

Noname manuscript No.  
(will be inserted by the editor)

# Subdiffusion with Time-Dependent Coefficients: Improved Regularity and Second-Order Time Stepping

Bangti Jin · Buyang Li · Zhi Zhou

Received: date / Accepted: date

**Abstract** This article concerns second-order time discretization of subdiffusion equations with time-dependent diffusion coefficients. High-order differentiability and regularity estimates are established for subdiffusion equations with time-dependent coefficients. Using these regularity results and a perturbation argument of freezing the diffusion coefficient, we prove that the convolution quadrature generated by the second-order backward differentiation formula, with proper correction at the first time step, can achieve second-order convergence for both nonsmooth initial data and incompatible source term. Numerical experiments are consistent with the theoretical results.

**Keywords** subdiffusion, time-dependent coefficient, second-order backward differentiation formula, convolution quadrature, perturbation argument, error estimate

---

The work of B. Jin is partially supported by UK EPSRC EP/T000864/1, that of B. Li is supported by 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 (Project No. 25300818).

---

B. Jin  
Department of Computer Science, University College London, Gower Street, London, WC1E 6BT, UK  
E-mail: b.jin@ucl.ac.uk, bangti.jin@gmail.com

B. Li  
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong.  
E-mail: buyang.li@polyu.edu.hk, libuyang@gmail.com

Z. Zhou  
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong.  
E-mail: zhizhou@polyu.edu.hk, zhizhou0125@gmail.com

## 1 Introduction

Let  $\Omega \subset \mathbb{R}^d$  ( $d = 1, 2, 3$ ) be a convex polygonal domain with a boundary  $\partial\Omega$ . Consider the following subdiffusion equation

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

where  $a(x, t) : \Omega \times (0, T) \rightarrow \mathbb{R}^{d \times d}$  is a positive definite matrix-valued function,  $f$  and  $u_0$  are the source term and initial value, respectively, and

$$\partial_t^\alpha u(x, t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \partial_s u(x, s) ds, \quad (1.2)$$

denotes the Caputo fractional time derivative of order  $\alpha \in (0, 1)$  [19, p. 70].

In recent years, there has been a growing interest in the mathematical and numerical analysis of subdiffusion models due to their diverse applications in describing subdiffusion processes arising from physics, engineering, biology and finance. In a subdiffusion process, the mean squared particle displacement grows only sublinearly with time, instead of growing linearly with time as in a normal diffusion process. At a microscopic level, such processes can be adequately described by continuous time random walk, and accordingly, at a macroscopical level, the probability density function of the particle appearing at certain time  $t$  and location  $x$  is described by a subdiffusion model of the form (1.1). We refer interested readers to [29, 30] for a long list of applications arising in biology and physics. In the physical literature, a time-dependent diffusion coefficient is often employed to study complex systems, e.g., turbulence system [13, 20, 9] and cooling process in geology [6, 10]; see also [32, 11] for its connection with birth-death processes.

The numerical analysis of the subdiffusion problem has been the topic of many recent investigations. In particular, a large number of time-stepping schemes for approximating the Caputo derivative have been developed. The most popular ones include convolution quadrature [5, 14, 16, 2], piecewise polynomial approximation [36, 23, 1, 27], and discontinuous Galerkin method [28]. For a given smooth source term  $f$  and initial value  $u_0$ , these schemes generally exhibit only first-order convergence due to the inherent weak singularity of the solution at  $t = 0$ . If the solution  $u$  is smooth, then higher-order convergence may be achieved, otherwise some modifications of the schemes [5, 14, 16] or locally refined meshes [28, 35, 22] (see also [3] for related works in the context of Volterra integral equations) can be used; see the recent survey [15] for further references. All these works focus on subdiffusion with a time-independent coefficient, i.e.,  $a(x, t) \equiv a(x)$ .

When the diffusion coefficient  $a(x, t)$  is time-dependent, the analysis of regularity of solutions and the development and convergence analysis of numerical schemes are rather limited, despite its obvious practical importance. Many existing analytical techniques, e.g., Laplace transform and separation of

variables, are not directly applicable, due to the time-dependency of the coefficient  $a(x, t)$ . Kubica and Yamamoto [21] proved the existence and uniqueness of a weak solution, and also several regularity results. In this work, we present new regularity estimates in Theorems 1 and 2. For example, for  $u_0 \in L^2(\Omega)$  and  $f \equiv 0$ , under suitable conditions on  $a(x, t)$ , there holds  $\|\frac{d^k}{dt^k}(t^k u(t))\|_{L^2(\Omega)} \leq c\|u_0\|_{L^2(\Omega)}$ . Such an estimate provides one crucial tool for the error analysis of high-order time-stepping schemes.

So far there are very few works on the numerical approximation of the model (1.1) [31, 18]. Mustapha [31] analyzed a spatially semidiscrete Galerkin finite element method (FEM) for the homogeneous problem, and showed optimal order convergence by a novel energy argument. In essence, the approach extends the argument in [26] for standard parabolic problems to the fractional case. In the authors' prior work [18], we developed a different approach to analyze the spatially semidiscrete Galerkin scheme, as well as a fully discrete scheme based on convolution quadrature (CQ) generated by backward Euler method (and L1 scheme), and showed optimal order convergence rates for both semidiscrete and fully discrete schemes (up to a logarithmic factor), based on a perturbation argument and new regularity results. However, the discrete scheme in [18] is only first order accurate in time. To the best of our knowledge, there is no proven second- or higher-order accurate time-stepping scheme for the subdiffusion model with a time-dependent coefficient and nonsmooth problem data in the literature. This contrasts sharply with the case of time-independent elliptic operators, for which there are several strategies for devising high-order schemes, e.g., initial correction [16]. These observations motivate the present work.

In this article, we propose a second-order time-stepping scheme for problem (1.1) with nonsmooth initial data and incompatible source term. It is based on the CQ generated by the second-order backward differentiation formula (BDF2), with suitable correction at the first step. The correction is inspired by the recent works [5, 14, 16] and essential for restoring the second-order convergence. Further, we present a complete error analysis in Section 4, and prove a convergence rate  $O(\tau^2)$  with  $\tau$  being the time stepsize, for any fixed  $t_n > 0$ , of the scheme for both nonsmooth initial data and incompatible source term. The error analysis relies heavily on new temporal regularity results for the model (1.1) in Section 3 and a refined perturbation argument, which substantially extends the prior work [18]. Specifically, the error analysis relies on suitable nonstandard bounds for problem data in the space  $\dot{H}^{-\gamma}(\Omega)$  (cf., Lemma 4 and Theorem 5), and perturbation estimates at both  $t = 0$  and  $t = t_m$  (cf., the proof of Lemma 7), which are substantially different from the one in [18] which only requires estimates at  $t = t_m$  for problem data in  $L^2(\Omega)$ . The new scheme, regularity results and time discretization errors represent the main contributions of this work.

In the context of the standard parabolic counterpart with  $L^2(\Omega)$ -initial data and zero forcing term, Luskin and Rannacher [26] analyzed a fully discrete scheme based on Galerkin FEM in space and the backward Euler method in time, and proved a first-order temporal convergence. Somewhat surprisingly,

Sammon [34] proved that for standard parabolic problems with  $L^2(\Omega)$  initial data, generally only second-order convergence can be achieved for a class of single step and linear multi-step time stepping schemes (by ignoring the errors at starting steps). The design and analysis of schemes with higher order accuracy remain largely elusive for standard parabolic models with time-dependent elliptic operators and nonsmooth data. Thus, the development and analysis of high-order time-stepping schemes for the model (1.1) with general problem data is still very challenging; see Section 2 for further discussions.

The rest of the paper is organized as follows. In Section 2, we describe the proposed time-stepping scheme. In Section 3, we prove new temporal regularity results, and in Section 4, we give a complete error analysis for both smooth and nonsmooth data. Finally in Section 5, we present numerical results to complement the error analysis. Throughout, the notation  $c$  denotes a generic constant which may differ at each occurrence, but it is always independent of the time stepsize  $\tau$ , but may depend on the final time  $T$ .

## 2 Derivation of the numerical scheme

In this section, we construct a second-order time-stepping scheme for problem (1.1) using CQ generated by BDF2 with initial correction, derived from a perturbation argument. For notational simplicity, we shall denote by  $v(t) = v(\cdot, t)$  for a function  $v$  defined on  $\Omega \times (0, T]$ .

Since the Riemann-Liouville derivative is equivalent to the Caputo one for functions with zero initial value, we rewrite problem (1.1) as

$${}^R\partial_t^\alpha(u(t) - u_0) + A(t)u(t) = f(t), \quad (2.1)$$

where the Riemann-Liouville derivative  ${}^R\partial_t^\alpha\varphi(t)$  is defined by  ${}^R\partial_t^\alpha\varphi(t) = \frac{d}{dt} \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \varphi(s) ds$ , and the time-dependent elliptic operator  $A(t) : H_0^1(\Omega) \cap H^2(\Omega) \rightarrow L^2(\Omega)$  is defined by

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

Let  $t_n = n\tau$ ,  $n = 0, 1, \dots, N$ , be a uniform partition of the interval  $[0, T]$  with a time stepsize  $\tau = T/N$ . BDF2-CQ approximates the Riemann-Liouville derivative  ${}^R\partial_t^\alpha\varphi(t)$  at the time  $t = t_n$  by

$$\bar{\partial}_\tau^\alpha\varphi^n := \frac{1}{\tau^\alpha} \sum_{j=0}^n b_j \varphi^{n-j} \quad \text{with } \varphi^n = \varphi(t_n), \quad (2.2)$$

where the weights  $\{b_j\}_{j=0}^\infty$  are the coefficients in the power series expansion

$$\delta_\tau(\zeta)^\alpha = \frac{1}{\tau^\alpha} \sum_{j=0}^\infty b_j \zeta^j \quad \text{with} \quad \delta_\tau(\zeta) := \frac{\zeta^2 - 4\zeta + 3}{2\tau} \quad (2.3)$$

If the function  $\varphi$  is smooth and has sufficiently many vanishing derivatives at  $t = 0$ , then BDF2-CQ is second-order accurate pointwise in time [24] [25, Theorem 3.1].

By employing (2.2) to discretize the term  ${}^R\partial_t^\alpha(u(t) - u_0)$  in (2.1), we obtain a BDF2-CQ scheme for (1.1): given  $u^0 = u_0$ , find  $u^n$  such that

$$\bar{\partial}_\tau^\alpha(u - u_0)^n + A(t_n)u^n = f(t_n), \quad n = 1, 2, \dots, N. \quad (2.4)$$

This scheme generally has only first-order accuracy, instead of second-order accuracy, due to the low regularity of the solution  $u(t)$  at  $t = 0$ , unless restrictive compatibility conditions on the initial data  $u_0$  and  $f$  are satisfied (which guarantee good solution regularity at  $t = 0$ ). This has been observed for many different time-stepping schemes for subdiffusion with a time-independent diffusion coefficient [16, 14, 5]. Hence, the vanilla BDF2-CQ scheme (2.4) has to be modified in order to achieve second-order convergence for general data.

In this work, we propose the following time-stepping scheme:

$$\begin{cases} \bar{\partial}_\tau^\alpha(u - u_0)^1 + A(t_1)u^1 + \frac{1}{2}A(0)u_0 = f(t_1) + \frac{1}{2}f(0), \\ \bar{\partial}_\tau^\alpha(u - u_0)^n + A(t_n)u^n = f(t_n), \quad n = 2, 3, \dots, N, \end{cases} \quad (2.5)$$

which is obtained by first rewriting problem (1.1) into

$${}^R\partial_t^\alpha(u - u_0) + A(0)u(t) = F(t) \quad \text{with} \quad F(t) = f(t) + (A(0) - A(t))u(t),$$

and then following [14, 16] to modify the first step as

$$\bar{\partial}_\tau^\alpha(u - u_0)^1 + A(0)u^1 + \frac{1}{2}A(0)u_0 = F(t_1) + \frac{1}{2}F(0).$$

Then substituting the expression of  $F(t)$  and collecting terms yield the correction in (2.5). In (2.5), the term  $A(0)u_0$  should be interpreted in a distributional sense for weak initial data, e.g.,  $u_0 \in L^2(\Omega)$ .

Note that  $F'(0)$  is generally not defined in  $L^2(\Omega)$ . Hence, the existing correction methods in [16] for higher-order BDFs cannot be applied directly. It is still very challenging to develop higher-order time discretization methods for problem (1.1) with nonsmooth problem data. This seems to be open even for the standard parabolic counterpart [34].

