TWO FULLY DISCRETE SCHEMES FOR FRACTIONAL DIFFUSION AND DIFFUSION-WAVE EQUATIONS WITH NONSMOOTH DATA

BANGTI JIN[*], RAYTCHO LAZAROV[†], AND ZHI ZHOU[†‡]

Abstract. We consider initial/boundary value problems for the subdiffusion and diffusion-wave equations involving a Caputo fractional derivative in time. We develop two fully discrete schemes based on the piecewise linear Galerkin finite element method in space and convolution quadrature in time with the generating function given by the backward Euler method/second-order backward difference method, and establish error estimates optimal with respect to the regularity of problem data. These two schemes are first and second-order accurate in time for both smooth and nonsmooth data. Extensive numerical experiments for two-dimensional problems confirm the convergence analysis and robustness of the schemes with respect to data regularity.

Key words. fractional diffusion, diffusion wave, finite element method, convolution quadrature, error estimate

AMS subject classifications. 65M60, 65N30, 65N15

1. Introduction.

1.1. Mathematical model. In this work, we develop robust numerical schemes for the subdiffusion and diffusion wave equations. Let $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) be a bounded convex polygonal domain with a boundary $\partial\Omega$, and $T > 0$ be a fixed time. Then the mathematical model is given by

$$^{C}\partial_t^\alpha u(x,t) - \Delta u(x,t) = f(x,t) \quad (x,t) \in \Omega \times (0,T), \tag{1.1}$$

where f is a given source term. Here $^{C}\partial_t^\alpha u$ denotes the Caputo fractional derivative with respect to time t of order α, and it is defined by [23, pp. 91, eq. (2.4.1)]

$$^{C}\partial_t^\alpha u(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-s)^{n-\alpha-1} \frac{d^n}{ds^n} u(s) ds, \quad n-1 < \alpha < n, \, n \in \mathbb{N}, \tag{1.2}$$

where Γ is the Gamma function. We shall also need the Riemann-Liouville fractional derivative $\partial_t^\alpha u$ defined by [23, pp. 70, eq. (2.1.5)]

$$\partial_t^\alpha u = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_0^t (t-s)^{n-\alpha-1} u(s) ds, \quad n-1 < \alpha < n, \, n \in \mathbb{N}. \tag{1.3}$$

Note that if $\alpha = 1$ and $\alpha = 2$, then equation (1.1) represents a parabolic and a hyperbolic equation, respectively. In this paper we focus on the fractional cases $0 < \alpha < 1$ and $1 < \alpha < 2$, with a Caputo derivative, which are known as the subdiffusion and diffusion-wave equation, respectively, in the literature. In analogy with Brownian motion for normal diffusion, the model (1.1) with $0 < \alpha < 1$ is the macroscopic counterpart of continuous time random walk [38, 13].

Throughout, equation (1.1) is subjected to the following boundary condition

$$u(x,t) = 0, \quad (x,t) \in \partial\Omega \times (0,T),$$

---

[*]Department of Computer Science, University College London, Gower Street, London, WC1E 6BT, UK (b.jin@ucl.ac.uk, bangti.jin@gmail.com)

[†]Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA (lazarov, zzhou@math.tamu.edu)

[‡]Department of Applied Mathematics and Applied Physics, Columbia University, 500 W. 120th St., New York, NY 10027, USA (zhizhou0125@gmail.com)

and the initial condition(s)

$$u(x,0) = v(x), \quad x \in \Omega, \quad \text{if} \quad 0 < \alpha < 1,$$
$$u(x,0) = v(x), \quad \partial_t u(x,0) = b(x), \quad x \in \Omega, \text{ if } 1 < \alpha < 2.$$

The model (1.1) has received much attention over the last few decades, since it can adequately capture the dynamics of physical processes involving anomalous transport mechanism. For example, the subdiffusion equation has been successfully used to describe thermal diffusion in media with fractal geometry [42], and highly heterogeneous aquifers [1, 15]. The diffusion wave equation can be used to model the propagation of mechanical waves in viscoelastic media [33, 34]. Further, we refer to [23, 44, 8] for physical applications and mathematical theory, and [2] for a comprehensive survey on numerical methods for fractional ordinary differential equations.

1.2. Review of existing schemes. Due to the excellent modeling capability of problem (1.1), its accurate numerical treatment has been the subject of numerous studies. A number of efficient schemes, notably based on finite difference in space and various discretizations in time, have been developed. Their error analysis is often based on Taylor expansion and the error bounds are expressed in terms of solution smoothness. The analysis of many existing methods for problem (1.1) requires that the solution be a C^2 or C^3-function in time, cf. Table 1.1, where τ is the time step size; see Section 2.3 for further details. However, such an assumption is not always valid since often the solution does not have the requisite regularity. For example, the C^2 in time regularity assumption on the solution does not hold for the homogeneous subdiffusion problem. Specifically, for initial data $v \in L^2(\Omega)$, $f = 0$ and $\alpha \in (0,1)$, the following estimate holds true [45, Theorem 2.1]:

$$\|{}^C\partial_t^\alpha u(t)\|_{L^2(\Omega)} \le c t^{-\alpha} \|v\|_{L^2(\Omega)}. \tag{1.4}$$

This shows that the α-th order Caputo derivative in time is already unbounded near $t = 0$. Hence, the convergence rates listed in Table 1.1 may not hold for nonsmooth data. Actually, the limited solution regularity underlies the main technical difficulty in developing robust numerical schemes and in carrying out a rigorous error analysis.

Table 1.1: The convergence rates for existing schemes for subdiffusion ($0 < \alpha < 1$). In the table, RL denotes the Riemann-Liouville fractional derivative, τ is the time step size, and $\bar{u}$ is the zero extension in time of u to $\mathbb{R}$.

method	rate	derivative	regularity assumption
Lin-Xu [29]	$O(\tau^{2-\alpha})$	Caputo	$\forall x \in \Omega$, u is C^2 in t
Zeng et al I [52]	$O(\tau^{2-\alpha})$	Caputo	$\forall x \in \Omega$, u is C^2 in t
Zeng et al II [52]	$O(\tau^{2-\alpha})$	Caputo	$\forall x \in \Omega$, u is C^2 in t
Li-Xu [26]	$O(\tau^2)$	Caputo	$\forall x \in \Omega$, u is C^3 in t
Gao et al [12]	$O(\tau^{3-\alpha})$	Caputo	$\forall x \in \Omega$, u is C^3 in t
L1 scheme [43, 24]	$O(\tau^{2-\alpha})$	RL	$\forall x \in \Omega$, u is C^2 in t
SBD [25]	$O(\tau^2)$	RL	$\forall x \in \Omega$, ${}^R_{-\infty}D_t^{3-\alpha}\bar{u}$ is L^1 in t

In the subdiffusion case, there are two predominant discretization techniques (in time): the L1-type approximation [43, 24, 29, 48, 27, 28, 12] and the Grünwald-Letnikov approximation [51, 4, 52]; see Table 1.1 for a summary. To the first group

belongs the method devised by Langlands and Henry [24]. They analyzed the discretization error for the Riemann-Liouville derivative. Also, Lin and Xu [29] (see also [43, pp. 140]) developed the L1 scheme (of finite difference nature) for the Caputo derivative and a Legendre spectral method in space, and analyzed the stability and convergence rates. It has a local truncation error $O(\tau^{2-\alpha})$. Li and Xu [27] developed a space-time spectral element method, but only for zero initial data; see also [28, 9]. In [26] a variant of the L1 approximation was analyzed, and a convergence rate $O(\tau^2)$ was established for C^3 solutions. Recently, a new L1-type formula based on quadratic interpolation was derived in [12] with a convergence rate $O(\tau^{3-\alpha})$ for smooth solutions. We also refer readers to [36, 40, 41] for studies on discontinuous Galerkin discretization of the Riemann-Liouville derivative in time, and [5, 6] for spectral methods, which merits exponential convergence for smooth solutions.

In the second group, Yuste and Acedo [51] suggested a Grünwald-Letnikov discretization of the Riemann-Liouville derivative and central finite difference in space, and provided a von Neumann type stability analysis. Zeng et al [52] developed two numerical schemes of the order $O(\tau^{2-\alpha})$ based on an integral reformulation of problem (1.1), a fractional linear multistep method in time and finite element method (FEM) in space, and analyzed their stability and convergence. However, the schemes are not robust with respect to data regularity; see Remark 2.3 and the comparative study in Section 4. Convolution quadrature [30, 31] provides a systematic framework for deriving high-order schemes for the Riemann-Liouville derivative, and has been the foundation of many works (see e.g., [51, 50, 4] for some early works). However, the error estimates in these works were derived under the assumption that the solution is sufficiently smooth. Further, all works, except [52], focus exclusively on the Riemann-Liouville derivative; and high-order methods were scarcely applied, despite that they can be conveniently analyzed even for nonsmooth data [32, 7].

The study on the diffusion wave equation is scarce. In [48], a Crank-Nicolson scheme was developed and its stability and convergence rate were shown. With $b = 0$ and $f = 0$, under suitable regularity assumptions, problem (1.1) can be rewritten as

$$\partial_t u = \frac{1}{\Gamma(\alpha - 1)} \int_0^t (t - s)^{\alpha - 2} \Delta u(s) ds,$$

with an initial condition $u(0) = v$. This model has been intensively studied [32, 37, 7], where convolution quadrature and Laplace transform method were analyzed. The error estimates derived in these works [32, 37, 7] cover the nonsmooth case. This model is closely connected to (1.1), but these problems have different smoothing properties in the inhomogeneous case [35, 7, 45]. To the best of our knowledge, convolution quadrature for (1.1) with a Caputo derivative and $1 < \alpha < 2$ has not been studied.

The excessive smoothness required in existing error analysis and the lack of convolution quadrature type schemes for the diffusion-wave equation with a Caputo derivative motivate us to revisit these issues. The goal of this work is to develop robust schemes based on convolution quadrature for the model (1.1) and to derive optimal error bounds that are expressed in terms of problem data, including nonsmooth data, e.g., $v \in L^2(\Omega)$, which is important in inverse problems and optimal control [21, 22].

1.3. Contributions and organization of the paper. In this work, we shall develop two fully discrete schemes for problem (1.1) based on convolution quadrature in time generated by backward Euler or second-order backward difference and the piecewise linear Galerkin FEM in space. This is achieved by reformulating problem

3

(1.1) using a Riemann-Liouville derivative. To the best of our knowledge, our application of convolution quadrature to the Caputo derivative is new, especially for the diffusion wave case, and for the first time a second-order scheme is obtained for problem (1.1) for both smooth and nonsmooth data. We shall establish optimal convergence rates in either case. This is in sharp contrast with existing works on the model (1.1), where the convergence analysis is mostly done under unverifiable solution regularity assumptions. The error analysis exploits an operator trick [10] and a general strategy developed in [7]. Extensive two-dimensional (in space) experiments confirm the convergence theory and their robustness with respect to data regularity.

To illustrate the features of our schemes, we describe one result from Theorem 3.5: for $0 < \alpha < 1$, $v \in L^2(\Omega)$, $v_h = P_h v$ (with P_h being the L^2 projection) and $f = 0$, for $n \geq 1$, the fully discrete approximation U_h^n obtained by the backward Euler method (with a mesh size h and time step size τ) satisfies the following error bound

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leq c(t_n^{-1}\tau + t_n^{-\alpha}h^2)\|v\|_{L^2(\Omega)}. \tag{1.5}$$

This estimate differs from those listed in Table 1.1 in several aspects:
  (a) For any fixed t_n, the scheme is first-order accurate in time.
  (b) It deteriorates near $t = 0$, whereas that in Table 1.1 are uniform in t. The factor t_n^{-1} reflects the singularity behavior (1.4) for initial data $v \in L^2(\Omega)$.
  (c) The scheme is robust with respect to the regularity of the initial data v in the sense that for fixed t_n, the first-order in time and second-order in space convergence rates hold for both smooth and nonsmooth data.

The rest of the paper is organized as follows. In Section 2, we develop two fully discrete schemes using the Galerkin FEM in space and convolution quadrature in time. The error analysis of the schemes is given in Section 3. In Section 4, we present extensive numerical experiments to illustrate the convergence behavior of the methods. A comparison with existing methods is also included. In Appendix A, we collect the solution theory for the diffusion wave equation. Throughout, the notation c denotes a generic constant, which may differ at different occurrences, but it is always independent of the solution u, the mesh size h and the time step size τ.

2. Fully discrete schemes. In this part, we develop two fully discrete schemes, using the standard Galerkin FEM in space and convolution quadrature in time.

2.1. Space semidiscrete Galerkin FEM. Let $\mathcal{T}_h$ be a shape regular and quasi-uniform triangulation of the domain Ω into d-simplexes, denoted by T and called finite elements. Then over the triangulation $\mathcal{T}_h$, we define a continuous piecewise linear finite element space X_h by

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

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

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

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

$$({}^C\partial_t^\alpha u_h, \chi) + a(u_h, \chi) = (f, \chi) \quad \forall \chi \in X_h, \tag{2.1}$$

4

where the bilinear form $a(u, \chi)$ and the initial data are given by $a(u, \chi) = (\nabla u, \nabla \chi)$, and $u_h(0) = v_h$ and if $1 < \alpha < 2$, also $\partial_t u_h(0) = b_h$. Here $v_h, b_h \in X_h$ are approximations to the initial data v and b, respectively. Following [49], we choose $v_h \in X_h$ (and similarly for $b_h \in X_h$) depending on the smoothness of the data: $v_h = R_h v$ if $v \in \dot{H}^2(\Omega)$ and $v_h = P_h v$ if $v \in L^2(\Omega)$.

