TWO FULLY DISCRETE SCHEMES FOR FRACTIONAL DIFFUSION
AND DIFFUSION-WAVE EQUATIONS

BANGTI JIN, RAYTCHO LAZAROV, AND ZHI ZHOU

ABSTRACT. We consider the initial/boundary value problem for the fractional dif- 
sion and diffusion-wave equations involving a Caputo fractional derivative in time. 
We develop two “simple” fully discrete schemes based on the Galerkin finite element 
method in space and convolution quadrature in time with the generating function 
given by the implicit backward Euler method/second-order backward difference 
method, and establish error estimates optimal with respect to the regularity of the 
initial data. These two schemes are first and second-order accurate in time for non-
smooth initial data. Extensive numerical experiments for one and two-dimensional 
problems confirm the convergence analysis. A detailed comparison with several 
popular time stepping schemes is also performed. The numerical results indicate 
that the proposed fully discrete schemes are accurate and robust for nonsmooth 
data, and competitive with existing schemes.

Keywords: fractional diffusion, diffusion wave, finite element method, convolution 
quadrature, error estimates

1. Introduction

1.1. Mathematical model. In this work, we consider the numerical solution of the frac-
tional diffusion and diffusion wave equation. 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

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

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 $\alpha$, and it is defined by [19, pp. 91, eq. (2.4.1)]

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

where $\Gamma$ is the Gamma function. We shall also need also the Riemann-Liouville fractional 
derivative $\partial_t^\alpha u$ defined by [19, pp. 70, eq. (2.1.5)]

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

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 case, i.e., $0 < \alpha < 1$ 
and $1 < \alpha < 2$, with the Caputo derivative, which is known as a fractional-diffusion and 
diffusion-wave equation, respectively.

Throughout, we assume that problem (1.1) is subject to the following boundary con-
ditions

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

and the initial conditions

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

Date: started Feb, 2014; today is July 30, 2014.
Fractional differential equations have received significant attention over the last few decades, since they can faithfully capture the dynamics of physical processes involving anomalous transport mechanism. The model \((1.1)\) is the macroscopic counterpart of the continuous time random walk \([35, 11]\). For example, fractional diffusion has been successfully used to describe diffusion in media with fractal geometry \([38]\), highly heterogeneous aquifer \([1]\) and underground environmental problem \([12]\). Mainardi \([30, 31]\) pointed out that the diffusion wave equation governs the propagation of mechanical diffusive waves in viscoelastic media. We refer to \([19, 40, 6]\) for interesting applications and mathematical theory, and \([2]\) for a comprehensive survey on numerical methods for fractional ordinary differential equations.

1.2. Review of existing numerical schemes and motivation for our study. Due to significant potentials of \((1.1)\) in practical applications, its accurate numerical solution has attracted immense interest in recent years. 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 certain smoothness of the solution. Many existing methods for problem \((1.1)\) requires that the solution is a \(C^2\) or \(C^3\)-function in time, cf. Table 1; in Section 2.3 we will give more details on several popular schemes and further comments. However, such analysis is not always realistic since often the solution does not have the required regularity. For example, the \(C^2\) in time regularity assumption on the solution does not hold for the homogeneous problem, i.e., \(f \equiv 0\). In case of the initial data \(v \in L^2(\Omega)\) and \(\alpha \in (0, 1)\), the best possible regularity is \(C^\alpha \partial_t^\alpha u(t) \in L^2(\Omega)\) for any fixed \(t > 0\) \([41]\, \text{Theorem 2.1}\):

\[
\|C^\alpha \partial_t^\alpha u(t)\|_{L^2(\Omega)} \leq Ct^{-\alpha}\|v\|_{L^2(\Omega)}.
\]

Similarly, if \(v \in H^1(\Omega)\), the following bound \(\|C^\alpha \partial_t^\alpha u(t)\|_{L^2(\Omega)} \leq Ct^{-\alpha/2}\|v\|_{H^1(\Omega)}\) holds. Obviously, in both cases the \(\alpha\)-th order Caputo derivative in time is not bounded near \(t = 0\) (not to mention higher-order derivatives). Hence, the convergence rates listed in Table 1 may not hold. The main technical difficulty in a rigorous error analysis for nonsmooth data stems from the non-locality of the time fractional derivative and from the lower regularity of the solution when modeling realistic cases.

| method               | rate       | derivative | regularity assumption |
|---------------------|------------|------------|-----------------------|
| Lin-Xu \([26, 44]\) | \(O(\tau^{-\alpha})\) | Caputo     | \(\forall x \in \Omega, u \in C^2\) in \(t\) |
| Zeng et al I \([49]\) | \(O(\tau^{-\alpha})\) | Caputo     | \(\forall x \in \Omega, u \in C^2\) in \(t\) |
| Zeng et al II \([49]\) | \(O(\tau^{-\alpha})\) | Caputo     | \(\forall x \in \Omega, u \in C^2\) in \(t\) |
| Li-Xu \([23]\)       | \(O(\tau^3)\) | Caputo     | \(\forall x \in \Omega, u \in C^3\) in \(t\) |
| Gao et al \([10]\)   | \(O(\tau^3)\) | Caputo     | \(\forall x \in \Omega, u \in C^3\) in \(t\) |
| L1 scheme \([39, 20]\) | \(O(\tau^3)\) | RL         | \(\forall x \in \Omega, u \in C^2\) in \(t\) |
| SBD \([21]\)         | \(O(\tau^3)\) | RL         | \(\forall x \in \Omega, \text{\(\frac{D}{\partial t}\)}^{3-\alpha}u \in L^1\) in \(t\) |

In the fractional diffusion case, there are two predominant approximations in time: the L1-type approximation \([39, 20, 26, 44, 24, 25, 10]\) and the Grünwald-Letnikov approximation \([47, 4, 49]\). A summary of the convergence rates and the required regularity of solution for a number of known schemes is given in Table 1.

To the first group, L1 approximations, belongs the method devised by Langlands and Henry \([20]\). They analyzed the discretization error for the Riemann-Liouville derivative. Also, Lin and Xu \([26]\) developed a numerical method based on a finite difference scheme for the Caputo derivative and a Legendre collocation spectral method in space, and analyzed
the stability and studied the convergence rates. The scheme has a local truncation error $O(\tau^{2-\alpha})$; see also [44]. Li and Xu [24] extended the work [26] and developed a space-time spectral element method, but only for the case of zero initial data; see also [25, 7]. In [23] a variant of the L1 approximation was analyzed, and a convergence rate $O(\tau^2)$ was established for $C^4$ solutions. Recently, Gao et al [10] derived a new L1-type formula based on quadratic interpolation with a convergence rate $O(\tau^{3-\alpha})$ for smooth solutions. To overcome the local singularity of the solution and to enhance the computational efficiency a nonuniform mesh in time was suggested in [48, 50]. We also refer interested readers to [32, 36, 37] for studies on piecewise constant and piecewise linear discontinuous Galerkin discretization of the Riemann-Liouville derivative in time.

In the second group, Yuste and Acedo [47] suggested a Grünwald-Letnikov discretization of the Riemann-Liouville derivative and the central finite difference in space, and provided a von Neumann type stability analysis of the scheme. Zeng et al [49] 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 the finite element method (FEM) in space, and analyzed their stability and convergence. The convolution quadrature due to Lubich [27, 28] provides a systematic strategy for deriving high-order schemes for the Riemann-Liouville derivative, and has been the foundation of many existing works (see e.g., [47, 46, 4] for some earlier works). However, the error estimates in these works were derived under the assumption that the solution is sufficiently smooth in time. Further, in the latter group, except [49], all works focus exclusively on the Riemann-Liouville derivative; and high-order methods were scarcely applied, despite the fact that these schemes can be conveniently analyzed, including the case of nonsmooth data [29, 5].

The study on the diffusion wave equation, $1 < \alpha < 2$, is scarce. Sun and Wu [44] derived an $L1$ type finite difference scheme and analyzed the stability and $L^\infty$ convergence by an energy method. This seems to be the only time stepping scheme for the diffusion wave equation, with rigorous analysis. It was applied further in [51, 22] for the diffusion wave equation, but with different methods for space discretization. We note that with $b = 0$ and under certain regularity assumption, problem (1.1) can be equivalently 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 [29, 33, 5], where the convolution quadrature and Laplace transform method were analyzed. The error estimates derived in these works [29, 33, 5] covers the nonsmooth case. In addition, in [5], the subdiffusion case was also briefly discussed. However, the convolution quadrature to the model (1.1) with a Caputo derivative with $1 < \alpha < 2$ has not been studied before.

The excessive smoothness required in the error analysis provided in the aforementioned studies and the lack of convolution quadrature type methods for the case of Caputo derivative motivate us to revisit these issues. The goal of this work is to develop new schemes based on convolution quadrature for the model (1.1) and to derive optimal error bounds for the approximations that are expressed in terms of the smoothness of the data, including nonsmooth data, i.e., $v \in L^2(\Omega)$ (and $b \in L^2(\Omega)$ if $1 < \alpha < 2$), which is important in inverse problems and optimal control [18].

1.3. **Our contributions and organization of the paper.** We develop two fully discrete schemes for the problem (1.1) based on convolution quadrature in time generated by backward Euler or second-order backward difference and the piecewise linear Galerkin finite element method (FEM) in space. This is achieved by a reformulation of problem (1.1) through the Riemann-Liouville fractional derivative. The time stepping schemes are of Grünwald-Letnikov type. To the best of our knowledge, the application to the Caputo derivative, especially for $1 < \alpha < 2$, is novel, and for the first time a second-order scheme is obtained for problem (1.1) with both smooth and nonsmooth data. Further, following
the general strategy [5], we prove that the fully discrete schemes are first and second-order accurate in time for both smooth and nonsmooth data. To verify the convergence theory, a number of experiments on one and two-dimensional (in space) problems are presented. The comparative study with several popular time stepping schemes indicates that the error estimates in these interesting studies are generally not valid for nonsmooth initial data.

To illustrate the distinct features of our schemes, we mention one convergence rate result below. In Theorem 3.3, we will establish that in case of $0 < \alpha < 1$ and the backward Euler method, if $v \in L^2(\Omega)$ and $v_h = P_h v$, then for $n \geq 1$

\begin{equation}
\|u(t_n) - U^n_h\|_{L^2(\Omega)} \leq C(\tau t_n^{-1} + h^2 t_n^{-\alpha})\|v\|_{L^2(\Omega)}.
\end{equation}

For data $v \in \dot{H}^1(\Omega)$ we refer to the error bound established in Corollary 3.1. This estimate differs from the known estimates listed in Table 1 in several aspects:

(a) For any fixed $t_n$, the time stepping scheme is first-order accurate.

(b) The error estimate in (1.5) deteriorates near $t = 0$, whereas that in Table 1 are uniform in $t$. The prefactor $t_n^{-1}$ in (1.5) reflects the singularity behavior (1.4) for initial data $v \in L^2(\Omega)$. A similar observation holds for the case $v \in \dot{H}^1(\Omega)$, cf. Corollary 3.1.

(c) The scheme is robust with respect to the regularity of the initial data in the sense that for fixed $t_n$ the first order in time and second order in space convergence rate hold for both smooth and nonsmooth initial data.

