\Big(\|f(0)\|_{L^2(\Omega)} + \int_0^{t_m} (t_m - \xi)^{\alpha-1}\|f'(\xi)\|_{L^2(\Omega)}\mathrm{d}\xi\Big),\end{aligned}$$

and similarly, by Theorem 3.1(ii), we deduce

$$\begin{aligned}\sum_{k=1}^m \|\mathrm{III}_{k,2}\|_{L^2(\Omega)} &\le c\tau \int_0^{t_m} (t_m - s)^{\alpha-1} \|A_h'(s) u_h(s)\|_{L^2(\Omega)}\mathrm{d}s \\ &\le c\tau\Big(\|f(0)\|_{L^2(\Omega)} + \int_0^{t_m} (t_m - s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)}\mathrm{d}s\Big).\end{aligned}$$

These estimates together with discrete Gronwall's inequality complete the proof. □

Remark 4.2. *The proof techniques in this section apply to other first-order methods, e.g., L1 scheme* [21]*, and similar error estimates can also be derived for these methods.*

Remark 4.3. *We briefly comment on the dependence of the constant c in error estimates on the fractional order α. At a few occasions, it can blow up as $\alpha \to 1^-$; see e.g., $\mathrm{I}_3(t_0)$ in Lemma 3.3, $\mathrm{I}_{4,2}(t_0)$ in Lemma 3.4 and $\mathrm{III}_{k,2}$ in Lemma 4.4. This phenomenon does not fully agree with the results for the continuous model. Such a blowup phenomenon appears also in some existing error analysis; see, e.g.,* [28, eq. (2.2)] *and* [32, Lemma 4.3]*, and it is of interest to further refine the estimates to fill in the gap.*

5. Numerical results

Now we present numerical examples to verify the theoretical results in Sections 3 and 4. We consider problem (1.1) with a time-dependent elliptic operator $A(t) = -(2+\cos(t))\Delta$ on the domain $\Omega = (0,1)$ and the following two sets of problem data:

(a) $u_0(x) = x^{-1/4} \in H^{1/4-\epsilon}(\Omega)$ with $\epsilon \in (0,1/4)$ and $f \equiv 0$.
(b) $u_0(x) = 0$ and $f = e^t(1+\chi_{(0,1/2)}(x))$.

Unless otherwise specified, the final time T is fixed at $T = 1$.

We divide the domain Ω into M subintervals of equal length $h = 1/M$. The numerical solutions are computed using the Galerkin FEM in space, and the backward Euler (BE) CQ or L1 scheme in time. To evaluate the convergence, we compute the spatial error e_s and temporal error e_t, respectively, defined by

$$e_s(t_N) = \|u_h(t_N) - u(t_N)\|_{L^2(\Omega)} \quad \text{and} \quad e_t(t_N) = \|u_h^N - u_h(t_N)\|_{L^2(\Omega)}.$$

Since the exact solution is unavailable, we compute reference solutions on a finer mesh: for the error e_s, we take the time step $\tau = 1/10000$ and mesh size $h = 1/1280$, and for the error e_t, take $h = 1/100$ and $\tau = 1/10000$, unless otherwise specified.

First we examine the spatial convergence of the semidiscrete Galerkin scheme (3.2). The spatial errors for case (a) are shown in Table 1, which indicates a steady $O(h^2)$ rate for the semidiscrete scheme (3.2), just as predicted by Theorem 3.2. The $O(h^2)$ rate holds for all three fractional orders and different terminal times. Since the initial data is nonsmooth, the spatial error $e_s(t_N)$ decreases with the time t_N, which is in good agreement with the regularity result in Theorem 2.2. To further illuminate the precise dependence of the spatial error $e_s(t_N)$ on t_N, in Table 2, we present the error e_s as the time $t_N \to 0$ for case (a). By repeating the argument for Theorem 3.2, there holds $e_s(t_N) \le c t_N^{-(2-\beta)\alpha/2} h^2 \|v\|_{\dot H^\beta(\Omega)}$, $0 \le \beta \le 2$. For case (a), this estimate predicts an exponent $7\alpha/8$ for the dependence on the time t_N, which gives the numbers shown in the bracket in Table 2. Table 2 indicates that the empirical rate agrees excellently with the predicted one, fully confirming the analysis. Similar observations hold also for the numerical results for the inhomogeneous problem in case (b), cf. Table 3. These results fully support the error analysis of the semidiscrete scheme in Section 3.