Upon introducing the discrete Laplacian $\Delta_h : X_h \to X_h$ defined by

$$-(\Delta_h \varphi, \chi) = (\nabla \varphi, \nabla \chi) \quad \forall \varphi, \chi \in X_h, \tag{2.2}$$

$f_h(t) = P_h f(t)$, and $A_h = -\Delta_h$, the semidiscrete scheme (2.1) can be rewritten into

$$^C\partial_t^\alpha u_h(t) + A_h u_h(t) = f_h(t), \ t > 0 \tag{2.3}$$

with $u_h(0) = v_h \in X_h$, and if $1 < \alpha < 2$, $\partial_t u_h(0) = b_h \in X_h$.

REMARK 2.1. *Upon minor modifications, our discussions extend to more general sectorial operators, including a strongly elliptic second-order differential operator $Au = -\nabla \cdot (a(x)\nabla u)$, where the conductivity tensor $a(x) : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is smooth and has a positive minimum eigenvalue $\lambda_{\min}(a(x)) \geq c_0$ for some $c_0 > 0$ almost everywhere, and the Riemann-Liouville derivative operator of order $\beta \in (1, 2)$ (with a zero Dirichlet boundary condition) [17]. Further, it is trivial to extend the fully discrete schemes to the multi-term subdiffusion/diffusion-wave problem.*

2.2. Fully discrete schemes. Now we develop two fully discrete schemes for problem (1.1). This is achieved by first reformulating problem (2.3) with a Riemann-Liouville derivative ∂_t^α and then applying convolution quadrature. Specifically, we rewrite the semidiscrete problem (2.3) using the defining relation of the Caputo derivative. Namely, for $n - 1 < \alpha < n$, there holds [23, pp. 91, equation (2.4.10)]

$$^C\partial_t^\alpha \varphi(t) := \partial_t^\alpha \left[\varphi(t) - \sum_{k=0}^{n-1} \frac{\varphi^{(k)}(0)}{k!} t^k \right].$$

In particular, for subdiffusion, $0 < \alpha < 1$, $^C\partial_t^\alpha \varphi = \partial_t^\alpha(\varphi(t) - \varphi(0))$, and diffusion-wave, $1 < \alpha < 2$, $^C\partial_t^\alpha \varphi(t) = \partial_t^\alpha(\varphi(t) - \varphi(0) - t\varphi'(0))$. Hence, for $t > 0$, the semidiscrete scheme (2.3) can be respectively recast into

$$\partial_t^\alpha (u_h - v_h) + A_h u_h = f_h, \quad \text{for} \quad 0 < \alpha < 1 \tag{2.4}$$

$$\partial_t^\alpha (u_h - v_h - t b_h) + A_h u_h = f_h, \quad \text{for} \quad 1 < \alpha < 2, \tag{2.5}$$

where $f_h = P_h f(t)$. The formulae (2.4) and (2.5) form the basis for time discretization, which is done in the elegant framework – convolution quadrature – developed in [7, Sections 2 and 3], initiated in [30, 31]. Below we describe this framework.

Let K be a complex valued or operator valued function which is analytic in a sector $\Sigma_\theta := \{z \in \mathbb{C} : |\arg z| \leq \theta\}$, $\theta \in (\pi/2, \pi)$ and bounded by

$$\|K(z)\| \leq M|z|^{-\mu} \quad \forall z \in \Sigma_\theta, \tag{2.6}$$

for some $\mu, M \in \mathbb{R}$. Then $K(z)$ is the Laplace transform of a distribution k on the real line, which vanishes for $t < 0$, has its singular support empty or concentrated at $t = 0$, and which is an analytic function for $t > 0$. For $t > 0$, the analytic function $k(t)$ is given by the inversion formula

$$k(t) = \frac{1}{2\pi i} \int_\Gamma K(z) e^{zt} dz, \quad t > 0,$$

5

where Γ is a contour lying in the sector of analyticity, parallel to its boundary and oriented with an increasing imaginary part. With ∂_t being time differentiation, we define $K(\partial_t)$ as the operator of (distributional) convolution with the kernel k: $K(\partial_t)g = k * g$ for a function $g(t)$ with suitable smoothness. Further, the convolution rule of Laplace transform gives the following associativity property: for the operators K_1 and K_2 (generated by the kernels k_1 and k_2), we have

$$K_1(\partial_t)K_2(\partial_t) = (K_1 K_2)(\partial_t). \tag{2.7}$$

For time discretization, we divide the interval $[0, T]$ into a uniform grid with a time step size $\tau = T/N$, $N \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_N = T$, $t_n = n\tau$, $n = 0, \ldots, N$. Then the convolution quadrature $K(\bar{\partial}_\tau)g(t)$ of $K(\partial_t)g(t)$ is given by [31]

$$K(\bar{\partial}_\tau)g(t) = \sum_{0 \leq j\tau \leq t} \omega_j g(t - j\tau), \quad t > 0, \tag{2.8}$$

where the quadrature weights $\{\omega_j\}_{j=0}^\infty$ are determined by $\sum_{j=0}^\infty \omega_j \xi^j = K(\delta(\xi)/\tau)$. Here δ is the quotient of the generating polynomials of a stable and consistent linear multistep method [14, Chapter 3]. In this work, we consider the backward Euler (BE) method and second-order backward difference (SBD) method, for which

$$\delta(\xi) = \begin{cases} (1 - \xi), & \text{BE}, \\ (1 - \xi) + (1 - \xi)^2/2, & \text{SBD}. \end{cases}$$

The weights $\{\omega_j\}$ can be computed efficiently via fast Fourier transform [44]. The associativity property is also valid for convolution quadrature:

$$K_1(\bar{\partial}_\tau)K_2(\bar{\partial}_\tau) = (K_1 K_2)(\bar{\partial}_\tau) \quad \text{and} \quad K_1(\bar{\partial}_\tau)(k * g) = (K_1(\bar{\partial}_\tau)k) * g. \tag{2.9}$$

Now we are ready to derive fully discrete schemes. Hence we rewrite the scheme (2.4) in a convolution form

$$u_h(t) = (\partial_t^\alpha + A_h)^{-1}\partial_t^\alpha v_h + (\partial_t^\alpha + A_h)^{-1}f_h. \tag{2.10}$$

Then the associativity property (2.9) yields the BE scheme for the case $0 < \alpha < 1$:

$$U_h^n = (\bar{\partial}_\tau^\alpha + A_h)^{-1}\bar{\partial}_\tau^\alpha v_h + (\bar{\partial}_\tau^\alpha + A_h)^{-1}f_h. \tag{2.11}$$

Equivalently, it can be stated as: find U_h^n for $n = 1, 2, \ldots, N$ such that

$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha v_h + F_h^n, \quad \text{with} \quad U_h^0 = v_h, \quad F_h^n = P_h f(t_n). \tag{2.12}$$

In the same manner we derive a fully discrete scheme for the diffusion-wave equation: find U_h^n for $n = 1, 2, ..., N$ such that

$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha v_h + (\bar{\partial}_\tau^\alpha t)b_h + F_h^n, \quad \text{with} \quad U_h^0 = v_h, \quad F_h^n = P_h f(t_n). \tag{2.13}$$

However, in view of the regularity result in Theorem A.4, we correct the last term F_h^n in the scheme (2.13) to $\bar{\partial}_\tau \partial_t^{-1} f_h(t_n)$, in order to obtain better error estimates, cf. Theorem 3.6 and the remark afterwards. Hence we arrive at the following corrected scheme: with $G_h^n = \partial_t^{-1} f_h(t_n)$, find U_h^n for $n = 1, \ldots, N$ such that

$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha v_h + (\bar{\partial}_\tau^\alpha t)b_h + \bar{\partial}_\tau G_h^n, \quad U_h^0 = v_h. \tag{2.14}$$

In the BE scheme, the first-order convergence remains valid even if $g(0) \neq 0$ [46, 32].

Next we turn to the SBD scheme. It is well known that the basic scheme (2.8) is only first-order accurate if $g(0) \neq 0$, e.g., for $g \equiv 1$ [31, Theorem 5.1] [7, Section 3]. Hence, to get a second-order convergence one has to correct (2.8) properly. We follow the approach proposed in [32, 7]. Using the notation $\partial_t^\beta u$, $\beta < 0$ for the Riemann-Liouville integral $\partial_t^\beta u = \frac{1}{\Gamma(-\beta)} \int_0^t (t-s)^{-\beta-1} u(s) ds$ and the identity

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

after splitting $f_h = f_{h,0} + \tilde{f}_h$, with $f_{h,0} = f_h(0)$ and $\tilde{f}_h = f_h - f_{h,0}$, we rewrite the semidiscrete scheme (2.10) as

$$u_h(t) = v_h - (\partial_t^\alpha + A_h)^{-1} A_h v_h + (\partial_t^\alpha + A_h)^{-1} (f_{h,0} + \tilde{f}_h)$$
$$= v_h - (\partial_t^\alpha + A_h)^{-1} \partial_t \partial_t^{-1} A_h v_h + (\partial_t^\alpha + A_h)^{-1} (\partial_t \partial_t^{-1} f_{h,0} + \tilde{f}_h).$$

Now with $\bar{\partial}_\tau^\alpha$ being convolution quadrature for the SBD formula we get

$$U_h^n = v_h - (\bar{\partial}_\tau^\alpha + A_h)^{-1} \bar{\partial}_\tau \partial_t^{-1} A_h v_h + (\bar{\partial}_\tau^\alpha + A_h)^{-1} (\bar{\partial}_\tau \partial_t^{-1} f_{h,0} + \tilde{f}_h). \qquad (2.15)$$

The purpose of keeping the operator ∂_t^{-1} intact is to achieve a second-order accuracy. Letting $1_\tau = (0, 3/2, 1, \ldots)$, using the identity $1_\tau = \bar{\partial}_\tau \partial_t^{-1} 1$ at grid points t_n [7], and the associativity (2.9), the scheme (2.15) can be rewritten as

$$(\bar{\partial}_\tau^\alpha + A_h)(U_h^n - v_h) = -1_\tau A_h v_h + 1_\tau f_{h,0} + \tilde{f}_h.$$

Hence the second-order fully discrete scheme for the subdiffusion case reads: with $F_h^n = P_h f(t_n)$ and $U_h^0 = v_h$, find U_h^n, $n \geq 1$ such that

$$\bar{\partial}_\tau^\alpha U_h^1 + A_h U_h^1 + \tfrac{1}{2} A_h U_h^0 = \bar{\partial}_\tau^\alpha U_h^0 + F_h^1 + \tfrac{1}{2} F_h^0,$$
$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha U_h^0 + F_h^n, \quad n = 2, \ldots, N. \qquad (2.16)$$

The modification at the first step maintains a second order convergence. Otherwise, due to the limited smoothing property of the model (1.1), the scheme can only achieve a first-order convergence, unlike that for the classical parabolic problem [49].

Similarly, for $1 < \alpha < 2$, we can derive a fully discrete scheme. In analogy with the scheme (2.14), we correct the basic scheme in order to obtain error estimates consistent with Theorem A.4. The corrected scheme reads:

$$\bar{\partial}_\tau^\alpha U_h^1 + A_h U_h^1 + \tfrac{1}{2} A_h U_h^0 = \bar{\partial}_\tau^\alpha U_h^0 + \bar{\partial}_\tau^\alpha(tb_h) + \bar{\partial}_\tau G_h^1 + \tfrac{1}{2} \bar{\partial}_\tau G_h^0,$$
$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha U_h^0 + \bar{\partial}_\tau^\alpha(tb_h) + \bar{\partial}_\tau G_h^n, \quad n = 2, \ldots, N, \qquad (2.17)$$

with $U_h^0 = v_h$ and $G_h^n = P_h \partial_t^{-1} f(t_n)$.

REMARK 2.2. *It is known that without a correction the SBD scheme in general is only first-order accurate. Lubich [30, 31] has developed various useful corrections to obtain second-order accuracy. Even though these facts are well understood in the numerical PDEs community, it seems not so in the community of FDEs.*

2.3. Review of some existing methods. Now we review several existing time stepping schemes in the subdiffusion case for the Caputo derivative. The first is the popular L1 approximation of the fractional derivative [43, 29]

$${}^C\partial_t^\alpha \varphi(t_n) \approx \frac{1}{\Gamma(1-\alpha)} \sum_{j=0}^{n-1} b_j \frac{\varphi(t_{j+1}) - \varphi(t_j)}{\tau^\alpha}$$

with $b_j := (j+1)^{1-\alpha} - j^{1-\alpha}$, $j = 0, 1, \ldots, n-1$. The local truncation error is of the order $O(\tau^{2-\alpha})$, if the function φ is twice continuously differentiable [29].