In [49], Zeng et al derived two schemes based on a similar idea. However, they are different from ours and their schemes are not robust with respect to data regularity; see Remark 2.3 and the comparison in Section 4. To the best of our knowledge, our construction represents the first application of convolution quadrature to the diffusion wave equation with a Caputo derivative. Further, we provide a complete analysis of the schemes, and establish optimal convergence rates for both smooth and nonsmooth data. This is in sharp contrast with existing works on the model (1.1), where the convergence analysis is often done under the assumption that the solution is sufficiently smooth, which cannot be verified directly by the problem 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 presented in Section 3. The analysis exploits the operator framework [5]. In Section 4, we present extensive numerical experiments to illustrate their convergence behavior. A comparison with several existing methods is also included. The numerical results confirm first and second-order convergence rates and their robustness with respect to regularity of the data. 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 $\tau$.

2. Two fully discrete schemes

In this part, we develop two fully discrete schemes, using the standard Galerkin FEM in space and convolution quadrature in time. The error analysis of the schemes will be presented in Section 3.

2.1. Space semidiscrete Galerkin FEM. Let $\mathcal{T}_h$ be a shape regular and quasi-uniform triangulation of the domain $\Omega$ 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

$$X_h = \{ v_h \in H^1_0(\Omega) : v_h|_T \text{ is a linear function, } \forall T \in \mathcal{T}_h \}.$$
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
\[
\begin{align*}
(P_h \varphi, \chi) &= (\varphi, \chi) \quad \forall \chi \in X_h, \\
(\nabla R_h \varphi, \nabla \chi) &= (\nabla \varphi, \nabla \chi) \quad \forall \chi \in X_h,
\end{align*}
\]
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
\[
(\partial_h^\alpha u_h, \chi) + a(u_h, \chi) = (f, \chi), \quad \forall \chi \in X_h
\]
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 $\partial_h^\alpha u_h(0) = b_h$, where $v_h \in X_h$ and $b_h \in X_h$ are approximations to the initial data $v$ and $b$, respectively. Following [45], 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 $v \in H^2(\Omega)$ 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
\[
- (\Delta_h \varphi, \chi) = (\nabla \varphi, \nabla \chi) \quad \forall \varphi, \chi \in X_h,
\]
and $f_h(t) = P_h f(t)$, the spatial semidiscrete scheme (2.1) can be rewritten into
\[
\partial_h^\alpha u_h(t) + A_h u_h(t) = f_h(t), \quad t > 0
\]
with $u_h(0) = v_h \in X_h$, and if $1 < \alpha < 2$, $\partial_h^\alpha u_h(0) = b_h \in X_h$, and $A_h = -\Delta_h$.

2.2. Fully discrete schemes. In this part, we develop two fully discrete schemes for problem (1.1). This is achieved by applying convolution quadrature [27, 28] to the Riemann-Liouville derivative $\partial_h^\alpha$ defined by (1.3). Specifically, we rewrite the semidiscrete problem (2.3) using the defining relation of the Caputo derivative. Namely, for $n - 1 < \alpha < n$ [19, pp. 91, equation (2.4.10)], we have
\[
\partial_h^\alpha \varphi(t) := \partial_h^\alpha \left[ \varphi(t) - \sum_{k=0}^{n-1} \frac{\varphi^{(k)}(0)}{k!} t^k \right].
\]
In particular, in the cases of subdiffusion, $0 < \alpha < 1$, and diffusion-wave, $1 < \alpha < 2$,
\[
\partial_h^\alpha \varphi = \begin{cases} 
\partial_h^\alpha (\varphi(t) - \varphi(0)), & 0 < \alpha < 1, \\
\partial_h^\alpha (\varphi(t) - \varphi(0) - t \varphi'(0)), & 1 < \alpha < 2.
\end{cases}
\]
Hence for $t > 0$ the spatial semidiscrete scheme (2.3) can be rewritten as
\[
\partial_h^\alpha (u_h - v_h) + A_h u_h = f_h, \quad \text{for } 0 < \alpha < 1
\]
and
\[
\partial_h^\alpha (u_h - v_h - tb_h) + A_h u_h = f_h, \quad \text{for } 1 < \alpha < 2,
\]
respectively, where $f_h = P_h f(t)$. The semidiscrete formulas (2.4) and (2.5) form the basis for the time discretization, which is done in the framework developed in [5, Sections 2 and 3], initiated in [27, 28]. Below we describe this framework.

Let $K$ be a complex valued or operator valued valued function which is analytic in a sector $\Sigma_\theta := \{ z \in \mathbb{C} : \arg z \leq \theta \}$, $\theta \in (0, \pi/2)$ and is bounded by
\[
\| K(z) \| \leq M |z|^{-\mu} \quad \forall z \in \Sigma_\theta,
\]
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
\[
k(t) = \frac{1}{2\pi i} \int_{\Gamma} K(z) e^{zt} dz, \quad t > 0,
\]
where $\Gamma$ is a contour lying in the sector of analyticity, parallel to its boundary and oriented with increasing imaginary part. With $\partial_t$ being time differentiation, we define $K(\partial_t)$ as the operator of (distributional) convolution with the kernel $k : K(\partial_t)g = k \ast g$ for a function $g(t)$ with suitable smoothness. Further, we note that the convolution rule of Laplace transforms gives the following associativity property: for the operators $K_1$ and $K_2$ (generated by the kernels $k_1$ and $k_2$), we have

$$K_1(\partial_t)K_2(\partial_t) = (K_1K_2)(\partial_t).$$

(2.7)

Now we describe the time discretization process. We divide the interval $[0, T]$ into a uniform grid with a time step size $\tau = T/N$, $N \in \mathbb{N}$, with $0 = t_0 < t_1 < \ldots < t_N = T$, and $t_n = n\tau$, $n = 0, \ldots, N$. Then the convolution quadrature $K(\partial_t)g(t)$ of $K(\partial_t)g(t)$ is defined by (see e.g. [28]):

$$K(\partial_t)g(t) = \sum_{0 \leq j \leq t} \omega_j g(t - j\tau), \quad t > 0,$$

(2.8)

where the quadrature weights $\{\omega_j\}_{j=0}^{\infty}$ are determined by the generating function: $\sum_{j=0}^{\infty} \omega_j e^{\xi j\tau} = K(\delta(\xi)/\tau)$. Here $\delta$ is the quotient of the generating polynomials of a stable and consistent linear multistep method (see e.g. [13, p. 27, Lemma 2.3]). In this work, we consider the backward Euler (BE) method and second-order backward difference (SBD) method, for which

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

We remark that the quadrature weights $\{\omega_j\}$ can be computed efficiently via the fast Fourier transform [40]. The associativity property is also valid for convolution quadratures:

$$K_1(\partial_t)K_2(\partial_t) = (K_1K_2)(\partial_t).$$

(2.9)

Next we derive the fully discrete schemes. First, we rewrite the semidiscrete scheme (2.4) in the form

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

(2.10)

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

$$U_h^n = (\partial_t^\alpha + A_h)^{-1} \partial_t^\beta v_h + (\partial_t^\alpha + A_h)^{-1} f_h.$$

(2.11)

It is equivalent to the following method: find $U_h^n$ for $n = 1, 2, \ldots, N$ such that

$$\partial_t^\alpha U_h^n + A_h U_h^n = \partial_t^\beta v_h + F_h^n, \quad \text{with} \quad U_h^0 = v_h, \quad F_h^n = P_h f(t_n).$$

(2.12)

In the same manner we derive the fully discrete scheme for the diffusion-wave equation, $1 < \alpha < 2$, for which it is to find $U_h^n$ for $n = 1, 2, \ldots$, such that

$$\partial_t^\alpha U_h^n + A_h U_h^n = \partial_t^\beta v_h + (\partial_t^\alpha t) b_h + F_h^n, \quad \text{with} \quad U_h^0 = v_h, \quad F_h^n = P_h f(t_n).$$

(2.13)

**Remark 2.1.** In the BE scheme, the term at $j = 0$ can be omitted since by construction $u_h(0) - v_h = 0$. Further, we note that the first-order convergence remains valid for the modified scheme even if the condition $\varphi(0) = 0$ does not hold [42, 29].

Next we turn to the SBD scheme. It is known that the convolution quadrature (2.8) is only first-order accurate if $g(0) \neq 0$, e.g., for $g \equiv 1$ [28, Theorem 5.1] [5, Section 3]. Hence, to get a second-order approximation one needs to introduce some correction. For this we follow the approach proposed in [29, 5]. 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

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

(2.14)
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

\[
 u_h = v_h - (\partial_\alpha^2 + A_h)^{-1} A_h v_h + (\partial_\alpha^2 + A_h)^{-1}(f_{h,0} + \tilde{f}_h)
\]

\[
 = v_h - (\partial_\alpha^2 + A_h)^{-1} \partial_t \partial_\alpha^{-1} A_h v_h + (\partial_\alpha^2 + A_h)^{-1}(\partial_t \partial_\alpha^{-1} f_{h,0} + \tilde{f}_h).
\]

Now with \( \partial_\alpha^2 \) being the convolution quadrature for the SBD formula we get

\[
 U_h^n = v_h - (\partial_\alpha^2 + A_h)^{-1} \partial_t \partial_\alpha^{-1} A_h v_h + (\partial_\alpha^2 + A_h)^{-1}(\partial_t \partial_\alpha^{-1} f_{h,0} + \tilde{f}_h).
\]

The purpose of keeping the operator \( \partial_\alpha^{-1} \) is to achieve a second-order accuracy. Letting \( 1_\tau = (0, 3/2, 1, \ldots) \), using the identity \( 1_\tau = \partial_t \partial_\alpha^{-1} \) at grid points \( t_n \) [5], and the associativity (2.9), the scheme (2.14) can be rewritten as

\[
 (\partial_\alpha^2 + A_h)(U_h^n - v_h) = -1, A_h v_h + 1_\tau f_{h,0} + \tilde{f}_h.
\]

Here the second-order fully discrete scheme for the subdiffusion case is: find \( U_h^n \), \( n \geq 1 \) such that

\[
 \partial_\alpha^2 U_h^n + A_h U_h^n + \frac{1}{2} A_h U_h^0 = \partial_\alpha^2 U_h^n + F_h^n + \frac{1}{2} F_h^0,
\]

\[
 \partial_\alpha^2 U_h^0 + A_h U_h^0 = \partial_\alpha^2 U_h^n + F_h^n, \quad 2 \leq n \leq N.
\]

Here \( F_h^n = P_h f(t_n) \) and \( U_h^0 = v_h \).

Similarly, in case of \( 1 < \alpha < 2 \), the fully discrete scheme reads

\[
 \partial_\alpha^2 U_h^n + A_h U_h^n + \frac{1}{2} A_h U_h^0 = \partial_\alpha^2 U_h^n + \partial_\alpha^2 (t h_n) + F_h^n + \frac{1}{2} F_h^0,
\]

\[
 \partial_\alpha^2 U_h^0 + A_h U_h^0 = \partial_\alpha^2 U_h^n + \partial_\alpha^2 (t h_n) + F_h^n, \quad 2 \leq n \leq N,
\]

with \( U_h^0 = v_h \) and \( F_h^n = P_h f(t_n) \). The purpose of the modification at the first step is to achieve a second order convergence. Otherwise, the scheme can only achieve a first-order convergence, unlike that for the classical parabolic problem [45].

