

NUMERICAL ANALYSIS OF FRACTIONAL-ORDER DIFFERENTIAL EQUATIONS WITH  
NONSMOOTH DATA

A Dissertation

by

ZHI ZHOU

Submitted to the Office of Graduate and Professional Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of  
DOCTOR OF PHILOSOPHY

Chair of Committee,	Raytcho Lazarov
Committee Members,	Bangti Jin
	Efendiev Yalchin
	Joseph Pasciak
	Peter Vălko
Head of Department,	Emil Straube

August 2015

Major Subject: Mathematics

Copyright 2015 Zhi Zhou

## ABSTRACT

This thesis is devoted to theoretical and experimental justifications of numerical methods for fractional differential equations, which have received significant attention over the past decades due to their extraordinary capability of modeling the dynamics of anomalous diffusion processes.

In recent years, a number of numerical schemes were developed, analyzed and tested. However, in most of these interesting works, the error estimates were established under the assumption that the solution is sufficiently smooth. Unfortunately, it has been shown that these assumptions are too restrictive for the solutions of fractional differential equations. The goal of this thesis is to develop robust numerical schemes and to establish optimal error bounds that are expressed directly in terms of the regularity of the problem data. We are especially interested in the case of nonsmooth data arising in many applications, e.g. inverse and control problems.

After some background introduction and preliminaries in Chapters 1 and 2, we analyze two semidiscrete schemes obtained by standard Galerkin finite element approximation and lumped mass finite element method in Chapter 3. The error bounds for approximate solutions of the homogeneous and inhomogeneous problems are established separately in terms of the smoothness of the data directly. In Chapter 4 we revisit the most popular fully discrete scheme based on L1-type approximation in time and Galerkin finite element method in space and show the first order convergence in time by the discrete Laplace transform technique, which fills the gap between the existing error analysis theory and numerical results. Two fully discrete schemes based on convolution quadrature are developed in Chapter 5. The error bounds are given using two different techniques, i.e., discretized operational calculus and discrete Laplace transform. Last, in Chapter 6, we summarize our work and mention possible future research topics.

In each chapter, the discussion is focused on the fractional diffusion model and then extended to some other fractional models. Throughout, numerical results for one- and two-dimensional examples will be provided to illustrate the convergence theory.

## ACKNOWLEDGEMENTS

I would like to express my profound gratitude to my advisor, Dr. Raytocho Lazarov. His wonderful guidance and insightful discussion with me were crucial to the successful completion of the research. I am also thankful to my co-advisor Dr. Bangti Jin for his constant support and encouragement. I would say that I was extremely lucky to meet them at Texas A&M University. They taught me not only the way to do scientific research, but also the way to become a professional scientist.

Dr. Joeseeph Pasciak taught me numerical analysis and iterative techniques. Dr. Vivette Girault and Dr. Jean-Luc Guermond taught me mathematical theory of finite element methods. Dr. William Rundell gave me the first insight on inverse problems. I would like to thank them and all the professors who taught me at Texas A&M University. I thank Dr. Paulo Lima-Filho, Dr. Peter Howard, Dr. Emil Straube, Ms. Stewart Monique, and all other staffs at the Department of Mathematics for their appropriate support.

I also wish to thank my collaborators Dr. Emilia Bazhlekova, Dr. Yikan Liu, Dr. Xiliang Lu, Dr. Joeseeph Pasciak, Dr. Dongwoo Sheen and Dr. Vidar Thomeé. They provided me many great ideas and showed me many invaluable insights. I am also grateful to Dr. Yalchin Efendiev and Dr. Peter Valko for serving on the committee and for their inputs and suggestions. I wish to extend my thank to all my friends at our math department for their help and collaboration.

Finally, I have to express my deep gratitude to my parents and family for their love and patience.

## TABLE OF CONTENTS

	Page
ABSTRACT . . . . .	ii
ACKNOWLEDGEMENTS . . . . .	iii
TABLE OF CONTENTS . . . . .	iv
LIST OF FIGURES . . . . .	vi
LIST OF TABLES . . . . .	vii
1. INTRODUCTION . . . . .	1
1.1 Introduction to anomalous diffusion . . . . .	1
1.1.1 Standard diffusion . . . . .	1
1.1.2 Anomalous diffusion . . . . .	2
1.2 Review on numerical methods for fractional diffusions . . . . .	4
1.3 Contributions and organization of the thesis . . . . .	6
2. PRELIMINARIES . . . . .	9
2.1 Fractional calculus . . . . .	9
2.2 Mittag-Leffler functions . . . . .	10
2.2.1 Two-parameter Mittag-Leffler functions . . . . .	10
2.2.2 Multinomial Mittag-Leffler functions . . . . .	11
2.3 Fractional diffusion model . . . . .	11
2.4 Multi-term fractional diffusion equation . . . . .	17
2.5 Diffusion-wave equation . . . . .	21
3. SPATIAL SEMIDISCRETE SCHEMES . . . . .	24
3.1 Spatial semidiscretization by Galerkin finite element method . . . . .	24
3.1.1 Error estimates for homogeneous problems . . . . .	27
3.1.2 Error estimates for inhomogeneous problems . . . . .	32
3.2 Spatial semidiscretization by lumped mass method . . . . .	34
3.2.1 Error estimates for the homogeneous problem . . . . .	36
3.2.2 Error estimates for inhomogeneous problems . . . . .	40
3.3 Numerical results . . . . .	45
3.3.1 Homogeneous problems . . . . .	45
3.3.2 Inhomogeneous problems . . . . .	49
3.4 Extension to multi-term fractional diffusion . . . . .	51
3.4.1 Error analysis . . . . .	51
3.4.2 Numerical results . . . . .	55
3.5 Extension to the diffusion-wave equation . . . . .	58
3.5.1 Error analysis . . . . .	58
3.5.2 Numerical results . . . . .	61
3.6 Conclusion and comments . . . . .	66

4. FULLY DISCRETE SCHEME BY L1 APPROXIMATION . . . . .	67
4.1 Fully discrete schemes by L1 time stepping . . . . .	67
4.1.1 Error estimate for nonsmooth initial data . . . . .	72
4.1.2 Error estimate for smooth initial data . . . . .	82
4.2 Extension to a more general sectorial operator . . . . .	85
4.3 Extension to multi-term fractional diffusion . . . . .	88
4.4 Numerical results . . . . .	91
4.5 Conclusion and comments . . . . .	96
5. FULLY DISCRETE SCHEMES BY CONVOLUTION QUADRATURE . . . . .	97
5.1 Fractional convolution quadrature . . . . .	97
5.1.1 Approximating fractional derivatives by convolution quadrature . . . . .	98
5.1.2 Discrete operational calculus framework . . . . .	99
5.2 Error analysis by discretized operational calculus . . . . .	102
5.2.1 Error analysis for BE method . . . . .	103
5.2.2 Error analysis for the SBD scheme . . . . .	105
5.3 Error analysis by discrete Laplace transform . . . . .	107
5.3.1 Error analysis for the BE method . . . . .	107
5.3.2 Error analysis for the SBD method . . . . .	110
5.4 Numerical results . . . . .	115
5.5 Extensions . . . . .	118
5.5.1 Extension to the diffusion-wave equation . . . . .	118
5.5.2 Extension to the multi-term fractional diffusion . . . . .	120
5.5.3 Numerical results . . . . .	121
5.6 Conclusions and comments . . . . .	125
6. CONCLUSIONS AND FUTURE WORK . . . . .	127
REFERENCES . . . . .	129

## LIST OF FIGURES

FIGURE	Page
3.1 Error plots for smooth initial data, Example (a): $\alpha = 0.1, 0.5, 0.9$ at $t = 0.1$ . . . . .	46
3.2 Numerical solutions of examples (b) and (c) with $h = 2^{-6}$ , $\alpha = 0.5$ , $\beta = 0.2$ at $t = 0.1$ . . .	57
3.3 Errors of scheme SBD for example (a): $N = 1000$ and $t = 0.1$ . . . . .	62
3.4 Numerical results for example (a2) using SBD method: $t = 0.1$ , $h = 2^{-13}$ , $N = 160$ . . . . .	63
3.5 Numerical results for diffusion-wave equations using SBD with $N = 1000$ . . . . .	65
3.6 Numerical solutions of examples (h) with $h = 2^{-6}$ and $N = 1000$ , $\alpha = 1.5$ at $t = 0.1$ and $t = 0.01$ . . . . .	66
5.1 Numerical solutions of examples (b) and (c) by SBD with $h = 2^{-7}$ and $N = 1000$ , $\alpha = 0.5$ at $t = 0.1$ . . . . .	118
5.2 Numerical results for examples (b) and (c) using the BE method with $h = 2^{-9}$ and $N = 10$ , $\alpha = 0.5$ for $t \rightarrow 0$ . . . . .	118

LIST OF TABLES

TABLE	Page
1.1 Convergence rates for existing schemes for subdiffusion case, $0 < \alpha < 1$ . In the table, RL denotes the Riemann-Liouville derivative, and $\bar{u}$ is the zero extension in time of $u$ to $\mathbb{R}$ . . . . .	4
3.1 Numerical results for example (a) at $t = 0.1$ , with $\alpha = 0.1, 0.5$ and $0.9$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	47
3.2 Numerical results for example (b), $\alpha = 0.5$ , at $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	48
3.3 Numerical results, i.e., errors $\ u(t) - u_h(t)\ _{\dot{H}^p(\Omega)}$ , $p = 0, 1$ , for example (c), $\alpha = 0.5$ , at $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	48
3.4 Numerical results, i.e., errors $\ u(t) - \bar{u}_h(t)\ _{\dot{H}^p(\Omega)}$ , $p = 0, 1$ , for example (c), $\alpha = 0.5$ , at $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	49
3.5 Numerical results for example (d) at $t = 1$ , with $\alpha = 0.1, 0.5$ and $0.9$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	50
3.6 Numerical results for example (e) at $t = 1$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	50
3.7 Numerical results for example (e) at $t = 1$ , with $f_h = P_h f$ , discretized on a uniform mesh, $h = 2^{-k}$ . . . . .	51
3.8 Numerical results for (a) with $\alpha = 0.5$ and $\beta = 0.2$ at $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh, $h = 2^{-k}$ and $\tau = t/10^4$ . . . . .	56
3.9 Numerical results for example (b) with $\alpha = 0.5$ and $\beta = 0.2$ at $t = 0.1, 0.01, 0.001$ for a uniform mesh with $h = 2^{-k}$ and $\tau = t/10^4$ . . . . .	57
3.10 Numerical results for example (c) with $\alpha = 0.5$ and $\beta = 0.2$ at $t = 0.1, 0.01, 0.001$ for a uniform mesh with $h = 2^{-k}$ and $\tau = t/10^4$ . . . . .	58
3.11 Numerical results example (e): $\alpha = 1.5$ , $h = 2^{-k}$ . . . . .	63
3.12 The $L^2$ -norm of the error for examples (a1), (a2), (b1) and (b2) with $\alpha = 1.5$ : $t \rightarrow 0$ , $h = 2^{-13}$ , and $N = 10^5$ . . . . .	64
3.13 Numerical results for example (h) with $\alpha = 1.5$ , at $t = 0.1$ with $h = 2^{-k}$ . . . . .	65
4.1 The $L^2$ -norm of the error for the fractional diffusion equation with initial data (a) and (b) at $t = 0.1$ with $h = 2^{-13}$ , $\tau = 1/N$ . . . . .	92

4.2	The $L^2$ -norm of the error for the fractional diffusion equation with initial data (a) and (b) with $\alpha = 0.5$ , $h = 2^{-13}$ and $N = 10$ . . . . .	93
4.3	The $L^2$ -norm of the error for the time-space fractional problem with initial data (a) at $t = 0.1$ with $h = 2^{-13}$ , $\tau \rightarrow 0$ . . . . .	94
4.4	The $L^2$ -norm of the error for the time-space fractional problem with initial data (b) with $\beta = 1.5$ , $h = 2^{-13}$ , $t = 0.1$ , 0.01 and 0.001. . . . .	94
4.5	The $L^2$ -norm of the error for the time-space fractional problem with initial data (a) and (b), $\alpha = 0.5$ and $\beta = 1.5$ , as $t \rightarrow 0$ with $h = 2^{-13}$ and $N = 5$ . . . . .	95
4.6	The $L^2$ -norm of the error for the multi-term fractional diffusion equation with initial data (a) and (b) with $\alpha_1 = 0.1$ at $t = 0.1$ with $h = 2^{-12}$ , $\tau = 1/N$ . . . . .	95
4.7	The $L^2$ -norm of the error for the the multi-term fractional diffusion equation with initial data (a) and (b) with $\alpha = 0.5$ , $\alpha_1 = 0.1$ , $h = 2^{-13}$ and $N = 10$ . . . . .	96
5.1	The $L^2$ -norm of the error for example (a) at $t = 0.1$ with $h = 2^{-12}$ . . . . .	116
5.2	The $L^2$ -norm of the error for examples (b) and (c) for $t = 0.1$ , $\alpha = 0.5$ , and $h = 2^{-9}$ . . .	117
5.3	The $L^2$ -norm of the error for examples (b) and (c) as $t \rightarrow 0$ with $h = 2^{-9}$ and $N = 10$ . .	117
5.4	The $L^2$ -norm of the error for example (a1) at $t = 0.1$ with $h = 2^{-13}$ . . . . .	122
5.5	The $L^2$ -norm of the error for example (a2) at $t = 0.1$ with $h = 2^{-13}$ . . . . .	122
5.6	The $L^2$ -norm of the error for example (a2) with $\alpha = 1.9$ at $t = 0.1$ with $h = 2^{-14}$ . . . . .	123
5.7	The $L^2$ -norm of the error for examples (a1), (a2), (b1) and (b2) with $\alpha = 1.5$ : $t \rightarrow 0$ , $h = 2^{-13}$ , and $N = 10$ . . . . .	123
5.8	The $L^2$ -norm of the error for examples (b1) and (b2) at $t = 0.1$ with $h = 2^{-13}$ . . . . .	124
5.9	The $L^2$ -norm of the error for example (c) at $t = 0.1$ with $\alpha = 1.5$ and $h = 2^{-9}$ . . . . .	124
5.10	Numerical results for muti-term time-fractional parabolic equation with $\alpha = 0.5$ , $\beta = 0.1$ at $t = 0.1$ with $h = 2^{-12}$ and $N = 5 \times 2^k$ . . . . .	125
5.11	The $L^2$ -norm of the error for muti-term time-fractional parabolic equation as $t \rightarrow 0$ with $h = 2^{-9}$ and $N = 10$ . . . . .	125



## 1. INTRODUCTION

In this chapter, we introduce anomalous diffusion and describe some existing numerical schemes for fractional diffusion equations in Sections 1.1 and 1.2, respectively. Then the contributions and organization of this thesis are given in Section 1.3.