The second scheme, due to Zeng et al [52], derived by applying convolution quadrature to a fractional integral reformulation, is given by

$$D^\alpha \varphi^n = L^\alpha(\Delta\varphi^n + f^n),$$

where the operators D^α and L^α are given by

$$D^\alpha \varphi^n = \tau^{-\alpha} \sum_{j=0}^{n} \omega_j(\varphi^{n-j} - \varphi^0) \quad \text{and} \quad L^\alpha \varphi^n = \frac{1}{2^\alpha} \sum_{j=0}^{n} \omega_j(-1)^j \varphi^{n-k},$$

with weights $\{\omega_j\}$ generated by the identity $(1-z)^\alpha = \sum_{j=0}^{\infty} \omega_j z^j$ [52, formula (3.13)]. The third scheme is obtained in the same spirit, but with $L^\alpha \varphi^n = (1 - \alpha/2)\varphi^n + \alpha/2\varphi^{n-1}$ [52, formula (3.14)]. Both the second and the third schemes converge at a rate $O(\tau^{2-\alpha})$, provided that φ is twice continuously differentiable, and at a rate $O(\tau^2)$, if further $\varphi'(0) = 0$. These results were given in Table 1.1.

REMARK 2.3. *The two schemes due to Zeng et al [52] are essentially a direct application of convolution quadrature based on the trapezoidal rule and Newton-Gregory formula to the fractional integral term. However, no correction for the first time step is incorporated, which will generally deteriorate the convergence.*

We have the following general comments on the schemes in Table 1.1. The convergence rates in Table 1.1 were mostly obtained by Taylor expansion, and thus requires high solution regularity. The SBD scheme for the Riemann-Liouville derivative was recently analyzed in [25], using Fourier transform, which uses substantially the zero extension $\bar\varphi$ of φ for $t < 0$. In particular, the assumption ${}^R_{-\infty}D_t^{3-\alpha}\bar\varphi \in L^1(\mathbb{R})$ requires $\varphi(0) = \varphi'(0) = 0$ and also $\varphi''(0) = 0$ for α close to zero. These conditions are restrictive, and do not hold for homogeneous problems [45].

Finally, for the diffusion wave equation, one scheme is the Crank-Nicolson scheme [48]. It approximates the Caputo derivative ${}^C\partial_t^\alpha \varphi(t_{n-1/2})$ by

$$ {}^C\partial_t^\alpha \varphi(t_{n-1/2}) \approx \frac{\tau^{-\alpha}}{\Gamma(3-\alpha)} \left[a_0 \delta_t \varphi^{n-1/2} - \sum_{j=1}^{n-1}(a_{n-j-1} - a_{n-j})\delta_t \varphi^{j-1/2} - a_{n-1}\tau\varphi'(0) \right], $$

where $\delta_t \varphi^{j-1/2} = \varphi^j - \varphi^{j-1}$ denotes central difference, and $a_j = (j+1)^{2-\alpha} - j^{2-\alpha}$. The local truncation error is $O(\tau^{3-\alpha})$ for C^3 functions [48].

3. Error analysis. Our goal in this section is to derive error estimates expressed directly in terms of problem data, verifying the robustness of the proposed schemes. This is done in two steps. First we bound the spatial error $u(t) - u_h(t)$, and then the temporal error $u_h(t_n) - U_h^n$. Throughout, let $\{(\lambda_j, \varphi_j)\}_{j=1}^{\infty}$ be the Dirichlet eigenpairs of $-\Delta$ on Ω, and $\{\varphi_j\}_{j=1}^{\infty}$ form an orthonormal basis in $L^2(\Omega)$. For any $q \geq 0$, we denote by $\dot H^q(\Omega) \subset L^2(\Omega)$ the Hilbert space induced by the norm $\|v\|_{\dot H^q(\Omega)}^2 = \sum_{j=1}^{\infty} \lambda_j^q(v, \varphi_j)^2$. Thus $\|v\|_{\dot H^0(\Omega)} = \|v\|$ is the norm in $L^2(\Omega)$, $\|v\|_{\dot H^1(\Omega)}$ the norm in $H_0^1(\Omega)$ and $\|v\|_{\dot H^2(\Omega)} = \|\Delta v\|$ is equivalent to the norm in $H_0^1(\Omega) \cap H^2(\Omega)$ [49].

3.1. Error analysis of the semidiscrete scheme. The semidiscrete scheme (2.3) for the subdiffusion case was already studied [19, 16, 18]. Hence we focus on the diffusion wave case. We employ an operator technique developed in [10] for the homogeneous problem, and an energy argument for the inhomogeneous problem.

First we derive an integral representation of the solution u to the homogeneous problem with $f = 0$ (see Appendix A for the solution theory). Since $u : (0, T] \to L^2(\Omega)$ can be analytically extended to the sector $\{z \in \mathbb{C}; z \neq 0, |\arg z| < \pi/2\}$ [45, Theorem 2.3], we apply the Laplace transform to (1.1) to deduce

$$z^\alpha \widehat{u}(z) + A\widehat{u}(z) = z^{\alpha-1}v + z^{\alpha-2}b,$$

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

$$u(t) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt}(z^\alpha I + A)^{-1}(z^{\alpha-1}v + z^{\alpha-2}b)\, dz, \tag{3.1}$$

where the contour $\Gamma_{\theta,\delta}$ is given by $\Gamma_{\theta,\delta} = \{z \in \mathbb{C} : |z| = \delta, |\arg z| \le \theta\} \cup \{z \in \mathbb{C} : z = \rho e^{\pm i\theta}, \rho \ge \delta\}$. Throughout, we choose the angle θ such that $\pi/2 < \theta < \min(\pi, \pi/\alpha)$ and hence $z^\alpha \in \Sigma_{\theta'}$ with $\theta' = \alpha\theta < \pi$ for all $z \in \Sigma_\theta := \{z \in \mathbb{C} : |\arg z| \le \theta\}$. Then there exists a constant c which depends only on θ and α such that

$$\|(z^\alpha I + A)^{-1}\| \le c|z|^{-\alpha}, \quad \forall z \in \Sigma_\theta. \tag{3.2}$$

Similarly, with $A_h = -\Delta_h$, the solution u_h to (2.3) can be represented by

$$u_h(t) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt}(z^\alpha I + A_h)^{-1}(z^{\alpha-1}v_h + z^{\alpha-2}b_h)\, dz. \tag{3.3}$$

The next lemma gives an important error estimate [10, 3].

LEMMA 3.1. *Let $\varphi \in L^2(\Omega)$, $z \in \Sigma_\theta$, $w = (z^\alpha I + A)^{-1}\varphi$, and $w_h = (z^\alpha I + A_h)^{-1}P_h\varphi$. Then there holds*

$$\|w_h - w\|_{L^2(\Omega)} + h\|\nabla(w_h - w)\|_{L^2(\Omega)} \le Ch^2\|\varphi\|_{L^2(\Omega)}. \tag{3.4}$$

Now we state an error estimate for the homogeneous problem (2.3).

THEOREM 3.2. *Let $1 < \alpha < 2$, and u and u_h be the solutions of problem (1.1) with $v \in \dot{H}^q(\Omega)$, $b \in \dot{H}^r(\Omega)$, $q, r \in [0, 2]$, and $f = 0$ and (2.3) with for $v_h = P_h v$, $b_h = P_h b$, $f_h = 0$, respectively. The following estimate on $e_h(t) = u(t) - u_h(t)$ holds*

$$\|e_h(t)\|_{L^2(\Omega)} + h\|\nabla e_h(t)\|_{L^2(\Omega)} \le ch^2(t^{-\alpha(2-q)/2}\|v\|_{\dot{H}^q(\Omega)} + t^{1-\alpha(2-r)/2}\|b\|_{\dot{H}^r(\Omega)}).$$

Proof. For $v, b \in L^2(\Omega)$, by (3.1) and (3.3), $e_h(t)$ can be represented as

$$e_h(t) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt}\left(z^{\alpha-1}(w^v - w_h^v) + z^{\alpha-2}(w^b - w_h^b)\right) dz,$$

with $w^v = (z^\alpha I + A)^{-1}v$, $w^b = (z^\alpha I + A)^{-1}b$, $w_h^v = (z^\alpha I + A_h)^{-1}P_h v$ and $w_h^b = (z^\alpha I + A_h)^{-1}P_h b$. By Lemma 3.1 and choosing $\delta = 1/t$ in $\Gamma_{\theta,\delta}$ we have

$$\|\nabla e_h(t)\|_{L^2(\Omega)} \le ch \left(\int_{-\theta}^{\theta} e^{\cos\psi} t^{-\alpha} d\psi + \int_{1/t}^{\infty} e^{rt\cos\theta} \rho^{\alpha-1} d\rho \right) \|v\|_{L^2(\Omega)}$$

$$+ ch \left(\int_{-\theta}^{\theta} e^{\cos\psi} t^{1-\alpha}\, d\psi + \int_{1/t}^{\infty} e^{rt\cos\theta} \rho^{\alpha-2}\, d\rho \right) \|b\|_{L^2(\Omega)}$$

$$\le ch(t^{-\alpha}\|v\|_{L^2(\Omega)} + t^{1-\alpha}\|b\|_{L^2(\Omega)}).$$

9

A similar argument yields the L^2-estimate. This shows the assertion for $v, b \in L^2(\Omega)$. Next for $v, b \in \dot{H}^2(\Omega)$, first we consider the choice $v_h = R_h v$ and $b_h = R_h b$. Then $e_h(t) = u(t) - u_h(t)$ is given by

$$e_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} z^{\alpha - 1} \left((z^\alpha I + A)^{-1} - (z^\alpha I + A_h)^{-1} R_h \right) v \, dz$$
$$+ \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} z^{\alpha - 2} \left((z^\alpha I + A)^{-1} - (z^\alpha I + A_h)^{-1} R_h \right) b \, dz.$$

Using the identity $z^\alpha (z^\alpha I + A)^{-1} = I - (z^\alpha I + A)^{-1} A$, we deduce

$$e_h(t) = \frac{1}{2\pi i} \left(\int_{\Gamma_{\theta, 1/t}} e^{zt} z^{-1} (w^v(z) - w_h^v(z)) \, dz + \int_{\Gamma_{\theta, 1/t}} e^{zt} z^{-1} (v - R_h v) \, dz \right)$$
$$+ \frac{1}{2\pi i} \left(\int_{\Gamma_{\theta, 1/t}} e^{zt} z^{-2} (w^b(z) - w_h^b(z)) \, dz + \int_{\Gamma_{\theta, 1/t}} e^{zt} z^{-2} (b - R_h b) \, dz \right) := \mathrm{I} + \mathrm{II},$$

where $w^v(z) = (z^\alpha I + A)^{-1} A v$ and $w_h^v(z) = (z^\alpha I + A_h)^{-1} A_h R_h v$, and $w^b(z)$ and $w_h^b(z)$ are defined analogously. Now Lemma 3.1 and the identity $A_h R_h = P_h A$ yield

$$\|w^v(t) - w_h^v(t)\|_{L^2(\Omega)} + h\|\nabla(w^v(t) - w_h^v(t))\|_{L^2(\Omega)} \leq ch^2 \|Av\|_{L^2(\Omega)}.$$

Consequently, we can bound the first term I (with $\delta = 1/t$)

$$\|\mathrm{I}\|_{L^2(\Omega)} \leq ch^2 \|Av\|_{L^2(\Omega)} \left| \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} z^{-1} \, dz \right|$$
$$\leq ch^2 \|Av\|_{L^2(\Omega)} \left(\int_{1/t}^\infty e^{rt \cos \theta} r^{-1} \, dr + \int_{-\theta}^\theta e^{\cos \psi} \, d\psi \right) \leq ch^2 \|v\|_{\dot{H}^2(\Omega)}.$$

We derive a bound on the second term II in a similar way:

$$\|\mathrm{II}\|_{L^2(\Omega)} \leq ch^2 \|Ab\|_{L^2(\Omega)} \left| \frac{1}{2\pi i} \int_\Gamma e^{zt} z^{-2} \, dz \right| \leq ch^2 t \|b\|_{\dot{H}^2(\Omega)},$$

and the L^2-error estimate follows. The H^1-estimate is established analogously. Last, for the choice $v_h = P_h v$ and $b_h = P_h b$, we have

$$E(t)v - E_h P_h v = E(t)v - E_h R_h v + E_h (R_h v - P_h v),$$

where E and E_h are continuous and semidiscrete solution operators, respectively (see Appendix A for the definitions). The first term is already bounded. By Theorem A.5 in Appendix A and approximation properties of P_h and R_h, there holds

$$\|E_h(t)(P_h v - R_h v)\|_{\dot{H}^p(\Omega)} \leq c\|P_h v - R_h v\|_{\dot{H}^p(\Omega)} \leq ch^{2-p} \|v\|_{\dot{H}^2(\Omega)}, \quad p = 0, 1.$$

The estimate for $b \in \dot{H}^2(\Omega)$ follows analogously. These estimates and interpolation complete the proof of the theorem. $\square$

For problem (1.1) with $f \in L^\infty(0, T; L^2(\Omega))$, we have the following result. The proof is identical to [18, Theorem 3.2], and hence omitted. The factor ℓ_h^2 reflects the limited smoothing property in space of the diffusion wave operator, cf. Theorem A.3.

THEOREM 3.3. *Let $1 < \alpha < 2$, u and u_h be the solutions of (1.1) with $v, b = 0$, $f \in L^\infty(0, T; L^2(\Omega))$, and (2.1) with $v_h = b_h = 0$, $f_h = P_h f$, respectively. Then with $\ell_h = |\ln h|$, there holds for $e_h(t) = u(t) - u_h(t)$*

$$\|e_h(t)\|_{L^2(\Omega)} + h\|\nabla e_h(t)\|_{L^2(\Omega)} \leq ch^2 \ell_h^2 \|f\|_{L^\infty(0, t; L^2(\Omega))}.$$

3.2. Error analysis for BE method. Now we derive L^2 error estimates for the fully discrete schemes (2.12) and (2.14) using the framework developed in [31, 7]. Alternatively, one can analyze the schemes by directly bounding the kernel function in the resolvent [32, 20]. We begin with an important result [31, Theorem 5.2].