### 3 Regularity of solutions

We assume that the diffusion coefficient  $a(x, t) : \Omega \times (0, T) \rightarrow \mathbb{R}^{d \times d}$  satisfies that for some real number  $\lambda \geq 1$ , integer  $K \geq 2$  and  $i, j = 1, \dots, d$ :

$$\lambda^{-1}|\xi|^2 \leq a(x, t)\xi \cdot \xi \leq \lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^d, \quad \forall (x, t) \in \Omega \times (0, T], \quad (3.1)$$

$$|\frac{\partial}{\partial t}a_{ij}(x, t)| + |\nabla_x \frac{\partial^k}{\partial t^k}a_{ij}(x, t)| \leq c, \quad \forall (x, t) \in \Omega \times (0, T], \quad k = 0, \dots, K + 1, \quad (3.2)$$

where  $\cdot$  and  $|\cdot|$  denote the standard Euclidean inner product and norm, respectively. Under these conditions, there holds  $D(A(t)) = H_0^1(\Omega) \cap H^2(\Omega)$  for all  $t \in [0, T]$ . By the complex interpolation method [38], this implies

$$D(A(t)^\gamma) = \dot{H}^{2\gamma}(\Omega), \quad \forall t \in [0, T], \quad \forall \gamma \in [0, 1],$$

where  $\dot{H}^{2\gamma}(\Omega) = (L^2(\Omega), H_0^1(\Omega) \cap H^2(\Omega))_{[\gamma]}$  denotes the complex interpolation space between  $L^2(\Omega)$  and  $H_0^1(\Omega) \cap H^2(\Omega)$ . Equivalently, it can be defined via spectral decomposition of the operator  $A(t)$  [37, Chapter 3]. Let  $\{(\lambda_j, \varphi_j)\}_{j=1}^\infty$  be the eigenpairs of  $A(t)$  with multiplicity counted and  $\{\varphi_j\}_{j=1}^\infty$  be an orthonormal basis in  $L^2(\Omega)$ . Then the space  $\dot{H}^\gamma(\Omega)$  can be defined as

$$\dot{H}^\gamma(\Omega) = \left\{ v \in L^2(\Omega) : \sum_{j=1}^\infty \lambda_j^\gamma (v, \varphi_j)^2 < \infty \right\}.$$

In particular,  $\dot{H}^2(\Omega) = H_0^1(\Omega) \cap H^2(\Omega)$ ,  $\dot{H}^1(\Omega) = H_0^1(\Omega)$  and  $\dot{H}^0(\Omega) = L^2(\Omega)$ . For  $\gamma \in [0, 2]$  we also denote by  $\dot{H}^{-\gamma}(\Omega)$  the dual space of  $\dot{H}^\gamma(\Omega)$ . Then the norm of  $\dot{H}^{-\gamma}(\Omega)$  satisfies

$$\|v\|_{\dot{H}^{-\gamma}(\Omega)} = \|A(t)^{-\frac{\gamma}{2}} v\|_{L^2(\Omega)} \quad \forall v \in \dot{H}^{-\gamma}(\Omega), \quad \forall t \in [0, T].$$

In this section, we prove the following regularity results.

**Theorem 1 (Homogeneous problem)** *If  $a(x, t)$  satisfies (3.1)-(3.2),  $u_0 \in \dot{H}^{2\gamma}(\Omega)$  with  $\gamma \in [0, 1]$  and  $f \equiv 0$ , then for all  $t \in (0, T]$  and  $k = 0, \dots, K$ , the solution  $u(t)$  to problem (1.1) satisfies*

$$\left\| \frac{d^k}{dt^k} (t^k u(t)) \right\|_{\dot{H}^{2\beta}(\Omega)} \leq ct^{-(\beta-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}, \quad \forall \beta \in [\gamma, 1].$$

**Theorem 2 (Inhomogeneous problem)** *If  $a(x, t)$  satisfies (3.1)-(3.2),  $u_0 \equiv 0$ , then for all  $t \in (0, T]$  and  $k = 0, \dots, K$ , the solution  $u(t)$  to problem (1.1) satisfies for any  $\beta \in [0, 1)$*

$$\begin{aligned} \left\| \frac{d^k}{dt^k} (t^k u(t)) \right\|_{\dot{H}^{2\beta}(\Omega)} &\leq c \sum_{j=0}^{k-1} t^{(1-\beta)\alpha+j} \|f^{(j)}(0)\|_{L^2(\Omega)} \\ &\quad + ct^k \int_0^t (t-s)^{(1-\beta)\alpha-1} \|f^{(k)}(s)\|_{L^2(\Omega)} ds, \end{aligned}$$

and similarly for  $\beta = 1$ ,

$$\left\| \frac{d^k}{dt^k} (t^k u(t)) \right\|_{\dot{H}^2(\Omega)} \leq c \sum_{j=0}^k t^j \|f^{(j)}(0)\|_{L^2(\Omega)} + ct^k \int_0^t \|f^{(k+1)}(s)\|_{L^2(\Omega)} ds.$$

*Remark 1* These regularity results are identical with that for subdiffusion with a time-independent elliptic operator [33, Theorems 2.1–2.2], [15, Theorem 2.1]. All the constants in Theorems 1 and 2 may grow with  $k$  and blow up as  $K \rightarrow \infty$ , but stay bounded for any finite  $K$ . Further, these constants are uniformly bounded as  $\alpha \rightarrow 1^-$ , similar to the prior estimates in [18, Remark 2.1].

Theorem 2 implies the following estimate for smooth initial data.

**Corollary 1** *If  $a(x, t)$  satisfies (3.1)-(3.2),  $u_0 \in \dot{H}^2(\Omega)$  and  $f \equiv 0$ , then for  $w(t) = u(t) - u_0$ , for all  $t \in (0, T]$  and  $k = 0, \dots, K$ , there holds*

$$\left\| \frac{d^k}{dt^k} (t^k w(t)) \right\|_{\dot{H}^{2\beta}(\Omega)} \leq ct^{(1-\beta)\alpha} \|u_0\|_{\dot{H}^2(\Omega)}, \quad \forall \beta \in [0, 1].$$

*Proof* The function  $w(t)$  satisfies  $\partial_t^\alpha w(t) + A(t)w(t) = -A(t)u_0$  with  $w(0) = 0$ . Then the assertion follows directly from Theorem 2.  $\square$

The rest of this section is devoted to the proof of Theorems 1 and 2.

### 3.1 Preliminaries

First, we recall some preliminary results [17] on the solution representation and smoothing properties of solution operators for subdiffusion with a time-independent coefficient, i.e.,

$$\partial_t^\alpha u(t) + A_* u(t) = g(t), \quad \forall t \in (0, T], \quad \text{with } u(0) = u_0, \quad (3.3)$$

where  $A_* = A(t_*)$ , for some fixed  $t_* \in [0, T]$  independent of  $t \in (0, T]$ . By means of Laplace transform, the solution  $u$  of (3.3) can be represented by (cf. [15, Section 2] and [17, Section 2])

$$u(t) = F_*(t)u_0 + \int_0^t E_*(t-s)g(s)ds, \quad (3.4)$$

where the operators  $F_*(t)$  and  $E_*(t)$  are respectively defined by

$$F_*(t) := \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} z^{\alpha-1} (z^\alpha + A_*)^{-1} dz, \quad (3.5)$$

$$E_*(t) := \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} (z^\alpha + A_*)^{-1} dz, \quad (3.6)$$

with the contour  $\Gamma_{\theta, \delta}$  (oriented with an increasing imaginary part):

$$\Gamma_{\theta, \delta} = \{z \in \mathbb{C} : |z| = \delta, |\arg z| \leq \theta\} \cup \{z \in \mathbb{C} : z = \rho e^{\pm i\theta}, \rho \geq \delta\}. \quad (3.7)$$

Throughout, we choose a fixed angle  $\theta \in (\frac{\pi}{2}, \pi)$  so that

$$z^\alpha \in \Sigma_{\alpha\theta} \quad \text{for } z \in \Sigma_\theta := \{z \in \mathbb{C} \setminus \{0\} : |\arg(z)| \leq \theta\}.$$

From the definitions (3.5) and (3.6), we deduce

$$A_* E_*(t) = (I - F_*(t))', \quad (3.8)$$

which follows by straightforward computation

$$(I - F_*(t))' = -\frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} z^\alpha (z^\alpha + A_*)^{-1} dz$$

$$= -\frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (I - A_*(z^\alpha + A_*)^{-1}) dz = A_* E_*(t).$$

The next lemma summarizes the smoothing properties of  $F_*(t)$  and  $E_*(t)$ , where  $\|\cdot\|$  denotes the operator norm from  $L^2(\Omega)$  to  $L^2(\Omega)$ .

**Lemma 1** *For any integer  $k = 0, 1, \dots$ , the operators  $F_*$  and  $E_*$  defined in (3.5)-(3.6) satisfy for any  $t \in (0, T]$*

- (i)  $t^{-\alpha} \|A_*^{-1}(I - F_*(t))\| + t^{1-\alpha} \|A_*^{-1} F'_*(t)\| \leq c;$
- (ii)  $t^{k+1-\alpha} \|E_*^{(k)}(t)\| + t^{k+1} \|A_* E_*^{(k)}(t)\| + t^{k+1+\alpha} \|A_*^2 E_*^{(k)}(t)\| \leq c;$
- (iii)  $t^k \|F_*^{(k)}(t)\| + t^{k+\alpha} \|A_* F_*^{(k)}(t)\| \leq c.$

*Proof* The assertions for  $k = 0, 1$  were already given in [18, Lemma 2.2]. The proof for  $k > 1$  is similar. For example, in part (i), by (3.8) and choosing  $\delta = t^{-1}$  in the contour  $\Gamma_{\theta,\delta}$  and letting  $\hat{z} = tz$ :

$$\begin{aligned} \|A_*^{-1} F'_*(t)\| &= \|E_*(t)\| \leq \frac{1}{2\pi} \int_{\Gamma_{\theta,\delta}} e^{\Re(z)t} \|(z^\alpha + A_*)^{-1}\| |dz| \\ &\leq ct^{\alpha-1} \frac{1}{2\pi} \int_{\Gamma_{\theta,1}} e^{\Re(\hat{z})} |\hat{z}|^{-\alpha} |d\hat{z}| \\ &\leq ct^{\alpha-1} \frac{1}{2\pi} \int_{\Gamma_{\theta,1}} e^{\cos(\theta)|\hat{z}|} (1 + |\hat{z}|^{-1}) |d\hat{z}| \leq ct^{\alpha-1}, \end{aligned}$$

and in part (iii) with  $k = 0$ ,  $\|F_*(t)\|$  can be bounded by

$$\begin{aligned} \|F_*(t)\| &\leq \frac{1}{2\pi} \int_{\Gamma_{\theta,\delta}} e^{\Re(z)t} |z|^{\alpha-1} \|(z^\alpha + A_*)^{-1}\| |dz| \\ &\leq \frac{1}{2\pi} \int_{\Gamma_{\theta,\delta}} e^{\Re(z)t} |z|^{-1} |dz| \leq c. \end{aligned}$$

The proof of (3.8) gives  $A_* E_*(t) = -\frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} z^\alpha (z^\alpha + A_*)^{-1} dz$ , and since  $\|A_*(z^\alpha + A_*)^{-1}\| \leq c$ , we deduce

$$\|A_*^2 E_*(t)\| \leq \frac{1}{2\pi} \int_{\Gamma_{\theta,\delta}} e^{\Re(z)t} |z|^\alpha |dz| \leq ct^{-1-\alpha}.$$

All other estimates can be proved similarly and the details are omitted. Note that all the constants  $c$  remain bounded as  $\alpha \rightarrow 1^-$ .  $\square$

The following perturbation estimate [18, Corollary 3.1] will be used extensively. In particular, it implies that  $\|A(s)^{-1} A(t)\| \leq c$  for any  $s, t \in [0, T]$ , and by interpolation also,  $\|A(s)^{-\beta} A(t)^\beta\| \leq c$  for any  $\beta \in (0, 1)$ .

**Lemma 2** *Under conditions (3.1)-(3.2), for any  $\beta \in [0, 1]$ , there holds*

$$\|(I - A(t)^{-1} A(s))v\|_{\dot{H}^{2\beta}(\Omega)} \leq c|t - s| \|v\|_{\dot{H}^{2\beta}(\Omega)}, \quad \forall v \in \dot{H}^{2\beta}(\Omega). \quad (3.9)$$



The following regularity results for problem (1.1) were proved in [18] (also see [21, 7] for related results under different assumptions).

**Theorem 3** *Under conditions (3.1)–(3.2), the solution  $u(t)$  of problem (1.1) satisfies the following estimates:*

(i) *If  $u_0 \in \dot{H}^{2\gamma}(\Omega)$ , with some  $\gamma \in [0, 1]$ , and  $f = 0$ , then*

$$\|u(t)\|_{\dot{H}^{2\gamma}(\Omega)} \leq ct^{-(1-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} \text{ and } \|u'(t)\|_{L^2(\Omega)} \leq ct^{\gamma\alpha-1} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}.$$

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

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

Theorem 3 is a special case of Theorems 1 and 2 corresponding to  $(k, \beta) = (0, 1)$  and  $(k, \beta) = (1, 0)$ , respectively. These results were used in [18] to prove first-order convergence of backward Euler CQ. But they are insufficient to prove second-order convergence of the corrected BDF2–CQ scheme (2.5), which requires the regularity results in Theorems 1 and 2 for  $k = 2$ . Below, we prove Theorems 1 and 2 for a general nonnegative integer  $k$ .