Table 1. Spatial errors e_s for example (a) with $\tau = 1/10000$ and $h = 1/M$.

T	α \ N	10	20	40	80	160	rate
1	0.25	1.44e-5	3.62e-6	9.06e-7	2.26e-7	5.62e-8	2.00 (2.00)
	0.50	1.02e-5	2.56e-6	6.40e-7	1.60e-7	3.97e-8	2.00 (2.00)
	0.75	5.18e-6	1.30e-6	3.25e-7	8.12e-8	2.02e-8	2.00 (2.00)
10^{-3}	0.25	6.26e-5	1.57e-5	3.93e-6	9.80e-7	2.44e-7	2.01 (2.00)
	0.50	2.12e-4	5.31e-5	1.33e-5	3.32e-6	8.24e-7	2.01 (2.00)
	0.75	5.99e-4	1.50e-4	3.75e-5	9.36e-6	2.33e-6	2.01 (2.00)

Next we turn to the temporal convergence, and present numerical results for both BE and L1 schemes, cf. Remark 4.2. The temporal errors e_t for case (a) at two time instances are given in Table 4, which

Table 2. Spatial errors e_s for example (a) with $h = 1/200$ and $N = 10000$, at $T = 10^{-k}$.

α \ k	2	3	4	5	6	7	rate
0.25	2.40e-6	4.04e-6	6.58e-6	1.05e-5	1.65e-5	2.66e-5	0.21 (0.22)
0.5	5.28e-6	1.31e-5	3.36e-5	9.02e-5	2.40e-4	6.40e-4	0.43 (0.44)
0.75	9.95e-6	3.90e-5	1.70e-4	7.45e-4	3.26e-3	1.39e-2	0.64 (0.66)

Table 3. Spatial errors e_s for example (b) at $T = 1$ with $\tau = 1/10000$ and $h = 1/M$.

α \ M	10	20	40	80	160	rate
0.25	2.03e-4	5.06e-5	1.27e-5	3.16e-6	7.85e-7	2.01 (2.00)
0.50	2.08e-4	5.19e-5	1.30e-5	3.24e-6	8.04e-7	2.01 (2.00)
0.75	2.13e-4	5.32e-5	1.33e-5	3.32e-6	8.25e-7	2.01 (2.00)

indicate an $O(\tau)$ convergence rate for both time stepping schemes. Further, the accuracy of both schemes is largely comparable. The convergence is very steady for both schemes, and the convergence rate is independent of the fractional order α and the final time t_N (so long as it is fixed). Further, it is observed that the error e_t decreases with the time t_N. To show the dependence of the temporal error $e_t(t_N)$ with the time t_N, in Table 5, we present $e_t(t_N)$ as the time t_N tends to zero. In view of Remark 4.1, there holds $e_t(t_N) \leq c\tau t_N^{-(1-\beta\alpha/2)}\|u_0\|_{\dot{H}^\beta(\Omega)}$, $0 < \beta \leq 2$. This estimate predicts a decay $O(N^{-\alpha/8})$ for case (a), which agrees excellently with the empirical rate (in the bracket) in Table 5, thereby confirming the sharpness of the error estimate. These observations hold also for the inhomogeneous problem in case (b), cf. Table 6. These numerical results fully support the error analysis of the fully discrete scheme in Section 4.

Table 4. Temporal errors e_t for example (a) with $h = 1/100$ and $\tau = T/N$.

T	method	α \ N	100	200	400	800	1600	rate
1	BE	0.25	5.43e-5	2.71e-5	1.35e-5	6.76e-6	3.38e-6	1.00 (1.00)
		0.50	9.49e-5	4.73e-5	2.36e-5	1.18e-5	5.90e-6	1.00 (1.00)
		0.75	9.01e-5	4.49e-5	2.24e-5	1.12e-5	5.59e-6	1.00 (1.00)
	L1	0.25	4.35e-5	2.17e-5	1.08e-5	5.41e-6	2.70e-6	1.00 (1.00)
		0.50	6.33e-5	3.15e-5	1.57e-5	7.84e-6	3.92e-6	1.00 (1.00)
		0.75	5.12e-5	2.54e-5	1.26e-5	6.29e-6	3.14e-6	1.01 (1.00)
10^{-3}	BE	0.25	2.00e-4	9.99e-5	4.99e-5	2.49e-5	1.25e-5	1.00 (1.00)
		0.50	8.16e-4	4.08e-4	2.04e-4	1.02e-4	5.10e-5	1.00 (1.00)
		0.75	7.58e-4	3.79e-4	1.89e-4	9.46e-5	4.73e-5	1.00 (1.00)
	L1	0.25	1.69e-4	8.43e-5	4.21e-5	2.10e-5	1.05e-5	1.00 (1.00)
		0.50	8.08e-4	3.99e-4	1.98e-4	9.84e-5	4.90e-5	1.01 (1.00)
		0.75	8.28e-4	4.11e-4	2.04e-4	1.02e-4	5.07e-5	1.01 (1.00)