**Remark 2.2.** It is known that without a correction the SBD scheme in general is only first-order accurate. Lubich [27, 28] developed various modifications for the first step to obtain second-order method. Here we have used his ideas. Even though these modifications are now well understood in the numerical PDEs community, it seems that this is not the case in the community of fractional differential equations. Inadvertent implementation can compromise the convergence rate [47, Section 3.2].

### 2.3. Review of some existing methods.

Now we review several existing time stepping schemes in the subdiffusion case for Caputo derivative. The first scheme is the popular L1 approximation of the fractional derivative [39, 44, 26]

\[
 C \partial_\alpha^2 \varphi(t_n) \approx \frac{1}{\Gamma(1-\alpha)} \sum_{j=0}^{n-1} b_j \frac{\varphi(t_{j+1}) - \varphi(t_j)}{\tau^\alpha}
\]

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 \( \varphi \) is twice continuously differentiable [44, 26].

The second scheme, developed by Zeng et al [49], is for the Caputo case. It is derived by applying convolution quadrature to the fractional integral reformulation, and the resulting scheme is given by

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

where the operators \( D^\alpha \) and \( L^\alpha \) 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 \( \{\omega_j\} \) are generated by the identity \( (1-z)^\alpha = \sum_{j=0}^{\infty} \omega_j z^j \), cf. [49, formula (3.13)]. The third scheme is obtained in the same spirit, but with \( L^\alpha \varphi^n = (1-\alpha/2) \varphi^n + \alpha/2 \varphi^{n-1} \), cf. [49, formula (3.14)]. The second and the third schemes converge at a rate
provided that $\varphi$ is twice continuously differentiable, and at a rate $O(\tau^2)$, if further $\varphi''(0) = 0$. We have summarized these results in Table 1.

**Remark 2.3.** The two schemes due to Zeng et al [49] 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 deteriorate the convergence, unless $\varphi'(0) = 0$. Hence, the expected second-order convergence does not hold for these two schemes in general. In contrast, we have applied the SBD method, and made necessary modifications to maintain the second-order convergence rate.

Further, we have the following comments on the methods given in Table 1.

**Remark 2.4.** The convergence rate results in Table 1 were mostly obtained by a Taylor series argument, and thus the error estimates are suboptimal with respect to the regularity of the solution. The analysis of our schemes in Section 3 below relies essentially on Laplace transform and is different from the analysis of existing schemes. Further, we note that our approach is quite different from the method developed in [34] for the case of space fractional diffusion equations.

**Remark 2.5.** The SBD scheme for the Riemann-Liouville derivative was recently analyzed [21], by means of a Fourier transform and uses substantially the zero extension $\tilde{\varphi}$ of $\varphi$ for $t < 0$. In particular, the assumption $\frac{R}{-\infty} \left[ D_t^{\alpha} \tilde{\varphi} \right] \in L^1(\mathbb{R})$ requires $\varphi(0) = \varphi'(0) = 0$ and also $\varphi''(0) = 0$ for $\alpha$ close to zero. These conditions are very restrictive, and do not hold even in the simplest case of a homogeneous problem [41]. In view of Remark 2.2, these conditions might be unnecessary for the SBD scheme.

Finally, for the diffusion wave equation, $1 < \alpha < 2$, there seems only one numerical scheme in the literature, i.e., the Crank-Nicolson method, proposed and analyzed in [44]. It approximates the fractional derivative $\partial_t^\alpha \varphi(t_{n-1/2})$ by

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

where $a_j = (j+1)^{2-\alpha} - j^{2-\alpha}$. The local truncation error of the method is $O(\tau^{3-\alpha})$, if the function $\varphi$ is three times continuously differentiable [44].

3. Error analysis of the numerical schemes

Now we analyze the fully discrete schemes derived in Section 2. Our goal and main achievements are error estimates that are expressed not by some assumed regularity of the solution but directly in terms of the initial data, including nonsmooth data. The analysis is done in two steps. First we study the semidiscrete method (2.3) and establish the appropriate error bounds for the error $u(t) - u_h(t)$. Then we analyze the error $u_h(t_n) - U_h^n$ between the solutions of the semidiscrete and the fully discrete problems.

To this end, we first introduce some notation. For $q \geq 0$, we denote by $H^q(\Omega) \subset L^2(\Omega)$ the Hilbert space induced by the norm $\|v\|_{H^q(\Omega)} = \sum_{j=1}^{\infty} \lambda^q_j (v, \varphi_j)^2$, $\{ (\lambda_j, \varphi_j) \}_{j=1}^{\infty}$ being the Dirichlet eigenpairs of $-\Delta$ on $\Omega$. The set $\{ \varphi_j \}_{j=1}^{\infty}$ forms an orthonormal basis in $L^2(\Omega)$. Thus $\|v\|_{H^q(\Omega)} = \|v\|$ is the norm in $L^2(\Omega)$, $\|v\|_{H^1(\Omega)}$ the norm in $H_0^1(\Omega)$ and $\|v\|_{H^2(\Omega)} = \|\Delta v\|$ is equivalent to the norm in $H_0^3(\Omega) \cap H^2(\Omega)$ [45].

### 3.1. Error analysis of the semidiscrete scheme

The semidiscrete scheme (2.3) for the subdiffusion case was already studied in [17, Section 3] (see also [16] for very weak data and [15] for inhomogeneous problem). Hence we focus on the diffusion wave problem, i.e., $1 < \alpha < 2$. We employ an operator technique developed in [8]. First we derive an integral representation...
representation of the solution. Since the solution \( u : (0, T] \to L^2(\Omega) \) can be analytically extended to the sector \( \{ z \in \mathbb{C} : z \neq 0, |\arg z| < \pi/2 \} \) [41, Theorem 2.3], we may apply the Laplace transform to (1.1) to deduce
\[
(3.1) \quad z^\alpha \hat{u}(z) + A \hat{u}(z) = z^{\alpha-1} v + z^{\alpha-2} b,
\]
with \( A = -\Delta \). Hence the solution \( u(t) \) can be represented by
\[
(3.2) \quad u(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha I + A)^{-1} (z^{\alpha-1} v + z^{\alpha-2} b) \, dz,
\]
where the contour \( \Gamma_{\theta,\delta} \) is given by
\[
\Gamma_{\theta,\delta} = \{ z \in \mathbb{C} : |z| = \delta, |\arg z| \leq \theta \} \cup \{ z \in \mathbb{C} : z = \rho e^{i\theta}, \rho \geq \delta \}.
\]
Throughout, we choose the angle \( \theta \) such that \( \pi/2 < \theta < \min(\pi, \pi/\alpha) \) and hence \( z^\alpha \in \Sigma_{\alpha'} \) with \( \theta' = \alpha \theta < \pi \) for all \( z \in \Sigma_{\theta} := \{ z \in \mathbb{C} : |\arg z| \leq \theta \} \). Then there exists a constant \( C \) which depend only on \( \theta \) and \( \alpha \) such that
\[
(3.3) \quad \|(z^\alpha I + A)^{-1}\| \leq C z^{-\alpha}, \quad \forall z \in \Sigma_{\theta}.
\]
Similarly, with \( A_h = -\Delta_h \), the solution \( u_h \) to (2.3) can be represented by
\[
(3.4) \quad u_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha + A_h)^{-1} (z^{\alpha-1} v_h + z^{\alpha-2} b_h) \, dz.
\]

The next lemma shows an important error estimate [8, 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
\[
(3.5) \quad \|w_h - w\|_{L^2(\Omega)} + h\|\nabla(w_h - w)\|_{L^2(\Omega)} \leq C h^2 \|\varphi\|_{L^2(\Omega)}.
\]

Now we can state an error estimate for the scheme (2.3) with \( v, b \in L^2(\Omega) \).

**Theorem 3.1.** Let \( u \) and \( u_h \) be the solutions of problem (1.1) and (2.3) with \( v, b \in L^2(\Omega) \), \( v_h = P_h v \) and \( b_h = P_h b \), respectively. Then for \( t > 0 \), there holds:
\[
(3.6) \quad \|u(t) - u_h(t)\|_{L^2(\Omega)} + h\|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq C h^2 \left( t^{-\alpha} \|v\|_{L^2(\Omega)} + t^{1-\alpha} \|b\|_{L^2(\Omega)} \right).
\]

**Proof.** By (3.2) and (3.4), the error \( e(t) := u(t) - u_h(t) \) can be represented as
\[
e(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} \left(z^{\alpha-1}(w^v - w_h^v) + z^{\alpha-2}(w^b - w_h^b)\right) \, dz,
\]
with \( w^v = (z^\alpha I + A)^{-1} v, w_h^v = (z^\alpha I + A_h)^{-1} P_h v \) and \( w_h^b = (z^\alpha I + A_h)^{-1} P_h b \). By Lemma 3.1 and choosing \( \delta = 1/t \) we have
\[
\|\nabla e(t)\|_{L^2(\Omega)} \leq C h \left( \int_{-\theta}^{\theta} e^{t \cos \psi} \rho^{-\alpha} \, d\psi + \int_{1/t}^{\infty} e^{t \cos \theta} \rho^{-\alpha-1} \, d\rho \right) \|v\|_{L^2(\Omega)}
+ C h \left( \int_{-\theta}^{\theta} e^{t \cos \psi} \rho^{-\alpha} \, d\psi + \int_{1/t}^{\infty} e^{t \cos \theta} \rho^{-\alpha-2} \, d\rho \right) \|b\|_{L^2(\Omega)}
\leq C h \left( t^{-\alpha} \|v\|_{L^2(\Omega)} + t^{1-\alpha} \|b\|_{L^2(\Omega)} \right).
\]
A similar argument yields the \( L^2 \)-estimate. \( \square \)

Next we turn to the case of smooth initial data, i.e., \( v, b \in \dot{H}^2(\Omega) \).