### 3.2 Proof of Theorems 1 and 2

The overall proof strategy is to employ a perturbation argument [17, 18] and then to properly resolve the singularity. Specifically, for any fixed  $t_* \in (0, T]$ , we rewrite problem (1.1) into

$$\begin{cases} \partial_t^\alpha u(t) + A_* u(t) = (A_* - A(t))u(t) + f(t), & \forall t \in (0, T], \\ u(0) = u_0. \end{cases} \quad (3.10)$$

By (3.4), the solution  $u(t)$  of (3.10) is given by

$$u(t) = F_*(t)u_0 + \int_0^t E_*(t-s)(f(s) + (A_* - A(s))u(s))ds. \quad (3.11)$$

The objective is to estimate the  $k$ th temporal derivative  $u^{(k)}(t) := \frac{d^k}{dt^k} u(t)$  in  $\dot{H}^{2\beta}(\Omega)$  for  $\beta \in [0, 1]$  using (3.11). However, direct differentiation of  $u(t)$  in (3.11) with respect to  $t$  leads to strong singularity that precludes the use of Gronwall's inequality in Lemma 10, in order to handle the perturbation term. To overcome the difficulty, we instead estimate  $\|(t^{k+1}u(t))^{(k)}\|_{\dot{H}^{2\beta}(\Omega)}$  using the expansion of  $t^{k+1} = [(t-s) + s]^{k+1}$  in the the following expression:

$$t^{k+1}u(t) = t^{k+1}F_*(t)u_0 + t^{k+1} \int_0^t E_*(t-s)f(s)ds \quad (3.12)$$

$$+ \sum_{m=0}^{k+1} \binom{m}{k+1} \int_0^t (t-s)^m E_*(t-s) (A_* - A(s)) s^{k+1-m} u(s) ds,$$

where  $\binom{m}{k+1}$  denotes binomial coefficients. One crucial part in the proof is to bound  $k$ th-order derivatives of the summands in (3.12).

Now we can give the proof of Theorem 1.

*Proof (of Theorem 1)* When  $k = 0$ , setting  $f = 0$  and  $t = t_*$  in (3.11) yields

$$A_*^\beta u(t_*) = A_*^\beta F_*(t_*) u_0 + \int_0^{t_*} A_*^\beta E_*(t_* - s) (A_* - A(s)) u(s) ds,$$

where  $\beta \in [\gamma, 1]$ . By Lemmas 1 and 2,

$$\begin{aligned} \|A_*^\beta u(t_*)\|_{L^2(\Omega)} &\leq \|A_*^{\beta-\gamma} F_*(t_*) A_*^\gamma u_0\|_{L^2(\Omega)} \\ &\quad + \int_0^{t_*} \|A_* E_*(t_* - s)\| \|A_*^\beta (I - A_*^{-1} A(s)) u(s)\|_{L^2(\Omega)} ds \\ &\leq c t_*^{-(\beta-\gamma)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)} + c \int_0^{t_*} (t_* - s) \|A_* E_*(t_* - s)\| \|A_*^\beta u(s)\|_{L^2(\Omega)} ds \\ &\leq c t_*^{-(\beta-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} + c \int_0^{t_*} \|A_*^\beta u(s)\|_{L^2(\Omega)} ds. \end{aligned}$$

This and Gronwall's inequality in Lemma 10 with  $\mu = (\beta - \gamma)\alpha$  yield

$$\|A_*^\beta u(t_*)\|_{L^2(\Omega)} \leq c(1 - (\beta - \gamma)\alpha)^{-1} t_*^{-(\beta-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}.$$

In particular, we have  $\|A_*^{\frac{\beta+\gamma}{2}} u(t_*)\|_{L^2(\Omega)} \leq c t_*^{-\frac{\beta-\gamma}{2}\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}$ , with  $c$  being bounded as  $\alpha \rightarrow 1^-$ . This estimate and Lemmas 1(ii) and 2 then imply

$$\begin{aligned} \|A_*^\beta u(t_*)\|_{L^2(\Omega)} &\leq \|A_*^{\beta-\gamma} F_*(t_*) A_*^\gamma u_0\|_{L^2(\Omega)} \\ &\quad + \int_0^{t_*} \|A_*^{\frac{\beta-\gamma}{2}} A_* E_*(t_* - s)\| \|A_*^{\frac{\beta+\gamma}{2}} (I - A_*^{-1} A(s)) u(s)\|_{L^2(\Omega)} ds \\ &\leq c t_*^{-(\beta-\gamma)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)} + c \int_0^{t_*} (t_* - s) \|A_*^{\frac{\beta-\gamma}{2}} A_* E_*(t_* - s)\| \|A_*^{\frac{\beta+\gamma}{2}} u(s)\|_{L^2(\Omega)} ds \\ &\leq c \left( t_*^{-(\beta-\gamma)\alpha} + \int_0^{t_*} (t_* - s)^{-\frac{\beta-\gamma}{2}\alpha} s^{-\frac{\beta-\gamma}{2}\alpha} ds \right) \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} \leq c t_*^{-(\beta-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}. \end{aligned}$$

Equivalently, we have

$$\|A_*^\beta t_* u(t_*)\|_{L^2(\Omega)} \leq c t_*^{1-(\beta-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)},$$

where  $c$  is bounded as  $\alpha \rightarrow 1^-$ . This proves the assertion for  $k = 0$ .

Next we prove the case  $1 \leq k \leq K$  using mathematical induction. Suppose that the assertion holds up to  $k-1 < K$ , and we prove it for  $k \leq K$ . Indeed, by Lemma 3 below,

$$\left\| A_*^\beta \frac{d^k}{dt^k} \int_0^t (t-s)^m E_*(t-s) (A_* - A(s)) s^{k+1-m} u(s) ds \Big|_{t=t_*} \right\|_{L^2(\Omega)}$$

$$\leq ct_*^{-(\beta-\gamma)\alpha+1} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} + c \int_0^{t_*} \|A_*^\beta(s^{k+1}u(s))^{(k)}\|_{L^2(\Omega)} ds,$$

where  $m = 0, 1, \dots, k+1$ . Meanwhile, the estimates in Lemma 1 imply

$$\|A_*^\beta(t^{k+1}F_*(t)u_0)^{(k)}\|_{L^2(\Omega)} \leq ct^{-(\beta-\gamma)\alpha+1} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}.$$

By applying  $A_*^\beta \frac{d^k}{dt^k}$  to (3.12) and using the last two estimates, we obtain

$$\begin{aligned} \|A_*^\beta(t^{k+1}u(t))^{(k)}|_{t=t_*}\|_{L^2(\Omega)} &\leq ct_*^{-(\beta-\gamma)\alpha+1} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} \\ &\quad + c \int_0^{t_*} \|A_*^\beta(s^{k+1}u(s))^{(k)}\|_{L^2(\Omega)} ds. \end{aligned}$$

Last, applying the standard Gronwall's inequality, we complete the induction step and also the proof of the theorem.  $\square$

In the proof of Theorem 1, we have used the following result.

**Lemma 3** *Under the conditions of Theorem 1, for  $m = 0, \dots, k+1$ , there holds*

$$\begin{aligned} &\left\| A_*^\beta \frac{d^k}{dt^k} \int_0^t (t-s)^m E_*(t-s) (A_* - A(s)) s^{k+1-m} u(s) ds \right\|_{L^2(\Omega)} \\ &\leq ct_*^{-(\beta-\gamma)\alpha+1} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} + c \int_0^{t_*} \|A_*^\beta(s^{k+1}u(s))^{(k)}\|_{L^2(\Omega)} ds. \end{aligned}$$

*Proof* Denote the integral on the left hand side by  $I_m(t)$ , and let  $v_m = t^m u(t)$  and  $W_m(t) = t^m E_*(t)$ . Direct computation using product rule and changing variables gives that for any  $0 \leq m \leq k$ , there holds

$$\begin{aligned} I_m^{(k)}(t) &= \frac{d^{k-m}}{dt^{k-m}} \int_0^t W_m^{(m)}(t-s) (A_* - A(s)) v_{k-m+1}(s) ds \\ &= \frac{d^{k-m}}{dt^{k-m}} \int_0^t W_m^{(m)}(s) (A_* - A(t-s)) v_{k-m+1}(t-s) ds \\ &= \int_0^t W_m^{(m)}(s) \frac{d^{k-m}}{dt^{k-m}} ((A_* - A(t-s)) v_{k-m+1}(t-s)) ds \\ &= \sum_{\ell=0}^{k-m} \binom{\ell}{k-m} \underbrace{\int_0^t W_m^{(m)}(s) (A_* - A(t-s))^{(k-m-\ell)} v_{k-m+1}^{(\ell)}(t-s) ds}_{I_{m,\ell}(t)}. \end{aligned}$$

Next we bound the integrand

$$\tilde{I}_{m,\ell}(s) := W_m^{(m)}(A_* - A(t_* - s))^{(k-m-\ell)} v_{k-m+1}^{(\ell)}(t_* - s)$$

of the integral  $I_{m,\ell}(t_*)$ . We shall distinguish between  $\beta \in [\gamma, 1)$  and  $\beta = 1$ . First we analyze the case  $\beta \in [\gamma, 1)$ . When  $\ell < k$ , by Lemmas 1(ii) and 2 and the induction hypothesis, we bound the integrand  $\tilde{I}_{m,\ell}(s)$  by

$$\|A_*^\beta \tilde{I}_{m,\ell}(s)\|_{L^2(\Omega)}$$

$$\begin{aligned}
&\leq \|A_*^\beta W_m^{(m)}(s)\| \|(A_* - A(t_* - s))^{(k-m-\ell)} v_{k-m+1}^{(\ell)}(t_* - s)\|_{L^2(\Omega)} \\
&\leq \begin{cases} cs^{(1-\beta)\alpha-1} s \|A_* v_{k-m+1}^{(k-m)}(t_* - s)\|_{L^2(\Omega)}, & \ell = k - m, \\ cs^{(1-\beta)\alpha-1} \|A_* v_{k-m+1}^{(\ell)}(t_* - s)\|_{L^2(\Omega)}, & \ell < k - m, \end{cases} \\
&\leq \begin{cases} cs^{(1-\beta)\alpha} (t_* - s)^{1-(1-\gamma)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)}, & \ell = k - m, \\ cs^{(1-\beta)\alpha-1} (t_* - s)^{k-m-\ell+1-(1-\gamma)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)}, & \ell < k - m. \end{cases}
\end{aligned}$$

Similarly for the case  $\ell = k$  (and thus  $m = 0$ ), there holds

$$\begin{aligned}
\|A_*^\beta \tilde{I}_{0,k}(s)\|_{L^2(\Omega)} &\leq \|A_* E_*(s)\| \|A_*^\beta (I - A_*^{-1} A(t_* - s)) v_{k+1}^{(k)}\|_{L^2(\Omega)} \\
&\leq c \|A_*^\beta v_{k+1}^{(k)}(t_* - s)\|_{L^2(\Omega)}.
\end{aligned}$$

Thus, for  $0 \leq m \leq k$  and  $\ell = k - m$ , upon integrating from 0 to  $t_*$ , we obtain

$$\|A_*^\beta I_m^{(k)}(t_*)\|_{L^2(\Omega)} \leq ct_*^{2+(\gamma-\beta)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)} + c \int_0^{t_*} \|A_*^\beta v_{k+1}^{(k)}(s)\|_{L^2(\Omega)} ds,$$

and similarly for  $0 \leq m \leq k$  and  $\ell < k - m$ ,

$$\begin{aligned}
\|A_*^\beta I_m^{(k)}(t_*)\|_{L^2(\Omega)} &\leq c((1-\beta)\alpha)^{-1} t_*^{1+(\gamma-\beta)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)} \\
&\quad + c \int_0^{t_*} \|A_*^\beta v_{k+1}^{(k)}(s)\|_{L^2(\Omega)} ds.
\end{aligned}$$

Meanwhile, for  $m = k + 1$ , we have

$$A_*^\beta I_{k+1}^{(k)}(t_*) = \int_0^{t_*} A_*^{\beta+1-\gamma} W_{k+1}^{(k)}(t_* - s) A_*^\gamma (I - A_*^{-1} A(s)) u(s) ds,$$

and consequently, by Lemmas 1(ii) and 2 and the induction hypothesis,

$$\begin{aligned}
&\|A_*^\beta I_{k+1}^{(k)}(t_*)\|_{L^2(\Omega)} \\
&\leq \int_0^{t_*} \|A_*^{\beta+1-\gamma} W_{k+1}^{(k)}(t_* - s)\| \|A_*^\gamma (I - A_*^{-1} A(s)) u(s)\|_{L^2(\Omega)} ds \\
&\leq c \int_0^{t_*} (t_* - s)^{1-(\beta-\gamma)\alpha} \|A_*^\gamma u(s)\|_{L^2(\Omega)} ds \leq ct_*^{2+(\gamma-\beta)\alpha} \|A_*^\gamma u_0\|_{L^2(\Omega)}.
\end{aligned}$$