LEMMA 3.4. *Let $K(z)$ be analytic in Σ_θ and (2.6) hold. Then for $g(t) = ct^{\beta-1}$, the convolution quadrature based on the BE method satisfies*

$$\|(K(\partial_t) - K(\bar{\partial}_\tau))g(t)\| \leq \begin{cases} ct^{\mu-1}\tau^\beta, & 0 < \beta \leq 1, \\ ct^{\mu+\beta-2}\tau, & \beta \geq 1. \end{cases}$$

First we state an error estimate for the homogeneous subdiffusion problem.

THEOREM 3.5. *Let u and U_h^n be the solutions of problem (1.1) with $v \in \dot{H}^q(\Omega)$, $q \in [0, 2]$, and $f = 0$ and (2.12)/(2.14) with $v_h = P_h v$ and $f_h = 0$, respectively. Then the following statements hold.*

(i) If $0 < \alpha < 1$, then

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

(ii) If $1 < \alpha < 2$, $b \in \dot{H}^r(\Omega)$, $r \in [0, 2]$, and $b_h = P_h b$, then

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leq c(t_n^{q\alpha/2-1}\tau + t_n^{(q-2)\alpha/2}h^2)\|v\|_{\dot{H}^q(\Omega)}$$
$$+ c(t_n^{r\alpha/2}\tau + t_n^{1+(r-2)\alpha/2}h^2)\|b\|_{\dot{H}^r(\Omega)}.$$

Proof. In view of the semidiscrete error estimates [19, Section 3] (where the log factor ℓ_h in the estimates in [19, Section 3] can be removed using the operator trick in Section 3.1), it suffices to bound $U_h^n - u_h(t_n)$. To this end, we denote for $z \in \Sigma_\theta$, $\theta \in (\pi/2, \pi)$, $G(z) = z^\alpha(z^\alpha I + A_h)^{-1}$. Then by (2.10) and (2.11), we have

$$U_h^n - u_h(t_n) = (G(\bar{\partial}_\tau) - G(\partial_t))v_h. \tag{3.5}$$

By (3.2), there holds $G(z) \leq c$ for $z \in \Sigma_\theta$. Hence, for $v \in L^2(\Omega)$, (3.5), Lemma 3.4 (with $\mu = 0$ and $\beta = 1$), and the $L^2(\Omega)$ stability of P_h give

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leq c\tau t_n^{-1}\|v_h\|_{L^2(\Omega)} \leq ct_n^{-1}\tau\|v\|_{L^2(\Omega)}. \tag{3.6}$$

For $v \in \dot{H}^2(\Omega)$, first consider the choice $v_h = R_h v$. Using the identity $G(z) = I - (z^\alpha I + A_h)^{-1}A_h$, with $G_s(z) = (z^\alpha I + A_h)^{-1}$, we have $U_h^n - u_h(t_n) = (G_s(\bar{\partial}_\tau) - G_s(\partial_t))A_h v_h$. Appealing to (3.2) and Lemma 3.4 (with $\mu = \alpha$ and $\beta = 1$) gives

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leq c\tau t_n^{\alpha-1}\|A_h v_h\|_{L^2(\Omega)} \leq c\tau t_n^{\alpha-1}\|v\|_{\dot{H}^2(\Omega)},$$

where the last line follows from $A_h R_h = P_h A$. The estimate holds also for the choice $v_h = P_h v$ in view of the $L^2(\Omega)$ stability of the scheme, cf. (3.6), and the argument in the proof of Theorem 3.2. The assertion now follows from interpolation. The case of $1 < \alpha < 2$ is analogous, and hence the proof is omitted. $\square$

Last we give error estimates for the BE method for problem (1.1) with $f \neq 0$ but $v = 0$ (also $b = 0$, if $1 < \alpha < 2$).

THEOREM 3.6. *Let u be the solution of problem (1.1) with homogeneous initial data and $f \in L^\infty(0, T; L^2(\Omega))$, and U_h^n be the solution to (2.12)/(2.14) with $f_h = P_h f$. Then with $\ell_h = |\ln h|$, the following statements hold:*

11

(i) *For $0 < \alpha < 1$, if $\int_0^t (t-s)^{\alpha-1} \|f'(s)\|_{L^2(\Omega)} ds < \infty$ for $t \in (0,T]$, then*

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \le c(h^2\ell_h^2\|f\|_{L^\infty(0,T;L^2(\Omega))} + t_n^{\alpha-1}\tau\|f(0)\|_{L^2(\Omega)}$$
$$+ \tau \int_0^{t_n} (t_n - s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)} ds).$$

(ii) *For $1 < \alpha < 2$, then*

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \le c(h^2\ell_h^2 + \tau)\|f\|_{L^\infty(0,T;L^2(\Omega))}.$$

Proof. With $G(z) = (z^\alpha I + A_h)^{-1}$, the semidiscrete solution u_h and fully discrete solution U_h^n are given by $u_h = G(\partial_t)f_h$ and $U_h^n = G(\partial_\tau)f_h$, respectively. Using the splitting $f_h(t) = f_h(0) + (1 * f'_h)(t)$ and the convolution relation (2.9), we have

$$u_h(t_n) - U_h^n = (G(\partial_t) - G(\partial_\tau))(f_h(0) + (1 * f'_h)(t_n))$$
$$= (G(\partial_t) - G(\partial_\tau))f_h(0) + ((G(\partial_t) - G(\partial_\tau))1) * f'_h(t_n)) := \mathrm{I} + \mathrm{II}.$$

Then Lemma 3.4 (with $\mu = \alpha$ and $\beta = 1$) yields a bound on the first term I

$$\|\mathrm{I}\|_{L^2(\Omega)} \le c\tau t_n^{\alpha-1}\|f_h(0)\|_{L^2(\Omega)} \le c\tau t_n^{\alpha-1}\|f(0)\|_{L^2(\Omega)}.$$

Likewise, the term II can be bounded using Lemma 3.4 and the L^2 stability of P_h by

$$\|\mathrm{II}\|_{L^2(\Omega)} \le \int_0^{t_n} \|((G(\partial_t) - G(\partial_\tau))1)(t_n - s)f'_h(s)\|_{L^2(\Omega)} ds$$
$$\le c\tau \int_0^{t_n} (t_n - s)^{\alpha-1}\|f'_h(s)\|_{L^2(\Omega)} ds \le c\tau \int_0^{t_n} (t_n - s)^{\alpha-1}\|f'(s)\|_{L^2(\Omega)} ds.$$

This shows assertion (i). For the scheme (2.14) with $v = b = 0$, $U_h^n = G(z)g_h$ with $g_h = \partial_t^{-1}f_h$ and $G(z) = z(z^\alpha I + A_h)^{-1}$. Hence the equality $g_h = 1 * f_h$, the convolution rule (2.9) and Lemma 3.4 with $\mu = \alpha - 1$ and $\beta = 1$ yield

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \le c\tau \int_0^{t_n} (t_n - s)^{\alpha-2}\|f(s)\|_{L^2(\Omega)} ds \le c_T\tau\|f\|_{L^\infty(0,T;L^2(\Omega))},$$

from which follows directly Assertion (ii), and this completes the proof. □

REMARK 3.1. *Assertion (i) in Theorem 3.6 holds also for the basic scheme (2.13). However, the corrected scheme (2.14) is uniformly first order in time for $v = b = 0$ and $f \in L^\infty(0,T;L^2(\Omega))$, which is consistent with the temporal regularity result in Theorem A.4. Hence, the correction in (2.14) gives better error estimates.*

3.3. Error analysis for the SBD scheme. Now we turn to the analysis of the SBD scheme. Like Lemma 3.4, the following estimate holds [31, Theorem 5.2].

LEMMA 3.7. *Let $K(z)$ be analytic in Σ_θ and (2.6) hold. Then for $g(t) = ct^{\beta-1}$, the convolution quadrature based on the SBD scheme satisfies*

$$\|(K(\partial_t) - K(\bar\partial_\tau))g(t)\| \le \begin{cases} ct^{\mu-1}\tau^\beta, & 0 < \beta \le 2, \\ ct^{\mu+\beta-3}\tau^2, & \beta \ge 2. \end{cases}$$

Now we state the following error estimates for the homogeneous problem.

THEOREM 3.8. *Let u and U_h^n be the solutions of problem (1.1) with $v \in \dot{H}^q(\Omega)$, $q \in [0,2]$, and $f = 0$ and (2.16)/(2.17) with $v_h = P_h v$ and $f_h = 0$, respectively. Then the following statements hold.*

12

(i) If $0 < \alpha < 1$, then

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \le c(\tau^2 t_n^{-2+q\alpha/2} + h^2 t_n^{-(2-q)\alpha/2})\|v\|_{\dot{H}^q(\Omega)}.$$

(ii) If $1 < \alpha < 2$, $b \in \dot{H}^r(\Omega)$, $r \in [0,2]$, and $b_h = P_h b$, then

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \le c(\tau^2 t_n^{-2+q\alpha/2} + h^2 t_n^{-(2-q)\alpha/2})\|v\|_{\dot{H}^q(\Omega)}$$
$$+ c(\tau^2 t_n^{r\alpha/2-1} + h^2 t_n^{1-(2-r)\alpha/2})\|b\|_{\dot{H}^r(\Omega)}.$$

Proof. We provide the proof only for $0 < \alpha < 1$, since that for $1 < \alpha < 2$ is identical. For $v \in L^2(\Omega)$, the difference between $u_h(t_n)$ and U_h^n is given by

$$u_h(t_n) - U_h^n = (G(\partial_t^\alpha) - G(\bar{\partial}_\tau))\partial_t^{-1}(A_h v_h)(t_n),$$

where $G(z) = -z(z^\alpha I + A_h)^{-1}A_h$. By (3.2) and the identity

$$G(z) = -z(z^\alpha I + A_h)^{-1}A_h = -zI + z^{\alpha+1}(z^\alpha I + A_h)^{-1}) \quad \forall z \in \Sigma_{\pi-\theta},$$

there holds $\|G(z)\| \le c|z|$, for $z \in \Sigma_\theta$. Then Lemma 3.7 (with $\mu = -1$ and $\beta = 2$) and the $L^2(\Omega)$ stability of P_h give

$$\|U_h^n - u_h(t_n)\|_{L^2(\Omega)} \le c\tau^2 t_n^{-2}\|v_h\|_{L^2(\Omega)} \le c\tau^2 t_n^{-2}\|v\|_{L^2(\Omega)}. \tag{3.7}$$

For smooth initial data $v \in \dot{H}^2(\Omega)$, we first set $U_h^0 = v_h = R_h v$. By setting $G_s(z) = -z(z^\alpha I + A_h)^{-1}$, the difference $U_h^n - u_h(t_n)$ can be written by

$$U_h^n - u_h(t_n) = (G_s(\bar{\partial}_\tau) - G_s(\partial_t))A_h v_h.$$

From (3.2), we deduce $\|G_s(z)\| \le M|z|^{1-\alpha}$ for all $z \in \Sigma_\theta$. Now Lemma 3.7 (with $\mu = \alpha - 1$ and $\beta = 2$) and the identity $A_h R_h = P_h A$ gives

$$\|U_h^n - u_h(t_n)\|_{L^2(\Omega)} \le c\tau^2 t_n^{\alpha-2}\|A_h v_h\|_{L^2(\Omega)} \le c\tau^2 t_n^{\alpha-2}\|v\|_{\dot{H}^2(\Omega)}.$$

Last, the desired assertion follows from the L^2 stability of the scheme (2.16) (as a direct consequence of (3.7), cf. the proof of Theorem 3.5) and interpolation. $\square$

Last we give error estimates for the inhomogeneous problem.