**Theorem 3.2.** Let \( u \) and \( u_h \) be the solutions of problems (1.1) and (2.3) with \( v, b \in \dot{H}^2(\Omega) \), \( v_h = R_h v \) and \( b_h = R_h b \), respectively. Then for \( t > 0 \), there holds
\[
(3.6) \quad \|u(t) - u_h(t)\|_{L^2(\Omega)} + h\|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq C h^2 \left( \|v\|_{\dot{H}^2(\Omega)} + t \|b\|_{\dot{H}^2(\Omega)} \right).
\]
Proof. Like before, the error $e(t) := u(t) - u_h(t)$ can be represented as

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

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

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

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

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

Consequently,

$$\|I\|_{L^2(\Omega)} \leq C h^2 \|Av\|_{L^2(\Omega)} \left( \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} z^{\alpha-1} \, dz \right) \leq C h^2 \|Av\|_{L^2(\Omega)} \left( \int_{1/2}^\infty e^{zt} \cos \theta t^{-1} \, dt + \int_0^\theta e^{zt} \cos \psi \, d\psi \right) \leq C h^2 \|v\|_{H^2(\Omega)}.$$

We derive a bound for $II$ in a similar way:

$$\|II\|_{L^2(\Omega)} \leq C h^2 \|Ab\|_{L^2(\Omega)} \left( \frac{1}{2\pi i} \int_1^\delta e^{zt} z^{\alpha-2} \, dz \right) \leq C h^2 \|b\|_{H^2(\Omega)}.$$

and the $L^2$-error estimate follows. The $H^1$-estimate is established analogously. \qed

Remark 3.1. For $v, b \in H^2(\Omega)$, we may also take $v_h = P_h v$ and $b_h = P_h b$. Then

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

where $E$ and $E_h$ are solution operators (see Appendix A for the definitions). The first term is already bounded in Theorem 3.2. By Theorem A.2, there holds

$$\|E_h(t)(P_h v - R_h v)\|_{H^p(\Omega)} \leq C \|P_h v - R_h v\|_{H^p(\Omega)} \leq C h^{2-p} \|v\|_{H^2(\Omega)}.$$ 

An estimate on $b_h$ follows analogously. Hence the error estimate (3.6) holds also for the choice $v_h = P_h v$ and $b_h = P_h b$. By Theorem 3.1 and interpolation, we deduce that for $v_h = P_h v$, $b_h = P_h b$, all $q, r \in [0,2]$, and $t > 0$, there holds

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} + h\|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq C h^2 (t^{-\alpha-2(q/2)} + t^{13-\alpha(q/2)}),$$

and the $H^1$-error estimate is established in a similar way.

3.2 Error analysis for BE method. Now we derive $L^2$ error estimates for the fully discrete schemes (2.12) and (2.13) using the technique developed in [28, 5]. Here we denote for $z \in \Sigma_\theta$, $\theta \in (0,1/2)$, $G(z) = z^\alpha (z^\alpha + A_h)^{-1}$. Then for the homogeneous problem ($f \equiv 0$), by (2.10) and (2.11), the difference between $U_h^n$ and $u_h(t_n)$ can be represented by

$$U_h^n - u_h(t_n) = (G(\partial_t) - G(\partial_t)) v_h.$$

For the error analysis, we need the following estimate [28, Theorem 5.2].
Lemma 3.2. Let $K(z)$ be analytic in $\Sigma_{\alpha-\theta}$ and (26) hold. Then for $g(t) = ct^{\beta-1}$, the convolution quadrature based on the BE method satisfies
\[
\| (K(\partial_t) - K(\partial_t)) g(t) \| \leq \left\{ \begin{array}{ll}
C t^{\mu-1} \tau^{\beta}, & 0 < \beta \leq 1, \\
C t^{\mu+\beta-2} \tau, & \beta \geq 1.
\end{array} \right.
\]

Then an estimate for $u_h(t_n) - u_h^n$ for $v \in L^2(\Omega)$ follows immediately.

Lemma 3.3 (0 < $\alpha$ < 1). Let $u_h$ and $u_h^n$ be the solutions of problem (2.3) and (2.12) with $v \in L^2(\Omega)$, $U_h^0 = v_h = P_h v$ and $f \equiv 0$, respectively. Then there holds
\[
\| u_h(t_n) - U_h^n \|_{L^2(\Omega)} \leq C \tau t_n^{-1} \| v_h \|_{L^2(\Omega)}, \quad n \geq 1.
\]

Proof. By (3.3), there holds $G(z) \leq C$ for $z \in \Sigma_{\theta}$. Hence (3.7) and Lemma 3.2 (with $\mu = 0$ and $\beta = 1$) give
\[
\| u_h(t_n) - U_h^n \|_{L^2(\Omega)} \leq C \tau t_n^{-1} \| v_h \|_{L^2(\Omega)}.
\]
and the desired result follows directly from the $L^2(\Omega)$ stability of $P_h$.

Remark 3.2. The stability of the fully discrete scheme follows from Lemma 3.3
\[
\| U_h^n \|_{L^2(\Omega)} \leq \| U_h^n - u_h(t_n) \|_{L^2(\Omega)} + \| u_h(t_n) \|_{L^2(\Omega)} \leq C \| U_h^n \|_{L^2(\Omega)}.
\]

Next we turn to the case of smooth initial data, i.e., $v \in \dot{H}^2(\Omega)$.

Lemma 3.4 (0 < $\alpha$ < 1). Let $u_h$ and $u_h^n$ be the solutions of problem (2.3) and (2.12) with $v \in \dot{H}^2(\Omega)$, $U_h^0 = v_h = R_h v$ and $f \equiv 0$, respectively. Then there holds
\[
\| u_h(t_n) - U_h^n \|_{L^2(\Omega)} \leq C \tau t_n^{-1} \| Av \|_{L^2(\Omega)}, \quad n \geq 1.
\]

Proof. Using the fact $G(z) = I - (z^\alpha + A_h)^{-1} A_h$ and denoting $G_z(z) = (z^\alpha + A_h)^{-1}$, then $U_h^n - u_h(t_n) = (G_z(\partial_t) - G_z(\partial_t)) A_h v_h$. Now using (3.3) and Lemma 3.2 (with $\mu = \alpha$ and $\beta = 1$) gives
\[
\| u_h(t_n) - U_h^n \|_{L^2(\Omega)} \leq C \tau t_n^{-1} \| A_h v_h \|_{L^2(\Omega)}.
\]
Now the desired result follows directly from the fact that $A_h R_h = P_h A$.

The error estimates for the fully discrete scheme (2.12) follows from the triangle inequality and [17, Section 3]. We note that the log factor in the estimates in [17, Section 3] can be removed using the operator trick in Section 3.1.

Theorem 3.3 (0 < $\alpha$ < 1). Let $u$ and $U_h^n$ be the solutions of problem (1.1) and (2.12) with $U_h^0 = v_h$ and $f \equiv 0$, respectively. Then the following estimates hold
(a) If $v \in \dot{H}^2(\Omega)$ and $v_h = R_h v$, then for $n \geq 1$
\[
\| u(t_n) - U_h^n \|_{L^2(\Omega)} \leq C (\tau t_n^{-2} + h^2) \| v \|_{\dot{H}^2(\Omega)}.
\]
(b) If $v \in L^2(\Omega)$ and $v_h = P_h v$, then for $n \geq 1$
\[
\| u(t_n) - U_h^n \|_{L^2(\Omega)} \leq C (\tau t_n^{-2} + h^2 t_n^{-\alpha}) \| v \|_{L^2(\Omega)}.
\]

Corollary 3.1. Using interpolation we can get the following bound for data of intermediate smoothness, i.e., for $v \in \dot{H}^1(\Omega)$:
\[
\| u(t_n) - U_h^n \|_{L^2(\Omega)} \leq C (\tau t_n^{-2} + h^2 t_n^{-\alpha/2}) \| v \|_{\dot{H}^1(\Omega)}.
\]

The diffusion wave case follows analogously.

Theorem 3.4 (1 < $\alpha$ < 2). Let $u$ and $U_h^n$ be the solutions of problem (1.1) and (2.13) with $U_h^0 = v_h$ and $f \equiv 0$, respectively. Then the following estimates hold
(a) If $v, b \in \dot{H}^2(\Omega)$ and $v_h = R_h v$ and $b_h = R_h b$, then
\[
\| u(t_n) - U_h^n \|_{L^2(\Omega)} \leq C \left( (\tau t_n^{\alpha-1} + h^2) \| v \|_{\dot{H}^2(\Omega)} + (\tau t_n^{\alpha} + h^2 t_n) \| b \|_{\dot{H}^2(\Omega)} \right).
\]
(b) If \( v, b \in L^2(\Omega) \) and \( v_h = P_h v \) and \( b_h = P_h b \), then
\[
\|u(t_h) - U^n_{\alpha,h}\|_{L^2(\Omega)} \leq C \left( (\tau t_n^{-\alpha} + h^2 t_n^{-\alpha}) \|v\|_{L^2(\Omega)} + (\tau + h^2 t_n^{-\alpha}) \|b\|_{L^2(\Omega)} \right).
\]

3.3. Error analysis for the SBD scheme. Now we turn to the analysis of the SBD scheme. Like Lemma 3.2, the following estimate holds [28, Theorem 5.2].

**Lemma 3.5.** Let \( K(z) \) be analytic in \( \Sigma_{\alpha-\beta} \) and (2.6) hold. Then for \( g(t) = ct^{\beta-1} \), the convolution quadrature based on the SBD satisfies
\[
\|(K(\partial_t) - K(\partial_t^h))g(t)\| \leq \left\{ \begin{array}{ll}
C \mu^{-1} \tau^{\beta}, & 0 < \beta \leq 2, \\
C \mu^{\beta} (\beta-2)^2, & \beta \geq 2.
\end{array} \right.
\]

Now we state the following result for the nonsmooth data, i.e., \( v \in L^2(\Omega) \).

**Lemma 3.6.** Let \( u_h \) and \( U^n_{\alpha,h} \) be the solutions of problem (2.3) and (2.15) with \( v \in L^2(\Omega) \), \( U^0_{\alpha} = v_h = P_h v \) and \( f \equiv 0 \), respectively. Then there holds
\[
\|u_h(t_n) - U^n_{\alpha,h}\|_{L^2(\Omega)} \leq C \tau^2 t_n^{-2} \|v\|_{L^2(\Omega)}.
\]

**Proof.** Like before, the difference between \( u_h(t_n) \) and \( U^n_{\alpha,h} \) can be represented by
\[
u_h(t_n) - U^n_{\alpha,h} = (G(\partial_t^h) - G(\partial_t))A_h v_h(t_n),
\]
where \( G(z) = -z(z^{\alpha} + A_h)^{-1}A_h \). By (3.3) and the identity
\[
G(z) = -z(z^{\alpha} + A_h)^{-1}A_h = -zI + z^{\alpha+1}(z^{\alpha} + A_h)^{-1}, \quad \forall z \in \Sigma_{\alpha-\beta},
\]
there holds \( \|G(z)\| \leq C|z| \), for \( z \in \Sigma_{\alpha} \). Then Lemma 3.5 (with \( \mu = 1 \) and \( \beta = 2 \)) gives
\[
\|U^n_{\alpha,h} - u_h(t_n)\|_{L^2(\Omega)} \leq C \tau^2 t_n^{-2} \|v_h\|_{L^2(\Omega)},
\]
and the desired result follows directly from the \( L^2(\Omega) \) stability of \( P_h \).

Next we turn to smooth initial data \( v \in H^2(\Omega) \).

**Lemma 3.7.** Let \( u_h \) and \( U^n_{\alpha,h} \) be the solutions of problem (2.3) and (2.15) with \( v \in H^2(\Omega) \), \( U^0_{\alpha} = v_h = R_h v \) and \( f \equiv 0 \), respectively. Then there holds
\[
\|u_h(t_n) - U^n_{\alpha,h}\|_{L^2(\Omega)} \leq C \tau^2 t_n^{-2} \|A_h v_h\|_{L^2(\Omega)}.
\]

**Proof.** By setting \( G_s(z) = -z(z^{\alpha} + A_h)^{-1} \), \( U^n_{\alpha,h} - u_h(t_n) \) can be written by
\[
u^n_{\alpha,h} - u_h(t_n) = (G_s(\partial_t^h) - G_s(\partial_t))A_h v_h.
\]
From (3.3), we deduce
\[
\|G_s(z)\| \leq M|z|^{1-\alpha} \quad \forall z \in \Sigma_{\alpha}.
\]
Now Lemma 3.5 (with \( \mu = \alpha - 1 \) and \( \beta = 2 \)) gives
\[
\|U^n_{\alpha,h} - u_h(t_n)\|_{L^2(\Omega)} \leq C \tau^2 t_n^{-2} \|A_h v_h\|_{L^2(\Omega)},
\]
and the desired estimate follows from the identity \( A_h R_h = P_h A \).