In the case  $0 \leq m \leq k$  and  $\ell < k - m$ , the preceding estimates require  $\beta \in [0, 1)$ . When  $0 \leq m \leq k$ ,  $\ell < k - m$  and  $\beta = 1$ , we apply the identity (3.8) and rewrite  $A_* I_{m,\ell}(t_*)$  as

$$A_* I_{m,\ell}(t_*) = \int_0^{t_*} (s^m (I - F_*(s))')^{(m)} (A_* - A(t_* - s))^{(k-m-\ell)} v_{k-m+1}^{(\ell)}(t_* - s) ds.$$

Then integration by parts and product rule yield

$$A_* I_{m,\ell}(t_*) = - \int_0^{t_*} D(s) (A_* - A(t_* - s))^{(k-m-\ell+1)} v_{k-m+1}^{(\ell)}(t_* - s) ds$$

$$\begin{aligned}
& - \int_0^{t_*} D(s)(A_* - A(t_* - s))^{(k-m-\ell)} v_{k-m+1}^{(\ell+1)}(t_* - s) ds \\
& - D(0)(A_* - A(t_* - s))^{(k-m-\ell)}|_{s=0} v_{k-m+1}^{(\ell)}(t_*), \tag{3.13}
\end{aligned}$$

with

$$D(s) = \begin{cases} I - F_*(s), & m = 0, \\ (s^m(I - F_*(s))')^{(m-1)}, & m > 0. \end{cases}$$

By Lemma 1(iii),  $\|D(s)\| \leq c$ , and thus the preceding argument with Lemmas 1 and 2 and the induction hypothesis allows bounding the integrand  $A_* \tilde{I}_{m,\ell}(s)$  of (3.13) by

$$\begin{aligned}
\|A_* \tilde{I}_{m,\ell}(s)\|_{L^2(\Omega)} & \leq c(t_* - s)^{k-\ell-(1-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)} \\
& + \begin{cases} c\|A_* v_{k+1}^{(k)}(t_* - s)\|_{L^2(\Omega)}, & \ell = k - 1, \\ c(t_* - s)^{k-1-\ell-(1-\gamma)\alpha} \|u_0\|_{\dot{H}^{2\gamma}(\Omega)}, & \ell < k - 1, \end{cases}
\end{aligned}$$

where for  $\ell = k - 1$ , we have  $m = 0$  and hence  $D(0) = 0$ .

Combining the last estimates and then integrating from 0 to  $t_*$  in  $s$ , we obtain the desired assertion of Lemma 3. All the estimates are based on Lemmas 1 and 2, and thus the constants  $c$  in Lemma 3 is bounded as  $\alpha \rightarrow 1^-$ .  $\square$

The proof of Theorem 2 is similar to that of Theorem 1. The lengthy and technical proof is deferred to Appendix B.

#### 4 Error analysis

In this section, we present error estimates for the scheme (2.5). To this end, let  $w(t) = u(t) - u(0)$ , which satisfies the equation

$$\begin{cases} \partial_t^\alpha w(t) + A(0)w(t) = g(t), & \forall t > 0, \\ w(0) = 0. \end{cases} \tag{4.1}$$

with

$$g(t) := (A(0) - A(t))w(t) - A(t)u_0 + f(t).$$

Then the error  $e^n := u^n - u(t_n)$  of the numerical solution  $u^n$  is given by

$$e^n = w^n - w(t_n), \quad \text{with } w^n = u^n - u_0. \tag{4.2}$$

We also introduce an intermediate solution  $\bar{w}^n$  defined by

$$\begin{cases} \bar{\partial}_\tau^\alpha \bar{w}^1 + A(0)\bar{w}^1 = g(t_1) + \frac{1}{2}g(t_0), \\ \bar{\partial}_\tau^\alpha \bar{w}^n + A(0)\bar{w}^n = g(t_n), & n = 2, 3, \dots, N. \end{cases} \tag{4.3}$$

which is the numerical approximation of (4.1) with the source  $g(t)$ . Using  $\bar{w}^n$ , we further decompose the error  $e^n$  into

$$e^n = (w^n - \bar{w}^n) + (\bar{w}^n - w(t_n)) =: \varrho^n + \vartheta^n,$$

where  $\vartheta^n$  is the error due to time discretization of problem (4.1) with a “time-independent” operator  $A(0)$ , and  $\varrho^n$  is the error between two numerical solutions due to the perturbation of the source term.

It suffices to estimate the two terms  $\varrho^n$  and  $\vartheta^n$ . The analysis for  $\vartheta^n$  will employ the following nonstandard error estimates.

**Lemma 4** *Let  $u(t)$  be the solution of problem (3.3) with  $u_0 \equiv 0$  and  $u^n$ , with  $u^0 = 0$ , defined by*

$$\begin{cases} \bar{\partial}_\tau^\alpha u^1 + A(t_*)u^1 = g(t_1) + \frac{1}{2}g(0), \\ \bar{\partial}_\tau^\alpha u^n + A(t_*)u^n = g(t_n), \end{cases} \quad n = 2, \dots, N.$$

*Then the following statements hold.*

(i) *If  $\beta, \gamma \in [0, 1)$  and  $\beta + \gamma < 1$ , then*

$$\begin{aligned} \|u(t_n) - u^n\|_{\dot{H}^{2\beta}(\Omega)} &\leq c\tau^2 \left( t_n^{(1-\beta)\alpha-2} \|g(0)\|_{L^2(\Omega)} + t_n^{(1-\beta)\alpha-1} \|g'(0)\|_{L^2(\Omega)} \right. \\ &\quad \left. + \int_0^{t_n} (t_{n+1} - s)^{(1-\beta-\gamma)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\gamma}(\Omega)} ds \right). \end{aligned}$$

(ii) *If  $\beta = 1$ , then*

$$\begin{aligned} \|u(t_n) - u^n\|_{\dot{H}^2(\Omega)} &\leq c\tau^2 \left( t_n^{-2} \|g(0)\|_{L^2(\Omega)} + t_n^{-1} \|g'\|_{C([0, \tau]; L^2(\Omega))} \right. \\ &\quad \left. + \int_\tau^{t_n} (t_{n+1} - s)^{-1} \|g''(s)\|_{L^2(\Omega)} ds \right). \end{aligned}$$

Lemma 4 can be proved using discrete Laplace transform (generating function technique) similarly as the error estimation for CQ-BDFk [16]. This type of error estimation yields an error bound directly from a contour integral, while the constant produced from a contour integral is bounded as  $\alpha \rightarrow 1^-$ . We will use Lemma 4 and a perturbation argument to bound  $\vartheta^n$  and  $\varrho^n$ , respectively, and derive error estimates for numerical solutions.

For the convenience of error analysis, we further split  $w(t)$  into  $w(t) = w_0(t) + w_1(t)$ , where  $w_0(t)$  and  $w_1(t)$  are respectively solutions of

$$\partial_t^\alpha w_0(t) + A(0)w_0(t) = (A(0) - A(t))w(t), \quad \text{with } w_0(0) = 0, \quad (4.4)$$

$$\partial_t^\alpha w_1(t) + A(0)w_1(t) = -A(t)u_0 + f(t), \quad \text{with } w_1(0) = 0. \quad (4.5)$$

Correspondingly, we split  $\bar{w}^n$  into  $\bar{w}^n = \bar{w}_0^n + \bar{w}_1^n$ , defined by  $\bar{w}_0^0 = 0$ ,

$$\bar{\partial}_\tau^\alpha \bar{w}_0^n + A(0)\bar{w}_0^n = (A(0) - A(t_n))w(t_n), \quad n = 1, 2, 3, \dots, N, \quad (4.6)$$

and  $\bar{w}_1^0 = 0$  and

$$\begin{cases} \bar{\partial}_\tau^\alpha \bar{w}_1^1 + A(0)\bar{w}_1^1 = -A(t_1)u_0 - \frac{1}{2}A(0)u_0 + f(t_1) + \frac{1}{2}f(t_0), \\ \bar{\partial}_\tau^\alpha \bar{w}_1^n + A(0)\bar{w}_1^n = -A(t_n)u_0 + f(t_n), \end{cases} \quad n = 2, 3, \dots, N, \quad (4.7)$$

The functions  $\bar{w}_0^n$  and  $\bar{w}_1^n$  approximate  $w_0(t_n)$  and  $w_1(t_n)$ , respectively.

#### 4.1 Error analysis for the homogeneous problem

Now we analyze the scheme (2.5) for the homogeneous problem with  $f \equiv 0$ . First, we bound the function  $g(t) = (A(0) - A(t))w(t)$  in equation (4.4).

**Lemma 5** *Let Assumptions (3.1)–(3.2) hold. For the function  $g(t) = (A(0) - A(t))w(t)$ , the following statements hold when  $f \equiv 0$ .*

- (i)  $u_0 \in \dot{H}^2(\Omega)$  and  $\beta \in [0, 1]$ , then  $\|g'(0)\|_{L^2(\Omega)} + t^{1-\alpha\beta}\|g''(t)\|_{\dot{H}^{-2\beta}(\Omega)} \leq c\|u_0\|_{\dot{H}^2(\Omega)}$ .
- (ii)  $u_0 \in L^2(\Omega)$ , then  $\|g'(t)\|_{\dot{H}^{-2}(\Omega)} + t\|g''(t)\|_{\dot{H}^{-2}(\Omega)} \leq c\|u_0\|_{L^2(\Omega)}$ .

*Proof* By Theorem 1 and triangle inequality,  $\|w(t)\|_{\dot{H}^2(\Omega)} \leq \|u(t)\|_{\dot{H}^2(\Omega)} + \|u_0\|_{\dot{H}^2(\Omega)} \leq c\|u_0\|_{\dot{H}^2(\Omega)}$ . Thus, by Lemma 2,

$$\begin{aligned} \|g'(t)\|_{L^2(\Omega)} &\leq \|(A(0) - A(t))w'(t)\|_{L^2(\Omega)} + \|A'(t)w(t)\|_{L^2(\Omega)} \\ &\leq ct\|w'(t)\|_{\dot{H}^2(\Omega)} + c\|w(t)\|_{\dot{H}^2(\Omega)} \leq c\|u_0\|_{\dot{H}^2(\Omega)}, \end{aligned}$$

Thus,  $\|g'(0)\|_{L^2(\Omega)} \leq c\|u_0\|_{\dot{H}^2(\Omega)}$ . Since  $g''(t) = (A(0) - A(t))w''(t) - 2A'(t)w'(t) - A''(t)w(t)$ , it follows from Corollary 1 and Theorem 1 that for  $\beta \in [0, 1]$

$$\begin{aligned} \|g''(t)\|_{\dot{H}^{-2\beta}(\Omega)} &= \|(A(0) - A(t))w''(t) - 2A'(t)w'(t) - A''(t)w(t)\|_{\dot{H}^{-2\beta}(\Omega)} \\ &\leq ct\|w''(t)\|_{\dot{H}^{-2-2\beta}(\Omega)} + c\|w'(t)\|_{\dot{H}^{-2-2\beta}(\Omega)} + c\|w(t)\|_{\dot{H}^{-2-2\beta}(\Omega)} \\ &\leq ct^{\alpha\beta-1}\|u_0\|_{\dot{H}^2(\Omega)}. \end{aligned}$$

Similarly, when  $u_0 \in L^2(\Omega)$ , repeating the preceding argument shows (ii).

The next lemma bounds  $\vartheta^n = \bar{w}^n - w(t_n)$ .

**Lemma 6** *Let conditions (3.1)–(3.2) hold, and  $w$  be the solution to problem (4.1) with  $f \equiv 0$ . Let  $\vartheta^n := \bar{w}^n - w(t_n)$ . Then there hold*

$$\begin{aligned} \|\vartheta^n\|_{\dot{H}^{2\beta}(\Omega)} &\leq c\tau^2 t_n^{\alpha(1-\beta)-2} \|u_0\|_{\dot{H}^2(\Omega)}, \quad \forall \beta \in [0, 1/2), \\ \|\vartheta^n\|_{L^2(\Omega)} &\leq c\tau^2 t_n^{-2} \ell_n \|u_0\|_{L^2(\Omega)}, \quad \text{with } \ell_n = \log(1 + t_n/\tau). \end{aligned}$$

*Proof* Using the decompositions  $w(t) = w_0(t) + w_1(t)$  and  $\bar{w}^n = \bar{w}_0^n + \bar{w}_1^n$  defined in (4.4)–(4.5) and (4.6)–(4.7), respectively, we have

$$\|\vartheta^n\|_{\dot{H}^{2\beta}(\Omega)} \leq \|\bar{w}_0^n - w_0(t_n)\|_{\dot{H}^{2\beta}(\Omega)} + \|\bar{w}_1^n - w_1(t_n)\|_{\dot{H}^{2\beta}(\Omega)}. \quad (4.8)$$

We discuss the cases  $u_0 \in \dot{H}^2(\Omega)$  and  $u_0 \in L^2(\Omega)$ , separately.