THEOREM 3.9. *Let u be the solution of problem (1.1) with homogeneous initial data and $f \in L^\infty(0,T;L^2(\Omega))$, and U_h^n be the solution to (2.16)/(2.17) with $f_h = P_h f$. Then with $\ell_h = |\ln h|$, the following statements hold.*

(i) For $0 < \alpha < 1$, if $\int_0^t (t-s)^{\alpha-1}\|f''(s)\|_{L^2(\Omega)}ds < \infty$ for $t \in [0,T]$, then

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \le c(h^2\ell_h^2\|f\|_{L^\infty(0,T;L^2(\Omega))} + t_n^{\alpha-2}\tau^2\|f(0)\|_{L^2(\Omega)}$$
$$+ t_n^{\alpha-1}\tau^2\|f'(0)\|_{L^2(\Omega)} + \tau^2 \int_0^{t_n} (t_n-s)^{\alpha-1}\|f''(s)\|_{L^2(\Omega)}\, ds).$$

(ii) For $1 < \alpha < 2$, if $\int_0^t (t-s)^{\alpha-2}\|f'(s)\|_{L^2(\Omega)}ds < \infty$ for $t \in [0,T]$, then

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \le ch^2\ell_h^2\|f\|_{L^\infty(0,T;L^2(\Omega))}$$
$$+ c\tau^2(t_n^{\alpha-2}\|f(0)\|_{L^2(\Omega)} + \int_0^{t_n} (t_n-s)^{\alpha-2}\|f'(s)\|_{L^2(\Omega)}\, ds).$$

13

Proof. By [18, Theorem 3.2] and Theorem 3.3, it suffices to bound $e_h^n = u_h(t_n) - U_h^n$. Upon letting $G_1(z) = z(z^\alpha I + A_h)^{-1}$ and $G_2(z) = (z^\alpha I + A_h)^{-1}$ and using the identity $\tilde{f}_h = f_h(0) + t f_h' + t * f_h''$, the solutions $u_h(t_n)$ and U_h^n are represented by

$$u_h(t_n) = G_1(\partial_t) t f_h(0) + G_2(\partial_t) t f_h'(0) + (G_2(\partial_t)t) * f_h'', \quad \text{and}$$
$$U_h^n = G_1(\partial_\tau) t f_h(0) + G_2(\partial_\tau) t f_h'(0) + (G_2(\partial_\tau)t) * f_h'',$$

respectively. Therefore,

$$u_h(t_n) - U_h^n = (G_1(\partial_t) - G_1(\partial_\tau)) \, t f_h(0) + (G_2(\partial_t) - G_2(\partial_\tau)) \, t f_h'(0)$$
$$+ ((G_2(\partial_t) - G_2(\partial_\tau))t) * f_h'' := \mathrm{I} + \mathrm{II}.$$

Lemma 3.7 (with $\mu = 1 - \alpha$ and $\beta = 2$) gives a bound on the first term I as $\|\mathrm{I}\|_{L^2(\Omega)} \leq c t_n^{\alpha-2} \tau^2$. The bound on the second term II follows from Lemma 3.7 (with $\mu = \alpha$ and $\beta = 2$) and the $L^2(\Omega)$ stability of P_h by

$$\|\mathrm{II}\|_{L^2(\Omega)} \leq \| (G_2(\partial_t) - G_2(\partial_\tau)) \, t f_h'(0) \|_{L^2(\Omega)}$$
$$+ \int_0^{t_n} \| (G_2(\partial_t) - G_2(\partial_\tau))t) \, (t_n - s) f_h''(s) \|_{L^2(\Omega)} \, ds$$
$$\leq c t_n^{\alpha-1} \tau^2 \|f'(0)\|_{L^2(\Omega)} + c \tau^2 \int_0^{t_n} (t_n - s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} \, ds.$$

This shows assertion (i). Assertion (ii) follows analogously from Lemma 3.4 with $\mu = \alpha - 1$ and $\beta = 2$. □

REMARK 3.2. *For subdiffusion with $v = 0$, the SBD scheme (2.16) is of uniformly second order if $f(0) = f'(0) = 0$ and $\int_0^t (t - s)^{\alpha-1} \|f''(s)\|_{L^2(\Omega)} ds < \infty$ for $t \in [0, T]$, and for diffusion-wave, the SBD scheme (2.17) is of uniformly second order if $f(0) = 0$ and $\int_0^t (t - s)^{\alpha-1} \|f'(s)\|_{L^2(\Omega)} ds < \infty$ for $t \in [0, T]$. These conditions are directly verifiable. It is noteworthy that the estimate in Theorem 3.9(i) holds also for an uncorrected SBD scheme for the diffusion wave problem.*

4. Numerical experiments and discussions. Now we illustrate the convergence and robustness of the schemes on several examples on the square domain $\Omega = (0, 1)^2$. For the examples below, the exact solution can be expressed as an infinite series involving the Mittag-Leffler function $E_{\alpha,\beta}(z)$ [45, 19], for which we employ an algorithm developed in [47]. In the computations, the domain $\Omega = (0, 1)^2$ is triangulated as follows. It is first divided into M^2 small equal squares, by partitioning the unit interval $(0, 1)$ into M equally spaced subintervals, and the diagonal of each small square is then connected to obtain a symmetric triangulation. Likewise, we fix the time step size τ at $\tau = t/N$, where t is the time of interest. To examine the spatial and temporal convergence rates separately, we take a small time step size τ (or mesh size h, respectively), so that the temporal (or spatial) discretization error is negligible. We measure the error $e^n =: u(t_n) - U_h^n$ by the normalized error $\|e^n\|_{L^2(\Omega)}/\|v\|_{L^2(\Omega)}$ and (also $\|e^n\|_{\dot{H}^1(\Omega)}/\|v\|_{L^2(\Omega)}$ for spatial convergence).

4.1. Subdiffusion. We consider the following three examples:
(a) $v = xy(1 - x)(1 - y) \in \dot{H}^2(\Omega)$ and $f = 0$;
(b) $v = \chi_{(0,1/2)\times(0,1)}(x, y) \in \dot{H}^{1/2-\epsilon}(\Omega)$ with $\epsilon \in (0, 1/2)$ and $f = 0$.
(c) $v = 0$, and $f = (1 + t^{0.2})\chi_{(0,1/2)\times(0,1)}(x, y)$.

Since the spatial convergence of the semidiscrete Galerkin scheme for subdiffusion was studied in [19, 16], we focus on the temporal convergence at $t = 0.1$. The

numerical results for cases (a) and (b) are shown in Tables 4.1 and 4.2, respectively, where `rate` refers to the empirical convergence rate of the errors, and the numbers in the bracket denote theoretical predictions. The BE and SBD schemes exhibit a very steady first and second-order convergence, respectively, for both smooth and nonsmooth data, which agree well with the convergence theory.

In Tables 4.1 and 4.2, we also include results by four existing schemes: the L1 scheme [29] (denoted by Lin-Xu), two schemes from Zeng et al [52] (Zeng I and Zeng II), where the theoretical rates in the bracket are for smooth solutions, all at a rate $O(\tau^{2-\alpha})$, cf. Table 1.1. For both cases, the L1 scheme only achieves a first-order convergence. This is not accidental: its best possible convergence rate for homogeneous problems is $O(\tau)$ [20]. The convergence of the Zeng I scheme strongly depends on α, and can fail to achieve an $O(\tau)$ rate for nonsmooth data. Their second scheme can only achieve a first-order convergence in either case. Hence, time stepping schemes derived under the assumption that the solution is smooth may be not robust with respect to data regularity, which motivates revisiting their analysis for nonsmooth data. In contrast, the schemes proposed in this work are robust. Further, we include the numerical results for the discontinuous Petrov-Galerkin scheme (denoted by DPG) of Mustapha et al [39], using piecewise linear functions. The DPG scheme was analyzed for graded meshes in [39], which compensates the solution singularity with local refinement, and the optimal convergence rates for uniform meshes remain unknown. Numerically, we observe that it merits second-order convergence for case (a), but for nonsmooth case (b), the convergence seems not so steady.

Table 4.1: The L^2-error for case (a) for $t = 0.1$, $\alpha = 0.5$, and $h = 1/512$.

$\alpha \backslash N$	method	10	20	40	80	160	320	rate
	BE	2.96e-4	1.46e-4	7.27e-5	3.63e-5	1.81e-5	9.05e-6	1.00 (1.00)
	SBD	2.94e-5	6.88e-6	1.66e-6	4.09e-7	1.01e-7	2.49e-8	2.02 (2.00)
0.1	Lin-Xu	2.76e-4	1.35e-4	6.71e-5	3.34e-5	1.66e-5	8.31e-6	1.02 (1.90)
	Zeng I	1.12e-2	5.84e-3	3.03e-3	1.57e-3	8.05e-4	4.12e-4	0.96 (1.90)
	Zeng II	2.40e-4	1.20e-4	5.98e-5	2.99e-5	1.49e-5	7.47e-6	1.00 (1.90)
	DPG	7.99e-1	6.77e-1	4.79e-1	2.33e-1	5.64e-2	1.52e-2	1.88 (−−)
	BE	3.33e-3	1.63e-3	8.05e-4	4.00e-4	1.99e-4	9.96e-5	1.00 (1.00)
	SBD	4.18e-4	9.70e-5	2.33e-5	5.71e-6	1.48e-6	3.45e-7	2.02 (2.00)
0.5	Lin-Xu	2.45e-3	1.17e-3	5.68e-4	2.80e-4	1.38e-4	6.88e-5	1.01 (1.50)
	Zeng I	1.25e-2	3.30e-3	9.01e-4	2.69e-4	9.14e-5	3.58e-5	1.85 (1.50)
	Zeng II	9.68e-4	5.15e-4	2.65e-4	1.35e-4	6.75e-5	3.39e-5	0.95 (1.50)
	DPG	1.58e-2	3.64e-3	9.35e-4	2.48e-4	6.64e-5	1.74e-5	1.93 (−−)
	BE	1.89e-2	9.42e-3	4.70e-3	2.35e-3	1.17e-3	5.85e-4	1.00 (1.00)
	SBD	2.53e-3	5.98e-4	1.45e-4	3.59e-5	8.88e-6	2.19e-6	2.02 (2.00)
0.9	Lin-Xu	1.78e-2	8.51e-3	4.08e-3	1.97e-3	9.50e-4	4.61e-4	1.04 (1.10)
	Zeng I	3.85e-4	1.45e-4	1.10e-4	6.59e-5	3.59e-5	1.87e-5	0.91 (1.10)
	Zeng II	6.07e-4	2.27e-4	1.65e-4	9.55e-5	5.10e-5	2.63e-5	0.93 (1.10)
	DPG	9.03e-4	2.13e-4	5.08e-5	1.21e-5	2.84e-6	5.96e-7	2.06 (−−)

If the mesh size h is small and the number N of time steps is fixed, then by Theorems 3.5(i) and 3.8(i), for both BE and SBD schemes there holds for $t_N \to 0$

$$\|U_h^N - u(t_N)\|_{L^2(\Omega)} \le C t_N^{q\alpha/2} N^{-1} \|v\|_{\dot{H}^q(\Omega)}. \tag{4.1}$$

Table 4.2: The L^2-error for case (b) for $t = 0.1$, $\alpha = 0.5$, and $h = 1/512$.

$\alpha\backslash N$	method	10	20	40	80	160	320	rate
	BE	1.86e-4	9.21e-5	4.58e-5	2.28e-5	1.14e-5	5.69e-6	1.00 (1.00)
	SBD	1.85e-5	4.34e-6	1.05e-6	2.58e-7	6.38e-8	1.57e-8	2.02 (2.00)
0.1	Lin-Xu	1.74e-4	8.51e-5	4.22e-5	2.10e-5	1.05e-5	5.23e-6	1.00 (1.90)
	Zeng I	1.21e-2	6.42e-3	3.39e-3	1.80e-3	9.42e-4	4.95e-4	0.93 (1.90)
	Zeng II	1.51e-4	7.53e-5	3.76e-5	1.88e-5	9.40e-6	4.70e-6	1.00 (1.90)
	DPG	8.98e-1	8.38e-1	7.41e-1	6.14e-1	4.93e-1	4.02e-1	0.29 (−−)
	BE	2.09e-3	1.02e-3	5.05e-4	2.51e-4	1.25e-4	6.25e-5	1.01 (1.00)
	SBD	2.64e-4	6.11e-5	1.47e-5	3.60e-6	8.87e-7	2.17e-7	2.02 (2.00)
0.5	Lin-Xu	1.52e-3	7.26e-4	3.54e-4	1.74e-4	8.64e-5	4.30e-5	1.00 (1.50)
	Zeng I	7.87e-2	4.61e-2	2.70e-2	1.56e-2	8.95e-3	5.05e-3	0.80 (1.50)
	Zeng II	6.04e-4	3.22e-4	1.65e-4	8.38e-5	4.22e-5	2.12e-5	1.00 (1.50)
	DPG	3.97e-1	3.06e-1	2.31e-1	1.72e-1	1.25e-1	8.71e-2	0.47 (−−)
	BE	1.13e-2	5.60e-3	2.78e-3	1.39e-3	6.92e-4	3.46e-4	1.00 (1.00)
	SBD	1.63e-3	3.79e-4	9.15e-5	2.25e-5	6.56e-6	1.37e-6	2.02 (2.00)
0.9	Lin-Xu	1.05e-2	5.00e-3	2.39e-3	1.15e-3	5.55e-4	2.69e-4	1.05 (1.10)
	Zeng I	1.39e-1	8.39e-2	4.46e-2	1.74e-2	3.15e-3	1.30e-4	−− (1.10)
	Zeng II	2.36e-2	1.97e-3	9.74e-5	5.56e-5	2.97e-5	1.54e-5	0.93 (1.10)
	DPG	1.94e-1	1.30e-1	8.21e-2	4.41e-2	1.58e-2	1.73e-3	−− (−−)

In Table 4.3 and Fig. 4.1 we show the L^2-norm of the error for cases (a) and (b), for fixed $N = 10$ and $t_N \to 0$ with $\alpha = 0.5$. In the table, the `rate` (with respect to t_N, for fixed N only) is computed from (4.1), and the theoretical decay rate is $t_N^{q\alpha/2}$. In the smooth case (a), the temporal error decreases like $O(t_N^{1/2})$, whereas in the nonsmooth case (b), it decays like $O(t_N^{1/8})$. Note that in case (b), the initial data $v \in \dot{H}^{1/2-\epsilon}(\Omega)$ for any $\epsilon > 0$, the formula (4.1) predicts an error decay rate $O(t^{\alpha/4}) = O(t^{1/8})$. Hence the empirical rates in Table 4.3 and Fig. 4.1 agree well with the theoretical predictions, thereby fully confirming the factor $t_n^{q\alpha/2-1}$ in Theorem 3.5(i) (and likewise the factor $t_n^{q\alpha/2-2}$ in Theorem 3.8(i)).