### 1.1 Introduction to anomalous diffusion

#### 1.1.1 Standard diffusion

Diffusion is one of the most fundamental phenomena found in nature. It describes the evolution process in time of the density  $u$  of some quantity (such as heat or concentration of chemicals). Let  $\Omega \in \mathbb{R}^3$  be a bounded domain and  $U \subset \Omega$  be any smooth subregion. Then the rate of change of the total quantity within  $U$  equals the negative of the flux through the boundary of  $U$ :

$$\frac{d}{dt} \int_U u \, dx = - \int_{\partial U} \mathbf{F} \cdot \nu \, dS = - \int_U \nabla \cdot \mathbf{F} \, dx,$$

where  $\mathbf{F}$  denotes the flux density. Since  $U$  is chosen arbitrarily, we have

$$u_t = -\nabla \cdot \mathbf{F}.$$

Further, the phenomenological Fick's first law states that the flux  $\mathbf{F}$  is proportional to the local density gradient but points in the opposite direction (since the flow is from the regions of high to low concentration) with the proportionality constant  $a > 0$  (a.k.a. diffusion coefficient):

$$\mathbf{F} = -a\nabla u.$$

Then we obtain the standard diffusion equation

$$\partial_t u = a\Delta u. \tag{1.1}$$

At a microscopic level, the diffusion process is related to the random motion of individual particles, and the use of the Laplace operator and the first-order derivative in the canonical diffusion model rests on a Gaussian assumption on the particle motion. Specifically, let  $u(x, t)$  be the probability density

function (pdf) of a particle which moves randomly in one dimension. Then  $u(x, t)$  should satisfies the following equations

$$u(x_j, t_i) = \frac{1}{2} (u(x_{j-1}, t_{i-1}) + u(x_{j+1}, t_{i-1})). \quad (1.2)$$

For small  $\Delta t$ , we have by Taylor's expansion

$$u(x_j, t_i) \approx u(x_j, t_{i-1}) + (\Delta t) \partial_t u(x_j, t_{i-1}). \quad (1.3)$$

Similarly, we have for small  $\Delta x$

$$u(x_{j\pm 1}, t_{i-1}) \approx u(x_j, t_{i-1}) \pm (\Delta x) \partial_x u(x_j, t_{i-1}) + \frac{(\Delta x)^2}{2} \partial_{xx} u(x_j, t_{i-1}). \quad (1.4)$$

Then plugging (1.3) and (1.4) into (1.2) yields

$$\partial_t u(x_j, t_{i-1}) \approx \frac{(\Delta x)^2}{2\Delta t} \partial_{xx} u(x_j, t_{i-1}).$$

Further, in Brownian motion the mean squared displacement of the particle is linear to time, i.e.,  $\frac{(\Delta x)^2}{2\Delta t} = a$  and hence we arrive at the standard diffusion equation (1.1) in one dimensional case.

### 1.1.2 Anomalous diffusion

Over the last two decades, a large body of literature has shown that diffusion processes in complex systems usually no longer follow Gaussian statistics. The anomalous diffusion in which the mean square variance grows faster (superdiffusion) or slower (subdiffusion) than that in a Gaussian process, offers a superior fit to experimental data observed in a number of important practical applications. For example, anomalous diffusion has been successfully used to describe diffusion in media with fractal geometry [65], highly heterogeneous aquifer [2] and underground environmental problem [20], wave propagation in viscoelastic media [51, 52].

At a microscopic level, the anomalous process can be described by a continuous time random walk (CTRW) model, which is based on the idea that the length of a given jump, as well as the waiting time elapsing between two successive jumps are drawn from a possibility density function  $\Psi(x, t)$ , called jump pdf. Then the jump length pdf and waiting time pdf can be respectively expressed by

$$\varphi(x) = \int_0^\infty \Psi(x, t) dt \quad \text{and} \quad \psi(t) = \int_{-\infty}^\infty \Psi(x, t) dx. \quad (1.5)$$

In case that jump length and the waiting time are independent random variables, we may find the decoupled form  $\Psi(x, t) = \varphi(x)\psi(t)$ . Then the characteristics waiting time

$$T = \int_0^{\infty} t\psi(t) dt$$

and the jump length variance

$$\Sigma^2 = \int_{-\infty}^{\infty} x^2\varphi(x) dx$$

may be finite or infinite. Now a CTRW process can be described by the equation [59, p. 17, eq. (22)]

$$\eta(x, t) = \int_0^t \int_{-\infty}^{\infty} \eta(y, s)\Psi(x - y, t - s) dy ds + u_0(x)\delta(t),$$

where  $\eta(x, t)$  denotes the pdf of just having arrived at position  $x$  at time  $t$ . Consequently, by setting

$$\Phi(t) = 1 - \int_0^t \psi(s) ds$$

assigned to the probability of no jump event during the time interval  $(0, t)$ , and  $u(x, t)$  be the pdf of arriving at position  $x$  at time  $t$ , and not having moved since. Then  $u(x, t)$  may be represented by

$$u(x, t) = \int_0^t \eta(x, s)\Phi(t - s) ds$$

and in Fourier-Laplace space by

$$u(z, \xi) = \frac{(1 - \psi(z))u_0(\xi)}{z(1 - \Psi(\xi, z))}.$$

Note that this framework can reproduce the standard diffusion model (1.1) by setting  $\psi(t) = \frac{\exp(-t/\tau)}{\tau}$  and  $\varphi(x) = \frac{\exp(-x^2/(4\sigma^2))}{(4\pi\sigma^2)^{1/2}}$  [59, Section 3.3], which shows that the anomalous diffusion is a generalized counterpart of standard diffusion.

Next we derive a fractional diffusion (subdiffusion) model through the so called fractal time random walk [21], where  $T$  diverges while  $\Sigma^2$  is kept finite. Here we consider the heavy-tailed waiting time pdf with the asymptotic behavior for  $\alpha \in (0, 1)$ ,  $\psi(t) \sim A_\alpha \tau^\alpha / t^{\alpha+1}$ ,  $A_\alpha = \alpha/\Gamma(1 - \alpha)$  [60, eq. (2)], and Gaussian jump length pdf  $\varphi(x) = \frac{\exp(-x^2/(4\sigma^2))}{(4\pi\sigma^2)^{1/2}}$  with

$$\psi(z) \sim 1 - (z\tau)^\alpha \quad \text{and} \quad \varphi(\xi) \sim 1 - \sigma^2\xi^2.$$

Then the pdf  $u(x, t)$  can be expressed in the Fourier Laplace space by

$$u(\xi, z) = z^{\alpha-1}(z^\alpha + K_\alpha \xi^2)^{-1} u_0(\xi),$$

where  $K_\alpha = \sigma^2/\tau^\alpha$ . We note that this is exactly the solution of the fractional diffusion (2.7) with  $K_\alpha = 1$ ,  $d = 1$ ,  $v = u_0$  and  $f \equiv 0$  in the Fourier-Laplace space.

Further, we refer interested readers to [51, 71, 60, 7] for the derivation of some other anomalous diffusion with time or space fractional derivatives and relevant physical explanations and [4] for a comprehensive survey on numerical methods for fractional ordinary differential equations.

## 1.2 Review on numerical methods for fractional diffusions

The excellent capabilities of fractional differential equations to accurately model such processes have generated considerable interest in deriving, analyzing and testing numerical methods for solving such problems. As a result, a number of numerical techniques were developed and their stability and convergence were investigated. The major body of this thesis is concerned with fractional diffusion model and hence we briefly introduce some popular numerical treatment for this model here. There are two predominant approximations in time: the L1-type approximation [66, 36, 45, 72, 41, 44, 18] and the Grünwald-Letnikov approximation [74, 9, 77]. The convergence rates and regularity assumption of a number of schemes are summarized in Table 1.1.

Table 1.1: Convergence rates for existing schemes for subdiffusion case,  $0 < \alpha < 1$ . In the table, RL denotes the Riemann-Liouville derivative, and  $\bar{u}$  is the zero extension in time of  $u$  to  $\mathbb{R}$ .

method	conv.rate	derivative	regularity assumption
L1 scheme [45, 72]	$O(\tau^{2-\alpha})$	Caputo	$\forall x \in \Omega, u$ is $C^2$ in $t$
Zeng et al I [77]	$O(\tau^{2-\alpha})$	Caputo	$\forall x \in \Omega, u$ is $C^2$ in $t$
Zeng et al II [77]	$O(\tau^{2-\alpha})$	Caputo	$\forall x \in \Omega, u$ is $C^2$ in $t$
Li-Xu [40]	$O(\tau^2)$	Caputo	$\forall x \in \Omega, u$ is $C^2$ in $t$
Gao et al [18]	$O(\tau^{3-\alpha})$	Caputo	$\forall x \in \Omega, u$ is $C^3$ in $t$
L1 scheme [66, 36]	$O(\tau^{2-\alpha})$	RL	$\forall x \in \Omega, u$ is $C^2$ in $t$
SBD [38]	$O(\tau^2)$	RL	$\forall x \in \Omega, {}^R_{-\infty}D_t^{3-\alpha}\bar{u}$ is $L^1$ in $t$

To the first group, L1 approximations, belongs the method devised by Langlands and Henry [36]. They analyzed the discretization error for the Riemann-Liouville derivative. Also, Lin and Xu [45] 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 [72]. Li and Xu [41] extended the work [45] and developed a space-time spectral element method, but only for the case of zero initial data; see also [44, 15]. In [40] a variant of the L1 approximation was analyzed, and a convergence rate  $O(\tau^2)$  was established for  $C^3$  solutions. Recently, Gao et al [18] 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 has been suggested in [76, 78]. We also refer interested readers to [54, 63, 64] for studies on piecewise constant and piecewise linear discontinuous Galerkin discretization of the Riemann-Liouville derivative in time.

In the second group methods, Yuste and Acedo [74] 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 [77] developed two numerical schemes of the order  $O(\tau^{2-\alpha})$  based on an integral reformulation of the subdiffusion problem, a fractional linear multistep method in time and the finite element method in space, and analyzed their stability and convergence. The convolution quadrature due to Lubich [46, 47] 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., [74, 75, 9] 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 [77], 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 [48, 10].

Note that the fractional diffusion operator has only very limited smoothing property, especially for  $t$  close to zero. For example, for the homogeneous equation with an initial data  $v \in \dot{H}^2(\Omega)$ , we have the following stability estimate [53, Theorem 4.2]

$$\|\partial_t u\|_{L^2(\Omega)} \leq ct^{\alpha-1} \|v\|_{\dot{H}^2(\Omega)}. \quad (1.6)$$

That is, the first order derivative in time is already unbounded, not to mention the high-order derivatives. In view of the limited smoothing regularity of the solution to the subdiffusion model, the high regularity required in the convergence analysis in these useful works is restrictive. This observation necessitates revisiting the convergence analysis of some of these existing schemes, especially for the case of nonsmooth problem data.

### 1.3 Contributions and organization of the thesis

In this thesis, we provide a complete analysis of the spatially semidiscrete and fully discrete schemes based on the L1 scheme and convolution quadrature. The schemes cover the subdiffusion model, the multi-term subdiffusion model, and the diffusion-wave model. The obtained error estimates are optimal with respect to the regularity of the problem data, including the case of nonsmooth data.

This thesis is organized as follows:

In Chapter 2, we state some preliminary related to the fractional-order differential equations. First, we introduce some basic definition of fractional calculus such as fractional integrals, Caputo and Riemann-Liouville type fractional derivatives and their connections. Since the Mittag-Leffler functions play a important role in our analysis, we also recall the definition and related properties. Finally we introduce some usefull fractional models and their smoothing property.

In Chapter 3 we study semidiscrete approximations for the fractional diffusion equation (2.7). Two semidiscrete schemes, i.e., the Galerkin FEM and lumped mass Galerkin FEM, using piecewise linear functions are discussed. We establish error estimates for the cases of initial data  $v \in \dot{H}^q(\Omega)$ ,  $q \in (-1, 2]$ . The case  $q = 2$  is referred to as smooth initial data, while the case  $q \in (-1, 0]$  is known as nonsmooth initial data.