*Case (i):*  $u_0 \in \dot{H}^2(\Omega)$ . Lemma 4(i) with  $g(t) = A(t)u_0$ , for  $\beta \in [0, 1/2)$ , implies

$$\|\bar{w}_1^n - w_1(t_n)\|_{\dot{H}^{2\beta}(\Omega)} \leq c\tau^2 t_n^{(1-\beta)\alpha-2} \|u_0\|_{\dot{H}^2(\Omega)}. \quad (4.9)$$

For  $g(t) = (A(0) - A(t))w(t)$  and any  $\beta \in [0, 1/2)$ , Lemmas 4(i) and 5 imply

$$\|\bar{w}_0^n - w_0(t_n)\|_{\dot{H}^{2\beta}(\Omega)}$$

$$\begin{aligned}
&\leq c\tau^2 t_n^{(1-\beta)\alpha-1} \|g'(0)\|_{L^2(\Omega)} + c\tau^2 \int_0^{t_n} (t_{n+1} - s)^{(1-2\beta)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} ds \\
&\leq c\tau^2 \left( t_n^{(1-\beta)\alpha-1} + \int_0^{t_n} (t_{n+1} - s)^{(1-2\beta)\alpha-1} s^{\alpha\beta-1} ds \right) \|u_0\|_{\dot{H}^2(\Omega)} \\
&\leq c\tau^2 t_n^{\alpha(1-\beta)-1} \|u_0\|_{\dot{H}^2(\Omega)}.
\end{aligned}$$

This and (4.9) yield the desired estimate for  $u_0 \in \dot{H}^2(\Omega)$ .

*Case (ii):*  $u_0 \in L^2(\Omega)$ . By Lemma 4(ii), we have

$$\|\bar{w}_1^n - w_1(t_n)\|_{L^2(\Omega)} \leq c\tau^2 t_n^{-2} \ell_n \|u_0\|_{L^2(\Omega)}.$$

Meanwhile, by Lemmas 4(ii) and 5, we have

$$\begin{aligned}
&\|\bar{w}_0^n - w_0(t_n)\|_{L^2(\Omega)} \\
&\leq c\tau^2 \left( t_n^{-1} \|g'\|_{C([0,\tau];\dot{H}^{-2}(\Omega))} + \int_\tau^{t_n} (t_{n+1} - s)^{-1} \|g''(s)\|_{\dot{H}^{-2}(\Omega)} ds \right) \\
&\leq c\tau^2 \left( t_n^{-1} + \int_\tau^{t_n} (t_{n+1} - s)^{-1} s^{-1} ds \right) \|u_0\|_{L^2(\Omega)} \leq c\tau^2 t_n^{-1} \ell_n \|u_0\|_{L^2(\Omega)}.
\end{aligned}$$

These two estimates give the second assertion, completing the proof.  $\square$

We need a temporally semidiscrete solution operator  $E_{\tau,m}^n$  defined by

$$E_{\tau,m}^n = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}^\tau} e^{zn\tau} (\delta_\tau(e^{-z\tau})^\alpha + A(t_m))^{-1} dz, \quad (4.10)$$

with the contour  $\Gamma_{\theta,\delta}^\tau$  given by

$$\Gamma_{\theta,\delta}^\tau := \{z \in \Gamma_{\theta,\delta} : |\Im(z)| \leq \pi/\tau\}, \quad (4.11)$$

oriented with an increasing imaginary part. The following smoothing property of the operator  $E_{\tau,m}^n$  holds [18, Lemma 4.3]: for any  $\beta \in [0, 1]$

$$\|A(t_m)^\beta E_{\tau,m}^n\| \leq c(t_n + \tau)^{(1-\beta)\alpha-1}, \quad n = 0, 1, \dots, N. \quad (4.12)$$

We have the following  $L^2(\Omega)$  stability for  $\varrho^n$ .

**Lemma 7** *Let conditions (3.1)-(3.2) be fulfilled, and  $u$  the solution to problem (1.1) with  $f \equiv 0$ . Let  $\varrho^n = w^n - \bar{w}^n$ . Then with  $\ell_n = \log(1 + t_n/\tau)$ , there holds*

$$\|\varrho^m\|_{L^2(\Omega)} \leq c\tau \sum_{k=1}^m \|\varrho^k\|_{L^2(\Omega)} + \begin{cases} c\tau^2 t_m^{\alpha-1} \|u_0\|_{\dot{H}^2(\Omega)}, & \text{if } u_0 \in \dot{H}^2(\Omega), \\ c\tau^2 t_m^{-1} \ell_m^2 \|u_0\|_{L^2(\Omega)}, & \text{if } u_0 \in L^2(\Omega). \end{cases}$$



*Proof* It follows from (2.5) and (4.3) that  $\varrho^n$  satisfies  $\varrho^0 = 0$  and

$$\begin{aligned}\bar{\partial}_\tau^\alpha \varrho^n + A(t_m) \varrho^n &= \bar{\partial}_\tau^\alpha (w^n - \bar{w}^n) + A(t_m)(w^n - \bar{w}^n) \\ &= (A(t_m) - A(t_n))w^n - (A(t_m) - A(0))\bar{w}^n - (A(0) - A(t_n))w(t_n) \\ &= (A(t_m) - A(t_n))\varrho^n - (A(t_n) - A(0))\vartheta^n, \quad n = 1, 2, \dots, N.\end{aligned}$$

Using the operator  $E_{\tau,m}^n$  in (4.10),  $\varrho^m$  is represented by

$$\varrho^m = \tau \sum_{k=1}^m E_{\tau,m}^{m-k} [(A(t_m) - A(t_k))\varrho^k - (A(t_k) - A(0))\vartheta^k].$$

Consequently, by triangle inequality,

$$\begin{aligned}\|\varrho^m\|_{L^2(\Omega)} &\leq \tau \sum_{k=1}^m \|E_{\tau,m}^{m-k} (A(t_m) - A(t_k))\varrho^k\|_{L^2(\Omega)} \\ &\quad + \tau \sum_{k=1}^m \|E_{\tau,m}^{m-k} (A(t_k) - A(0))\vartheta^k\|_{L^2(\Omega)} := \text{I} + \text{II}.\end{aligned}$$

For the term I, by (4.12) with  $\beta = 1$  and Lemma 2, we have

$$\|A(t_m)E_{\tau,m}^{m-k}\| \|(I - A(t_m)^{-1}A(t_k))\varrho^k\|_{L^2(\Omega)} \leq ct_{m-k+1}^{-1} t_{m-k} \|\varrho^k\|_{L^2(\Omega)},$$

and thus

$$\text{I} \leq c\tau \sum_{k=1}^m \|\varrho^k\|_{L^2(\Omega)}. \quad (4.13)$$

For the term II, we discuss the cases  $u_0 \in \dot{H}^2(\Omega)$  and  $u_0 \in L^2(\Omega)$  separately.

*Case (i):*  $u_0 \in \dot{H}^2(\Omega)$ . The estimate (4.12) with  $\beta = \frac{3}{4}$ , Lemmas 2 and 6 with  $\beta = \frac{1}{4}$  imply that  $\text{II}_{m,k} = \|E_{\tau,m}^{m-k} (A(t_k) - A(0))\vartheta^k\|_{L^2(\Omega)}$  is bounded by

$$\text{II}_{m,k} \leq ct_{m-k+1}^{\frac{\alpha}{4}-1} t_k \|\vartheta^k\|_{\dot{H}^{\frac{1}{2}}(\Omega)} \leq c\tau^2 t_{m-k+1}^{\frac{\alpha}{4}-1} t_k^{\frac{3\alpha}{4}-1} \|u_0\|_{\dot{H}^2(\Omega)}$$

and further, since  $\tau \sum_{k=1}^m t_{m-k+1}^{\frac{\alpha}{4}-1} t_k^{\frac{3\alpha}{4}-1} \leq ct_m^{\alpha-1}$ , there holds

$$\text{II} \leq \tau \sum_{k=1}^m \text{II}_{m,k} \leq c\tau^2 t_m^{\alpha-1} \|u_0\|_{\dot{H}^2(\Omega)}.$$

*Case (ii):*  $u_0 \in L^2(\Omega)$ . By (4.12) and Lemmas 6 and 2,

$$\begin{aligned}\text{II}_{m,k} &\leq \|E_{\tau,m}^{m-k} A(t_m)\| \|A(t_m)^{-1}A(0)\| \|(I - A(0)^{-1}A(t_k))\vartheta^k\|_{L^2(\Omega)} \\ &\leq ct_{m-k+1}^{-1} t_k \|\vartheta^k\|_{L^2(\Omega)} \leq c\tau^2 \ell_m t_{m-k+1}^{-1} t_k^{-1} \|u_0\|_{L^2(\Omega)}.\end{aligned}$$

This and the inequality  $\tau \sum_{k=1}^m t_{m-k+1}^{-1} t_k^{-1} \leq ct_m^{-1} \ell_m$  yield

$$\text{II} \leq c\tau^2 t_m^{-1} \ell_m^2 \|u_0\|_{L^2(\Omega)}.$$

In either case, combining the bounds on I and II gives the desired assertion.  $\square$

Now we can derive error estimates for the homogeneous problem.

**Theorem 4** *Let  $u$  and  $u^n$  be the solutions to problems (1.1) and (2.5) with  $f \equiv 0$ , respectively. Then with  $\ell_n = \log(1 + t_n/\tau)$ , there holds*

$$\|u(t_n) - u^n\|_{L^2(\Omega)} \leq \begin{cases} c\tau^2 t_n^{\alpha-2} \|u_0\|_{\dot{H}^2(\Omega)}, & \text{if } u_0 \in \dot{H}^2(\Omega), \\ c\tau^2 t_n^{-2} \ell_n^2 \|u_0\|_{L^2(\Omega)}, & \text{if } u_0 \in L^2(\Omega). \end{cases}$$

*Proof* It follows directly from Lemma 7 that

$$\|\varrho^m\|_{L^2(\Omega)} \leq c\tau \sum_{k=1}^m \|\varrho^k\|_{L^2(\Omega)} + \begin{cases} c\tau^2 t_m^{\alpha-1} \|u_0\|_{\dot{H}^2(\Omega)}, & \text{if } u_0 \in \dot{H}^2(\Omega), \\ c\tau^2 t_m^{-1} \ell_m^2 \|u_0\|_{L^2(\Omega)}, & \text{if } u_0 \in L^2(\Omega). \end{cases}$$

Thus, by the discrete Gronwall's inequality from Lemma 11 (with  $\mu = 1 - \alpha$ ) and Lemma 12,

$$\|\varrho^m\|_{L^2(\Omega)} \leq \begin{cases} c\tau^2 t_m^{\alpha-1} \|u_0\|_{\dot{H}^2(\Omega)}, & \text{if } u_0 \in \dot{H}^2(\Omega), \\ c\tau^2 t_m^{-1} \ell_m^2 \|u_0\|_{L^2(\Omega)}, & \text{if } u_0 \in L^2(\Omega). \end{cases}$$

This, Lemma 6 and the triangle inequality complete the proof. The preceding estimates are based on Lemma 7 and Lemmas 11–12. In particular, applying Lemma 11 to the case  $u_0 \in \dot{H}^2(\Omega)$  yields a constant  $c$  depending on  $1/\alpha$ . Therefore, the constants  $c$  in Theorem 4 is bounded as  $\alpha \rightarrow 1^-$ .  $\square$

*Remark 2* The error estimate for  $u_0 \in \dot{H}^2(\Omega)$  in Theorem 4 is identical with that for the case of a time-independent elliptic operator, and that for nonsmooth initial data is also nearly identical, up to the factor  $\ell_n^2$  [14]. The  $\ell_n$  factor is also present for backward Euler convolution quadrature [18] for subdiffusion, and backward Euler method [26] and general single-step and multi-step methods [34] for standard parabolic problems with a time-dependent coefficient.

#### 4.2 Error analysis for the inhomogeneous problem

Now we analyze the scheme (2.5) for  $u_0 \equiv 0$ . We need the following inequality.

**Lemma 8** *For any  $\beta \in (0, 1/2)$  and  $s \in [0, t_m]$ , the following inequality holds*

$$\tau \sum_{k=1}^m t_{m-k+1}^{\beta\alpha-1} (t_{k+1} - s)^{(1-2\beta)\alpha-1} \chi_{[0, t_k]}(s) \leq c(t_m - s)^{(1-\beta)\alpha-1}.$$

*Proof* We denote the left-hand side by  $I(s)$ . For any  $s \in [t_{i-1}, t_i]$ ,  $i \leq m$ ,

$$\begin{aligned} I(s) &= \tau \sum_{k=i}^m t_{m-k+1}^{\beta\alpha-1} (t_{k+1} - s)^{(1-2\beta)\alpha-1} \\ &\leq \tau \sum_{k=i}^m t_{m-k+1}^{\beta\alpha-1} t_{k+1-i}^{(1-2\beta)\alpha-1} \leq c t_{m-i+1}^{(1-\beta)\alpha-1} \leq c(t_m - s)^{(1-\beta)\alpha-1}. \end{aligned}$$

This completes the proof of the lemma.  $\square$

The next result gives a bound on  $g(t) = (A(0) - A(t))w(t)$  when  $u_0 \equiv 0$ .