Table 4.3: The L^2-error for cases (a) and (b) as $t \to 0$ with $h = 1/512$ and $N = 10$.

t	method	1e-3	1e-4	1e-5	1e-6	1e-7	1e-8	rate
(a)	BE	6.16e-3	2.64e-3	8.93e-4	2.88e-4	9.20e-5	2.93e-5	0.49 (0.50)
	SBD	4.52e-4	1.55e-4	4.98e-5	1.59e-5	5.04e-6	1.60e-6	0.50 (0.50)
(b)	BE	5.86e-3	4.61e-3	3.32e-3	2.51e-3	1.92e-3	1.51e-3	0.12 (0.13)
	SBD	5.44e-4	3.81e-4	2.85e-4	2.14e-4	1.60e-4	1.19e-4	0.13 (0.13)

Last, we examine the inhomogeneous problem, i.e., case (c). The numerical results are given in Table 4.4, where the last two rows were obtained by correcting the right hand side f, cf. (2.14). The BE scheme converges at the expected $O(\tau)$ rate, but the SBD scheme only converges at a rate $O(\tau^{1.18})$. The latter is attributed to the insufficient temporal regularity of the right hand side f: only the first order derivative is integrable, but not high-order ones. One can also employ the correction in (2.14), which seems to restore the second-order convergence, cf. the last row of Table 4.4. However, the mechanism behind this remedy is still unknown.
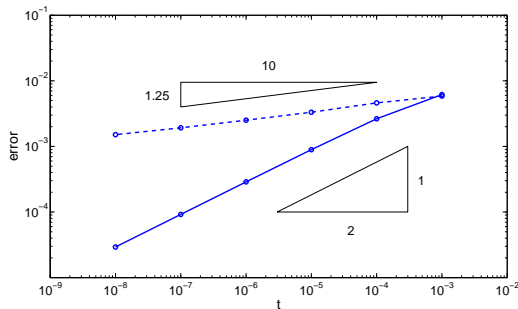
Fig. 4.1: Numerical results for cases (a) and (b) using the BE scheme with $h = 1/512$ and $N = 10$, $\alpha = 0.5$ for $t \to 0$, where the solid and dashed lines stand for cases (a) and (b), respectively.

Table 4.4: The L^2-error for case (c) at $t = 0.1$, with $\alpha = 0.5$ and $h = 1/512$.

method\N	10	20	40	80	160	rate
BE (2.12)	9.07e-4	4.34e-5	2.10e-5	1.02e-5	5.02e-6	1.04 (1.00)
SBD (2.16)	3.35e-6	3.29e-6	1.79e-6	8.32e-7	3.38e-7	1.18 (––)
BE	2.40e-4	1.18e-4	5.85e-5	2.91e-5	1.45e-5	1.00 (1.00)
SBD	1.99e-5	4.61e-6	1.11e-6	2.68e-7	6.32e-8	2.06 (2.00)

4.2. Diffusion-wave equation. We consider the following four examples for the diffusion wave equation (with $\epsilon \in (0, 1/2)$):

(d) $v = xy(1 - x)(1 - y) \in \dot{H}^2(\Omega)$, $b = 0$ and $f = 0$

(e) $v = \chi_{(0,1/2] \times (0,1)}(x, y) \in \dot{H}^{1/2-\epsilon}(\Omega)$, $b = 0$ and $f = 0$.

(f) $v = 0$, $b = \chi_{(0,1/2] \times (0,1)}(x, y) \in \dot{H}^{1/2-\epsilon}(\Omega)$, $f = 0$.

(g) $v = 0$, $b = 0$, and $f = (1 + t^{0.2})\chi_{(0,1/2] \times (0,1)}(x, y)$.

Numerical results for examples (d) and (e). First we briefly examine the convergence of the semidiscrete Galerkin scheme. The numerical results for cases (d) and (e) are shown in Fig. 4.2 and Table 4.5, respectively. We observe a convergence rate $O(h^2)$ and $O(h)$ in the L^2- and H^1-norm, respectively, for both smooth and nonsmooth data. For nonsmooth data, the error deteriorates as t approaches zero, due to the weak singularity of the solution at t close to zero, cf. Theorem A.1, which we examine more closely next by verifying the prefactors in Theorem 3.3. For the smooth case (d), the error essentially stays unchanged with time t, whereas for the nonsmooth case (e) it deteriorates like $O(t^{-0.83})$ as $t \to 0$, cf. Table 4.6. These observations agree well with the theory: by Theorem 3.2, as $t \to 0$, there holds for any $\epsilon \in (0, 1/2)$

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} \le ct^{-\alpha(3+2\epsilon)/4}h^2\|v\|_{\dot{H}^{1/2-\epsilon}(\Omega)}.$$

Hence, the numerical results fully confirm the error estimates in Theorem 3.2.

Next we examine temporal convergence. The numerical results for case (d) are given in Table 4.7. The rates $O(\tau)$ and $O(\tau^2)$ are observed for the BE and SBD schemes, respectively, and they hold also for case (e), cf. Table 4.8. Hence, the proposed schemes exhibit a steady convergence for both smooth and nonsmooth data, verifying their robustness, cf. Theorems 3.5(ii) and 3.8(ii). Note that if the spatial error is negligible, then for fixed N and $t_N \to 0$, (4.1) holds, by Theorem 3.5(ii),
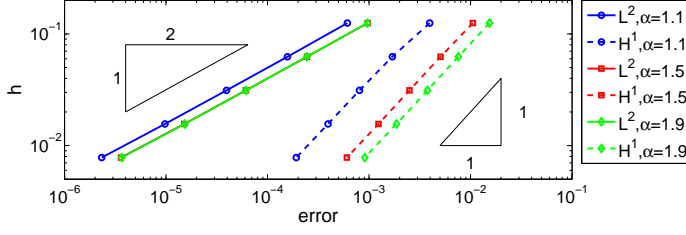
17

Fig. 4.2: The convergence of the semidiscrete Galerkin scheme for case (d) at $t = 0.1$, computed with the SBD scheme with $N = 1000$.

Table 4.5: Numerical results for case (e): $\alpha = 1.5$, $h = 2^{-k}$, $N = 1000$.

t	k	4	5	6	7	8	rate
0.1	L^2-norm	1.51e-1	4.00e-3	1.00e-3	2.40e-4	4.83e-5	2.05 (2.00)
	H^1-norm	3.27e-1	1.40e-1	6.51e-2	3.11e-2	1.39e-2	1.10 (1.00)
0.01	L^2-norm	5.71e-2	2.74e-2	9.29e-3	2.44e-3	5.07e-4	1.92 (2.00)
	H^1-norm	3.06e0	2.42e0	1.31e0	6.94e-1	2.55e-1	1.04 (1.00)
0.005	L^2-norm	9.87e-2	4.32e-2	1.64e-2	4.94e-3	1.07e-3	1.78 (2.00)
	H^1-norm	6.86e0	4.60e0	2.87e0	1.37e0	5.72e-1	1.00 (1.00)

Table 4.6: The L^2-error for cases (d)-(f) with $\alpha = 1.1$: $t \to 0$, $h = 2^{-7}$, $N = 10^3$.

t	1	1e-1	1e-2	1e-3	1e-4	1e-5	rate
(d)	1.66e-7	3.99e-6	5.14e-5	1.41e-4	9.16e-5	8.71e-5	0.02 (0)
(e)	1.17e-7	2.58e-6	1.07e-4	1.69e-4	9.94e-4	6.04e-3	-0.78 (-0.83)
(f)	3.71e-6	1.50e-5	3.68e-6	2.83e-6	1.33e-6	5.40e-7	0.22 (0.18)

which allows one to verify the temporal regularity in Theorem A.2. In Table 4.9, we present the results for the BE scheme for $\alpha = 1.1$. The L^2-norm of the error decays at a rate $O(t^{1.10})$ for (d) and $O(t^{0.28})$ for (e), respectively, as $t \to 0$, which concurs with the theoretical ones, thereby confirming the factor $t_n^{q\alpha/2-1}$ in Theorem 3.5(ii).

In Tables 4.7 and 4.8, we present also the results by the Crank-Nicolson (CN) scheme, which converges at a rate $O(\tau^{3-\alpha})$ for C^3 solutions [48]. It achieves the desired rate in either case, even though by Theorem A.2, the solution u does not have the requisite regularity. These observations call for further analysis of the scheme.

It is known that as the fractional order α increases from one to two, the model (1.1) changes from a diffusion equation to a wave one [11]. This transition can be observed numerically: for α close to one, the solution is diffusive and very smooth, whereas for α close to two, the plateau in the initial data v is well preserved, reflecting a "finite" speed of wave propagation, cf. Fig. 4.3. The small oscillations in Fig. 4.3 are not numerical artifacts: the L^2 projection $P_h v$ is oscillatory, and the numerical solution inherits the feature. Further, the closer is α to two, the slower is the decay of the solution (for t close to zero), showing again the wave feature.

Numerical results for example (f). Similar to cases (d) and (e), we observe the expected $O(h)$ and $O(h^2)$ convergence for the H^1- and L^2-norm of the error, respectively, cf. Fig. 4.4. An $O(\tau)$ and $O(\tau^2)$ convergence for the BE and SBD scheme,

18

Table 4.7: The L^2-error for case (d) at $t = 0.1$ with $h = 1/512$.

α	N	10	20	40	80	160	320	rate
	BE	2.90e-2	1.49e-2	7.57e-3	3.81e-3	1.91e-3	9.59e-4	1.00 (1.00)
1.1	SBD	6.35e-4	2.02e-4	5.49e-5	1.41e-5	3.45e-6	7.50e-7	2.06 (2.00)
	CN	3.40e-4	7.17e-5	1.44e-5	2.55e-6	2.62e-7	1.52e-7	2.36(1.90)
	BE	4.53e-3	2.43e-3	1.26e-3	6.44e-4	3.26e-4	1.64e-4	0.99 (1.00)
1.5	SBD	1.25e-3	3.26e-4	8.22e-5	2.03e-5	4.51e-6	6.72e-7	2.06 (2.00)
	CN	1.21e-3	4.41e-4	1.59e-4	5.72e-5	2.09e-5	7.83e-6	1.45(1.50)
	BE	9.64e-3	4.91e-3	2.48e-3	1.24e-3	6.23e-4	3.11e-4	1.00 (1.00)
1.9	SBD	4.41e-4	1.25e-4	3.38e-5	8.58e-6	2.10e-6	7.66e-7	2.00 (2.00)
	CN	8.06e-3	3.83e-3	1.80e-3	8.46e-4	3.95e-4	1.84e-4	1.10 (1.10)

Table 4.8: The L^2-error for case (e) at $t = 0.1$ with $h = 1/512$.

α	N	10	20	40	80	160	320	rate
	BE	2.16e-2	1.09e-2	5.47e-3	2.74e-3	1.37e-3	6.86e-4	1.00 (1.00)
1.1	SBD	3.47e-3	8.04e-4	1.94e-4	4.76e-5	1.17e-5	2.76e-6	2.02 (2.00)
	CN	8.66e-2	2.62e-2	1.47e-3	1.69e-5	4.21e-6	9.67e-7	2.06 (1.90)
	BE	3.82e-2	2.10e-2	1.11e-2	5.76e-3	2.93e-3	1.48e-3	1.00 (1.00)
1.5	SBD	1.18e-2	2.78e-3	6.64e-4	1.61e-4	3.92e-5	9.02e-6	2.06 (2.00)
	CN	1.00e-2	2.66e-3	9.90e-4	3.60e-4	1.29e-4	4.58e-5	1.50 (1.50)
	BE	1.11e-1	7.73e-2	5.15e-2	3.25e-2	1.93e-2	1.09e-2	0.83 (1.00)
1.9	SBD	7.36e-2	4.11e-2	1.85e-2	5.96e-3	1.56e-3	3.88e-4	1.95 (2.00)
	CN	6.76e-2	4.42e-2	2.68e-2	1.51e-2	7.92e-3	3.95e-3	1.00 (1.10)

Table 4.9: The L^2-error for cases (d)-(e) with $\alpha = 1.1$: $t \to 0$, $h = 1/512$, and $N = 10$.

t	1e0	1e-1	1e-2	1e-3	1e-4	1e-5	rate
(d)	4.44e-4	6.33e-4	5.93e-5	1.82e-6	7.94e-8	3.97e-9	1.30 (1.10)
(e)	2.56e-4	3.46e-3	1.35e-3	7.35e-4	3.77e-4	4.28e-5	0.32 (0.28)
(f)	3.21e-5	1.06e-4	6.09e-6	3.04e-7	1.58e-8	5.42e-10	1.32 (1.28)

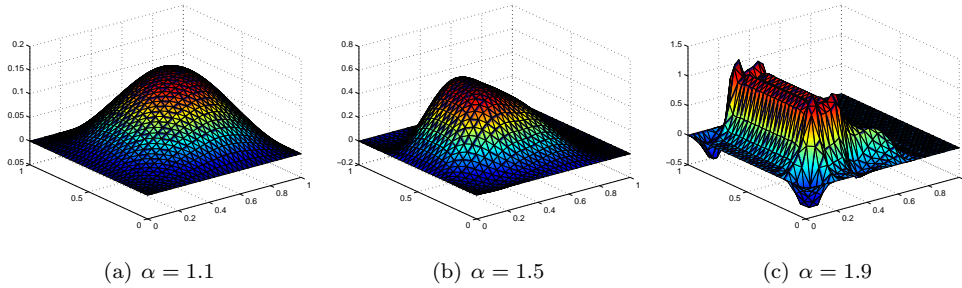(a) $\alpha = 1.1$ (b) $\alpha = 1.5$ (c) $\alpha = 1.9$