In the past, the initial value problem for a standard parabolic equation, i.e.  $\alpha = 1$ , has been thoroughly studied in all these cases. The proof of these results exploits the smoothing property of the parabolic problem via its representation 2.10 using the solution operator  $E(t) = \exp(t\Delta)$ . In this chapter we establish analogous estimates for fractional diffusion. The main difficulty in the error analysis stems from limited smoothing properties. Note that the solution operator is defined through the Mittag-Leffler function, which decays only linearly at infinity, cf. Lemma 2.2.1, whereas the exponential function in the standard parabolic case decays exponentially for  $t = 1$ . The difficulty is overcome by exploiting the mapping property of the discrete solution operators.

The main results of this chapter are as follows. Firstly, in case of smooth initial data, we derived an

error bound uniformly in  $t \geq 0$  (cf. Theorem 3.1.1) for the homogeneous problem ( $f \equiv 0$ ), as is in the case of the standard parabolic problem. Secondly, for quasi-uniform meshes we derived a nonsmooth data error estimate, which deteriorates for  $t$  approaching 0 (cf. Theorem 3.1.2).

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^2\ell_h t^{-\alpha}\|v\|_{L^2(\Omega)}, \quad \ell_h = |\ln h|, \quad t > 0.$$

It is similar to the parabolic counterpart but derived for quasi-uniform meshes and with an additional log-factor  $\ell_h$  [73, Theorem 3.4]. Next we improve the error estimate by getting rid of the log factor using the technique in [16]. Then in Theorem 3.1.4 and 3.1.5, we derive almost optimal error estimates for the inhomogeneous problem ( $v \equiv 0$ ,  $f \neq 0$ ).

Further, we study the more practical lumped mass scheme. We show the same convergence rate for initial data  $v \in \dot{H}^2(\Omega)$  (cf. Theorem 3.2.1), and an almost optimal error estimate for the gradient in the case of data  $v \in \dot{H}^1(\Omega)$  and  $v \in L^2(\Omega)$  (see Theorem 3.2.2). For nonsmooth data  $v \in L^2$ , for general quasi-uniform meshes, we are only able to establish a suboptimal  $L^2$ -error bound of order  $O(h\ell_h t^{-\alpha})$ , cf. (3.39). For a class of special triangulations satisfying condition (3.40), an almost optimal estimate (3.41) holds, analogous to its parabolic counterpart [8, Theorem 4.1]. Then the almost optimal error estimates with respect to smoothness of the source data were established analogously (cf. Theorems 3.2.3–3.2.6).

In Chapter 4, we revisit the fully discrete schemes for the fractional diffusion equation (2.7) with  $f \equiv 0$  based on Galerkin FEM in space and L1-stepping scheme in time. The goal of this work is to fill the gap between the theoretical convergence order of  $O(\tau^{2-\alpha})$  [45, 72] and the empirical first-order convergence. In Theorem 4.1.2, we present an optimal  $O(\tau)$  convergence rate for the fully discrete scheme based on the L1 scheme (4.8) in time and the Galerkin finite element method in space for both smooth and nonsmooth data, i.e.,  $v \in L^2(\Omega)$  and  $Av \in L^2(\Omega)$  ( $A = -\Delta$  with a homogeneous Dirichlet boundary condition), respectively. For example, for  $v \in L^2(\Omega)$  and  $U_h^0 = P_h v$ , for the fully discrete solution  $U_h^n$ , there holds

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

This estimate differs from the known estimates listed in Table 1.1 in several aspects:

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

- (b) The error estimate deteriorates near  $t = 0$ , whereas that in Table 1.1 are uniform in  $t$ . The prefactor  $t_n^{-1}$  in reflects the singularity behavior for initial data  $v \in L^2(\Omega)$ .
- (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.

Surprisingly, for both  $v \in L^2(\Omega)$  and  $Av \in L^2(\Omega)$ , the error estimate deteriorates as time  $t$  approaches zero, but for any fixed time  $t_n > 0$ , it can achieve a first-order convergence. Extensive numerical experiments confirm the optimality of the convergence rates. Our estimates were derived using the generating function technique developed by [48] for convolution quadrature and the interesting recent work of [55] on a piecewise constant discontinuous Galerkin method. The proof essentially boils down to some delicate estimates of the kernel function, which involves the polylogarithmic function.

In Chapter 5, we develop two fully discrete schemes for the subdiffusion problem based on convolution quadrature in time generated by backward Euler (BE) or second-order backward difference (SBD) and the piecewise linear Galerkin FEM in space. This is achieved by a reformulation through the Riemann-Liouville fractional derivative. The time stepping schemes are of Grünwald-Letnikov type we mentioned in Section 1.2. Further, following the general strategy [10], we prove that the fully discrete schemes are first and second-order accurate in time for both smooth and nonsmooth data in Section 5.2. In Section 5.3, we establish the same error bounds using the generating function technique established in Chapter 4. For example, in Theorem 5.2.2, we establish that in case of the second-order backward difference method, if  $v \in L^2(\Omega)$  and  $v_h = P_h v$ , then for  $n \geq 1$

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

To verify the convergence theory, a number of experiments on one and two-dimensional (in space) problems are presented.

Finally, we summarize the main results in the thesis and discuss possible future research topics in Chapter 6. In each chapter the argument is first focused on the fractional diffusion equation (2.7) and then extended to some more generalized fractional models, e.g. multi-term time-fractional diffusion and diffusion-wave equations.



## 2. PRELIMINARIES\*

In this chapter, we present preliminaries related to the fractional differential equations. In Sections 2.1 and 2.2, we briefly describe fundamentals of fractional calculus and collect some useful facts on the Mittag-Leffler function, respectively. Next in Section 2.3 we introduce the fractional diffusion model which plays the key role in this thesis. Solution operators and their smoothing properties are also provided. Finally, we extend the solution theory to the more generalized multi-term counterpart and the diffusion-wave model respectively in Sections 2.4 and 2.5.

### 2.1 Fractional calculus

First we introduce some notations in fractional calculus. For any  $\beta > 0$  with  $n - 1 < \beta < n$ ,  $n \in \mathbb{N}$ , the left-sided Caputo fractional derivative  ${}^C_0D_x^\beta u$  and the left-sided Riemann-Liouville fractional derivative  ${}^R_0D_x^\beta u$  of order  $\beta$  of a function  $u \in C^n[0, 1]$  is defined by [33, pp. 70, 92]:

$${}^C_0D_x^\beta u = {}_0I_x^{n-\beta} \left( \frac{d^n u}{dx^n} \right), \quad \text{and} \quad {}^R_0D_x^\beta u = \frac{d^n}{dx^n} \left( {}_0I_x^{n-\beta} u \right), \quad (2.1)$$

respectively. Here  ${}_0I_x^\gamma$  for  $\gamma > 0$  is the left-sided Riemann-Liouville fractional integral operator of order  $\gamma$  defined by

$$({}_0I_x^\gamma f)(x) = \frac{1}{\Gamma(\gamma)} \int_0^x (x-t)^{\gamma-1} f(t) dt, \quad (2.2)$$

where  $\Gamma(\cdot)$  is Euler's Gamma function defined by  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ . The right-sided versions of fractional-order integrals and derivatives are defined analogously, i.e.,

$${}_xI_1^\gamma f = \frac{1}{\Gamma(\gamma)} \int_x^1 (t-x)^{\gamma-1} f(t) dt,$$

---

\*The solution theory for homogeneous problems ( $f \equiv 0$ ) given in Section 2.3 is reprinted with permission from "Error estimates for a semidiscrete finite element method for fractional order parabolic equations", 2013, SIAM Journal of Numerical Analysis, 51 (1), 445–466, Copyright [2013] by Society for Industrial and Applied Mathematics. The discussion for inhomogeneous problems ( $f \neq 0$ ) given in section 2.3 is reprinted with permission from "Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion" by Bangti Jin, Raytcho Lazarov, Joeseeph Pasciak and Zhi Zhou, 2015, IMA Journal of Numerical Analysis, 35 (2), 561–588, Copyright [2015] by Oxford University Press. Portions of Section 2.4 is reprinted with permission from "The Galerkin finite element method for a multi-term time-fractional parabolic equations", by Bangti Jin, Raytcho Lazarov, Yikan Liu and Zhi Zhou, 2015, Journal of Computational Physics, 281 825–843, Copyright [2015] by Elsevier. The solution theory stated in Section 2.4 is reprinted with permission from "On two schemes for fractional diffusion and diffusion wave equations" by Bangti Jin, Raytcho Lazarov and Zhi Zhou, submitted to SIAM Journal of Scientific Computing, Copyright [2015] by Society for Industrial and Applied Mathematics.

and

$${}_x^C D_1^\beta u = (-1)^n {}_x I_1^{n-\beta} \left( \frac{d^n u}{dx^n} \right), \quad {}_x^R D_1^\beta u = (-1)^n \frac{d^n}{dx^n} \left( {}_x I_1^{n-\beta} u \right).$$

Next, we recall the relation between the Caputo and Riemann-Liouville derivatives. Namely, for  $n-1 < \alpha < n$  [33, pp. 91, equation (2.4.10)], we have

$${}_0^C D_x^\alpha \varphi(t) := {}_0^R D_x^\alpha \left[ \varphi(t) - \sum_{k=0}^{n-1} \frac{\varphi^{(k)}(0)}{k!} t^k \right], \quad (2.3)$$

for a smooth function  $\varphi(t) \in C^n[0, 1]$ .

## 2.2 Mittag-Leffler functions

In this section, we recall the Mittag-Leffler function which plays a very important role in the theory of fractional differential equations.

### 2.2.1 Two-parameter Mittag-Leffler functions

The two-parameter Mittag-Leffler function  $E_{\alpha, \beta}(z)$  is defined by

$$E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k\alpha + \beta)} \quad z \in \mathbb{C} \quad (2.4)$$

It generalizes the exponential function in the sense that  $E_{1,1}(z) = e^z$ . There are several important properties of the Mittag-Leffler function  $E_{\alpha, \beta}(z)$ , mostly derived by M. Djrbashian [11, Chapter 1]. The estimate (2.5) below can be found in [33, pp. 43] or [67, Theorem 1.4], while (2.6) is discussed in [33, Lemma 2.33].

**Lemma 2.2.1.** *Let  $0 < \alpha < 2$  and  $\beta \in \mathbb{R}$  be arbitrary, and  $\frac{\alpha\pi}{2} < \mu < \min(\pi, \alpha\pi)$ . Then there exists a constant  $C = C(\alpha, \beta, \mu) > 0$  such that*

$$|E_{\alpha, \beta}(z)| \leq \frac{C}{1 + |z|} \quad \mu \leq |\arg(z)| \leq \pi. \quad (2.5)$$

Moreover, for  $\lambda > 0$ ,  $\alpha > 0$ , and  $t > 0$  we have

$${}^C \partial_t^\alpha E_{\alpha, 1}(-\lambda t^\alpha) = -\lambda E_{\alpha, 1}(-\lambda t^\alpha) \quad \text{and} \quad E_{\alpha, 1}(-\lambda t^\alpha) = 1 - \lambda t^\alpha E_{\alpha, 1+\alpha}(-\lambda t^\alpha). \quad (2.6)$$

### 2.2.2 Multinomial Mittag-Leffler functions

Now we recall the multinomial Mittag-Leffler function, introduced in [19]. For  $0 < \beta < 2$ ,  $0 < \beta_i < 1$  and  $z_i \in \mathbb{C}$ ,  $i = 1, \dots, m$ , the multinomial Mittag-Leffler function  $E_{(\beta_1, \dots, \beta_m), \beta}(z_1, \dots, z_m)$  is given by

$$E_{(\beta_1, \dots, \beta_m), \beta}(z_1, \dots, z_m) = \sum_{k=0}^{\infty} \sum_{\substack{l_1 + \dots + l_m = k \\ l_1 \geq 0, \dots, l_m \geq 0}} (k; l_1, \dots, l_m) \frac{\prod_{i=1}^m z_i^{l_i}}{\Gamma(\beta + \sum_{i=1}^m \beta_i l_i)},$$

where the notation  $(k; l_1, \dots, l_m)$  denotes the multinomial coefficient, i.e.,

$$(k; l_1, \dots, l_m) = \frac{k!}{l_1! \dots l_m!} \quad \text{with } k = \sum_{i=1}^m l_i.$$

We shall need the following two important lemmas on the function  $E_{(\beta_1, \dots, \beta_m), \beta}(z_1, \dots, z_m)$ , recently obtained in [42, Section 2.1].