The proof for the diffusion-wave equation is identical and hence omitted.

**Theorem 3.5.** Let \( u \) and \( U^n_{\alpha,h} \) be the solutions of problem (1.1) and (2.15)/(2.16) with \( U^0_{\alpha} = P_h v \) and \( f \equiv 0 \), respectively. Then the following estimates hold for \( 0 \leq q, r \leq 2 \).

(a) If \( 0 < \alpha < 1 \), then
\[
\|u(t_n) - U^n_{\alpha,h}\|_{L^2(\Omega)} \leq C (\tau^2 t_n^{-2q/\alpha} + h^2 t_n^{-(2-q)\alpha/2}) \|v\|_{H^q(\Omega)}.
\]

(b) If \( 1 < \alpha < 2 \) and \( b_h = P_h b \), then
\[
\|u(t_n) - U^n_{\alpha,h}\|_{L^2(\Omega)} \leq C \left( \tau^2 t_n^{-2q/\alpha} + h^2 t_n^{-(2-q)\alpha/2} \right) \|v\|_{H^q(\Omega)}
+ C \left( \tau^2 t_n^{r-1} + h^2 t_n^{-(2-r)\alpha/2} \right) \|b\|_{H^r(\Omega)}.
\]
4. Numerical experiments

In this section, we present numerical results to illustrate the efficiency and accuracy of our fully discrete schemes, and to verify the convergence theory in Section 3.

The exact solution of each example can be written explicitly as an infinite series involving the Mittag-Leffler function $E_{\alpha,\beta}(z)$ ([17] and Appendix A), for which we employ an algorithm developed in [43]. We measure the error $e^n = u(t_n) - U^n_h$ by the normalized errors $\|e^n\|_{L^2(\Omega)}/\|v\|_{L^2(\Omega)}$ and $\|e^n\|_{H^1(\Omega)}/\|v\|_{L^2(\Omega)}$. In the computations, we divide the unit interval $\Omega = (0, 1)$ into $M = 2^n$ equally spaced subintervals with a mesh size $h = 1/M$. For the two-dimensional examples, the domain $\Omega = (0, 1)^2$ is first divided into $M^2$ small equal squares, and we obtain a symmetric triangulation by connecting the diagonal of each small square. Likewise, we fix the time step size $\tau = t/N$ in order to examine the spatial and temporal convergence separately, we take a small time step size (or mesh size $h$), respectively, so that the temporal (or spatial) discretization error is negligible. We describe results for the cases of subdiffusion and diffusion-wave separately.

4.1. Subdiffusion ($0 < \alpha < 1$).

We consider the following three examples:

(a) $\Omega = (0, 1)$, and $v = 1 \in \dot{H}^{1/2-\epsilon}(\Omega)$ with $\epsilon > 0$;
(b) $\Omega = (0, 1)^2$, and $v = xy(1-x)(1-y) \in \dot{H}^2(\Omega)$;
(c) $\Omega = (0, 1)^2$, and $v = \chi_{(0,1)^2 \times (0,1)} \in \dot{H}^{1/2-\epsilon}(\Omega)$ with $\epsilon > 0$.

Since the spatial convergence of the semidiscrete scheme for the subdiffusion equation was already studied in [17, 16], we focus on the temporal convergence rate at $t = 0.1$.

Numerical results for example (a). The numerical results for example (a) are shown in Table 2. For comparison we have also included the numerical results obtained by the three existing schemes described in Section 2. In all tables, rate refers to the empirical convergence rate of the errors when the time step size $\tau$ (or the mesh size $h$) halves, and the numbers in the bracket denote theoretical convergence rates. The proposed schemes, BE (backward Euler) and SBD (second-order backward difference), achieve first and second-order convergence, respectively, independently of the fractional order $\alpha$. This is in excellent agreement with the theory. Theoretically, the finite difference method due to Lin and Xu [26] achieves a convergence rate $O(\tau^{2-\alpha})$ for smooth solutions. However, for nonsmooth initial data (such as those provided in example (a)), it can only achieve a first-order convergence. As to the two schemes due to Zeng et al [49], the convergence of their first scheme strongly depends on the fractional order $\alpha$, and fails to achieve a first-order rate for $\alpha$ close to unity. Their second scheme, which theoretically is $O(\tau^{2-\alpha})$ accurate, can only achieve a first-order convergence for nonsmooth data. These results indicate that existing time stepping schemes may not work well for nonsmooth data, whereas our fully discrete schemes are robust and accurate.

Numerical results for examples (b) and (c). In Table 3, we report the numerical results for examples (b) and (c) with $\alpha = 0.5$; see also Figure 1 for the solution profiles. For both examples, a convergence rate $O(\tau)$ and $O(\tau^2)$ is observed for the BE and the SBD scheme, respectively, which agrees with our convergence theory in Section 3. Further, if the spatial error is negligible, $N$ is fixed and $t_N \to 0$, then we deduce from Theorem 3.3 and interpolation that

$$\|U^N_h - u(t_N)\|_{L^2(\Omega)} \leq C t^{4\alpha/2}_N N^{-1}\|e\|_{H^\alpha(\Omega)}; \tag{4.1}$$

In Table 4 and Figure 2 we show the $L^2$-norm of the error for examples (b) and (c), for fixed $N = 10$ and $t_N \to 0$ with $\alpha = 0.5$. It is observed that in the smooth case (b), the temporal error decreases like $O(\tau^{1/2})$, whereas in the nonsmooth case (c), it decays like $O(\tau^{1/8})$. Note that in example (c), the initial data $v \in \dot{H}^{1/2-\epsilon}(\Omega)$ for any $\epsilon > 0$, (4.1) predicts an error decay rate $O(\tau^{\epsilon/2}) = O(\tau^{1/8})$. Hence the empirical rates in Table 4 and Figure 2 agree well with the theoretical prediction.
Table 2. The $L^2$-norm of the error for example (a) at $t = 0.1$ with $h = 2^{-12}$.

| $\alpha$, $N$ | method | 10 | 20 | 40 | 80 | 160 | 320 | rate |
|---------------|--------|----|----|----|----|-----|-----|------|
| 0.1           | BE     | 4.75e-4 | 2.35e-4 | 1.17e-4 | 5.82e-5 | 2.91e-5 | 1.45e-5 | 1.00 (1.00) |
|               | SBD    | 4.69e-5 | 1.10e-5 | 2.66e-6 | 6.56e-7 | 1.66e-7 | 4.44e-8 | 1.98 (2.00) |
|               | Lin-Xu | 4.38e-4 | 2.19e-4 | 1.09e-4 | 5.41e-5 | 2.70e-5 | 1.34e-5 | 1.00      |
|               | Zeng I | 1.04e-2 | 5.38e-3 | 2.77e-3 | 1.42e-3 | 7.28e-4 | 3.72e-4 | 0.96      |
|               | Zeng II| 3.89e-4 | 1.94e-4 | 9.70e-5 | 4.84e-5 | 2.42e-5 | 1.21e-5 | 1.00      |
| 0.5           | BE     | 5.10e-3 | 2.51e-3 | 1.24e-3 | 6.20e-4 | 3.09e-4 | 1.54e-4 | 1.00 (1.00) |
|               | SBD    | 5.97e-4 | 1.39e-4 | 3.34e-5 | 8.22e-6 | 2.04e-6 | 5.14e-7 | 2.00 (2.00) |
|               | Lin-Xu | 4.15e-3 | 1.96e-3 | 9.50e-4 | 4.65e-4 | 2.29e-4 | 1.14e-5 | 1.02      |
|               | Zeng I | 5.12e-2 | 3.04e-2 | 1.80e-2 | 1.07e-2 | 6.33e-3 | 3.75e-3 | 0.75      |
|               | Zeng II| 1.62e-3 | 8.52e-4 | 4.36e-4 | 2.20e-4 | 1.11e-4 | 5.56e-5 | 1.00      |
| 0.9           | BE     | 4.15e-3 | 1.96e-3 | 9.50e-4 | 4.65e-4 | 2.29e-4 | 1.14e-5 | 1.00      |
|               | SBD    | 5.97e-4 | 1.39e-4 | 3.34e-5 | 8.22e-6 | 2.04e-6 | 5.14e-7 | 2.00 (2.00) |
|               | Lin-Xu | 9.94e-2 | 6.77e-2 | 4.60e-2 | 3.10e-2 | 2.07e-2 | 1.35e-2 | 0.55      |
|               | Zeng I | 1.71e-2 | 1.75e-3 | 3.09e-4 | 1.60e-3 | 1.06e-3 | 5.30e-4 | 1.00      |
|               | Zeng II| 1.74e-2 | 8.72e-3 | 4.31e-3 | 2.13e-3 | 1.04e-3 | 5.18e-4 | 1.01      |

Table 3. The $L^2$-norm of the error for examples (b) and (c) for $t = 0.1$, $\alpha = 0.5$, and $h = 2^{-9}$.

| method | 5  | 10  | 20  | 40  | 80  | rate |
|--------|----|-----|-----|-----|-----|------|
| (b)    | BE | 7.00e-3 | 3.34e-3 | 1.63e-3 | 8.05e-4 | 4.00e-4 | 1.00 (1.00) |
|        | SBD| 2.00e-3 | 4.20e-4 | 9.79e-5 | 2.42e-5 | 6.54e-6 | 1.98 (2.00) |
| (c)    | BE | 4.39e-3 | 2.09e-3 | 1.02e-3 | 5.05e-4 | 2.51e-4 | 1.01 (1.00) |
|        | SBD| 1.25e-3 | 2.64e-4 | 6.13e-5 | 1.49e-5 | 3.79e-6 | 2.05 (2.00) |

Figure 1. Numerical solutions of examples (b) and (c) by SBD with $h = 2^{-7}$ and $N = 1000$, $\alpha = 0.5$ at $t = 0.1$.

Table 4. The $L^2$-norm of the error for examples (b) and (c) as $t \to 0$ with $h = 2^{-9}$ and $N = 10$.

| $t$   | 1e-3 | 1e-4 | 1e-5 | 1e-6 | 1e-7 | 1e-8 | rate |
|-------|------|------|------|------|------|------|------|
| (b)   | 6.16e-3 | 2.64e-3 | 8.93e-4 | 2.88e-4 | 9.20e-5 | 2.93e-5 | 0.49 (0.50) |
| (c)   | 5.86e-3 | 4.61e-3 | 3.32e-3 | 2.51e-3 | 1.92e-3 | 1.51e-3 | 0.12 (0.13) |
4.2. Diffusion-wave equation $(1 < \alpha < 2)$. We consider the following five examples (with $\epsilon \in (0, 1/2)$):

(d) $\Omega = (0, 1)$, $v = x(1 - x) \in \dot{H}^{3/2-\epsilon}(\Omega)$ and $b = 0$

(e) $\Omega = (0, 1)$, $v = 1 \in \dot{H}^{3/2-\epsilon}(\Omega)$ and $b = 0$.