Fig. 4.3: The profile of the solutions to case (e) at $t = 0.1$ with three different α values, computed using the SBD scheme with $h = 1/64$ and $N = 160$.

respectively, is observed, cf. Table 4.10. To examine more closely the solution singu-

larity, we appeal to Theorem 3.5(ii) to deduce that for fixed N

$$\|u(t_N) - U_h^N\|_{L^2(\Omega)} \le c(N^{-1}t_N^{1+\alpha r/2} + h^2 t_N^{1-\alpha(2-r)/2})\|b\|_{\dot{H}^r(\Omega)}.$$

Hence, if the temporal error is negligible, the formula predicts a decay $O(t_N^{1-\alpha(2-r)/2})$ as $t_N \to 0$. Since $b \in \dot{H}^{1/2-\epsilon}(\Omega)$ for case (f), for small $\epsilon \in (0, 1/4)$, it predicts $O(t_N^{0.18-0.55\epsilon})$ for $\alpha = 1.1$, which is confirmed by the last row of Table 4.6. Likewise, if the spatial error is negligible, for fixed N, the formula predicts a decay $O(t_N^{1+\alpha r/2})$ as $t_N \to 0$. For $\alpha = 1.1$, it predicts a rate $O(t_N^{1.28-0.55\epsilon})$, which agrees well with the last row of Table 4.9. These results confirm the error estimate in Theorem 3.5(ii).
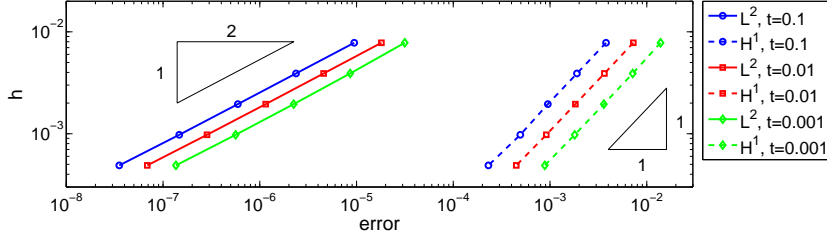

Fig. 4.4: Error plots for case (f) with $\alpha = 1.5$, by the SBD scheme with $N = 1000$.

Table 4.10: The L^2-error for case (f) at $t = 0.1$ with $h = 1/512$.

$\alpha\backslash N$	method	10	20	40	80	160	320	rate
1.1	BE	1.12e-3	5.73e-4	2.90e-4	1.46e-4	7.30e-5	3.66e-5	1.00 (1.00)
	SBD	1.07e-4	2.54e-5	6.20e-6	1.53e-6	3.76e-7	8.83e-8	2.04 (2.00)
1.5	BE	2.82e-3	1.46e-3	7.48e-4	3.78e-4	1.90e-4	9.53e-5	0.99 (1.00)
	SBD	2.37e-4	6.47e-5	1.65e-5	4.11e-6	1.01e-7	2.30e-8	2.02 (2.00)
1.9	BE	3.65e-3	2.04e-3	1.11e-3	5.94e-4	3.12e-4	1.61e-4	0.94 (1.00)
	SBD	1.02e-3	4.13e-4	1.42e-4	4.06e-5	1.05e-5	2.93e-6	1.89 (2.00)

Numerical results for example (g). Now we present numerical results for case (g) in Table 4.11, where the first and last two rows are for the basic and the corrected schemes, respectively. For the BE scheme, both variants (2.13) and (2.14) can achieve the desired $O(\tau)$ convergence, and the errors are comparable. However, for the SBD scheme, the basic variant converges only at a suboptimal rate $O(\tau^{1.36})$. It concurs with Theorem 3.9(i), since the right hand side f is not regular enough. The correction indeed restores the desired convergence rate. These observations clearly show the crucial role of proper initial correction in high-order schemes, and in particular, an inadvertent implementation can compromise the accuracy.

Table 4.11: The L^2-error for case (g) at $t = 0.1$ with $\alpha = 1.5$ and $h = 1/512$.

method$\backslash N$	10	20	40	80	160	rate
BE (2.13)	4.20e-4	2.20e-4	1.14e-4	5.81e-5	2.97e-5	0.96 (1.00)
SBD (basic)	2.40e-4	9.25e-5	3.66e-5	1.45e-5	5.43e-6	1.36 (−−)
BE (2.14)	4.71e-4	2.43e-4	1.24e-4	6.25e-5	3.14e-5	0.99 (1.00)
SBD (2.17)	7.64e-5	1.92e-5	4.77e-6	1.17e-6	2.73e-7	2.04 (2.00)

5. Conclusions. In this paper we develop two robust fully discrete schemes for the subdiffusion and diffusion wave equations. The schemes employ a Galerkin finite element method in space and the convolution quadrature generated by the backward Euler method and second-order backward difference. We provide a complete error analysis of the schemes, and derive optimal error estimates for both smooth and nonsmooth initial data. In particular, the schemes achieve a first-order and second-order convergence in time. We present extensive numerical experiments to illustrate the accuracy and robustness of the schemes. The experimental findings fully verify the convergence theory. Further, we compare our schemes with several existing time stepping schemes developed for smooth solutions, and find that existing ones may be not robust with respect to data regularity.

There are several questions deserving further investigation. First, in view of the solution singularity for nonsmooth data, it is of much practical interest to develop time stepping schemes using a nonuniform mesh in time and provide rigorous error analysis, including a posterior analysis. Second, our experiments indicate that existing time stepping schemes may yield only suboptimal convergence for nonsmooth data. This motivates revisiting these schemes for nonsmooth data, especially sharp error estimates. Last, it is important to study more complex models, e.g., variable coefficients in time and nonlinear models. The case of time dependent coefficients represents one of the major challenges in applying convolution quadrature, due to a lack of complete solution theory and loss of convolution structure.

Appendix A. The solution theory for the diffusion-wave equation. In the convergence analysis, the regularity of the solution to problem (1.1) plays an important role. The solution theory for $\alpha \in (0,1)$ with nonsmooth data is now well understood [45, 19, 18, 16]. Below we describe briefly the theory for $\alpha \in (1,2)$ following these works. Using the Dirichlet eigenpairs $\{(\lambda_j, \varphi_j)\}_{j=1}^{\infty}$ of the negative Laplacian $-\Delta$, the solution u to problem (1.1) with $1 < \alpha < 2$ is given by

$$u(x,t) = E(t)v + \widetilde{E}(t)b + \int_0^t \overline{E}(t-s)f(s)ds,$$

where the operators $E(t)$, $\widetilde{E}(t)$ and $\overline{E}(t)$ are given by $E(t)v = \sum_{j=1}^{\infty} E_{\alpha,1}(-\lambda_j t^\alpha)(v, \varphi_j)\varphi_j(x)$, $\widetilde{E}(t)\chi = \sum_{j=1}^{\infty} t E_{\alpha,2}(-\lambda_j t^\alpha)(\chi, \varphi_j)\varphi_j(x)$, $\overline{E}(t)\chi = \sum_{j=1}^{\infty} t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j t^\alpha)(\chi, \varphi_j)\varphi_j(x)$, respectively, where the Mittag-Leffler function $E_{\alpha,\beta}(z)$ is defined by $E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k\alpha+\beta)}$, $z \in \mathbb{C}$ [23, pp. 42]. It satisfies the following differentiation formula

$$\frac{d^m}{dt^m} t^{\beta-1} E_{\alpha,\beta}(-\lambda t^\alpha) = t^{\beta-1-m} E_{\alpha,\beta-m}(-\lambda t^\alpha) \tag{A.1}$$

and the following asymptotics: for $\frac{\alpha\pi}{2} < \mu < \min(\pi, \alpha\pi)$ [23, pp. 43]

$$|E_{\alpha,\beta}(z)| \leq c(1+|z|)^{-1} \qquad \mu \leq |\arg(z)| \leq \pi. \tag{A.2}$$

First we give a stability result for the homogeneous problem.

THEOREM A.1. *The solution $u(t)$ to problem (1.1) with $f \equiv 0$ satisfies for $t > 0$*

$$\|({}^C\partial_t^\alpha)^\ell u(t)\|_{\dot{H}^p(\Omega)} \leq c(t^{-\alpha(\ell+(p-q)/2)}\|v\|_{\dot{H}^q(\Omega)} + t^{1-\alpha(\ell+(p-r)/2)}\|b\|_{\dot{H}^r(\Omega)}),$$

where for $\ell = 0$, $0 \le q, r \le p \le 2$ and for $\ell = 1$, $0 \le p \le q, r \le 2$ and $q, r \le p + 2$.

Proof. First we discuss the case $\ell = 0$. By the triangle inequality and (A.2),

$$\|E(t)v\|^2_{\dot{H}^p(\Omega)} = \sum_{j=1}^{\infty} \lambda_j^p |(v, \varphi_j) E_{\alpha,1}(-\lambda_j t^\alpha)|^2 \le \sum_{j=1}^{\infty} t^{-\alpha(p-q)} \frac{c\lambda_j^{p-q} t^{(p-q)\alpha}}{(1 + \lambda_j t^\alpha)^2} \lambda_j^q (v, \varphi_j)^2$$

$$\le t^{-\alpha(p-q)} \sup_j \frac{c\lambda_j^{p-q} t^{(p-q)\alpha}}{(1 + \lambda_j t^\alpha)^2} \sum_{j=1}^{\infty} \lambda_j^q (v, \varphi_j)^2 \le ct^{-\alpha(p-q)} \|v\|^2_{\dot{H}^q(\Omega)},$$

where we have used $(\lambda_j t^\alpha)^{p-q}/(1 + \lambda_j t^\alpha)^2 \le c$ for $0 \le q \le p \le 2$. Similarly, one deduces $\|\widetilde{E} b\|^2_{\dot{H}^p(\Omega)} \le ct^{2-\alpha(p-r)} \|b\|^2_{\dot{H}^r(\Omega)}$. Thus the assertion for $\ell = 0$ follows by the triangle inequality. Now we consider the case $\ell = 1$. It follows from (A.2) that

$$\|\partial_t^\alpha E(t)v\|^2_{\dot{H}^p(\Omega)} = \sum_{j=1}^{\infty} \lambda_j^{2+p} (E_{\alpha,1}(-\lambda_j t^\alpha)(v, \varphi_j))^2$$

$$\le ct^{-\alpha(2+p-q)} \sum_{j=1}^{\infty} \frac{(\lambda_j t^\alpha)^{2+p-q}}{(1 + \lambda_j t^\alpha)^2} \lambda_j^q (v, \varphi_j)^2 \le ct^{-\alpha(2+p-q)} \|v\|^2_{\dot{H}^q(\Omega)}.$$

A similar estimate for $\|\partial_t^\alpha \widetilde{E}(t)b\|_{\dot{H}^p(\Omega)}$ holds, and this completes the proof. $\square$

The next result gives temporal regularity for the homogeneous problem.

THEOREM A.2. *If $v \in \dot{H}^q(\Omega)$, $b \in \dot{H}^r(\Omega)$, $q, r \in [0, 2]$, and $f \equiv 0$, then for $m \ge 1$*

$$\|\partial_t^m u\|_{L^2(\Omega)} \le ct^{q\alpha/2-m} \|v\|_{\dot{H}^q(\Omega)} + ct^{r\alpha/2-m+1} \|b\|_{\dot{H}^r(\Omega)}.$$

Proof. By (A.1), we have $\frac{d^m}{dt^m} E_{\alpha,1}(-\lambda t^\alpha) = -\lambda t^{\alpha-m} E_{\alpha,\alpha+1-m}(-\lambda t^\alpha)$. Hence, by (A.2), we deduce

$$\|\partial_t^m u\|^2_{L^2(\Omega)} = \left\| \sum_{j=1}^{\infty} \frac{d^m}{dt^m} E_{\alpha,1}(-\lambda_j t^\alpha)(v, \varphi_j) \varphi_j \right\|^2_{L^2(\Omega)}$$

$$= \sum_{j=1}^{\infty} (\lambda_j t^\alpha)^{2-q} t^{q\alpha-2m} E_{\alpha,\alpha-m+1}(-\lambda_j t^\alpha)^2 (v, \varphi_j)^2 \lambda_j^q$$

$$\le ct^{q\alpha-2m} \sup_j \frac{(\lambda_j t^\alpha)^{2-q}}{(1 + \lambda_j t^\alpha)^2} \sum_{j=1}^{\infty} (v, \varphi_j)^2 \lambda_j^q \le ct^{q\alpha-2m} \|v\|_{\dot{H}^q(\Omega)}.$$

This shows the assertion on v. The assertion on b follows analogously. $\square$

Now we turn to inhomogeneous problems. We have the following stability result.