**Lemma 2.2.2.** *Let  $0 < \beta < 2$ ,  $0 < \beta_i < 1$ ,  $\beta_1 > \max\{\beta_2, \dots, \beta_m\}$  and  $\frac{\beta_1 \pi}{2} < \mu < \beta_1 \pi$ . Assume that there is  $K > 0$  such that  $-K \leq z_i < 0$ ,  $i = 2, \dots, m$ . Then there exists a constant  $C = C(\beta_1, \dots, \beta_m, \beta, K, \mu) > 0$  such that*

$$E_{(\beta_1, \dots, \beta_m), \beta}(z_1, \dots, z_m) \leq \frac{C}{1 + |z_1|}, \quad \mu \leq |\arg(z_1)| \leq \pi.$$

**Lemma 2.2.3.** *Let  $0 < \beta < 2$ ,  $0 < \beta_i < 1$  and  $z_i \in \mathbb{C}$ ,  $i = 1, \dots, m$ . Then we have*

$$\frac{1}{\Gamma(\beta_0)} + \sum_{i=1}^m z_i E_{(\beta_1, \dots, \beta_m), \beta_0 + \beta_i}(z_1, \dots, z_m) = E_{(\beta_1, \dots, \beta_m), \beta_0}(z_1, \dots, z_m).$$

### 2.3 Fractional diffusion model

In this section, we consider the initial/boundary value problem for the fractional diffusion (subdiffusion) equation for  $u(x, t)$ :

$$\begin{aligned} {}^C \partial_t^\alpha u - \Delta u &= f(x, t), & \text{in } \Omega \quad T \geq t > 0, \\ u &= 0, & \text{on } \partial\Omega \quad T \geq t > 0, \\ u(0) &= v, & \text{in } \Omega. \end{aligned} \tag{2.7}$$

Here  ${}^C\partial_t^\alpha u$  ( $0 < \alpha < 1$ ) denotes the left-sided Caputo fractional derivative of order  $\alpha$  with respect to  $t$  as defined in (2.1). This model has been studied extensively from different aspects due to its extraordinary capability of modeling anomalous diffusion phenomena in highly heterogeneous aquifers and complex viscoelastic materials [2, 65]. It is the fractional analogue of the classical diffusion equation: with  $\alpha = 1$ , it recovers the latter, and thus inherits some of its analytical properties. However, it differs considerably from the latter in the sense that, due to the presence of the nonlocal fractional derivative term, it has limited smoothing property in space and slow asymptotic decay in time [28, 68].

Next, we state some important regularity results related to the fractional diffusion model (2.7). To this end, we shall need some notation. For  $s \geq -1$ , we denote by  $\dot{H}^s(\Omega) \subset H^{-1}(\Omega)$  the Hilbert space induced by the norm:

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

with  $\{\lambda_j\}_{j=1}^{\infty}$  and  $\{\varphi_j\}_{j=1}^{\infty}$  being respectively the eigenvalues and the  $L^2(\Omega)$ -orthonormal eigenfunctions of the Laplace operator  $-\Delta$  on the domain  $\Omega$  with a homogeneous Dirichlet boundary condition. Here  $(\cdot, \cdot)$  denotes the duality between  $\dot{H}^s(\Omega)$  and  $\dot{H}^{-s}(\Omega)$  for  $s \in [-1, 0]$  and the  $L^2$ -inner product if the function is in  $L^2(\Omega)$ . Then  $\{\varphi_j\}_{j=1}^{\infty}$  and  $\{\lambda_j^{1/2}\varphi_j\}_{j=1}^{\infty}$  form orthonormal basis in  $L^2(\Omega)$  and  $H^{-1}(\Omega)$ , respectively. Further,  $\|v\|_{\dot{H}^0(\Omega)} = \|v\|_{L^2(\Omega)} = (v, v)^{1/2}$  is the norm in  $L^2(\Omega)$  and  $\|v\|_{\dot{H}^{-1}(\Omega)} = \|v\|_{H^{-1}(\Omega)}$  is the norm in  $H^{-1}(\Omega)$ . Besides, it is easy to verify that  $\|v\|_{\dot{H}^1(\Omega)} = \|\nabla v\|_{L^2(\Omega)}$  is also the norm in  $H_0^1(\Omega)$  and  $\|v\|_{\dot{H}^2(\Omega)} = \|\Delta v\|_{L^2(\Omega)}$  is equivalent to the norm in  $H^2(\Omega) \cap H_0^1(\Omega)$  [73, Section 3.1]. Note that  $\dot{H}^s(\Omega)$ ,  $s \geq -1$ , form a Hilbert scale of interpolation spaces. Motivated by this, we denote  $\|\cdot\|_{H_0^s(\Omega)}$  to be the norm on the interpolation scale between  $H_0^1(\Omega)$  and  $L^2(\Omega)$  when  $s$  is in  $[0, 1]$  and  $\|\cdot\|_{H_0^s(\Omega)}$  to be the norm on the interpolation scale between  $L^2(\Omega)$  and  $H^{-1}(\Omega)$  when  $s$  is in  $[-1, 0]$ . Then,  $\|\cdot\|_{H_0^s(\Omega)}$  and  $\|\cdot\|_{\dot{H}^s(\Omega)}$  are equivalent for  $s \in [-1, 0]$  by interpolation.

Now we give a representation of the solution of problem (2.7) using the Dirichlet eigenpairs  $\{(\lambda_j, \varphi_j)\}$ . First, we introduce the operator  $E(t)$ :

$$E(t)v = \sum_{j=1}^{\infty} E_{\alpha,1}(-\lambda_j t^\alpha) (v, \varphi_j) \varphi_j(x). \quad (2.8)$$

This is the solution operator to problem (2.7) with a homogeneous right hand side so that for  $f(x, t) \equiv 0$  we have  $u(t) = E(t)v$ . This fact follows from an eigenfunction expansion and (2.6) (see e.g., [68]). Further, for the inhomogeneous equation with vanishing initial data  $v \equiv 0$ , we shall use the operator

defined for  $\chi \in L^2(\Omega)$  by

$$\bar{E}(t)\chi = \sum_{j=1}^{\infty} t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j t^\alpha) (\chi, \varphi_j) \varphi_j(x). \quad (2.9)$$

The operators  $E(t)$  and  $\bar{E}(t)$  are used to represent the solution  $u(x, t)$  of (2.7):

$$u(x, t) = E(t)v + \int_0^t \bar{E}(t-s)f(s)ds. \quad (2.10)$$

It was shown in [68, Theorem 2.2] that if  $f(x, t) \in L^\infty((0, T); L^2(\Omega))$  and  $v \in L^2(\Omega)$ , then there is a unique solution  $u(x, t) \in L^2((0, T); \dot{H}^2(\Omega))$ . For the solution of the homogeneous equation (2.7), we have the following stability estimates, essentially established in [68, Theorem 2.1], and slightly extended in the theorem below; see also [53] for related regularity estimates. Since these estimates will play a key role in the error analysis of the FEM approximations, we sketch the proof.

**Theorem 2.3.1.** *The solution  $u(t) = E(t)v$  to problem (2.7) with  $f \equiv 0$  satisfies for  $q \in (-1, 2]$*

$$\|({}^C \partial_t^\alpha)^\ell u(t)\|_{\dot{H}^p(\Omega)} \leq C t^{-\alpha(\ell + \frac{p-q}{2})} \|v\|_{\dot{H}^q(\Omega)}, \quad t > 0, \quad (2.11)$$

where for  $\ell = 0$ ,  $0 \leq p - q \leq 2$  and for  $\ell = 1$ ,  $-2 \leq p - q \leq 0$ ,  $p \geq -1$ .

*Proof.* We first discuss the case  $\ell = 0$ . By Lemma 2.2.1 and (2.7) we have for  $0 \leq p - q \leq 2$

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

where the last line follows from the inequality  $\sup_{j \in \mathbb{N}} \frac{(\lambda_j t^\alpha)^{p-q}}{(1 + \lambda_j t^\alpha)^2} \leq C$  for  $0 \leq p - q \leq 2$ . The estimate for the case  $\ell = 1$  follows from the identity  $\|{}^C \partial_t^\alpha E(t)v\|_{\dot{H}^p(\Omega)} = \|E(t)v\|_{\dot{H}^{p+2}(\Omega)}$ . It remains to show that (2.8) satisfies also the initial condition in the sense that  $\lim_{t \rightarrow 0^+} \|E(t)v - v\|_{\dot{H}^q(\Omega)} = 0$ . By identity (2.6) and Lemma 2.2.1, we deduce

$$\|E(t)v - v\|_{\dot{H}^q(\Omega)}^2 = \sum_{j=1}^{\infty} \lambda_j^2 t^{2\alpha} \left| E_{\alpha, 1+\alpha}(-\lambda_j t^\alpha) \right|^2 \lambda_j^q |(v, \varphi_j)|^2 \leq C \|v\|_{\dot{H}^q(\Omega)}^2 < \infty.$$

Upon noting the identity  $\lim_{t \rightarrow 0^+} (1 - E_{al,1}(-\lambda_j t^\alpha)) = 0$ , we deduce that for all  $j$

$$\lim_{t \rightarrow 0^+} \lambda_j t^\alpha E_{\bar{\alpha}, 1+\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}) = 0.$$

Hence, the desired assertion follows by Lebesgue's dominated convergence theorem.  $\square$

**Lemma 2.3.1.** *For any  $t > 0$ , we have for  $q \geq -1$ ,  $0 \leq p - q \leq 4$*

$$\|\bar{E}(t)\chi\|_{\dot{H}^p(\Omega)} \leq C t^{-1+\alpha(1+(q-p)/2)} \|\chi\|_{\dot{H}^q(\Omega)}.$$

*Proof.* The definition of  $\bar{E}$  in (2.9) and Lemma 2.2.1 yield

$$\begin{aligned} \|\bar{E}(t)\chi\|_{\dot{H}^p(\Omega)}^2 &= \sum_{j=1}^{\infty} \lambda_j^p |t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j t^\alpha)|^2 |(\chi, \varphi_j)|^2 \\ &= t^{-2+(2+q-p)\alpha} \sum_{j=1}^{\infty} (\lambda_j t^\alpha)^{p-q} |E_{\alpha,\alpha}(-\lambda_j t^\alpha)|^2 \lambda_j^q |(\chi, \varphi_j)|^2 \\ &\leq C t^{-2+(2+q-p)\alpha} \sum_{j=1}^{\infty} \frac{(\lambda_j t^\alpha)^{p-q}}{(1+(\lambda_j t^\alpha)^2)^2} \lambda_j^q |(\chi, \varphi_j)|^2 \\ &\leq C t^{-2+(2+q-p)\alpha} \sum_{j=1}^{\infty} \lambda_j^q |(\chi, \varphi_j)|^2 \leq C t^{-2+(2+q-p)\alpha} \|\chi\|_{\dot{H}^q(\Omega)}, \end{aligned}$$

where in the last line we have used the inequality  $\sup_j \frac{(\lambda_j t^\alpha)^{p-q}}{(1+(\lambda_j t^\alpha)^2)^2} \leq C$  for  $0 \leq p - q \leq 4$ .  $\square$

Next we state some stability estimates for the solution  $u$  to problem (2.7) for  $v \equiv 0$  and  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ . These estimates will be essential for the convergence analysis. The first estimate in Theorem 2.3.2 in the case  $q = 0$  was already established in [68, Theorem 2.1, part (i)]. Below this bound is extended for the whole range of  $q$ .

**Theorem 2.3.2.** *Assume that  $v \equiv 0$  and  $f \in L^2(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ . Then the representation (2.10) satisfies the differential equation in (2.7) and*

$$\|u\|_{L^2(0, T; \dot{H}^{q+2}(\Omega))} + \|\partial_t^\alpha u\|_{L^2(0, T; \dot{H}^q(\Omega))} \leq C \|f\|_{L^2(0, T; \dot{H}^q(\Omega))}. \quad (2.12)$$

*If  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ , then  $u \in L^\infty(0, T; \dot{H}^{q+2-\epsilon}(\Omega))$  for any  $0 < \epsilon < 1$ , and there holds*

$$\|u(t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} \leq C \epsilon^{-1} t^{\epsilon\alpha/2} \|f\|_{L^\infty(0, t; \dot{H}^q(\Omega))}. \quad (2.13)$$

Hence, (2.10) is a solution to problem (2.7).

*Proof.* By the complete monotonicity of the function  $E_{\alpha,1}(-t^\alpha)$  (with  $\alpha \in (0, 1)$ ) [68, Lemma 3.3], i.e.,

$$(-1)^n \frac{d^n}{dt^n} E_{\alpha,1}(-t^\alpha) \geq 0 \quad \text{for all } t > 0, \quad n = 0, 1, \dots,$$

and Lemma 2.2.1, we deduce  $E_{\alpha,\alpha}(-\eta) \geq 0, \eta \geq 0$ . Therefore, for  $t > 0$

$$\begin{aligned} \int_0^t |t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n t^\alpha)| dt &= \int_0^t t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n t^\alpha) dt \\ &= -\frac{1}{\lambda_n} \int_0^t \frac{d}{dt} E_{\alpha,1}(-\lambda_n t^\alpha) dt \\ &= \frac{1}{\lambda_n} (1 - E_{\alpha,1}(-\lambda_n t^\alpha)) \leq \frac{1}{\lambda_n}. \end{aligned} \tag{2.14}$$

Meanwhile, by the differentiation formula [33, pp. 140-141], we get

$$\begin{aligned} A_n &:= \partial_t^\alpha \int_0^t (f(\cdot, \tau), \varphi_n) (t - \tau)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n (t - \tau)^\alpha) d\tau \\ &= (f(\cdot, t), \varphi_n) - \lambda_n \int_0^t (f(\cdot, \tau), \varphi_n) (t - \tau)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n (t - \tau)^\alpha) d\tau. \end{aligned}$$

By means of Young's inequality for convolution, we deduce

$$\begin{aligned} \|A_n\|_{L^2(0,T)}^2 &\leq C_1 \int_0^T |(f(\cdot, t), \varphi_n)|^2 dt + C_2 \int_0^T |(f(\cdot, t), \varphi_n)|^2 dt \left( \int_0^T |\lambda_n t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n t^\alpha)| dt \right)^2 \\ &\leq C \int_0^T |(f(\cdot, t), \varphi_n)|^2 dt. \end{aligned}$$

Thus there holds

$$\begin{aligned} \|\partial_t^\alpha u\|_{L^2(0,T;\dot{H}^q(\Omega))}^2 &= \sum_{n=1}^{\infty} \int_0^T \lambda_n^q |\partial_t^\alpha \int_0^t (f(\cdot, \tau), \varphi_n) (t - \tau)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n (t - \tau)^\alpha) d\tau|^2 dt \\ &\leq C \sum_{n=1}^{\infty} \int_0^T \lambda_n^q |(f(\cdot, t), \phi_n)|^2 dt = C \|f\|_{L^2(0,T;\dot{H}^q(\Omega))}^2. \end{aligned}$$

Now using equation (2.7) and the triangle inequality, we also get  $\|\Delta u\|_{L^2(0,T;\dot{H}^q(\Omega))} \leq C \|f\|_{L^2(0,T;\dot{H}^q(\Omega))}$ .