(f) $\Omega = (0, 1)$, $v = 0$ and $b = x\chi_{[0,1/2]} + (1 - x)\chi_{[1/2,1]} \in \dot{H}^{3/2-\epsilon}(\Omega)$.

(g) $\Omega = (0, 1)$, $v = 0$ and $b = x^{-1/4} \in H^{1/4-\epsilon}(\Omega)$.

(h) $\Omega = (0, 1)^2$, $v = \sin(2\pi x)y(1 - y)$ and $b = \chi_{(0,1/2)x(0,1)}$.

The first two examples have a vanishing initial condition $b = 0$, and the next two examples have a vanishing initial condition $v = 0$. These examples allow us to examine the solution behavior with respect to the initial data $v$ and $b$ separately.

Numerical results for examples (d) and (e). From Table 5 and Figure 3 for example (d) we observe a temporal convergence rate $O(\tau)$ and $O(\tau^2)$ for the BE and SBD scheme, respectively, and a spatial convergence rate $O(h^2)$ and $O(h)$ in the $L^2$-norm and $H^1$-norm, respectively. These results are in full agreement with the analysis in Section 3. These observations remain valid for nonsmooth data, cf. Tables 6 and 8. Further, the spatial error deteriorates slightly as $t \to 0$. We note that in Table 6, the convergence of the SBD for $\alpha = 1.9$ is only of order $O(h^{1.05})$. The culprit of the apparent suboptimal convergence is due to the large time step $\tau$ at the beginning. Asymptotically, it remains second-order convergence, cf. Table 7. In Tables 5 and 6, we present also the results by the popular Crank-Nicolson scheme [44], which converges at a rate $O(\tau^{3-\alpha})$ for smooth solutions [44]. For smooth data, i.e., example (d), the theoretical rate does hold for large $\alpha$ values, e.g., $\alpha \geq 1.5$; but for small $\alpha$ values, e.g., $\alpha = 1.1$, the Crank-Nicolson scheme fails to achieve the theoretical rate, despite the fact that the initial data $v$ is fairly smooth. In the case of nonsmooth data, the Crank-Nicolson scheme can only achieve a first-order convergence, due to a lack of solution regularity.

It is widely accepted that as the fractional order $\alpha$ increases from one to two, the model (1.1) transit from the classical diffusion equation to the wave equation [9]. This transition can be observed numerically: for $\alpha$ values close to unity, the solution is very diffused and thus smooth, whereas for $\alpha$ values close to two, the plateau in the initial data $v$ is well preserved, reflecting a “finite” speed of wave propagation, cf. Figure 4. Further, the closer is the fractional order $\alpha$ to 2, the slower is the decay of the solution (for $t$ close to zero).

Like before, if the spatial error is negligible then for fixed $N$ and $t_N \to 0$, the error decay estimate (4.1) holds. In Table 9, we present the results for the BE scheme for the case $\alpha = 1.5$. The $L^2$-norm of the error decays at the theoretical rate $O(t^{3/2})$ for the example (d) and $O(t^{3/8})$ for the example (e), respectively, thereby confirming the estimate (4.1). Finally, in Table 10 we show the $L^2$-norm of the error for examples (d) and (e), for fixed $h = 2^{-13}$ and $t \to 0$. We observe that in the case of smooth data the error essentially stays unchanged, whereas in the case of nonsmooth data the error deteriorates like $O(t^{-1.13})$ as
Table 5. The $L^2$-norm of the error for example (d) at $t = 0.1$ with $h = 2^{-13}$.

| $\alpha$ | $N$   | 10   | 20   | 40   | 80   | 160  | rate |
|----------|-------|------|------|------|------|------|------|
| 1.1      | BE    | 1.34e-2 | 6.83e-3 | 3.45e-3 | 1.73e-3 | 8.68e-4 | 0.99 (1.00) |
|          | SBD   | 4.76e-4 | 1.05e-4 | 2.47e-5 | 5.97e-6 | 1.46e-6 | 2.03 (2.00) |
|          | CN    | 9.00e-5 | 2.75e-5 | 8.95e-6 | 2.88e-6 | 8.94e-7 | 1.46   |
| 1.5      | BE    | 4.95e-3 | 2.39e-3 | 1.32e-3 | 6.69e-4 | 3.36e-4 | 0.98 (1.00) |
|          | SBD   | 1.95e-3 | 4.39e-4 | 1.67e-4 | 6.00e-5 | 2.15e-5 | 1.99   |
|          | CN    | 4.33e-3 | 2.03e-3 | 9.67e-4 | 4.54e-5 | 2.12e-5 | 1.99   |
| 1.9      | BE    | 5.38e-3 | 2.72e-3 | 1.37e-3 | 6.89e-4 | 3.45e-4 | 0.90 (1.00) |
|          | SBD   | 3.08e-3 | 8.81e-5 | 2.41e-5 | 6.28e-6 | 1.60e-6 | 1.95 (2.00) |
|          | CN    | 4.33e-3 | 2.03e-3 | 9.67e-4 | 4.54e-5 | 2.12e-5 | 1.99   |

$t \to 0$. This is in excellent agreement with the theory: in view of Remark 3.1, the spatial error deteriorates as $t \to 0$ like

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

for any $\epsilon > 0$.

Table 6. The $L^2$-norm of the error for example (e) at $t = 0.1$ with $h = 2^{-13}$.

| $\alpha$ | $N$   | 10   | 20   | 40   | 80   | 160  | 320  | rate |
|----------|-------|------|------|------|------|------|------|------|
| 1.1      | BE    | 1.21e-2 | 6.16e-3 | 3.11e-3 | 1.56e-3 | 7.82e-4 | 3.92e-4 | 1.00 (1.00) |
|          | SBD   | 1.01e-3 | 1.78e-4 | 3.88e-5 | 9.13e-6 | 2.21e-6 | 5.35e-7 | 2.07 (2.00) |
|          | CN    | 9.66e-2 | 6.67e-2 | 4.59e-2 | 3.14e-2 | 2.10e-2 | 1.35e-2 | 0.63   |
| 1.5      | BE    | 2.58e-2 | 1.39e-2 | 7.30e-3 | 3.75e-3 | 1.90e-3 | 9.58e-4 | 0.98 (1.00) |
|          | SBD   | 1.40e-3 | 1.35e-3 | 4.09e-4 | 1.06e-4 | 2.81e-5 | 7.12e-6 | 2.02 (2.00) |
|          | CN    | 5.48e-2 | 3.49e-2 | 2.17e-2 | 1.28e-2 | 6.39e-3 | 1.01e-3 | 0.97   |
| 1.9      | BE    | 7.11e-2 | 5.00e-2 | 3.37e-2 | 2.17e-2 | 1.31e-2 | 7.63e-3 | 0.74 (1.00) |
|          | SBD   | 4.82e-2 | 2.78e-2 | 1.36e-2 | 5.26e-3 | 1.62e-3 | 4.25e-4 | 1.68 (2.00) |
|          | CN    | 5.48e-2 | 3.49e-2 | 2.04e-2 | 1.10e-2 | 5.17e-3 | 2.94e-3 | 0.92   |

Table 7. The $L^2$-norm of the error for example (e) with $\alpha = 1.9$ at $t = 0.1$ with $h = 2^{-14}$.

| $N$ | 80 | 160 | 320 | 640 | 1280 | rate |
|-----|----|-----|-----|-----|-----|------|
| BE  | 2.17e-2 | 1.32e-2 | 7.63e-3 | 4.21e-3 | 2.24e-3 | 0.91 (1.00) |
| SBD | 5.28e-3 | 1.62e-3 | 4.25e-4 | 1.06e-4 | 2.88e-5 | 1.94 (2.00) |
Figure 4. Numerical results for example (e) using SBD method: \( t = 0.1, h = 2^{-13}, N = 160. \)

Table 8. Numerical results example (e): \( \alpha = 1.5, h = 2^{-k}, N = 10^4. \)

| \( t \) | \( k \) | 7 | 8 | 9 | 10 | 11 | rate |
|------|------|---|---|---|----|----|------|
| 0.1  | \( L^2 \)-norm | 1.23e-4 | 3.98e-5 | 7.67e-6 | 1.90e-6 | 4.53e-7 | 2.01 (2.00) |
|      | \( H^1 \)-norm | 2.84e-2 | 1.42e-2 | 7.07e-3 | 3.51e-3 | 1.71e-3 | 1.01 (1.00) |
| 0.01 | \( L^2 \)-norm | 1.58e-3 | 4.05e-4 | 1.01e-4 | 2.51e-5 | 6.00e-6 | 2.01 (2.00) |
|      | \( H^1 \)-norm | 9.68e-1 | 1.92e-1 | 9.46e-2 | 4.67e-2 | 2.29e-2 | 1.02 (1.00) |
| 0.001| \( L^2 \)-norm | 1.32e-2 | 4.28e-3 | 1.28e-3 | 3.30e-4 | 7.97e-5 | 1.92 (2.00) |
|      | \( H^1 \)-norm | 5.72e0 | 2.84e0 | 1.37e0 | 6.42e-1 | 3.07e-2 | 1.00 (1.00) |

Table 9. The \( L^2 \)-norm of the error for examples (d), (e), (f) and (g) with \( \alpha = 1.5; t \rightarrow 0, h = 2^{-k}, \) and \( N = 10. \)

| \( t \) | 1 | 1e-1 | 1e-2 | 1e-3 | 1e-4 | 1e-5 | rate |
|------|---|------|------|------|------|------|------|
| (d)  | 1.71e-2 | 4.95e-3 | 2.84e-4 | 9.52e-6 | 3.04e-7 | 1.13e-8 | 1.47 (1.50) |
| (e)  | 1.54e-2 | 2.58e-2 | 1.06e-2 | 4.48e-3 | 1.87e-3 | 7.30e-4 | 0.38 (0.38) |
| (f)  | 2.20e-2 | 1.56e-3 | 1.17e-5 | 8.79e-8 | 6.61e-10 | 5.73e-12 | 2.11 (2.13) |
| (g)  | 1.75e-2 | 1.67e-3 | 1.02e-4 | 6.42e-5 | 4.14e-6 | 4.18e-7 | 1.15 (1.19) |

Table 10. The \( L^2 \)-norm of the error for examples (d), (e), (f) and (g) with \( \alpha = 1.5; t \rightarrow 0, h = 2^{-k}, \) and \( N = 10^4. \)

| \( t \) | 1 | 1e-1 | 1e-2 | 1e-3 | 1e-4 | 1e-5 | rate |
|------|---|------|------|------|------|------|------|
| (d)  | 1.31e-9 | 4.85e-9 | 5.72e-9 | 5.86e-9 | 5.88e-9 | 5.89e-9 | -0.02 (0) |
| (e)  | 1.19e-9 | 2.30e-8 | 3.04e-7 | 4.06e-6 | 5.37e-5 | 5.66e-4 | -1.12 (-1.13) |
| (f)  | 9.91e-10 | 6.62e-10 | 1.57e-10 | 3.72e-11 | 8.81e-12 | 2.04e-12 | 0.63 (0.63) |
| (g)  | 4.81e-10 | 1.84e-9 | 3.59e-9 | 7.16e-9 | 1.43e-8 | 2.75e-8 | -0.30 (-0.31) |

Numerical results for examples (f) and (g). Similarly to the results for examples (d) and (e), we observe a first-order and second-order convergence for the BE and SBD scheme, respectively, cf. Table 11, and the the first- and second-order convergence for the \( H^1 \)- and \( L^2 \)-norm of the error, cf. Figure 5. All the convergence rates are independent of the fractional order \( \alpha \). For the nonsmooth case, i.e., example (g), we are particularly interested in the errors for \( t \) close to zero, thus we also plot the error at \( t = 0.01 \) and \( 0.001. \) These results fully confirm the analysis in Section 3. Further, by Theorem 3.4

\[
\|u(t_N) - U^N_N\|_{L^2(\Omega)} \leq C \left( N^{-1} \delta_1 + h^2 \delta_1^{2-\alpha/2} \right) \|b\|_{H^{r}(\Omega)},
\]
Hence, if we fix $\alpha = 1.5$, $N = 10$ and let $t \to 0$, this estimate predicts a behavior $O(t^{17/8-3\varepsilon/4})$ and $O(t^{19/16-3\varepsilon/4})$ for examples (f) and (g), respectively, if the spatial error is negligible, which agree with the results in Tables 9 and 10.