THEOREM A.3. *For problem (1.1) with $v = b = 0$ and $f \in L^\infty(0, T; \dot{H}^q(\Omega))$, $-1 \le q \le 1$, there holds for any $\epsilon \in (0, 1)$*

$$\|u(t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} \le c\epsilon^{-1} t^{\epsilon\alpha/2} \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

Proof. Using (A.2), we deduce that for $q \ge -1$ and $0 \le p - q \le 2$, $\|\overline{E}(t)\psi\|_{\dot{H}^p(\Omega)} \le ct^{-1+(1+(q-p)/2)\alpha} \|\chi\|_{\dot{H}^q(\Omega)}$. Consequently, the desired estimate follows by

$$\|u(t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} = \left\| \int_0^t \overline{E}(t-s) f(s) ds \right\|_{\dot{H}^{q+2-\epsilon}(\Omega)} \le \int_0^t \|\overline{E}(t-s) f(s)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} ds$$

$$\le c \int_0^t (t-s)^{\epsilon\alpha/2-1} \|f(s)\|_{\dot{H}^q(\Omega)} ds \le c\epsilon^{-1} t^{\epsilon\alpha/2} \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

□

Last we state a temporal regularity result for the inhomogeneous problem.

THEOREM A.4. *If $f \in W^{1,\infty}(0,T;L^2(\Omega))$, and $v = b = 0$, then there holds*

$$\|\partial_t^m u\|_{L^2(\Omega)} \le c_T t^{\alpha-m}\|f\|_{W^{m-1,\infty}(0,T;L^2(\Omega))}, \quad m = 1, 2. \tag{A.3}$$

Proof. It follows from (A.1) and (A.2) that for $m \ge 1$

$$\|\partial_t^m \overline{E}(t)\psi\|_{L^2(\Omega)} \le c t^{\alpha-m-1}\|\psi\|_{L^2(\Omega)}. \tag{A.4}$$

For the case $m = 1$, using the representation $u(t) = \int_0^t \overline{E}(t-s)f(s)ds$, we deduce $u'(t) = \int_0^t \overline{E}'(t-s)f(s)ds$, and thus

$$\|\partial_t \overline{E}(t)u\|_{L^2(\Omega)} \le \int_0^t \|\partial_t \overline{E}(t-s)f(s)\|_{L^2(\Omega)}ds \le c \int_0^t (t-s)^{\alpha-2}\|f(s)\|_{L^2(\Omega)}ds$$
$$\le c t^{\alpha-1}\|f\|_{L^\infty(0,T;L^2(\Omega))}.$$

Using the convolution relation $t(f*g)' = f*g + (tf')*g + f*(tg')$ [35, Lemma 5.2] and (A.4), we deduce

$$t\|\partial_t^2 u\|_{L^2(\Omega)} \le c \sum_{p+q\le 1} \int_0^t \|(t-s)^p \partial_t^{p+1}\overline{E}(t-s)(s^q f^{(q)}(s))\|_{L^2(\Omega)}\,ds$$

$$\le c \sum_{p+q\le 1} \int_0^t (t-s)^{\alpha-2}s^q\|f^{(q)}(s)\|_{L^2(\Omega)}\,ds \le c_T t^{\alpha-1}\|f\|_{W^{1,\infty}(0,T;L^2(\Omega))},$$

from which the desired assertion follows. □

Like before, the solution u_h to the semidiscrete scheme (2.3) is given by

$$u_h(t) = E_h(t)v_h + \widetilde{E}_h(t)b_h + \int_0^t \overline{E}_h(t-s)f(s)ds,$$

where the operators E_h, $\widetilde{E}_h$ and $\overline{E}_h$ are given by $E_h(t)v_h = \sum_{j=1}^N E_{\alpha,1}(-\lambda_j^h t^\alpha)(v_h, \varphi_j^h)$ $\varphi_j^h(x)$, $\widetilde{E}_h(t)\chi_h = \sum_{j=1}^N tE_{\alpha,2}(-\lambda_j^h t^\alpha)(\chi_h, \varphi_j^h)\varphi_j^h(x)$, $\overline{E}_h(t)\chi_h = \sum_{j=1}^N t^{\alpha-1}E_{\alpha,\alpha}(-\lambda_j^h t^\alpha)$ $(\chi_h, \varphi_j^h)\varphi_j^h(x)$, respectively, with $\{(\lambda_j^h, \varphi_j^h)\}_{j=1}^N$ being the eigenpairs of the discrete Laplacian $-\Delta_h$. Then the following discrete counterpart of Theorem A.1 holds, where $||| \cdot |||$ denotes the discrete norm defined on X_h, induced by $-\Delta_h$ [19]. The proof is identical with that for Theorem A.1 and hence omitted.

THEOREM A.5. *The solution $u_h(t) = E_h(t)v_h + \widetilde{E}_h(t)b_h$ to problem (2.3) with $f_h \equiv 0$ satisfies for $t > 0$ and $0 \le q, r \le p \le 2$*

$$|||u_h(t)|||_{\dot{H}^p(\Omega)} \le c(t^{-\alpha(p-q)/2}|||v_h|||_{\dot{H}^q(\Omega)} + t^{1-\alpha(p-r)/2}|||b_h|||_{\dot{H}^r(\Omega)}).$$

REFERENCES

[1] E ERIC ADAMS AND LYNN W. GELHAR, *Field study of dispersion in a heterogeneous aquifer: 2. spatial moments analysis*, Water Res. Research, 28 (1992), pp. 3293–3307.

[2] DUMITRU BALEANU, KAI DIETHELM, ENRICO SCALAS, AND JUAN J. TRUJILLO, *Fractional Calculus*, World Scientific, Hackensack, NJ, 2012.

[3] EMILIA BAZHLEKOVA, BANGTI JIN, RAYTCHO LAZAROV, AND ZHI ZHOU, *An analysis of the Rayleigh–Stokes problem for a generalized second-grade fluid*, Numer. Math., 131 (2015), pp. 1–31.

[4] CHANG-MING CHEN, FAWANG LIU, I TURNER, AND V ANH, *A Fourier method for the fractional diffusion equation describing sub-diffusion*, J. Comput. Phys., 227 (2007), pp. 886–897.

[5] FENG CHEN, QINWU XU, AND JAN S. HESTHAVEN, *A multi-domain spectral method for time-fractional differential equations*, J. Comput. Phys., 293 (2015), pp. 157–172.

[6] SHENG CHEN, JIE SHEN, AND LI-LIAN WANG, *Generalized Jacobi functions and their applications to fractional differential equations*, Math. Comput., (2015), p. in press.

[7] EDUARDO CUESTA, CHRISTIAN LUBICH, AND CESAR PALENCIA, *Convolution quadrature time discretization of fractional diffusion-wave equations*, Math. Comp., 75 (2006), pp. 673–696.

[8] KAI DIETHELM, *The Analysis of Fractional Differential Equations*, Lecture Notes in Mathematics, Springer, 2004.

[9] NEVILLE J. FORD, JINGYU XIAO, AND YUBIN YAN, *A finite element method for time fractional partial differential equations*, Fract. Calc. Appl. Anal., 14 (2011), pp. 454–474.

[10] HIROSHI FUJITA AND TAKASHI SUZUKI, *Evolution problems*, in Handbook of Numerical Analysis, Vol. II, Handb. Numer. Anal., II, North-Holland, Amsterdam, 1991, pp. 789–928.

[11] YASUHIRO FUJITA, *Integrodifferential equation which interpolates the heat and the wave equation*, Osaka J. Math., 27 (1990), pp. 309–321.

[12] GUANG-HUA GAO, ZHI-ZHONG SUN, AND HONG-WEI ZHANG, *A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications*, J. Comput. Phys., 259 (2014), pp. 33–50.

[13] RUDOLF GORENFLO, FRANCESCO MAINARDI, DANIELE MORETTI, AND PAOLO PARADISI, *Time fractional diffusion: a discrete random walk approach*, Nonlin. Dyn., 29 (2002), pp. 129–143.

[14] ERNST HAIRER, SYVERT P. NØRSETT, AND GERHARD WANNER, *Solving Ordinary Differential Equations. I*, Springer-Verlag, Berlin, second ed., 1993. Nonstiff Problems.

[15] YUKO HATANO AND NAOMICHI HATANO, *Dispersive transport of ions in column experiments: An explanation of long-tailed profiles*, Water Res. Research, 34 (1998), pp. 1027–1033.

[16] BANGTI JIN, RAYTCHO LAZAROV, JOSEPH PASCIAK, AND ZHI ZHOU, *Galerkin FEM for fractional order parabolic equations with initial data in H^{-s}, $0 \leq s \leq 1$*. LNCS 8236 (Proc. 5th Conf. Numer. Anal. Appl. (June 15-20, 2012)), Springer, pp. 24–37, 2013.

[17] ———, *Error analysis of a finite element method for the space-fractional parabolic equation*, SIAM J. Numer. Anal., 52 (2014), pp. 2272–2294.

[18] ———, *Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion*, IMA J. Numer. Anal., 35 (2015), pp. 561–582.

[19] BANGTI JIN, RAYTCHO LAZAROV, AND ZHI ZHOU, *Error estimates for a semidiscrete finite element method for fractional order parabolic equations*, SIAM J. Numer. Anal., 51 (2013), pp. 445–466.

[20] ———, *An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data. IMA J. Numer. Anal.*, in press, 2015.

[21] BANGTI JIN AND WILLIAM RUNDELL, *An inverse problem for a one-dimensional time-fractional diffusion problem*, Inverse Problems, 28 (2012), pp. 075010, 19.

[22] ———, *A tutorial on inverse problems for anomalous diffusion processes*, Inverse Problems, 31 (2015), pp. 035003, 40.

[23] ANATOLY A. KILBAS, HARI MOHAN SRIVASTAVA, AND JUAN J. TRUJILLO, *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam, 2006.

[24] T.A.M. LANGLANDS AND BRUCE I. HENRY, *The accuracy and stability of an implicit solution method for the fractional diffusion equation*, J. Comput. Phys., 205 (2005), pp. 719–736.

[25] CHANGPIN LI AND HENGFEI DING, *Higher order finite difference method for the reaction and anomalous-diffusion equation*, Appl. Math. Model., 38 (2014), pp. 3802–3821.

[26] WULAN LI AND DA XU, *Finite central difference/finite element approximations for parabolic integro-differential equations*, Computing, 90 (2010), pp. 89–111.

[27] XIANJUAN LI AND CHUANJU XU, *A space-time spectral method for the time fractional diffusion equation*, SIAM J. Numer. Anal., 47 (2009), pp. 2108–2131.

[28] YUMIN LIN, XIANJUAN LI, AND CHUANJU XU, *Finite difference/spectral approximations for the fractional cable equation*, Math. Comp., 80 (2011), pp. 1369–1396.

[29] YUMIN LIN AND CHUANJU XU, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225 (2007), pp. 1533–1552.

[30] Christian Lubich, *Discretized fractional calculus*, SIAM J. Math. Anal., 17 (1986), pp. 704–719.

[31] ———, *Convolution quadrature and discretized operational calculus. I*, Numer. Math., 52 (1988), pp. 129–145.

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

[33] Francesco Mainardi, *Fractional relaxation-oscillation and fractional diffusion-wave phenomena*, Chaos, Solitons & Fractals, 7 (1996), pp. 1461–1477.

[34] ———, *Fractional Calculus and Waves in Linear Viscoelasticity*, Imperial College Press, London, 2010.

[35] William McLean, *Regularity of solutions to a time-fractional diffusion equation*, ANZIAM J., 52 (2010), pp. 123–138.

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

[37] William McLean and Vidar Thomée, *Maximum-norm error analysis of a numerical solution via Laplace transformation and quadrature of a fractional-order evolution equation*, IMA J. Numer. Anal., 30 (2010), pp. 208–230.

[38] Elliott W Montroll and George H Weiss, *Random walks on lattices. II*, J. Math. Phys., 6 (1965), pp. 167–181.

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

[40] Kassem Mustapha and William McLean, *Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations*, SIAM J. Numer. Anal., 51 (2013), pp. 491–515.

[41] Kassem Mustapha and Dominik Schötzau, *Well-posedness of hp-version discontinuous Galerkin methods for fractional diffusion wave equations*, IMA J. Numer. Anal., 34 (2014), pp. 1426–1446.

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

[43] Keith B Oldham and Jerome Spanier, *The Fractional Calculus*, Academic Press, New York, 1974.

[44] Igor Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, CA, 1999.

[45] Kenichi Sakamoto and Masahiro Yamamoto, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, J. Math. Anal. Appl., 382 (2011), pp. 426–447.

[46] Jesús María Sanz-Serna, *A numerical method for a partial integro-differential equation*, SIAM J. Numer. Anal., 25 (1988), pp. 319–327.

[47] Hansjörg Seybold and Rudolf Hilfer, *Numerical algorithm for calculating the generalized Mittag-Leffler function*, SIAM J. Numer. Anal., 47 (2008/09), pp. 69–88.

[48] Zhi-Zhong Sun and Xiaonan Wu, *A fully discrete scheme for a diffusion wave system*, Appl. Numer. Math., 56 (2006), pp. 193–209.

[49] Vidar Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2006.

[50] Santos B. Yuste, *Weighted average finite difference methods for fractional diffusion equations*, J. Comput. Phys., 216 (2006), pp. 264–274.

[51] Santos B. Yuste and Luis Acedo, *An explicit finite difference method and a new von Neumann-type stability analysis for fractional diffusion equations*, SIAM J. Numer. Anal., 42 (2005), pp. 1862–1874.

[52] Fanhai Zeng, Changpin Li, Fawang Liu, and Ian Turner, *The use of finite difference/element approaches for solving the time-fractional subdiffusion equation*, SIAM J. Sci. Comput., 35 (2013), pp. A2976–A3000.