This shows the first assertion. By Lemma 2.3.1 we have

$$\begin{aligned} \|u(t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} &= \left\| \int_0^t \bar{E}(t-s)f(s)ds \right\|_{\dot{H}^{q+2-\epsilon}(\Omega)} \leq \int_0^t \|\bar{E}(t-s)f(s)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} ds \\ &\leq C \int_0^t (t-s)^{\epsilon\alpha/2-1} \|f(s)\|_{\dot{H}^q(\Omega)} ds \leq C\epsilon^{-1}t^{\epsilon\alpha/2} \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \end{aligned}$$

which shows estimate (2.20). Finally, it follows directly from this that the representation  $u$  satisfies also the initial condition  $u(0) = 0$ , i.e., for any  $\epsilon \in (0, 1)$ ,  $\lim_{t \rightarrow 0^+} \|u(t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} = 0$ , and thus it is indeed a solution of the initial value problem (2.7).  $\square$

**Remark 2.3.1.** *The first estimate in Theorem 2.3.2 can be improved to*

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

for any  $r \in (1, \infty)$ . The proof is essentially identical. The  $\epsilon$  factor in the estimate reflects the limited smoothing property of the fractional differential operator.

**Remark 2.3.2.** *The condition  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$  can be weakened to  $f \in L^r(0, T; \dot{H}^q(\Omega))$  with  $r > 1/\alpha$ . This follows from Lemma 2.3.1 and the Cauchy-Schwarz inequality with  $r'$ ,  $1/r' + 1/r = 1$*

$$\begin{aligned} \|u(\cdot, t)\|_{\dot{H}^q(\Omega)} &\leq \int_0^t \|\bar{E}(t-s)f(s)\|_{\dot{H}^q(\Omega)} ds \leq \int_0^t (t-s)^{\alpha-1} \|f(s)\|_{\dot{H}^q(\Omega)} ds \\ &\leq \frac{C}{1+r'(\alpha-1)} t^{1+r'(\alpha-1)} \|f\|_{L^r(0,t;\dot{H}^q(\Omega))}, \end{aligned}$$

where  $1+r'(\alpha-1) > 0$  by the condition  $r > 1/\alpha$ . It follows from this that the initial condition  $u(0) = 0$  holds in the following sense:  $\lim_{t \rightarrow 0^+} \|u(t)\|_{\dot{H}^q(\Omega)} = 0$ . Hence for any  $\alpha \in (1/2, 1)$  the representation formula (2.10) remains a legitimate solution under the weaker condition  $f \in L^2(0, T; \dot{H}^q(\Omega))$ .

**Remark 2.3.3.** *In the error analysis in Chapter 3 we have restricted our discussion to the case  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ . Nonetheless, we note that for  $p = 0, 1$  the  $L^2(0, T; \dot{H}^p(\Omega))$ -norm estimate of the error below remain valid under the weakened regularity condition on the source term  $f$ .*



## 2.4 Multi-term fractional diffusion equation

In this section, we consider the the following more general fractional model:

$$\begin{aligned}
 P(\partial_t)u - \Delta u &= f, & \text{in } \Omega & \quad T \geq t > 0, \\
 u &= 0, & \text{on } \partial\Omega & \quad T \geq t > 0, \\
 u(0) &= v, & \text{in } \Omega, &
 \end{aligned}
 \tag{2.15}$$

where  $\Omega$  denotes a bounded convex polygonal domain in  $\mathbb{R}^d$  ( $d = 1, 2, 3$ ) with a boundary  $\partial\Omega$ ,  $f$  is the source term, and the initial data  $v$  is a given function on  $\Omega$  and  $T > 0$  is a fixed value. Here the differential operator  $P(\partial_t)$  is defined by

$$P(\partial_t) = {}^C\partial_t^\alpha + \sum_{i=1}^m b_i {}^C\partial_t^{\alpha_i},$$

where  $0 < \alpha_m \leq \dots \leq \alpha_1 < \alpha < 1$  are the orders of the left-sided Caputo fractional derivatives,  $b_i > 0$ ,  $i = 1, 2, \dots, m$ .

There are only few mathematical studies on the model (2.15). Luchko [49] established a maximum principle for problem (2.15), and constructed a generalized solution for the case  $f \equiv 0$  using the multinomial Mittag-Leffler function. Li and Yamamoto [43] established existence, uniqueness, and the Hölder regularity of the solution using a fixed point argument for problem (2.15) with variable coefficients  $\{b_i\}$ . Very recently, Li et al [42] showed the uniqueness and continuous dependence of the solution on the initial value  $v$  and the source term  $f$ , by exploiting refined properties of the multinomial Mittag-Leffler function.

Upon denoting  $\vec{\alpha} = (\alpha, \alpha - \alpha_1, \dots, \alpha - \alpha_m)$ , we introduce the following solution operator

$$E(t)v = \sum_{j=1}^{\infty} (1 - \lambda_j t^\alpha E_{\vec{\alpha}, 1+\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})) (v, \varphi_j) \varphi_j. \tag{2.16}$$

This operator is motivated by a separation of variables [50, 49]. Then for problem (2.15) with a homogeneous right hand side, i.e.,  $f \equiv 0$ , we have  $u(x, t) = E(t)v$ . However, the representation (2.16) is not always convenient for analyzing its smoothing property. We derive an alternative representation

of the solution operator  $E$  using Lemma 2.2.3:

$$\begin{aligned} E(t)v &= \sum_{j=1}^{\infty} E_{\bar{\alpha},1}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})(v, \varphi_j) \varphi_j \\ &+ \sum_{i=1}^m b_i t^{\alpha-\alpha_i} \sum_{j=1}^{\infty} E_{\bar{\alpha},1+\alpha-\alpha_i}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})(v, \varphi_j) \varphi_j. \end{aligned} \quad (2.17)$$

Besides, we define the following operator  $\bar{E}$  for  $\chi \in L^2(\Omega)$  by

$$\bar{E}(t)\chi = \sum_{j=1}^{\infty} t^{\alpha-1} E_{\bar{\alpha},\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})(\chi, \varphi_j) \varphi_j. \quad (2.18)$$

The operators  $E(t)$  and  $\bar{E}(t)$  can be used to represent the solution  $u$  of (2.15) as:

$$u(t) = E(t)v + \int_0^t \bar{E}(t-s)f(s)ds. \quad (2.19)$$

The operator  $\bar{E}$  has the following smoothing property. The proof is identical to the one of Lemma 2.3.1 and hence omitted.

**Lemma 2.4.1.** *For any  $t > 0$  and  $\chi \in \dot{H}^q(\Omega)$ ,  $q \in (-1, 2]$ , there holds for  $0 \leq p - q \leq 4$*

$$\|\bar{E}(t)\chi\|_{\dot{H}^p(\Omega)} \leq Ct^{-1+\alpha(1+(q-p)/2)} \|\chi\|_{\dot{H}^q(\Omega)}.$$

First we recall known regularity results in the literature. In [43], Li and Yamamoto showed that in the case of variable coefficients  $\{b_i(x)\}$ , there exists a unique mild solution  $u \in C((0, T]; \dot{H}^\gamma(\Omega)) \cap C([0, T]; L^2(\Omega))$  and  $u \in C([0, T]; \dot{H}^\gamma(\Omega)) \cap L^\infty(0, T; \dot{H}^2(\Omega))$  when  $v \in L^2(\Omega)$ ,  $f = 0$  and  $v = 0$ ,  $f \in L^\infty(0, T]; L^2(\Omega)$ , respectively, with  $\gamma \in [0, 1)$ . These results were recently refined in [42] for the case of constant coefficients, i.e., problem (2.15). In particular, it was shown that for  $v \in \dot{H}^q(\Omega)$ ,  $0 \leq q \leq 1$ , and  $f = 0$ ,  $u \in L^{1/(1-q/2)}(0, T; H^2(\Omega) \cap H_0^1(\Omega))$ ; and for  $v = 0$  and  $f \in L^r(0, T; \dot{H}^q(\Omega))$ ,  $0 \leq q \leq 2$ ,  $r \geq 1$ ,  $u \in L^r(0, T; \dot{H}^{q+2-\gamma}(\Omega))$  for some  $\gamma \in (0, 1]$ . Here we follow the approach in [42], and extend these results to a slightly more general setting of  $v \in \dot{H}^q(\Omega)$ ,  $-1 < q \leq 2$ , and  $f \in L^2(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ . The nonsmooth case, i.e.,  $-1 < q \leq 0$ , arises commonly in related inverse problems and optimal control problems.

We shall derive the solution regularity to the homogeneous problem, i.e.,  $f \equiv 0$ , and the inhomogeneous problem, i.e.,  $v \equiv 0$ , separately. These estimates will be essential for the error analysis of

the spatial semidiscrete Galerkin scheme in next chapter. First we consider the homogeneous problem with initial data  $v \in \dot{H}^q(\Omega)$ ,  $-1 < q \leq 2$ .

**Theorem 2.4.1.** *Let  $u(t) = E(t)v$  be the solution to problem (2.15) with  $f \equiv 0$  and  $v \in \dot{H}^q(\Omega)$ ,  $q \in (-1, 2]$ . Then there holds*

$$\|(P(\partial_t))^\ell u(t)\|_{\dot{H}^p(\Omega)} \leq Ct^{-\alpha(\ell+(p-q)/2)} \|v\|_{\dot{H}^q(\Omega)}, \quad t > 0,$$

where for  $\ell = 0$ ,  $0 \leq p - q \leq 2$  and for  $\ell = 1$ ,  $-2 \leq p - q \leq 0$ ,  $p \geq -1$ .

*Proof.* We show that (2.17) represents indeed the weak solution to problem (2.15) with  $f \equiv 0$  and further it satisfies the desired estimate. We first discuss the case  $\ell = 0$ . By Lemma 2.2.2 and (2.17) we have for  $0 \leq p - q \leq 2$

$$\begin{aligned} \|E(t)v\|_{\dot{H}^p(\Omega)}^2 &= \sum_{j=1}^{\infty} \lambda_j^p \left( E_{\bar{\alpha},1}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}) \right. \\ &\quad \left. + \sum_{i=1}^m b_i t^{\alpha-\alpha_i} E_{\bar{\alpha},1+\alpha-\alpha_i}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}) \right)^2 (v, \varphi_j)^2 \\ &\leq Ct^{-(p-q)\alpha} \sum_{j=1}^{\infty} \frac{(\lambda_j t^\alpha)^{p-q}}{(1+\lambda_j t^\alpha)^2} \lambda_j^q |(v, \varphi_j)|^2 \leq Ct^{-(p-q)\alpha} \|v\|_{\dot{H}^q(\Omega)}^2, \end{aligned}$$

where the last line follows from the inequality  $\sup_{j \in \mathbb{N}} \frac{(\lambda_j t^\alpha)^{p-q}}{(1+\lambda_j t^\alpha)^2} \leq C$  for  $0 \leq p - q \leq 2$ . The estimate for the case  $\ell = 1$  follows from the identity  $\|P(\partial_t)E(t)v\|_{\dot{H}^p(\Omega)} = \|E(t)v\|_{\dot{H}^{p+2}(\Omega)}$ . It remains to show that (2.17) satisfies also the initial condition in the sense that  $\lim_{t \rightarrow 0^+} \|E(t)v - v\|_{\dot{H}^q(\Omega)} = 0$ . By identity (2.16) and Lemma 2.2.2, we deduce

$$\begin{aligned} \|E(t)v - v\|_{\dot{H}^q(\Omega)}^2 &= \sum_{j=1}^{\infty} \lambda_j^2 t^{2\alpha} \left| E_{\bar{\alpha},1+\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}) \right|^2 \lambda_j^q |(v, \varphi_j)|^2 \\ &\leq C \|v\|_{\dot{H}^q(\Omega)}^2 < \infty. \end{aligned}$$

Using Lemma 2.2.3, we rewrite the term  $\lambda_j t^\alpha E_{\bar{\alpha},1+\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})$  as

$$\begin{aligned} &\lambda_j t^\alpha E_{\bar{\alpha},1+\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}) \\ &= (1 - E_{\bar{\alpha},1}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})) \\ &\quad - \sum_{i=1}^m b_i t^{\alpha-\alpha_i} E_{\bar{\alpha},1+\alpha-\alpha_i}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}). \end{aligned}$$

Upon noting the identity  $\lim_{t \rightarrow 0^+} (1 - E_{\bar{\alpha}, 1}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})) = 0$ , and the boundedness of  $E_{\bar{\alpha}, 1+\alpha-\alpha_i}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})$  from Lemma 2.2.2, we deduce that for all  $j$

$$\lim_{t \rightarrow 0^+} \lambda_j t^\alpha E_{\bar{\alpha}, 1+\alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}) = 0.$$

Hence, the desired assertion follows by Lebesgue's dominated convergence theorem.  $\square$

Now we turn to the inhomogeneous problem with a nonsmooth right hand side, i.e.,  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ , and a zero initial data  $v \equiv 0$ .