**Lemma 9** *Let  $g(t) = (A(0) - A(t))w(t)$  (with  $u_0 \equiv 0$ ). Then there holds*

$$\|g'(0)\|_{L^2(\Omega)} \leq c\|f(0)\|_{L^2(\Omega)}, \quad (4.14)$$

and further, for any  $\beta \in (0, 1/2)$

$$\begin{aligned} & \tau \sum_{k=1}^m t_{m-k+1}^{\beta\alpha-1} t_k \int_0^{t_k} (t_k - s)^{(1-2\beta)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} \\ & \leq ct_m^{\alpha-1} \|f(0)\|_{L^2(\Omega)} + t_m^\alpha \|f'(0)\|_{L^2(\Omega)} + t_m \int_0^{t_m} (t_m - s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds. \end{aligned} \quad (4.15)$$

*Proof* It follows from Lemma 2 that

$$\begin{aligned} \|g'(t)\|_{L^2(\Omega)} & \leq \|(A(0) - A(t))w'(t)\|_{L^2(\Omega)} + \|A'(t)w(t)\|_{L^2(\Omega)} \\ & \leq ct\|u'(t)\|_{\dot{H}^2(\Omega)} + c\|u(t)\|_{\dot{H}^2(\Omega)}. \end{aligned}$$

Then by Theorem 2,  $\|g'(0)\|_{L^2(\Omega)} \leq c\|f(0)\|_{L^2(\Omega)}$ , showing the estimate (4.14).

Next, by Lemma 8, the left hand side (LHS) of (4.15) is bounded by

$$\begin{aligned} \text{LHS} & \leq t_m \int_0^{t_m} \left( \tau \sum_{k=1}^m t_{m-k+1}^{\beta\alpha-1} (t_k - s)^{(1-2\beta)\alpha-1} \chi_{[0, t_k]}(s) \right) \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} ds \\ & \leq ct_m \int_0^{t_m} (t_m - s)^{(1-\beta)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} ds. \end{aligned}$$

Since  $g''(t) = (A(0) - A(t))u''(t) - 2A'(t)u'(t) - A''(t)u(t)$ , Theorem 2 implies

$$\begin{aligned} \|g''(t)\|_{\dot{H}^{-2\beta}(\Omega)} & \leq ct\|u''(t)\|_{\dot{H}^{2-2\beta}(\Omega)} + c\|u'(t)\|_{\dot{H}^{2-2\beta}(\Omega)} + c\|u(t)\|_{\dot{H}^{2-2\beta}(\Omega)} \\ & \leq ct^{\beta\alpha-1} \|f(0)\|_{L^2(\Omega)} + ct^{\beta\alpha} \|f'(0)\|_{L^2(\Omega)} \\ & \quad + ct \int_0^t (t-s)^{\beta\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds. \end{aligned}$$

Combining the last two estimates yields the desired assertion.  $\square$

Now we can derive error estimates for the inhomogeneous problem.

**Theorem 5** *Let  $u$  and  $u^n$  be the solutions to (1.1) and (2.5) with  $u_0 = 0$  and  $f \in C^1([0, T]; L^2(\Omega))$  and  $\int_0^t (t-s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds < \infty$ , respectively. Then under conditions (3.1)–(3.2), there holds*

$$\begin{aligned} \|u(t_n) - u^n\|_{L^2(\Omega)} & \leq c\tau^2 \left( t_n^{\alpha-2} \|f(0)\|_{L^2(\Omega)} + t_n^{\alpha-1} \|f'(0)\|_{L^2(\Omega)} \right. \\ & \quad \left. + \int_0^{t_n} (t_n - s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds \right). \end{aligned}$$

*Proof* The overall proof strategy is similar to Theorem 4. First, we bound  $\vartheta^n := \bar{w}^n - w(t_n)$ . By Lemma 4(i), for any  $\beta \in [0, 1/2)$ , there holds

$$\|\bar{w}_1^n - w_1(t_n)\|_{H^{2\beta}(\Omega)} \leq c\tau^2 R(t_n).$$

with  $R(t_n)$  defined by

$$\begin{aligned} R(t_n) &= t_n^{(1-\beta)\alpha-2} \|f(0)\|_{L^2(\Omega)} + t_n^{(1-\beta)\alpha-1} \|f'(0)\|_{L^2(\Omega)} \\ &\quad + \int_0^{t_n} (t_{n+1} - s)^{(1-\beta)\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds. \end{aligned}$$

Meanwhile, for any  $\beta \in [0, 1/2)$ , by Lemma 4(i) and (4.14), with  $g(t) = (A(0) - A(t))u(t)$ ,

$$\begin{aligned} &\|\bar{w}_0^n - w_0(t_n)\|_{\dot{H}^{2\beta}(\Omega)} \\ &\leq c\tau^2 \left( t_n^{(1-\beta)\alpha-1} \|f(0)\|_{L^2(\Omega)} + \int_0^{t_n} (t_{n+1} - s)^{(1-2\beta)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} ds \right). \end{aligned}$$

Thus, by the splitting (4.8) and triangle inequality, for any  $\beta \in [0, 1/2)$ ,

$$\|\vartheta^n\|_{\dot{H}^{2\beta}(\Omega)} \leq c\tau^2 R(t_n) + c\tau^2 \int_0^{t_n} (t_{n+1} - s)^{(1-2\beta)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} ds.$$

Next we bound  $\varrho^n := w^n - \bar{w}^n$ , by repeating the argument for Lemma 7. The term I can be bounded as (4.13). Further, by (4.12) and Lemma 2, for any  $\beta \in (0, 1/2)$ ,

$$\text{II} \leq \tau \sum_{k=1}^m \|E_{\tau,m}^{m-k}(A(t_k) - A(0))\vartheta^k\|_{L^2(\Omega)} \leq c \sum_{k=1}^m t_{m-k+1}^{\beta\alpha-1} t_k \|\vartheta^k\|_{\dot{H}^{2\beta}(\Omega)}.$$

Then the preceding bound on  $\vartheta^n$  implies

$$\text{II} \leq c\tau^3 \sum_{k=1}^m t_{m-k+1}^{\beta\alpha-1} t_k \left( R(t_k) + \int_0^{t_k} (t_{k+1} - s)^{(1-2\beta)\alpha-1} \|g''(s)\|_{\dot{H}^{-2\beta}(\Omega)} ds \right).$$

This and (4.15) imply

$$\begin{aligned} \|\varrho^m\|_{L^2(\Omega)} &\leq c\tau \sum_{k=1}^m \|\varrho^k\|_{L^2(\Omega)} + c\tau^2 \left( t_m^{\alpha-1} \|f(0)\|_{L^2(\Omega)} + t_m^\alpha \|f'(0)\|_{L^2(\Omega)} \right. \\ &\quad \left. + t_m \int_0^{t_n} (t_n - s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds \right). \end{aligned}$$

Thus, by the discrete Gronwall's inequality from Lemma 11 with  $\mu = 1 - \alpha$ ,

$$\begin{aligned} \|\varrho^m\|_{L^2(\Omega)} &\leq c\tau^2 \left( t_m^{\alpha-1} \|f(0)\|_{L^2(\Omega)} + t_m^\alpha \|f'(0)\|_{L^2(\Omega)} \right. \\ &\quad \left. + t_m \int_0^{t_n} (t_n - s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds \right), \end{aligned}$$

where the constant  $c$  depends on  $\alpha$  as  $O(\alpha^{-1})$ . This and the bound on  $\vartheta^n$  with  $\beta = 0$  complete the proof.  $\square$

*Remark 3* The error estimate in Theorem 5 is identical with that for the subdiffusion model with a time-independent diffusion coefficient [14].

*Remark 4* In the proof of Theorems 4 and 5, (discrete) Gronwall's inequality was employed a few times to bound  $\varrho^m$ . This leads to a dependence on  $\alpha$  as  $1/\alpha$ , which is nevertheless uniformly bounded on  $\alpha$  for  $\alpha \rightarrow 1^-$ . Further, the constants in the bounds on  $\vartheta$  are also bounded. Thus, the constants in Theorems 4 and 5 are bounded as the fractional order  $\alpha \rightarrow 1^-$ . We refer to [4] for an in-depth discussion and many further references on the important issue of  $\alpha$ -robustness.

## 5 Numerical results and discussions

Now we present numerical results to illustrate the convergence behavior of the scheme (2.5). To this end, we consider the domain  $\Omega = (0, 1)$  and the subdiffusion model (1.1) with a time-dependent diffusion operator  $A(t) = -(2 + \cos(t))\Delta$ . We consider the following three examples:

- (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  $f = e^t(1 + \chi_{(0, \frac{1}{2})}(x))$ .
- (c)  $u_0(x) = 0$  and  $f = t^{0.5}x(1-x)$ .

To discretize the problem, we divide the domain  $\Omega$  into  $M$  subintervals of equal length  $h = 1/M$ . The numerical solutions are computed by the standard Galerkin FEM (with P1 element) in space, and BDF2-CQ in time. Since the spatial convergence was already studied in [18], we only study the temporal convergence below. To this end, we fix a small spatial mesh size  $h = 1/1000$  so that the spatial discretization error is negligible, and compute the  $L^2(\Omega)$  error:

$$e(t_N) = \|u_h^N - u_h(t_N)\|_{L^2(\Omega)}.$$

Since the exact semidiscrete solution  $u_h(t)$  is unavailable, we compute the reference solutions on a finer temporal mesh with a time stepsize  $\tau = 1/5000$ .

The numerical results for the homogeneous case (a) by the schemes (2.4) and (2.5) are presented in Tables 1 and 2, respectively. It is clearly observed that the vanilla BDF2-CQ scheme (2.4) can only achieve a first-order convergence, whereas the corrected scheme (2.5) achieves the desired second-order convergence. The convergence is fairly robust with respect to the fractional order  $\alpha$ , despite the low regularity of the initial data  $u_0$ . Further, the error is larger when the time  $t_N$  gets closer to zero, which agrees well with the regularity theory in that the second-order temporal derivative of the solution has strong singularity at  $t = 0$ , cf. Theorem 1.

The numerical results for Examples (b) and (c) are presented in Tables 3–5, where the source term  $f$  is smooth and nonsmooth in time, respectively. Note that for Example (c), the corrected and uncorrected schemes are identical, since  $f(0) \equiv 0$ . The observations from Example (a) remain valid for the inhomogeneous problems: the correction at the first step in the scheme (2.5)

**Table 1** Temporal errors  $e$  for Example (a), uncorrected BDF2-CQ (2.4) with  $\tau = 1/N$ .

$t_N$	$\alpha \backslash N$	10	20	40	80	160	320	rate
1	0.25	2.96e-4	1.49e-4	7.45e-5	3.73e-5	1.86e-5	9.32e-6	1.00
	0.50	4.12e-4	2.12e-4	1.07e-4	5.37e-5	2.69e-5	1.35e-5	1.00
	0.75	3.00e-4	1.62e-4	8.34e-5	4.22e-5	2.12e-5	1.06e-5	1.00
$10^{-3}$	0.25	1.16e-3	5.80e-4	2.89e-4	1.45e-4	7.23e-5	3.62e-5	1.00
	0.50	5.49e-3	2.70e-3	1.34e-3	6.59e-4	3.34e-4	1.67e-4	1.00
	0.75	5.18e-3	2.54e-3	1.26e-3	6.28e-4	3.13e-4	1.57e-4	1.00

**Table 2** Temporal errors  $e$  for Example (a), corrected BDF2-CQ (2.5) with  $\tau = 1/N$ .

$t_N$	$\alpha \backslash N$	10	20	40	80	160	320	rate
1	0.25	4.20e-5	9.77e-6	2.36e-6	5.79e-7	1.44e-7	3.57e-8	2.01
	0.50	8.82e-5	2.04e-5	4.91e-6	1.20e-6	2.98e-7	7.41e-8	2.01
	0.75	1.01e-4	2.34e-5	5.60e-6	1.37e-6	3.38e-7	8.41e-8	2.01
$10^{-3}$	0.25	1.50e-4	3.49e-5	8.44e-6	2.07e-6	5.14e-7	1.28e-7	2.01
	0.50	4.77e-4	1.13e-4	2.74e-5	6.77e-6	1.68e-6	4.19e-7	2.00
	0.75	3.68e-4	8.67e-5	2.11e-5	5.21e-6	1.29e-6	3.22e-7	2.00

can restore the desired second-order convergence, whereas the vanilla BDF2-CQ scheme (2.4) can only give a first-order convergence, and the convergence rate does not depend on the fractional order  $\alpha$ .

The second-order convergence of the scheme (2.5) in Theorem 5 requires suitable temporal regularity of the source  $f$ , i.e.,  $\int_0^t (t-s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds < \infty$ , in the absence of which, the convergence rate suffers from a loss. This is clearly observed from the numerical results in Table 5 for Example (c), where the source term  $f$  does not satisfy the condition. Actually, by means of interpolation, the theoretical convergence rate is  $O(\tau^{3/2})$ . The corrected scheme (2.5) can achieve a convergence rate  $O(\tau^{3/2})$ , which agrees well with the theoretical one and is faster than the first-order convergence as exhibited by the scheme (2.4). These numerical results show clearly the robustness and efficiency of the corrected scheme (2.5).