**Table 11.** The $L^2$-norm of the error for examples (f) and (g) at $t = 0.1$ with $h = 2^{-13}$.

| $\alpha$ | $N$ | $h_{10}$ | $h_{20}$ | $h_{40}$ | $h_{80}$ | $h_{160}$ | $h_{320}$ | rate |
|---------|-----|-----------|-----------|-----------|-----------|-----------|-----------|------|
|         |     | BE        | SBD       | BE        | SBD       | BE        | SBD       |      |
| (f)     |     |           |           |           |           |           |           |      |
| 1.1     | 10  | 1.99e-3   | 2.34e-5  | 1.01e-3   | 8.38e-7   | 1.83e-7   | 4.00e-8   | 7.15e-9 | 1.00 (1.00) |
| 1.5     | 1.56e-3 | 7.83e-4  | 1.96e-4   | 9.81e-5   | 4.91e-5   | 1.01e-4   | 2.31 (2.00) |
| 1.9     | 8.78e-4 | 4.39e-4  | 2.20e-4   | 5.52e-5   | 2.76e-5   | 1.00 (1.00) |
| (g)     |     |           |           |           |           |           |           |      |
| 1.1     | 1.58e-3 | 8.04e-4  | 2.04e-4   | 1.02e-4   | 5.12e-4   | 1.29e-4   | 6.41e-5   | 2.31 (2.00) |
| 1.5     | 1.68e-3 | 8.71e-4  | 2.25e-4   | 1.13e-4   | 5.68e-5   | 1.10 (1.00) |
| 1.9     | 2.92e-3 | 1.70e-3  | 1.51e-5   | 5.54e-5   | 1.92 (2.00) |

**Figure 5.** Numerical results for diffusion-wave equations using SBD with $N = 1000$.

Numerical results for example (h). Finally, we present numerical solutions of the two-dimensional example. The temporal and spatial errors are showed in Tables 12 and 13, respectively. The empirical results fully confirm our analysis. In Figure 6, we plot the solution profiles at $t = 0.01$ and $t = 0.1$.

4.3. **Numerical results for multi-term time-fractional diffusion equation.** Our theory can be easily extended to the more general multi-term counterpart, i.e.,

$$
(C^{\alpha} \partial_t^{\alpha} + \sum_{j=1}^{m} \beta_j C^{\beta_j} \partial_t^{\beta_j}) u(x, t) - \Delta u(x, t) = f(x, t) \quad (x, t) \in \Omega \times (0, T).
$$
Table 12. The \( L^2 \)-norm of the error for example (h) at \( t = 0.1 \) with \( \alpha = 1.5 \) and \( h = 2^{-9} \).

| \( N \) | 5 | 10 | 20 | 40 | 80 | rate  |
|-------|---|----|----|----|----|-------|
| BE    | 1.11e-1 | 5.93e-2 | 3.09e-2 | 1.58e-2 | 7.97e-3 | 0.97 (1.00) |
| SBD   | 1.60e-2 | 4.70e-3 | 1.29e-3 | 3.30e-4 | 8.23e-5 | 1.95 (2.00) |

Table 13. Numerical results for example (h) with \( \alpha = 1.5 \), at \( t = 0.1 \) and \( t = 0.01 \) with \( h = 2^{-k} \) and \( N = 1000 \) (SBD).

| \( t \) | \( k \) | 7 | 8 | 9 | 10 | 11 | rate  |
|-------|------|---|---|---|----|----|-------|
| \( 0.1 \) | \( L^2 \)-norm | 2.06e-2 | 5.31e-3 | 1.34e-3 | 3.33e-4 | 7.97e-5 | 2.02 (2.00) |
|       | \( H^1 \)-norm | 4.45e-1 | 2.19e-1 | 1.19e-1 | 5.39e-2 | 2.63e-2 | 1.02 (1.00) |
| \( 0.01 \) | \( L^2 \)-norm | 1.65e-2 | 5.74e-3 | 1.47e-3 | 3.80e-4 | 9.26e-5 | 1.99 (2.00) |
|       | \( H^1 \)-norm | 5.14e-1 | 3.32e-1 | 1.59e-1 | 7.86e-2 | 3.81e-2 | 1.02 (1.00) |

Figure 6. Numerical solutions of examples (h) with \( h = 2^{-6} \) and \( N = 1000 \), \( \alpha = 1.5 \) at \( t = 0.1 \) and \( t = 0.01 \).

with \( 1 > \alpha > \beta_1 > \beta_2 > ... > \beta_m > 0 \) and constant coefficients \( b_j > 0, j = 1, 2, ..., m \). Error estimates for the space semidiscrete Galerkin FEM for problem (4.2) with nonsmooth initial data were established in [14]. Using the same argument from Section 3, we deduce that the error estimates stated in Theorems 3.3 and 3.5 remain valid. Now we illustrate the scheme for \( \alpha = 0.5, \beta = 0.1 \) and \( b = 1 \) with the nonsmooth initial data \( v = 1 \in H^{1/2-\epsilon}(\Omega) \). The results indicate that the BE and SBD schemes yield a \( O(\tau) \) and \( O(\tau^2) \) convergence for (4.2), respectively, cf. Tables 14 and 15.

Table 14. Numerical results for multi-term time-fractional parabolic equation with \( \alpha = 0.5, \beta = 0.1 \) at \( t = 0.1 \) with \( h = 2^{-12} \) and \( N = 5 \times 2^k \).

| \( N \) | 10 | 20 | 40 | 80 | 160 | rate |
|-------|----|----|----|----|-----|-----|
| BE    | 4.47e-3 | 2.20e-3 | 1.09e-3 | 5.43e-4 | 2.71e-4 | 1.00 (1.00) |
| SBD   | 5.12e-4 | 1.19e-4 | 2.87e-5 | 7.06e-6 | 1.76e-6 | 2.00 (2.00) |
Table 15. The $L^2$-norm of the error for multi-term time-fractional parabolic equation as $t \to 0$ with $h = 2^{-9}$ and $N = 10$.

| $t$   | 1e-3 | 1e-4 | 1e-5 | 1e-6 | 1e-7 | 1e-8 | rate |
|-------|------|------|------|------|------|------|------|
| BE    | 3.79e-3 | 2.74e-3 | 2.09e-3 | 1.57e-3 | 1.17e-3 | 8.79e-4 | 0.13 (0.13) |
| SBD   | 3.00e-4 | 2.39e-4 | 1.80e-4 | 1.36e-4 | 1.05e-4 | 8.46e-5 | 0.11 (0.13) |

5. Conclusions

In this paper develop two simple fully discrete schemes for the fractional diffusion 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 the backward second-order difference method. We provide a complete error analysis of the schemes, and derived 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 are in excellent agreement with the convergence theory. Further, we compare our schemes with several popular time stepping schemes developed for smooth initial data, and find that these schemes may only yield a slow convergence. Hence, the proposed fully schemes are competitive with existing schemes.

There are several questions deserving further investigation. First, in view of the solution singularity for nonsmooth data, it is of practical interest to develop time stepping schemes using a nonuniform in time mesh and provide rigorous error analysis. Second, our experiments indicate that existing time stepping schemes may yield only suboptimal convergence for nonsmooth data. This motivates revisiting these popular schemes for nonsmooth data, especially sharp error estimates. Last, it is natural to look into extensions to inhomogeneous problem and the multi-term time fractional diffusion or diffusion wave problem, which involves multiple time fractional derivatives.

Appendix A. Solution theory for the diffusion-wave equation

In our convergence analysis, we encounter problem (1.1) with initial data of intermediate smoothness. The solution theory for the subdiffusion equation with nonsmooth and very weak initial data is now well understood [41, 17, 15, 16]. In this part we recall briefly the solution theory for the diffusion-wave equation with nonsmooth initial data, following [41]. Using the Dirichlet eigenpairs $\{(\lambda_j, \varphi_j)\}_{j=1}^\infty$, the solution $u$ to problem (1.1) with $1 < \alpha < 2$ and $f \equiv 0$ can be written as

$$u(x,t) = E(t)v + \tilde{E}(t)b,$$

where the operators $E(t)$ and $\tilde{E}(t)$ are given by

$$E(t)v = \sum_{j=1}^\infty E_{\alpha,1}(-\lambda_j t^\alpha)(v_j) \varphi_j(x), \quad \tilde{E}(t)\chi = \sum_{j=1}^\infty t E_{\alpha,2}(-\lambda_j t^\alpha)(\chi_j) \varphi_j(x),$$

where the Mittag-Leffler function $E_{\alpha,\beta}(z)$ is defined by $E_{\alpha,\beta}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(k\alpha+\beta)}$, $z \in \mathbb{C}$. Then the following stability result holds, slightly extending [41, Theorem 2.3].