**Theorem 2.4.2.** *For  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ , and  $v \equiv 0$ , the representation (2.19) belongs to  $L^\infty(0, T; \dot{H}^{q+2-\epsilon}(\Omega))$  for any  $\epsilon \in (0, 1/2)$  and satisfies*

$$\|u(\cdot, t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} \leq C \epsilon^{-1} t^{\epsilon\alpha/2} \|f\|_{L^\infty(0, t; \dot{H}^q(\Omega))}. \quad (2.20)$$

Hence, it is a solution to problem (2.15) with a homogeneous initial data  $v = 0$ .

*Proof.* By construction, it satisfies the governing equation. By Lemma 2.4.1, we have

$$\begin{aligned} \|u(\cdot, t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} &= \left\| \int_0^t \bar{E}(t-s) f(s) ds \right\|_{\dot{H}^{q+2-\epsilon}(\Omega)} \leq \int_0^t \|\bar{E}(t-s) f(s)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} ds \\ &\leq C \int_0^t (t-s)^{\epsilon\alpha/2-1} \|f(s)\|_{\dot{H}^q(\Omega)} ds \leq C \epsilon^{-1} t^{\epsilon\alpha/2} \|f\|_{L^\infty(0, t; \dot{H}^q(\Omega))} \end{aligned}$$

which shows the desired estimate. Further, it satisfies the initial condition  $u(0) = 0$ , i.e., for any  $\epsilon > 0$ ,  $\lim_{t \rightarrow 0^+} \|u(\cdot, t)\|_{\dot{H}^{q+2-\epsilon}(\Omega)} = 0$ , and thus it is indeed a solution of (2.15).  $\square$

Next we extend Theorem 2.4.2 to allow less regular right hand sides  $f \in L^2(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ . Then the function  $u(x, t)$  satisfies also the differential equation as an element in the space  $L^2(0, T; \dot{H}^{q+2}(\Omega))$ . However, it may not satisfy the homogeneous initial condition  $u(x, 0) = 0$ . In Remark 2.3.2 below, we argue that a weaker class of source term that produces a legitimate weak solution of (2.15) is  $f \in L^r(0, T; \dot{H}^q(\Omega))$  with  $r > 1/\alpha$  and  $-1 < q \leq 1$ . Obviously, for  $1/2 < \alpha < 1$ , it does give a solution  $u(x, t) \in L^2(0, T; \dot{H}^{q+2}(\Omega))$ . To this end, we introduce the shorthand notation

$$\bar{E}_{\bar{\alpha}}^j(t) = t^{\alpha-1} E_{\bar{\alpha}, \alpha}(-\lambda_j t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m}).$$

The function  $\bar{E}_{\bar{\alpha}}^j(t)$  is completely monotone [5].

**Lemma 2.4.2.** *The function  $\bar{E}_\alpha^j(t)$  for  $j \in \mathbb{N}$  has the following properties:*

$$\bar{E}_\alpha^j(t) \text{ is completely monotone and } \int_0^T |\bar{E}_\alpha^j(t)| dt < \frac{1}{\lambda_j}.$$

**Theorem 2.4.3.** *For  $f \in L^2(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ , the representation (2.19) belongs to  $L^2(0, T; \dot{H}^{q+2}(\Omega))$  and satisfies the a priori estimate*

$$\|u\|_{L^2(0, t; \dot{H}^{q+2}(\Omega))} + \|P(\partial_t)u\|_{L^2(0, t; \dot{H}^q(\Omega))} \leq C\|f\|_{L^2(0, t; \dot{H}^q(\Omega))}. \quad (2.21)$$

*Proof.* By Young's inequality for the convolution  $\|k * f\|_{L^p} \leq \|k\|_{L^1}\|f\|_{L^p}$ ,  $k \in L^1$ ,  $f \in L^p$ ,  $p \geq 1$ , and Lemma 2.4.2, we deduce

$$\left\| \int_0^t \bar{E}_\alpha^n(t - \tau) f_n(\tau) d\tau \right\|_{L^2(0, T)}^2 \leq \left( \int_0^T |\bar{E}_\alpha^n(t)| dt \right)^2 \left( \int_0^T |f_n(t)|^2 dt \right) \leq \frac{1}{\lambda_n^2} \int_0^T |f_n(t)|^2 dt.$$

Hence,

$$\begin{aligned} \|u\|_{L^2(0, T; \dot{H}^{q+2}(\Omega))}^2 &= \sum_{n=1}^{\infty} \lambda_n^{q+2} \left\| \int_0^t \bar{E}_\alpha^n(t - \tau) f_n(\tau) d\tau \right\|_{L^2(0, T)}^2 \\ &\leq \sum_{n=1}^{\infty} \lambda_n^q \int_0^T |f_n(t)|^2 dt = \|f\|_{L^2(0, T; \dot{H}^q(\Omega))}^2. \end{aligned}$$

The estimate on  $\|P(\partial_t)u\|_{L^2(0, t; \dot{H}^q(\Omega))}$  follows analogously. This completes the proof.  $\square$

## 2.5 Diffusion-wave equation

In previous sections, the orders of the fractional derivatives are less than 1. It is also interesting to consider the case of  $\alpha \in (1, 2)$  which represents the process of superdiffusion:

$$\begin{aligned} {}^C\partial_t^\alpha u - \Delta u &= 0, \quad \text{in } \Omega \quad T \geq t > 0, \\ u &= 0, \quad \text{on } \partial\Omega \quad T \geq t > 0, \\ u(0) &= v, \quad u'(0) = b, \quad \text{in } \Omega, \end{aligned} \quad (2.22)$$

where the domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ . In [51, 52] Mainardi pointed out that the diffusion wave equation governs the propagation of mechanical diffusive waves in viscoelastic media.

This model has been extensively studied in [68]. Here we recall briefly the solution theory for the diffusion-wave equation with nonsmooth initial data. Using the Dirichlet eigenpairs  $\{(\lambda_j, \varphi_j)\}_{j=1}^{\infty}$ , the

solution  $u$  to problem (2.22) 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

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

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 estimates hold, which slightly extend [68, Theorem 2.3].

**Theorem 2.5.1** ( $1 < \alpha < 2$ ). *The solution  $u(t) = E(t)v + \tilde{E}(t)b$  to problem (2.22) with  $f \equiv 0$  satisfies*

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

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 Lemma 2.2.1, we deduce

$$\begin{aligned} \|E(t)v\|_{\dot{H}^p(\Omega)}^2 &= \sum_{j=1}^{\infty} \lambda_j^p |(v, \varphi_j) E_{\alpha,1}(-\lambda_j t^\alpha)|^2 \leq \sum_{j=1}^{\infty} t^{-\alpha(p-q)} \frac{C \lambda_j^{p-q} t^{(p-q)\alpha}}{(1 + \lambda_j t^\alpha)^2} \lambda_j^q (v, \varphi_j)^2 \\ &\leq t^{-\alpha(p-q)} \sup_j \frac{C \lambda_j^{p-q} t^{(p-q)\alpha}}{(1 + \lambda_j t^\alpha)^2} \sum_{j=1}^{\infty} \lambda_j^q (v, \varphi_j)^2 \leq C t^{-\alpha(p-q)} \|v\|_{\dot{H}^q(\Omega)}^2, \end{aligned}$$

where we have used the fact that, in view of Young's inequality,  $\frac{(\lambda_j t^\alpha)^{p-q}}{(1 + \lambda_j t^\alpha)^2} \leq C$  for  $0 \leq q \leq p \leq 2$ .

Similarly, one deduces

$$\|\tilde{E}b\|_{\dot{H}^p(\Omega)}^2 \leq C t^{2-\alpha(p-r)} \|v\|_{\dot{H}^r(\Omega)}^2.$$

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 (2.5) that

$$\begin{aligned} \|\partial_t^\alpha E(t)v\|_{\dot{H}^p(\Omega)}^2 &= \sum_{j=1}^{\infty} \lambda_j^{2+p} (E_{\alpha,1}(-\lambda_j t^\alpha)(v, \varphi_j))^2 \\ &\leq C t^{-\alpha(2+p-q)} \sum_{j=1}^{\infty} \frac{(\lambda_j t^\alpha)^{2+p-q}}{(1 + \lambda_j t^\alpha)^2} \lambda_j^q (v, \varphi_j)^2 \leq C t^{-\alpha(2+p-q)} \|v\|_{\dot{H}^q(\Omega)}^2. \end{aligned}$$

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

### 3. SPATIAL SEMIDISCRETE SCHEMES\*

The goal of this chapter is to develop optimal error estimates with respect to the regularity of the data for the semidiscrete Galerkin and the lumped mass Galerkin FEMs for the fractional diffusion (2.7) on convex polygonal domains  $\Omega \in \mathbb{R}^d$ ,  $d = 1, 2, 3$ .

The rest of the paper is organized as follows. In Section 3.1 we derive error estimates for the standard Galerkin FEM for both homogeneous and inhomogeneous problems. Then the error analysis for the lumped mass FEM are established in Section 3.2. In Section 3.3 we present numerical results on various one or two-dimensional examples, including both smooth and non-smooth data, which confirm our theoretical study. The semidiscrete schemes for multi-term fractional and diffusion-wave models are analyzed in Sections 3.4 and 3.5, respectively. Throughout, the notation  $c$ , with or without a subscript, denotes a generic constant, which may differ at different occurrences, but it is always independent of the mesh size  $h$  and the solution  $u$ .

#### 3.1 Spatial semidiscretization by Galerkin finite element method

Now we first describe numerical schemes using the standard notation from [73]. Let  $\{\mathcal{T}_h\}_{0 < h < 1}$  be a family of regular partitions of the domain  $\Omega$  into  $d$ -simplexes, called finite elements, with  $h$  denoting the maximum diameter. Throughout, we assume that the triangulation  $\mathcal{T}_h$  is quasi-uniform. That is, the diameter of the inscribed disk in the finite element  $\tau \in \mathcal{T}_h$  is bounded from below by  $h$ , uniformly on  $\mathcal{T}_h$ . The approximate solution  $u_h$  will be sought in the finite element space  $V_h \equiv V_h(\Omega)$  of continuous piecewise linear functions over the triangulation  $\mathcal{T}_h$

$$V_h = \{\chi \in H_0^1(\Omega) : \chi \text{ is a linear function over } \tau, \forall \tau \in \mathcal{T}_h\}.$$

---

\*Portions in Sections 3.1–3.3 for homogeneous equations with  $f \equiv 0$  is reprinted with permission from "Error estimates for a semidiscrete finite element method for fractional order parabolic equations", 2013, SIAM Journal of Numerical Analysis, 51 (1), 445–466, Copyright [2013] by Society for Industrial and Applied Mathematics. Portions in Sections 3.1 is reprinted with permission from "Error estimates for finite element methods for space-fractional equations", 2014, SIAM Journal of Numerical Analysis, 52 (5), 2272–2294, Copyright [2014] by Society for Industrial and Applied Mathematics. The discussion for inhomogeneous equations with  $f \neq 0$  is reprinted with permission from "Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion" by Bangti Jin, Raytcho Lazarov, Joeseeph Pasciak and Zhi Zhou, 2015, IMA Journal of Numerical Analysis, 35 (2), 561–588, Copyright [2015] by Oxford University Press. Part of Section 3.4 is reprinted with permission from "The Galerkin finite element method for a multi-term time-fractional parabolic equations", by Bangti Jin, Raytcho Lazarov, Yikan Liu and Zhi Zhou, 2015, Journal of Computational Physics, 281 825–843, Copyright [2015] by Elsevier.



To describe the scheme, we need the  $L^2(\Omega)$  projection  $P_h : \dot{H}^s(\Omega) \rightarrow V_h$  with  $s \in [-1, 0]$  and Ritz projection  $R_h : \dot{H}^1(\Omega) \rightarrow V_h$ , respectively, defined by

$$\begin{aligned} (P_h \psi, \chi) &= (\psi, \chi) \quad \forall \chi \in V_h, \\ (\nabla R_h \psi, \nabla \chi) &= (\nabla \psi, \nabla \chi) \quad \forall \chi \in V_h. \end{aligned}$$

The operators  $R_h$  and  $P_h$  satisfy the following approximation property [73, Lemma 1.1] or [12, Theorems 3.16 and 3.18].

**Lemma 3.1.1.** *Let the mesh be quasi-uniform. Then the operator  $R_h$  satisfies:*

$$\|R_h \psi - \psi\|_{L^2(\Omega)} + h \|\nabla(R_h \psi - \psi)\|_{L^2(\Omega)} \leq Ch^r \|\psi\|_{\dot{H}^r(\Omega)} \quad \text{for } \psi \in \dot{H}^r(\Omega), \quad r \in [1, 2].$$

Further, for  $s \in [0, 1]$  we have

$$\|(I - P_h)\psi\|_{H^s(\Omega)} \leq Ch^{r-s} \|\psi\|_{\dot{H}^r(\Omega)} \quad \text{for } \psi \in \dot{H}^r(\Omega), \quad r \in [1, 2].$$

In addition, by duality  $P_h$  is stable on  $\dot{H}^s(\Omega)$  for  $s \in [-1, 0]$ .