**Table 3** Temporal errors  $e$  for Example (b), uncorrected BDF2-CQ (2.4) with  $\tau = 1/N$ .

$t_N$	$\alpha \backslash N$	10	20	40	80	160	320	rate
1	0.25	2.22e-5	1.15e-5	5.81e-6	2.92e-6	1.46e-6	7.33e-7	1.00
	0.50	3.09e-5	1.64e-5	8.35e-6	4.21e-6	2.11e-6	1.06e-6	1.00
	0.75	2.36e-5	1.28e-5	6.57e-6	3.32e-6	1.67e-6	8.36e-6	1.00
$10^{-3}$	0.25	9.12e-5	4.56e-5	2.28e-5	1.14e-5	5.70e-6	2.85e-6	1.00
	0.50	4.32e-4	2.12e-4	1.05e-4	5.26e-5	2.62e-5	1.31e-5	1.00
	0.75	3.54e-4	1.74e-4	8.61e-5	4.28e-5	2.14e-5	1.07e-5	1.00

## Acknowledgements

The authors are grateful to two anonymous referees for their constructive comments which have led to an improvement in the quality of the paper.

**Table 4** Temporal errors  $e$  for Example (b), corrected BDF2-CQ (2.5) with  $\tau = 1/N$ .

$t_N$	$\alpha \backslash N$	10	20	40	80	160	320	rate
1	0.25	4.46e-6	1.04e-6	2.51e-7	6.16e-8	1.53e-8	3.80e-9	2.01
	0.50	8.51e-6	1.96e-6	4.70e-7	1.15e-7	2.85e-8	7.09e-9	2.01
	0.75	8.13e-6	1.85e-6	4.40e-7	1.07e-7	2.64e-8	6.56e-9	2.01
$10^{-3}$	0.25	1.18e-5	2.75e-6	6.64e-7	1.63e-7	4.05e-8	1.01e-8	2.01
	0.50	3.70e-5	8.75e-6	2.13e-6	5.26e-7	1.31e-7	3.26e-8	2.00
	0.75	5.66e-6	1.38e-6	3.42e-7	8.52e-8	2.12e-8	5.30e-9	2.00

**Table 5** Temporal errors  $e$  for Example (c), corrected BDF2-CQ (2.5) with  $\tau = 1/N$ .

$t_N$	$\alpha \backslash N$	50	100	200	400	800	1600	rate
1	0.25	2.31e-8	8.65e-9	3.18e-9	1.15e-9	4.11e-10	1.44e-10	1.52
	0.50	2.77e-8	1.11e-8	4.24e-9	1.58e-9	5.73e-10	2.02e-10	1.50
	0.75	6.59e-9	4.94e-9	2.40e-9	1.01e-9	3.93e-10	1.45e-10	1.44
$10^{-3}$	0.25	3.65e-9	1.27e-9	4.41e-10	1.54e-10	5.37e-11	1.84e-11	1.54
	0.50	1.72e-8	5.92e-9	2.05e-9	7.15e-10	2.48e-10	8.51e-11	1.55
	0.75	1.24e-8	4.37e-9	1.55e-9	5.45e-10	1.91e-10	6.59e-11	1.54

## A Gronwall's inequalities

In this appendix, we collect several useful Gronwall's inequalities. The following generalized Gronwall's inequality is useful [12, Exercise 4, p. 190].

**Lemma 10** *Let the function  $\varphi(t) \geq 0$  be continuous for  $0 < t \leq T$ . If*

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

*for some constants  $a, b \geq 0$ ,  $0 \leq \mu, \beta < 1$ , then there is a constant  $c = c(b, T, \beta)$  such that*

$$\varphi(t) \leq \frac{ac}{1-\mu} t^{-\mu}, \quad 0 < t \leq T.$$

The next result is a discrete analogue of Lemma 10 [8, Lemma 7.1].

**Lemma 11** *Let  $\varphi^n \geq 0$  for  $0 \leq t_n \leq T$ . If*

$$\varphi^n \leq at_n^{-\mu} + b\tau \sum_{j=1}^n \varphi^j, \quad 0 < t_n \leq T,$$

*for some constants  $a, b \geq 0$ , and  $b\tau < 1/2$ ,  $0 \leq \mu < 1$ , then there is constant  $c = c(b, T)$  such that*

$$\varphi^n \leq \frac{ac}{1-\mu} t_n^{-\mu}, \quad 0 < t_n \leq T.$$

*Proof* This lemma is a special case of [8, Lemma 7.1], but without explicit dependence on  $\alpha$ . We give a short proof for completeness following [37, p. 258]. Let  $\sigma^n = \tau \sum_{j=1}^n \varphi^j$ , and  $\phi^n = at_n^{-\mu}$ . Then

$$\tau^{-1}(\sigma^n - \sigma^{n-1}) \leq \phi^n + b\sigma^n,$$

i.e.,  $\sigma^n \leq (1 - b\tau)^{-1} \sigma^{n-1} + (1 - b\tau)^{-1} \tau \phi^n$ . Consequently, since the time interval is finite,

$$\sigma^n \leq \tau \sum_{j=1}^n (1 - b\tau)^{j-n-1} \phi^j \leq e^{2bT} \tau \sum_{j=1}^n \phi^j \leq \frac{ae^{2bT}}{1-\mu} t_n^{1-\mu},$$

since  $(1-x)^{-1} \leq e^{2x}$  for  $x \in [0, 1/2]$ . This directly shows the desired assertion.  $\square$

The next result is a variant of Lemma 11 with log factors [37, p. 258].

**Lemma 12** *Let  $\varphi^n \geq 0$  for  $0 \leq t_n \leq T$ . With  $\ell_n = \log(1 + t_n/\tau)$ , if*

$$\varphi^n \leq at_n^{-1}\ell_n^p + b\tau \sum_{j=1}^n \varphi^j, \quad 0 < t_n \leq T,$$

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

$$\varphi^n \leq cat_n^{-1}\ell_n^p, \quad 0 < t_n \leq T.$$

## B Proof of Theorem 2

In this part, we prove Theorem 2, by considering  $(t^k u(t))^{(k)}$ , instead of  $(t^{k+1} u(t))^{(k)}$  for the proof of Theorem 1. We begin with a bound on  $\frac{d^k}{dt^k} (t^k \int_0^t E(t-s; t_*) f(s) ds)$ , which follows from straightforward but lengthy computation.

**Lemma 13** *Let  $k \geq 1$ . Then for any  $\beta \in [0, 1)$ , there holds*

$$\begin{aligned} & \left\| A_*^\beta \frac{d^k}{dt^k} \left( t^k \int_0^t E_*(t-s) f(s) ds \right) \right\|_{L^2(\Omega)} \Big|_{t=t_*} \\ & \leq c \sum_{m=0}^{k-1} t_*^{(1-\beta)\alpha+m} \|f^{(m)}(0)\|_{L^2(\Omega)} + ct_*^k \int_0^{t_*} (t_*-s)^{(1-\beta)\alpha-1} \|f^{(k)}(s)\|_{L^2(\Omega)} ds, \end{aligned}$$

and further,

$$\begin{aligned} & \left\| A_* \frac{d^k}{dt^k} \left( t^k \int_0^t E_*(t-s) f(s) ds \right) \right\|_{L^2(\Omega)} \Big|_{t=t_*} \\ & \leq c \sum_{m=0}^k t_*^m \|f^{(m)}(0)\|_{L^2(\Omega)} + ct_*^k \int_0^{t_*} \|f^{(k+1)}(s)\|_{L^2(\Omega)} ds. \end{aligned}$$

*Proof* Let  $I(t) = \frac{d^k}{dt^k} (t^k \int_0^t E_*(t-s) f(s) ds)$ . It follows from the elementary identity

$$\frac{d^m}{dt^m} \int_0^t E_*(s) f(t-s) ds = \sum_{\ell=0}^{m-1} E_*^{(\ell)}(t) f^{(m-1-\ell)}(0) + \int_0^t E_*(s) f^{(m)}(t-s) ds$$

and direct computation that

$$I(t) = \sum_{m=0}^k \binom{m}{k}^2 t^m \left( \sum_{\ell=0}^{m-1} E_*^{(\ell)}(t) f^{(m-1-\ell)}(0) + \int_0^t E_*(s) f^{(m)}(t-s) ds \right).$$

Consequently, by Lemma 1, for  $\beta \in [0, 1)$ ,

$$\begin{aligned} \|A_*^\beta I(t^*)\|_{L^2(\Omega)} & \leq c \sum_{m=0}^k t_*^m \sum_{\ell=0}^{m-1} \|A_*^\beta E_*^{(\ell)}(t_*)\| \|f^{(m-1-\ell)}(0)\|_{L^2(\Omega)} \\ & \quad + c \sum_{m=0}^k t_*^m \int_0^{t_*} \|A_*^\beta E_*(s)\| \|f^{(m)}(t_*-s)\|_{L^2(\Omega)} ds \\ & \leq c \sum_{m=0}^k t_*^m \sum_{\ell=0}^{m-1} t_*^{(1-\beta)\alpha-1-\ell} \|f^{(m-1-\ell)}(0)\|_{L^2(\Omega)} \end{aligned}$$



$$\begin{aligned}
& + c \sum_{m=0}^k t_*^m \int_0^{t_*} s^{(1-\beta)\alpha-1} \|f^{(m)}(t_* - s)\|_{L^2(\Omega)} ds \\
& \leq c \sum_{m=0}^{k-1} t_*^{(1-\beta)\alpha+m} \|f^{(m)}(0)\|_{L^2(\Omega)} + c \sum_{m=0}^k t_*^m \int_0^{t_*} s^{(1-\beta)\alpha-1} \|f^{(m)}(t_* - s)\|_{L^2(\Omega)} ds.
\end{aligned}$$

Next we simplify the second summation. The following elementary identity: for  $m < k$ ,

$$f^{(m)}(s) = \sum_{j=0}^{k-m-1} f^{(m+j)}(0) \frac{s^j}{j!} + \frac{1}{(k-m)!} \int_0^s (s-\xi)^{k-m-1} f^{(k)}(\xi) d\xi, \quad (\text{B.1})$$

and the semigroup property of the Riemann-Liouville fractional integral operator imply

$$\begin{aligned}
& \int_0^{t_*} (t_* - s)^{(1-\beta)\alpha-1} \|f^{(m)}(s)\|_{L^2(\Omega)} ds \leq \int_0^{t_*} (t_* - s)^{(1-\beta)\alpha-1} \\
& \times \left( \sum_{j=0}^{k-m-1} \|f^{(m+j)}(0)\|_{L^2(\Omega)} \frac{s^j}{j!} + \frac{1}{(k-m)!} \int_0^s (s-\xi)^{k-m-1} \|f^{(k)}(\xi)\|_{L^2(\Omega)} d\xi \right) ds \\
& \leq c \sum_{j=0}^{k-m-1} t_*^{(1-\beta)\alpha+j} \|f^{(m+j)}(0)\|_{L^2(\Omega)} + c t_*^{k-m} \int_0^{t_*} (t_* - s)^{(1-\beta)\alpha-1} \|f^{(k)}(s)\|_{L^2(\Omega)} ds.
\end{aligned}$$

Combining these estimates gives the desired assertion for  $\beta \in [0, 1)$ . For  $\beta = 1$ , by the identity (3.8) and integration by parts (and the identity  $I - F_*(0) = 0$ ),

$$\begin{aligned}
A_* \int_0^{t_*} E_*(s) f^{(m)}(t_* - s) ds &= \int_0^{t_*} (I - F_*(s))' f^{(m)}(t_* - s) ds \\
&= (I - F_*(t_*)) f^{(m)}(0) + \int_0^{t_*} (I - F_*(s)) f^{(m+1)}(t_* - s) ds,
\end{aligned}$$

and thus

$$\begin{aligned}
\|A_* I(t_*)\|_{L^2(\Omega)} &\leq c \sum_{m=0}^k t_*^m \left( \sum_{\ell=0}^{m-1} \|A_* E_*^{(\ell)}(t_*)\| \|f^{(m-1-\ell)}(0)\|_{L^2(\Omega)} \right. \\
&\quad \left. + \|f^{(m)}(0)\| + \int_0^{t_*} \|f^{(m+1)}(s)\|_{L^2(\Omega)} ds \right).
\end{aligned}$$

Then repeating the preceding argument completes the proof.  $\square$

Now we can present the proof of Theorem 2.