Table 5. Temporal errors e_t for example (a) with $\alpha = 0.5$, $h = 10^{-3}$ and $N = 5$, at $T = 10^{-k}$.

α \ k		3	4	5	6	7	8	rate
0.5	BE	1.65e-2	1.06e-2	8.79e-3	7.45e-3	6.33e-3	5.39e-3	0.07 (0.06)
	L1	1.91e-2	1.41e-2	1.06e-2	8.95e-3	7.59e-3	6.46e-3	0.07 (0.06)
0.8	BE	1.61e-2	1.23e-2	9.52e-3	7.44e-3	5.84e-3	4.56e-3	0.11 (0.10)
	L1	1.83e-2	1.40e-2	1.08e-2	8.46e-3	6.64e-3	5.19e-3	0.11 (0.10)

TABLE 6. Temporal errors e_t for example (b) at $T = 1$ with $h = 1/100$ and $\tau = T/N$.

method	α \ N	100	200	400	800	1600	rate
BE	0.25	3.26e-6	1.63e-6	8.15e-7	4.07e-7	2.04e-7	1.00 (1.00)
	0.50	4.76e-6	2.37e-6	1.18e-6	5.92e-7	2.96e-7	1.00 (1.00)
	0.75	2.76e-6	1.37e-6	6.84e-7	3.41e-7	1.71e-7	1.00 (1.00)
L1	0.25	2.33e-6	1.17e-6	5.85e-7	2.93e-7	1.47e-7	1.00 (1.00)
	0.50	3.25e-6	1.64e-6	8.29e-7	4.18e-7	2.10e-7	0.99 (1.00)
	0.75	1.85e-6	9.89e-7	5.22e-7	2.72e-7	1.41e-7	0.94 (1.00)

Acknowledgements

The research of B. Li is partially supported by a Hong Kong RGC grant (Project No. 15300817), and that of Z. Zhou by a start-up grant from the Hong Kong Polytechnic University and Hong Kong RGC grant No. 25300818.

References

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

[2] A. A. Alikhanov. A new difference scheme for the time fractional diffusion equation. *J. Comput. Phys.*, 280:424–438, 2015.

[3] E. G. Bajlekova. *Fractional Evolution Equations in Banach Spaces.* PhD thesis, Eindhoven University of Technology, Eindhoven, Netherlands, 2001.

[4] S. D. Eidelman and A. N. Kochubei. Cauchy problem for fractional diffusion equations. *J. Diff. Eq.*, 199(2):211–255, 2004.

[5] C. M. Elliott and S. Larsson. Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation. *Math. Comp.*, 58(198):603–630, S33–S36, 1992.

[6] H. Fujita and T. Suzuki. Evolution problems. In *Handbook of Numerical Analysis, Vol. II*, pages 789–928. North-Holland, Amsterdam, 1991.

[7] G.-H. Gao, Z.-Z. Sun, and H.-W. Zhang. A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications. *J. Comput. Phys.*, 259:33–50, 2014.

[8] D. Henry. *Geometric Theory of Semilinear Parabolic Equations.* Springer-Verlag, Berlin-New York, 1981.

[9] B. Jin, R. Lazarov, J. Pasciak, and Z. Zhou. Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion. *IMA J. Numer. Anal.*, 35(2):561–582, 2015.

[10] B. Jin, R. Lazarov, and Z. Zhou. Numerical methods for time-fractional evolution equations with nonsmooth data: a concise overview. Preprint, arXiv:1805.11309.

[11] B. Jin, R. Lazarov, and Z. Zhou. Error estimates for a semidiscrete finite element method for fractional order parabolic equations. *SIAM J. Numer. Anal.*, 51(1):445–466, 2013.