Now the semidiscrete Galerkin FEM for problem (2.7) reads: find  $u_h(t) \in V_h$  such that

$$({}^C \partial_t^\alpha u_h, \chi) + a(u_h, \chi) = (f, \chi), \quad \forall \chi \in V_h, \quad T \geq t > 0, \quad u_h(0) = v_h, \quad (3.1)$$

where  $a(u, w) = (\nabla u, \nabla w)$  for  $u, w \in H_0^1(\Omega)$ , and  $v_h \in V_h$  is an approximation of the initial data  $v$ . The choice of  $v_h$  will depend on the smoothness of the initial data  $v$ . Following Thomée [73], we shall take  $v_h = R_h v$  in case of smooth initial data, i.e.,  $q = 2$ , and  $v_h = P_h v$  in case of nonsmooth initial data, i.e.,  $q \leq 0$ .

Upon introducing the discrete Laplacian  $\Delta_h : V_h \rightarrow V_h$  defined by

$$-(\Delta_h \psi, \chi) = (\nabla \psi, \nabla \chi) \quad \forall \psi, \chi \in V_h, \quad (3.2)$$

and  $f_h = P_h f$ , we may write the spatially discrete problem (3.1) as

$${}^C \partial_t^\alpha u_h(t) - \Delta_h u_h(t) = f_h(t) \quad \text{for } t \geq 0 \quad \text{with } u_h(0) = v_h. \quad (3.3)$$

Now we give a representation of the solution of (3.3) using the eigenvalues and eigenfunctions  $\{\lambda_j^h\}_{j=1}^N$  and  $\{\varphi_j^h\}_{j=1}^N$  of the discrete Laplacian  $-\Delta_h$ . First we introduce the discrete analogues of (2.8) and (2.9) for  $t > 0$ :

$$E_h(t)v_h = \sum_{j=1}^N E_{\alpha,1}(-\lambda_j^h t^\alpha)(v_h, \varphi_j^h)\varphi_j^h, \quad (3.4)$$

$$\bar{E}_h(t)f_h = \sum_{j=1}^N t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j^h t^\alpha)(f_h, \varphi_j^h)\varphi_j^h. \quad (3.5)$$

Then the solution  $u_h(x, t)$  of the discrete problem (3.3) can be expressed by:

$$u_h(x, t) = E_h(t)v_h + \int_0^t \bar{E}_h(t-s)f_h(s) ds. \quad (3.6)$$

Now we introduce the discrete norm  $||| \cdot |||_{\dot{H}^p(\Omega)}$  on  $V_h$  for any  $p \in \mathbb{R}$

$$|||\psi|||_{\dot{H}^p(\Omega)}^2 = \sum_{j=1}^N (\lambda_j^h)^p (\psi, \varphi_j^h)^2 \quad \psi \in V_h. \quad (3.7)$$

Clearly, the norm  $||| \cdot |||_{\dot{H}^p(\Omega)}$  is well defined for all real  $p$ . By the very definition of the discrete Laplacian  $-\Delta_h$  we have  $|||\psi|||_{\dot{H}^1(\Omega)} = \|\psi\|_{\dot{H}^1(\Omega)}$  and also  $|||\psi|||_{\dot{H}^0(\Omega)} = \|\psi\|$  for any  $\psi \in V_h$ . So there will be no confusion in using  $\|\psi\|_{\dot{H}^p(\Omega)}$  instead of  $|||\psi|||_{\dot{H}^p(\Omega)}$  for  $p = 0, 1$  and  $\psi \in V_h$ .

We shall state some smoothing properties of the operator  $E_h(t)$ , which are discrete analogues of those formulated in (2.11).

**Lemma 3.1.2.** *Let  $E_h(t)$  be defined by (3.4) and  $v_h \in V_h$ . Then*

$$|||(C\partial_t^\alpha)^\ell u_h(t)|||_{\dot{H}^p(\Omega)} = |||(C\partial_t^\alpha)^\ell E_h(t)v_h|||_{\dot{H}^p(\Omega)} \leq Ct^{-\alpha(\ell + \frac{p-q}{2})} |||v_h|||_{\dot{H}^q(\Omega)}, \quad t > 0, \quad (3.8)$$

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

The following estimates are crucial for the a priori error analysis in the sequel.

**Lemma 3.1.3.** *Let  $\bar{E}_h$  be defined by (3.5) and  $\psi \in V_h$ . Then we have for all  $t > 0$ ,*

$$|||\bar{E}_h(t)\psi|||_{\dot{H}^p(\Omega)} \leq \begin{cases} Ct^{-1+\alpha(1+\frac{q-p}{2})} |||\psi|||_{\dot{H}^q(\Omega)}, & p-2 \leq q \leq p, \\ Ct^{-1+\alpha} |||\psi|||_{\dot{H}^q(\Omega)}, & p < q. \end{cases} \quad (3.9)$$

*Proof.* The proof for the case of  $p-2 \leq q \leq p$  is analogue to the proof of Lemma 2.3.1 and the assertion for  $p < q$  follows from the fact that  $\{\lambda_j^h\}$  are bounded away from zero independent of  $h$ .  $\square$

The following estimate is the discrete analogue of Theorem 2.3.2.

**Lemma 3.1.4.** *Let  $u_h$  be the solution of (3.3) with  $v_h \equiv 0$ . Then for arbitrary  $p > -1$*

$$\| \partial_t^\alpha u_h(t) \|_{L^2(0,T;\dot{H}^p(\Omega))}^2 + \| u_h(t) \|_{L^2(0,T;\dot{H}^{p+2}(\Omega))}^2 \leq \| f_h \|_{L^2(0,T;\dot{H}^p(\Omega))}^2, \quad (3.10)$$

and

$$\| u_h(t) \|_{\dot{H}^{p+2-\epsilon}(\Omega)} \leq C \epsilon^{-1} t^{\epsilon\alpha/2} \| f_h \|_{L^\infty(0,t;\dot{H}^p(\Omega))}, \quad 0 < \epsilon < 1. \quad (3.11)$$

Further, we shall need the following inverse inequality.

**Lemma 3.1.5.** *For any  $l > s$ , there exists a constant  $C$  independent of  $h$  such that*

$$\| \psi \|_{\dot{H}^l(\Omega)} \leq C h^{s-l} \| \psi \|_{\dot{H}^s(\Omega)} \quad \forall \psi \in V_h. \quad (3.12)$$

*Proof.* For quasi-uniform triangulations  $\mathcal{T}_h$  the inverse inequality  $\| \psi \|_{\dot{H}^1(\Omega)} \leq C h^{-1} \| \psi \|_{L^2(\Omega)}$  holds for all  $\psi \in V_h$ . By the definition of  $-\Delta_h$  this implies  $\max_{1 \leq j \leq N} \lambda_j^h \leq C/h^2$ . Thus, for the norm  $\| \cdot \|_{\dot{H}^p(\Omega)}$  defined in (3.7), there holds for any real  $l > s$

$$\| \psi \|_{\dot{H}^l(\Omega)}^2 \leq C \max_j (\lambda_j^h)^{l-s} \sum_{j=1}^N (\lambda_j^h)^s (\psi, \varphi_j^h)^2 \leq C h^{2(s-l)} \| \psi \|_{\dot{H}^s(\Omega)}^2.$$

$\square$

### 3.1.1 Error estimates for homogeneous problems

First, we consider homogeneous problems, i.e.,  $f \equiv 0$ . To derive error estimates, first we consider the case of smooth initial data, i.e.,  $v \in \dot{H}^2(\Omega)$ . To this end, we split the error  $u_h(t) - u(t)$  into two terms:

$$u_h - u = (u_h - R_h u) + (R_h u - u) := \vartheta + \varrho.$$

By Lemma 3.1.1 and Theorem 2.3.1, we have for any  $t > 0$

$$\| \varrho(t) \|_{L^2(\Omega)} + h \| \nabla \varrho(t) \|_{L^2(\Omega)} \leq C h^2 t^{-(1-q/2)\alpha} \| v \|_{\dot{H}^q(\Omega)} \quad v \in \dot{H}^q(\Omega), \quad q = 1, 2. \quad (3.13)$$

So it suffices to derive proper estimates for  $\vartheta(t)$ , which is given below.

**Lemma 3.1.6.** *The function  $\vartheta(t) := u_h(t) - R_h u(t)$  satisfies for  $p = 0, 1$*

$$\|\vartheta(t)\|_{\dot{H}^p(\Omega)} \leq Ch^{2-p} \|v\|_{\dot{H}^2(\Omega)}.$$

*Proof.* Using the identity  $\Delta_h R_h = P_h \Delta$ , we note that  $\vartheta$  satisfies

$${}^C \partial_t^\alpha \vartheta(t) - \Delta_h \vartheta(t) = -P_h {}^C \partial_t^\alpha \varrho(t) \quad \text{for } t > 0,$$

with  $\vartheta(0) = 0$ . By the representation (3.6),

$$\vartheta(t) = - \int_0^t \bar{E}_h(t-s) P_h {}^C \partial_t^\alpha \varrho(s) ds.$$

Then by Lemmas 3.9 and 3.1.1, and Theorem 2.3.1 we have for  $p = 0, 1$

$$\begin{aligned} \|\vartheta(t)\|_{\dot{H}^p(\Omega)} &\leq \int_0^t \|\bar{E}_h(t-s) P_h {}^C \partial_t^\alpha \varrho(s)\|_{\dot{H}^p(\Omega)} ds \\ &\leq C \int_0^t (t-s)^{(1-p/2)\alpha-1} \|{}^C \partial_t^\alpha \varrho(s)\|_{L^2(\Omega)} ds \\ &\leq Ch^{2-p} \int_0^t (t-s)^{(1-p/2)\alpha-1} \|{}^C \partial_t^\alpha u(s)\|_{\dot{H}^{2-p}(\Omega)} ds \\ &\leq Ch^{2-p} \int_0^t (t-s)^{(1-p/2)\alpha-1} s^{-(1-p/2)\alpha} ds \|v\|_{\dot{H}^2(\Omega)} \leq Ch^{2-p} \|v\|_{\dot{H}^2(\Omega)}, \end{aligned}$$

which is the desired result.  $\square$

Using (3.13), Lemma 3.1.6 and the triangle inequality, we arrive at our first estimate, which is formulated in the following Theorem:

**Theorem 3.1.1.** *Let  $v \in \dot{H}^2(\Omega)$  and  $f \equiv 0$ , and  $u$  and  $u_h$  be the solutions of (2.7) and (3.1) with  $v_h = R_h v$ , respectively. Then*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h \|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^2 \|v\|_{\dot{H}^2(\Omega)}.$$

Now we turn to the nonsmooth case, i.e.,  $v \in \dot{H}^q(\Omega)$  with  $-1 < q \leq 1$ . Since the Ritz projection  $R_h$  may be not well-defined in these cases, we use instead the  $L^2(\Omega)$ -projection  $v_h = P_h v$  and split the

error  $u_h - u$  into:

$$u_h - u = (u_h - P_h u) + (P_h u - u) := \tilde{\vartheta} + \tilde{\varrho}. \quad (3.14)$$

By Lemma 3.1.1 and Theorem 2.3.1 we have for  $-1 < q \leq 1$

$$\begin{aligned} \|\tilde{\varrho}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\varrho}(t)\|_{L^2(\Omega)} &\leq Ch^{2+\min(0,q)}\|u(t)\|_{\dot{H}^{2+\min(0,q)}(\Omega)} \\ &\leq Ch^{2+\min(0,q)}t^{-\alpha(1-\max(q/2,0))}\|v\|_{\dot{H}^q(\Omega)}. \end{aligned}$$

Thus, we only need to estimate the term  $\tilde{\vartheta}(t)$ , which is stated in the following lemma.