Theorem A.1 (1 < $\alpha < 2$). The solution $u(t) = E(t)v + \tilde{E}(t)b$ to problem (1.1) with $f \equiv 0$ satisfies

$$\| (C^{\alpha}_{\ell} u(t)) \|_{\dot{H}^r(\Omega)} \leq C \left( t^{-\alpha(\ell+(p-q)/2)} \| v \|_{\dot{H}^r(\Omega)} + t^{1-\alpha(\ell+(p-r)/2)} \| b \|_{\dot{H}^r(\Omega)} \right), \quad t > 0,$$

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


Proof. First we discuss the case $\ell = 0$. By the triangle inequality and the following bound on $E_{\alpha,\beta}(z)$ for $\frac{\pi}{2} < \mu < \min(\pi, \alpha \pi)$,
\[ |E_{\alpha,\beta}(z)| \leq C(1 + |z|)^{-1} \quad \mu \leq |\arg(z)| \leq \pi, \]
we deduce
\[ \|E(t)v\|_H^2 \leq \sum_{j=1}^{\infty} \lambda_j^\alpha (v, \varphi_j) = \sum_{j=1}^{\infty} t^{-\alpha} \frac{C\lambda_j^{\alpha - q}(p-q)}{(1 + \lambda_j\epsilon)^2} \lambda_j^\alpha (v, \varphi_j)^2 \]
\[ \leq t^{-\alpha\epsilon} \sup_j \frac{C\lambda_j^{\alpha - q}(p-q)}{(1 + \lambda_j\epsilon)^2} \sum_{j=1}^{\infty} \lambda_j^\alpha (v, \varphi_j)^2 \leq C t^{-\alpha\epsilon} \|v\|^2_{H^\alpha(\Omega)}, \]
where we have used the fact that, in view of Young’s inequality, $\frac{(\lambda_j\epsilon)^{\epsilon\alpha}}{(1 + \lambda_j\epsilon)^2} \leq C$ for $0 \leq \epsilon \leq p < 2$. Similarly, one deduces
\[ \|	ilde{E}b\|^2_{HR(\Omega)} \leq Ct^{-\alpha\epsilon} \|b\|^2_{H^\alpha(\Omega)}. \]
Thus the assertion for $\ell = 0$ follows by the triangle inequality. Now we consider the case $\ell = 1$. It follows from the representation formula and (A.1) that
\[ \|\partial_t^\ell E(t)v\|_H^2 = \sum_{j=1}^{\infty} \lambda_j^{2\ell + \alpha} (E_{\alpha, 1}(-\lambda_j t^\alpha)(v, \varphi_j))^2 \]
\[ \leq Ct^{-\alpha(2\ell + \alpha)} \frac{C\lambda_j^{2\ell + \alpha}}{(1 + \lambda_j\epsilon)^2} \sum_{j=1}^{\infty} \lambda_j^{2\ell + \alpha} (v, \varphi_j)^2 \leq C t^{-\alpha(2\ell + \alpha)} \|v\|^2_{H^\alpha(\Omega)}. \]
A similar estimate for $\|\partial_t^\ell \tilde{E}(t)b\|_{HR(\Omega)}$ holds, and this completes the proof. \qed

The solution to the semidiscrete scheme (2.3) can also be expressed as
\[ u_h(t) = E_h(t)v_h + \tilde{E}_h(t)b_h, \]
where the operators $E_h$ and $\tilde{E}_h$ are given by
\[ E_h(t)v_h = \sum_{j=1}^{N_h} E_{\alpha, 1}(-\lambda_j^h t^\alpha) (v_h, \varphi_j^h) \varphi_j^h(x), \quad \tilde{E}_h(t)\chi_h = \sum_{j=1}^{N_h} t E_{\alpha, 2}(-\lambda_j^h t^\alpha) (\chi_h, \varphi_j^h) \varphi_j^h(x), \]
with $\{\lambda_j^h, \varphi_j^h\}_{j=1}^{N_h}$ being the eigenvalue and eigenfunctions of the discrete Laplacian $-\Delta_h$. The following discrete counterpart of Theorem A.1 holds, where $\|\cdot\|$ denotes the discrete norm defined on the finite element space $X_h$ [17].

**Theorem A.2** (1 < $\alpha < 2$). The solution $u_h(t) = E_h(t)v_h + \tilde{E}_h(t)b_h$ to problem (2.3) with $f \equiv 0$ satisfies for $t > 0$ and $0 \leq q, r < p < 2$
\[ \|u_h(t)\|_{HR(\Omega)} \leq C \left(t^{-\alpha\epsilon}/2\|v_h\|_{HR(\Omega)} + t^{1-\alpha\epsilon}/2\|b_h\|_{HR(\Omega)} \right), \]

**References**

[1] E. E. Adams and L. W. Gelhar. Field study of dispersion in a heterogeneous aquifer: 2. spatial moments analysis. *Water Res. Research*, 28(12):3293–3307, 1992.

[2] D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo. *Fractional Calculus*. World Scientific, Hackensack, NJ, 2012. Models and numerical methods.

[3] E. Bazhlekov, B. Jin, R. Lazarov, and Z. Zhou. An analysis of the Rayleigh-Stokes problem for the generalized second grade fluid. preprint, arXiv:1401.8049, 2014.

[4] C.-M. Chen, F. Liu, I. Turner, and V. Anh. A Fourier method for the fractional diffusion equation describing sub-diffusion. *J. Comput. Phys.*, 227(2):886–897, 2007.

[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] K. Diethelm. *The Analysis of Fractional Differential Equations*. Lecture Notes in Mathematics. Springer, 2004.

[7] N. J. Ford, J. Xiao, and Y. Yan. A finite element method for time fractional partial differential equations. *Pract. Calc. Appl. Anal.*, 14(3):454–474, 2011.
[8] H. Fujita and T. Suzuki. Evolution problems. In Handbook of Numerical Analysis, Vol. II, Handb. Numer. Anal., II, pages 789–928. North-Holland, Amsterdam, 1991.

[9] Y. Fujita. Integrodifferential equation which interpolates the heat and the wave equation. Osaka J. Math., 27(2):309–321, 1990.

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

[11] R. Gorenflo, F. Mainardi, D. Moretti, and P. Paradisi. Time fractional diffusion: a discrete random walk approach. Nonlin. Dyn., 29(1–4):129–143, 2002.

[12] Y. Hatano and N. Hatano. Dispersive transport of ions in column experiments: An explanation of long-tailed profiles. Water Res. Research, 34(5):1027–1033, 1998.

[13] A. Iskra. A first course in the numerical analysis of differential equations. Number 44. Cambridge University Press, 2009.

[14] B. Jin, R. Lazarov, Y. Liu, and Z. Zhou. The Galerkin finite element method for a multi-term time-fractional diffusion equation. preprint, arXiv:1404.2953, 2014.

[15] B. Jin, R. Lazarov, J. Pasciak, and Z. Zhou. Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion. IMA J. Numer. Anal., in press, available as arXiv:1307.1068, 2013.

[16] B. Jin, R. Lazarov, J. Pasciak, and Z. 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] B. Jin, R. Lazarov, and Z. Zhou. Error estimates for a semidiscrete finite element method for fractional order parabolic equations. SIAM J. Numer. Anal., 51(4):445–466, 2013.

[18] B. Jin and W. Hundell. An inverse problem for a one-dimensional time-fractional diffusion problem. Inverse Problems, 28(7):075010, 19, 2012.

[19] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. Theory and Applications of Fractional Differential Equations. Elsevier, Amsterdam, 2006.

[20] T. Langlands and B. Henry. The accuracy and stability of an implicit solution method for the fractional diffusion equation. J. Comput. Phys., 205(2):719–736, 2005.

[21] C. Li and H. Ding. Higher order finite difference method for the reaction and anomalous-diffusion equation. Appl. Math. Model., page in press, 2014.

[22] L. Li, D. Xu, and N. Luo. Alternating direction implicit Galerkin finite element method for the two-dimensional fractional diffusion-wave equation. J. Comput. Phys., 255:471–485, 2013.

[23] W. Li and D. Xu. Finite central difference/finite element approximations for parabolic integro-differential equations. Comput. & Math. with Applic., 51(3–4):89–111, 2010.

[24] X. Li and C. Xu. A space-time spectral method for the time fractional diffusion equation. SIAM J. Numer. Anal., 47(3):2108–2131, 2009.

[25] Y. Lin, X. Li, and C. Xu. Finite difference/spectral approximations for the fractional cable equation. J. Math. Anal. Appl., 383(1):136–1396, 2011.

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

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

[28] C. Lubich. Convolution quadrature and discretized operational calculus. I. Numer. Math., 52(2):129–145, 1988.

[29] C. Lubich, I. H. Sloan, and V. Thomée. Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term. Math. Comp., 65(213):1–17, 1996.

[30] R. Mainardi. Fractional relaxation-oscillation and fractional diffusion-wave phenomena. Chaos, Solitons & Fractals, 7(9):1461–1477, 1996.

[31] F. Mainardi. Fractional Calculus and Waves in Linear Viscoelasticity. Imperial College Press, 2010.

[32] W. McLean and V. 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(1):208–230, 2010.

[33] M. M. Meerschaert and C. Tadjeran. Finite difference approximations for fractional advection-dispersion flow equations. J. Comput. Appl. Math., 172(1):65–77, 2004.

[34] E. W. Montroll and G. H. Weiss. Random walks on lattices. II. J. Math. Phys., 6:167–181, 1965.

[35] K. Mustapha and W. McLean. Piecewise-linear, discontinuous galerkin method for a fractional diffusion equation. Numer. Algor., 56:159–184, 2011.

[36] K. Mustapha and W. McLean. Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations. SIAM J. Numer. Anal., 51(1):491–515, 2013.

[37] R. Nigmatulin. The realization of the generalized transfer equation in a medium with fractal geometry. Phys. Stat. Sol. B, 133:425–430, 1986.

[38] K. B. Oldham and J. Spanier. The Fractional Calculus. Academic Press, New York, 1974.

[39] I. Podlubny. Fractional Differential Equations. Academic Press, San Diego, CA, 1999.

[40] 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.

[41] J. M. Sanz-Serna. A numerical method for a partial integro-differential equation. SIAM J. Numer. Anal., 25(2):319–327, 1988.
[43] H. Seybold and R. Hilfer. Numerical algorithm for calculating the generalized Mittag-Leffler function. *SIAM J. Numer. Anal.*, 47(1):69–88, 2008/09.
[44] Z.-Z. Sun and X. Wu. A fully discrete scheme for a diffusion wave system. *Appl. Numer. Math.*, 56(2):193–209, 2006.
[45] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2006.
[46] S. B. Yuste. Weighted average finite difference methods for fractional diffusion equations. *J. Comput. Phys.*, 216(1):264–274, 2006.
[47] S. B. Yuste and L. Acedo. An explicit finite difference method and a new von Neumann-type stability analysis for fractional diffusion equations. *SIAM J. Numer. Anal.*, 42(5):1862–1874, 2005.
[48] S. B. Yuste and J. Quintana-Murillo. A finite difference method with non-uniform timesteps for fractional diffusion equations. *Comput. Phys. Comm.*, 183(12):2594–2600, 2012.
[49] F. Zeng, C. Li, F. Liu, and I. Turner. The use of finite difference/element approaches for solving the time-fractional subdiffusion equation. *SIAM J. Sci. Comput.*, 35(6):A2976–A3000, 2013.
[50] Y.-N. Zhang, Z.-Z. Sun, and H.-L. Liao. Finite difference methods for the time fractional diffusion equation on non-uniform meshes. *J. Comput. Phys.*, 265:195–210, 2014.
[51] Y.-N. Zhang, Z.-Z. Sun, and X. Zhao. Compact alternating direction implicit scheme for the two-dimensional fractional diffusion-wave equation. *SIAM J. Numer. Anal.*, 50(3):1535–1555, 2012.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, 900 UNIVERSITY AVE, RIVERSIDE, CA 92521, USA (bangti.jin@gmail.com)

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368, USA (lazarov@math.tamu.edu, zzhou@math.tamu.edu)