[12] B. Jin, R. Lazarov, and Z. Zhou. Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. *SIAM J. Sci. Comput.*, 38(1):A146–A170, 2016.

[13] B. Jin, B. Li, and Z. Zhou. Numerical analysis of nonlinear subdiffusion equations. *SIAM J. Numer. Anal.*, 56(1):1–23, 2018.

[14] S. Karaa. Semidiscrete finite element analysis of time fractional parabolic problems: a unified approach. *SIAM J. Numer. Anal.*, 56(3):1673–1692, 2018.

[15] S. Karaa, K. Mustapha, and A. K. Pani. Finite volume element method for two-dimensional fractional subdiffusion problems. *IMA J. Numer. Anal.*, 37(2):945–964, 2017.

[16] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and Applications of Fractional Differential Equations.* Elsevier Science B.V., Amsterdam, 2006.

[17] A. Kubica and M. Yamamoto. Initial-boundary value problems for fractional diffusion equations with time-dependent coeffcients. *Fract. Calc. Appl. Anal.*, 21(2):276–311, 2018.

[18] P. C. Kunstmann, B. Li, and C. Lubich. Runge-Kutta time discretization of nonlinear parabolic equations studied via discrete maximal parabolic regularity. *Found. Comput. Math.*, 2017, DOI: 10.1007/s10208-017-9364-x.

[19] H. Lee, J. Lee, and D. Sheen. Laplace transform method for parabolic problems with time-dependent coefficients. *SIAM J. Numer. Anal.*, 51(1):112–125, 2013.

[20] B. Li and W. Sun. Regularity of the diffusion-dispersion tensor and error analysis of FEMs for a porous media flow. *SIAM J. Numer. Anal.*, 53:1418–1437, 2015.

[21] Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.*, 225(2):1533–1552, 2007.

[22] C. Lubich. Discretized fractional calculus. *SIAM J. Math. Anal.*, 17(3):704–719, 1986.

[23] Y. Luchko. Maximum principle for the generalized time-fractional diffusion equation. *J. Math. Anal. Appl.*, 351(1):218–223, 2009.
[24] M. Luskin and R. Rannacher. On the smoothing property of the Galerkin method for parabolic equations. *SIAM J. Numer. Anal.*, 19(1):93–113, 1982.
[25] W. McLean. Regularity of solutions to a time-fractional diffusion equation. *ANZIAM J.*, 52(2):123–138, 2010.
[26] W. McLean and K. Mustapha. Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. *Numer. Algorithms*, 52(1):69–88, 2009.
[27] W. McLean and K. Mustapha. Time-stepping error bounds for fractional diffusion problems with non-smooth initial data. *J. Comput. Phys.*, 293:201–217, 2015.
[28] K. Mustapha. FEM for time-fractional diffusion equations, novel optimal error analyses. *Math. Comp.*, 87(313):2259–2272, 2018.
[29] K. Sakamoto and M. Yamamoto. 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.
[30] P. Sammon. Fully discrete approximation methods for parabolic problems with nonsmooth initial data. *SIAM J. Numer. Anal.*, 20(3):437–470, 1983.
[31] G. Savaré. $A(\Theta)$-stable approximations of abstract Cauchy problems. *Numer. Math.*, 65(3):319–335, 1993.
[32] M. Stynes, E. O'Riordan, and J. L. Gracia. Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. *SIAM J. Numer. Anal.*, 55(2):1057–1079, 2017.
[33] Z.-Z. Sun and X. Wu. A fully discrete difference scheme for a diffusion-wave system. *Appl. Numer. Math.*, 56(2):193–209, 2006.
[34] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin, second edition, 2006.
[35] Y. Yan, M. Khan, and N. J. Ford. 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.
[36] S. B. Yuste and L. Acedo. An explicit finite difference method and a new von Neumann-type stability analysis for fractional diffusion equations. *SIAM J. Numer. Anal.*, 42(5):1862–1874, 2005.
[37] R. Zacher. A De Giorgi–Nash type theorem for time fractional diffusion equations. *Math. Ann.*, 356(1):99–146, 2013.

Department of Computer Science, University College London, Gower Street, London, WC1E 2BT, UK.
E-mail address: b.jin@ucl.ac.uk,bangti.jin@gmail.com

Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
E-mail address: bygli@polyu.edu.hk

Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
E-mail address: zhizhou@polyu.edu.hk