*Proof* Similar to Theorem 1, the proof is based on mathematical induction. Let  $v_k(t) = t^k u(t)$  and  $W_k(t) = t^k E_*(t)$ . For  $k = 0$ , by the representation (3.11), we have

$$A_*^\beta u(t_*) = \int_0^{t_*} A_*^\beta E_*(t_* - s) f(s) ds + \int_0^{t_*} A_*^\beta E_*(t_* - s) (A_* - A(s)) u(s) ds.$$

Then for  $\beta \in [0, 1)$ , by Lemma 1(ii) and 2 there holds

$$\begin{aligned}
\|A_*^\beta u(t_*)\|_{L^2(\Omega)} &\leq \int_0^{t_*} \|A_*^\beta E_*(t_* - s)\| \|f(s)\|_{L^2(\Omega)} ds \\
&\quad + \int_0^{t_*} \|A_* E_*(t_* - s)\| \|A_*^\beta (I - A_*^{-1} A(s)) u(s)\|_{L^2(\Omega)} ds \\
&\leq c \int_0^{t_*} (t_* - s)^{(1-\beta)\alpha-1} \|f(s)\|_{L^2(\Omega)} ds + c \int_0^{t_*} \|A_*^\beta u(s)\|_{L^2(\Omega)} ds.
\end{aligned}$$

The case  $\beta = 1$  follows similarly from the identity (3.8) and integration by parts:

$$\|A_* u(t_*)\|_{L^2(\Omega)} \leq c\|f(0)\|_{L^2(\Omega)} + c \int_0^{t_*} \|f'(s)\|_{L^2(\Omega)} ds + c \int_0^{t_*} \|A_* u(s)\|_{L^2(\Omega)} ds.$$

Then standard Gronwall's inequality gives the assertion for the case  $k = 0$ . Now suppose the assertion holds up to  $k - 1 < K$ , and we prove it for  $k \leq K$ . Now note that

$$v_k^{(k)}(t) = \frac{d^k}{dt^k} \left( t^k \int_0^t E_*(t-s)f(s)ds \right) + \sum_{m=0}^k \binom{k}{m} \frac{d^k}{dt^k} \int_0^t W_m(t-s)(A_* - A(s))v_{k-m}(s)ds.$$

This, Lemmas 13 and 14 and triangle inequality give

$$\begin{aligned} \|A_*^\beta v_k^{(k)}(t)|_{t=t_*}\|_{L^2(\Omega)} &\leq c \int_0^{t_*} \|A_*^\beta v_k^{(k)}(s)\|_{L^2(\Omega)} ds \\ &+ \begin{cases} c \sum_{m=0}^{k-1} t_*^{(1-\beta)\alpha+m} \|f^{(m)}(0)\|_{L^2(\Omega)} + ct_*^k \int_0^{t_*} (t_* - s)^{(1-\beta)\alpha-1} \|f^{(k)}(s)\|_{L^2(\Omega)} ds, & \beta \in [0, 1), \\ c \sum_{m=0}^k t_*^m \|f^{(m)}(0)\|_{L^2(\Omega)} + ct_*^k \int_0^{t_*} \|f^{(k+1)}(s)\|_{L^2(\Omega)} ds, & \beta = 1. \end{cases} \end{aligned}$$

This and the standard Gronwall's inequality complete the induction step.  $\square$

The following result is needed in the proof of Theorem 2.

**Lemma 14** *Under the conditions in Theorem 2, for any  $\beta \in [0, 1]$  and  $m = 0, \dots, k$ , there holds*

$$\begin{aligned} &\left\| A_*^\beta \frac{d^k}{dt^k} \int_0^t (t-s)^{k-m} E_*(t-s)(A_* - A(s))s^m u(s)ds \Big|_{t=t_*} \right\|_{L^2(\Omega)} \\ &\leq c_0 \int_0^{t_*} \|A_*^\beta (s^k u(s))^{(k)}\|_{L^2(\Omega)} ds \\ &+ \begin{cases} c \sum_{m=0}^{k-1} t_*^{(1-\beta)\alpha+m} \|f^{(m)}(0)\|_{L^2(\Omega)} + ct_*^k \int_0^{t_*} (t_* - s)^{(1-\beta)\alpha-1} \|f^{(k)}(s)\|_{L^2(\Omega)} ds, & \beta \in [0, 1), \\ c \sum_{m=0}^k t_*^m \|f^{(m)}(0)\|_{L^2(\Omega)} + ct_*^k \int_0^{t_*} \|f^{(k+1)}(s)\|_{L^2(\Omega)} ds, & \beta = 1. \end{cases} \end{aligned}$$

*Proof* Let  $v_k = t^k u(t)$  and  $W_k(t) = t^k E_*(t)$ . By the induction hypothesis and (B.1), for  $\ell < m$ , we have

$$\begin{aligned} \|A_* v_m^{(\ell)}(s)\|_{L^2(\Omega)} &\leq cs^{m-\ell} \left( \sum_{j=0}^{\ell} s^j \|f^{(j)}(0)\|_{L^2(\Omega)} + s^\ell \int_0^s \|f^{(\ell+1)}(\xi)\|_{L^2(\Omega)} d\xi \right) \\ &\leq cs^{m-\ell} \left( \sum_{j=0}^{k-1} s^j \|f^{(j)}(0)\|_{L^2(\Omega)} + s^{k-1} \int_0^s \|f^{(k)}(\xi)\|_{L^2(\Omega)} d\xi \right). \quad (\text{B.2}) \end{aligned}$$

We denote the term in the bracket by  $T(s; f, k)$ . Now similar to the proof of Lemma 3, let  $I_m(t)$  be the integral on the left hand side. Then in view of the identity

$$I_m^{(k)}(t) = \sum_{\ell=0}^{k-m} \binom{\ell}{k-m} \underbrace{\int_0^t W_m^{(m)}(A_* - A(t_* - s))^{(k-m-\ell)} v_{k-m}^{(\ell)}(s) ds}_{I_{m,\ell}(t)},$$

it suffices to bound the integrand  $\tilde{I}_{m,\ell}(s)$  of the integral  $I_{m,\ell}(t_*)$ ,  $\ell = 0, 1, \dots, k-m$ . Below we discuss the cases  $\beta \in [0, 1)$  and  $\beta = 1$  separately, due to the difference in singularity, as in the proof of Lemma 3.

*Case (i):*  $\beta \in [0, 1)$ . For the case  $\ell < k$ , Lemmas 1(ii) and 2 lead to

$$\begin{aligned} \|A_*^\beta \tilde{I}_{m,\ell}(s)\|_{L^2(\Omega)} &\leq \|A_*^\beta W_m^{(m)}\| \| (A_* - A(t_* - s))^{(k-m-\ell)} v_{k-m}^{(\ell)}(t_* - s) \|_{L^2(\Omega)} \\ &\leq \begin{cases} cs^{(1-\beta)\alpha-1} s \|A_* v_{k-m}^{(k-m)}(t_* - s)\|_{L^2(\Omega)}, & \ell = k-m, \\ cs^{(1-\beta)\alpha-1} \|A_* v_{k-m}^{(\ell)}(t_* - s)\|_{L^2(\Omega)}, & \ell < k-m, \end{cases} \\ &\leq \begin{cases} cs^{(1-\beta)\alpha} T(t_* - s; f, k), & \ell = k-m, \\ cs^{(1-\beta)\alpha-1} (t_* - s)^{k-m-\ell} T(t_* - s; f, k), & \ell < k-m, \end{cases} \end{aligned}$$

where the last step is due to (B.2). Note that for  $\ell < k$ , the derivation requires  $\beta \in [0, 1)$ . Similarly, for the case  $\ell = k$  (and thus  $m = 0$ ),

$$\|A_*^\beta \tilde{I}_{0,k}(s)\|_{L^2(\Omega)} \leq c \|A_*^\beta v_k^{(k)}(t_* - s)\|_{L^2(\Omega)}. \quad (\text{B.3})$$

*Case (ii):*  $\beta = 1$ . Note that for  $\ell < k$ , the derivation in case (i) requires  $\beta \in [0, 1)$ . When  $\ell < k$  and  $\beta = 1$ , using the identity (3.13) and Lemma 1 and repeating the argument of Lemma 3, we obtain

$$\|A_* \tilde{I}_{m,\ell}(s)\|_{L^2(\Omega)} \leq \begin{cases} c(t_* - s)^{k-m-\ell-1} T(t_* - s; f, k), & \ell < k-1, \\ cT(t_* - s; f, k) + c \|A_* v_k^{(k)}(t_* - s)\|_{L^2(\Omega)}, & \ell = k-1, \end{cases}$$

Combining the preceding estimates, integrating from 0 to  $t_*$  in  $s$  and then applying standard Gronwall's inequality complete the proof.  $\square$

## References

1. A. A. Alikhanov. A new difference scheme for the time fractional diffusion equation. *J. Comput. Phys.*, 280:424–438, 2015.
2. L. Banjai and M. López-Fernández. Efficient high order algorithms for fractional integrals and fractional differential equations. *Numer. Math.*, 141(2):289–317, 2019.
3. H. Brunner. *Collocation Methods for Volterra Integral and Related Functional Differential Equations*. Cambridge University Press, Cambridge, 2004.
4. H. Chen and M. Stynes. Blow-up of error estimates in time-fractional initial-boundary value problems. Preprint, 2019.
5. E. Cuesta, C. Lubich, and C. Palencia. Convolution quadrature time discretization of fractional diffusion-wave equations. *Math. Comp.*, 75(254):673–696, 2006.
6. M. H. Dodson. Closure temperature in cooling geochronological and petrological systems. *Contrib. Mineral. Petrol.*, 40(3):259–274, 1973.
7. H. Dong and D. Kim.  $L_p$ -estimates for time fractional parabolic equations with coefficients measurable in time. *Adv. Math.*, 345:289–345, 2019.
8. 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.
9. K. S. Fa and E. K. Lenzi. Time-fractional diffusion equation with time dependent diffusion coefficient. *Phys. Rev. E*, 72:011107, 4 pp., 2005.
10. R. Garra, A. Giusti, and F. Mainardi. The fractional dodson diffusion equation: a new approach. *Ricerche di Matematica*, 67(2):899–909, 2018.
11. R. Garra, E. Orsingher, and F. Polito. Fractional diffusions with time-varying coefficients. *J. Math. Phys.*, 56(9):093301, 17 pp., 2015.
12. D. Henry. *Geometric Theory of Semilinear Parabolic Equations*. Springer-Verlag, Berlin-New York, 1981.

13. H. G. E. Hentschel and I. Procaccia. Relative diffusion in turbulent media: The fractal dimension of clouds. *Phys. Rev. A*, 29(3):1461–1471, 1984.
14. 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.
15. B. Jin, R. Lazarov, and Z. Zhou. Numerical methods for time-fractional evolution equations with nonsmooth data: A concise overview. *Comput. Methods Appl. Mech. Engrg.*, 346:332–358, 2019.
16. B. Jin, B. Li, and Z. Zhou. Correction of high-order BDF convolution quadrature for fractional evolution equations. *SIAM J. Sci. Comput.*, 39(6):A3129–A3152, 2017.
17. B. Jin, B. Li, and Z. Zhou. Numerical analysis of nonlinear subdiffusion equations. *SIAM J. Numer. Anal.*, 56(1):1–23, 2018.
18. B. Jin, B. Li, and Z. Zhou. Subdiffusion with a time-dependent coefficient: analysis and numerical solution. *Math. Comp.*, 88(319):2157–2186, 2019.
19. A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and Applications of Fractional Differential Equations*. Elsevier Science B.V., Amsterdam, 2006.
20. J. Klafter, A. Blumen, and M. F. Shlesinger. Stochastic pathway to anomalous diffusion. *Phys. Rev. A*, 35(7):3081–3085, 1987.
21. A. Kubica and M. Yamamoto. Initial-boundary value problems for fractional diffusion equations with time-dependent coefficients. *Fract. Calc. Appl. Anal.*, 21(2):276–311, 2018.
22. H. Liao, W. McLean, and J. Zhang. A discrete Grönwall inequality with applications to numerical schemes for subdiffusion problems. *SIAM J. Numer. Anal.*, 57(1):218–237, 2019.
23. Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.*, 225(2):1533–1552, 2007.
24. C. Lubich. Discretized fractional calculus. *SIAM J. Math. Anal.*, 17(3):704–719, 1986.
25. C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.*, 52(2):129–145, 1988.
26. 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.
27. C. Lv and C. Xu. Error analysis of a high order method for time-fractional diffusion equations. *SIAM J. Sci. Comput.*, 38(5):A2699–A2724, 2016.
28. W. McLean and K. Mustapha. Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. *Numer. Algorithms*, 52(1):69–88, 2009.
29. R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.*, 16:24128, 37 pp., 2014.
30. R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339(1):1–77, 2000.
31. K. Mustapha. FEM for time-fractional diffusion equations, novel optimal error analyses. *Math. Comp.*, 87(313):2259–2272, 2018.
32. E. Orsingher and F. Polito. Some results on time-varying fractional partial differential equations and birth-death processes. In *Proceedings of the XIII International Conference on Eventological Mathematics and Related Fields*, pages 23–27, Krasnoyarsk, 2009.
33. 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.
34. P. Sammon. Fully discrete approximation methods for parabolic problems with nonsmooth initial data. *SIAM J. Numer. Anal.*, 20(3):437–470, 1983.
35. 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.
36. Z.-Z. Sun and X. Wu. A fully discrete difference scheme for a diffusion-wave system. *Appl. Numer. Math.*, 56(2):193–209, 2006.
37. V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin, 2nd edition, 2006.
38. H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland Publishing Co., Amsterdam-New York, 1978.