**Lemma 3.1.7.** *Let  $\tilde{\vartheta}(t) = u_h(t) - P_h u(t)$ . Then for  $p = 0, 1$ ,  $-1 < q \leq 1$ , there holds (with  $\ell_h = |\ln h|$ )*

$$\|\tilde{\vartheta}(t)\|_{\dot{H}^p(\Omega)} \leq Ch^{\min(q,0)+2-p}\ell_h t^{-\alpha(1-\max(q/2,0))}\|v\|_{\dot{H}^q(\Omega)}.$$

*Proof.* Obviously,  $P_h^C \partial_t^\alpha \tilde{\varrho} = {}^C \partial_t^\alpha P_h (P_h u - u) = 0$  and using the identity  $\Delta_h R_h = P_h \Delta$ , we get the following problem for  $\tilde{\vartheta}$ :

$${}^C \partial_t^\alpha \tilde{\vartheta}(t) - \Delta_h \tilde{\vartheta}(t) = -\Delta_h (R_h u - P_h u)(t), \quad t > 0, \quad \tilde{\vartheta}(0) = 0. \quad (3.15)$$

Using (3.6),  $\tilde{\vartheta}(t)$  can be represented by

$$\tilde{\vartheta}(t) = -\int_0^t \bar{E}_h(t-s) \Delta_h (R_h u - P_h u)(s) ds. \quad (3.16)$$

Let  $A = \bar{E}_h(t-s) \Delta_h (R_h u - P_h u)(s)$ . Then by Lemma 3.1.3, there holds for  $p = 0, 1$ :

$$\begin{aligned} \|A\|_{\dot{H}^p(\Omega)} &\leq C(t-s)^{\epsilon\alpha/2-1} \|\Delta_h (R_h u - P_h u)(s)\|_{\dot{H}^{p-2+\epsilon}(\Omega)} \\ &\leq C(t-s)^{\epsilon\alpha/2-1} \|(R_h u - P_h u)(s)\|_{\dot{H}^{p+\epsilon}(\Omega)}. \end{aligned}$$

Then by (3.12), Theorem 2.3.1, Lemma 3.1.1 we have for  $p = 0, 1$  and  $0 \leq q \leq 1$

$$\begin{aligned} \|A\|_{\dot{H}^p(\Omega)} &\leq Ch^{\min(q,0)+2-p-\epsilon}(t-s)^{\epsilon\alpha/2-1} \|u(s)\|_{2+\min(0,q)} \\ &\leq Ch^{\min(q,0)+2-p-\epsilon}(t-s)^{\epsilon\alpha/2-1} s^{-(1-\max(q/2,0))\alpha} \|v\|_{\dot{H}^q(\Omega)}. \end{aligned}$$

Then plugging the estimate into (3.16) yields

$$\begin{aligned} \|\tilde{\vartheta}\|_{\dot{H}^p(\Omega)} &\leq Ch^{2-p-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} s^{-(1-\max(q/2,0))\alpha} ds \|v\|_{\dot{H}^q(\Omega)} \\ &\leq C\epsilon^{-1} h^{2-p-\epsilon} t^{-\alpha(1-\max(q/2,0))} \|v\|_{\dot{H}^q(\Omega)}. \end{aligned}$$

Now with the choice  $\epsilon = 1/\ell_h$ , we obtain the desired estimate.  $\square$

Now the triangle inequality yields an error estimate for nonsmooth initial data.

**Theorem 3.1.2.** *Let  $f \equiv 0$ ,  $u$  and  $u_h$  be the solutions of (2.7) with  $v \in \dot{H}^q(\Omega)$ ,  $-1 < q \leq 1$ , and (3.1) with  $v_h = P_h v$ , respectively. Then with  $\ell_h = |\ln h|$ , there holds*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^2 \ell_h t^{-\alpha(1-\max(q/2,0))} \|v\|_{\dot{H}^q(\Omega)}.$$

**Remark 3.1.1.** *In case of  $v \in \dot{H}^2(\Omega)$ , the initial approximation  $v_h$  can be chosen by  $P_h v$  also. In fact*

$$\|E(t)v - E_h(t)P_h v\|_{L^2(\Omega)} \leq \|E(t)v - E_h(t)P_h v\|_{L^2(\Omega)} + \|E_h(t)P_h v - E_h(t)R_h v\|_{L^2(\Omega)} := I + II.$$

The estimate for the first term is derived in Theorem 3.1.2, while the second term can be bounded by Lemma 3.1.1 and 3.1.2

$$II \leq \|P_h v - R_h v\|_{L^2(\Omega)} \leq Ch^2 \|v\|_{\dot{H}^2(\Omega)}.$$

Then we obtain the same results as Theorem 3.1.1 with  $v_h = P_h v$ . Now interpolation yields that

$$\|u - u_h\|_{\dot{H}^p(\Omega)} \leq Ch^2 \ell_h t^{-\alpha(1-\max(q/2,0))} \|v\|_{\dot{H}^q(\Omega)}, \quad \text{with } p = 0, 1 \quad \text{and } q \in (-1, 2],$$

where  $u$  and  $u_h$  be the solutions of (2.7) with  $f \equiv 0$ ,  $v \in \dot{H}^q(\Omega)$ ,  $-1 < q \leq 2$ , and (3.1) with  $v_h = P_h v$ .

Next we note that the log factor in the estimates in Theorem 3.1.2 can be removed using the operator trick in [16, 26]. To this end, we first derive an integral representation for the solution. Since the solution  $u : (0, T] \rightarrow L^2(\Omega)$  can be analytically extended to the sector  $\{z \in \mathbb{C}; z \neq 0, |\arg z| < \pi/2\}$  [68], we may apply the Laplace transform to (2.7) to deduce

$$z^\alpha \widehat{u}(z) + A\widehat{u}(z) = z^{\alpha-1} v, \tag{3.17}$$

with  $A = -\Delta$ . Hence the solution  $u(t)$  can be represented by

$$u(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha I + A)^{-1} z^{\alpha-1} v \, dz, \quad (3.18)$$

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

$$\Gamma_{\theta,\delta} = \{z \in \mathbb{C} : |z| = \delta, |\arg z| \leq \theta\} \cup \{z \in \mathbb{C} : z = \rho e^{\pm i\theta}, \rho \geq \delta\}. \quad (3.19)$$

Throughout, the angle is fixed that  $\theta \in (\pi/2, \pi)$  and hence  $z^\alpha \in \Sigma_\theta$  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

$$\|(z^\alpha I + A)^{-1}\| \leq C z^{-\alpha}, \quad \forall z \in \Sigma'_\theta \quad \text{with} \quad \Sigma'_\theta = \Sigma_\theta \setminus \{0\}.$$

Similarly, with  $A_h = -\Delta_h$ , the solution  $u_h$  to (3.1) can be represented by

$$u_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha + A_h)^{-1} z^{\alpha-1} v_h \, dz. \quad (3.20)$$

The next lemma shows an important error estimate [16, 6].

**Lemma 3.1.8.** *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*

$$\|w_h - w\|_{L^2(\Omega)} + h \|\nabla(w_h - w)\|_{L^2(\Omega)} \leq Ch^2 \|\varphi\|_{L^2(\Omega)}. \quad (3.21)$$

Now we can state an error estimate for the scheme (3.1) with nonsmooth initial data.

**Theorem 3.1.3.** *Let  $f \equiv 0$ ,  $u$  and  $u_h$  be the solutions of problem (2.7) and (3.1) with  $v \in L^2(\Omega)$ ,  $v_h = P_h v$ , respectively. Then for  $t > 0$ , there holds:*

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} + h \|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq Ch^2 t^{-\alpha} \|v\|_{L^2(\Omega)}.$$

*Proof.* By (4.2) and (4.6), the error  $e(t) := u(t) - u_h(t)$  can be represented as

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

with  $w^v = (z^\alpha I + A)^{-1}v$  and  $w_h^v = (z^\alpha I + A_h)^{-1}P_h v$ . By Lemma 3.1.8 and choosing  $\delta = 1/t$  we have

$$\begin{aligned}\|\nabla e(t)\|_{L^2(\Omega)} &\leq Ch \left( \int_{-\theta}^{\theta} e^{\cos \psi} t^{-\alpha} d\psi + \int_{1/t}^{\infty} e^{r t \cos \theta} \rho^{\alpha-1} d\rho \right) \|v\|_{L^2(\Omega)} \\ &\leq Ch t^{-\alpha} \|v\|_{L^2(\Omega)}.\end{aligned}$$

A similar argument yields the  $L^2$ -estimate.  $\square$

**Remark 3.1.2.** *If the initial data is very weak, i.e.,  $v \in \dot{H}^q(\Omega)$  with  $-1 < q < 0$ , Then Theorem 3.1.3 and the argument of [25, Theorem 2] yield the following optimal error estimate for the semidiscrete finite element approximation (3.1)*

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} + h \|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq Ch^{2+q} t^{-(1-\alpha)} \|v\|_{\dot{H}^q(\Omega)}. \quad (3.22)$$

### 3.1.2 Error estimates for inhomogeneous problems

In this part, we consider the case of  $v \equiv 0$  and  $f \neq 0$ . First, in Theorem 3.1.4 we establish error bounds in  $L^2(0, T; L^2(\Omega))$ - and  $L^2(0, T; H^1(\Omega))$ -norms:

**Theorem 3.1.4.** *Let  $v \equiv 0$ ,  $f \in L^\infty(0, T; L^2(\Omega))$ , and  $u$  and  $u_h$  be the solutions of (2.7) and (3.1) with  $f_h = P_h f$ , respectively. Then*

$$\|u_h - u\|_{L^2(0, T; L^2(\Omega))} + h \|\nabla(u_h - u)\|_{L^2(0, T; L^2(\Omega))} \leq Ch^2 \|f\|_{L^2(0, T; L^2(\Omega))}.$$

*Proof.* We use the splitting (3.53). By Lemma 3.1.1 and Theorem 2.3.2

$$\|\tilde{\varrho}\|_{L^2(0, T; L^2(\Omega))} + h \|\nabla \tilde{\varrho}\|_{L^2(0, T; L^2(\Omega))} \leq Ch^2 \|u\|_{L^2(0, T; L^2(\Omega))} \leq Ch^2 \|f\|_{L^2(0, T; L^2(\Omega))}.$$

By (3.16), and Lemmas 3.1.4 and 3.1.1, and Theorem 2.3.2, we have for  $p = 0, 1$ :

$$\begin{aligned}\int_0^T \|\tilde{\vartheta}(t)\|_p^2 dt &\leq C \int_0^T \|\Delta_h(R_h u - P_h u)(t)\|_{\dot{H}^{p-2}(\Omega)}^2 dt \\ &\leq C \int_0^T \|(R_h u - P_h u)(t)\|_{\dot{H}^p(\Omega)}^2 dt \\ &\leq Ch^{4-2p} \|u\|_{L^2(0, T; \dot{H}^2(\Omega))}^2 \leq Ch^{4-2p} \|f\|_{L^2(0, T; L^2(\Omega))}^2.\end{aligned}$$

Combing the preceding two estimates yields the desired assertion.  $\square$



Now we turn to error estimates in  $L^\infty(0, T; L^2(\Omega))$ - and  $L^\infty(0, T; H^1(\Omega))$ -norms. By Lemma 3.1.1 and Theorem 2.3.2, the following estimate holds for  $\tilde{\varrho}$ :

$$\|\tilde{\varrho}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\varrho}(t)\|_{L^2(\Omega)} \leq Ch^{2-\epsilon}\|u(t)\|_{\dot{H}^{2-\epsilon}(\Omega)} \leq C\epsilon^{-1}h^{2-\epsilon}\|f\|_{L^\infty(0,t;L^2(\Omega))}.$$

Now the choice  $\ell_h = |\ln h|$ ,  $\epsilon = 1/\ell_h$  yields

$$\|\tilde{\varrho}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\varrho}(t)\|_{L^2(\Omega)} \leq C\ell_h h^{2+q}\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}. \quad (3.23)$$

Thus, it suffices to bound the term  $\tilde{\vartheta}$ , which is shown in the following lemma.

**Lemma 3.1.9.** *Let  $\tilde{\vartheta}(t)$  be defined by (3.16). Then for  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 0$ , we have*

$$\|\tilde{\vartheta}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\vartheta}(t)\|_{L^2(\Omega)} \leq Ch^2\ell_h^2\|f\|_{L^\infty(0,t;L^2(\Omega))} \quad \text{with} \quad \ell_h = |\ln h|.$$

*Proof.* By (3.16) and Lemma 3.1.3, we deduce that for  $p = 0, 1$

$$\begin{aligned} \|\tilde{\vartheta}(t)\|_{\dot{H}^p(\Omega)} &\leq \int_0^t \|\bar{E}_h(t-s)\Delta_h(R_h u - P_h u)(s)\|_{\dot{H}^p(\Omega)} ds \\ &\leq C \int_0^t (t-s)^{\epsilon\alpha/2-1} \|\Delta_h(R_h u - P_h u)(s)\|_{\dot{H}^{p-2+\epsilon}(\Omega)} ds \\ &\leq C \int_0^t (t-s)^{\epsilon\alpha/2-1} \|R_h u(s) - P_h u(s)\|_{\dot{H}^{p+\epsilon}(\Omega)} ds := A. \end{aligned}$$

Further, we apply the inverse estimate from Lemma 3.1.5 for  $R_h u - P_h u$  and the bounds in Lemma 3.1.1, for  $P_h u - u$  and  $R_h u - u$ , respectively, to deduce

$$\begin{aligned} A &\leq Ch^{-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} \|R_h u(s) - P_h u(s)\|_{\dot{H}^p(\Omega)} ds \\ &\leq Ch^{2+q-p-2\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} \|u(s)\|_{\dot{H}^{2+q-\epsilon}(\Omega)} ds. \end{aligned}$$

Further, by applying estimate (2.20) and choosing  $\epsilon = 1/\ell_h$  we get

$$A \leq C\epsilon^{-1}h^{2+q-p-2\epsilon}\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \int_0^t (t-s)^{\epsilon\alpha/2-1} t^{\epsilon\alpha/2} ds \leq Ch^{2+q-p}\ell_h^2\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

This completes the proof of the lemma. □

Now we can state an almost optimal error estimate for the approximation  $u_h$ .

**Theorem 3.1.5.** *Let  $v \equiv 0$ ,  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 0$ , and  $u$  and  $u_h$  be the solutions of (2.7) and (3.1) with  $f_h = P_h f$ , respectively. Then with  $\ell_h = |\ln h|$ , there holds*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^{2+q}\ell_h^2\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

**Remark 3.1.3.** *In comparison with the  $L^2(0, T; \dot{H}^p(\Omega))$ -norm estimates, the  $L^\infty(0, T; \dot{H}^p(\Omega))$ -norm estimates suffer from the factor  $\ell_h^2$ . This is due to Lemma 3.1.3 and the regularity estimate in Theorem 2.3.2, reflecting the limited smoothing property of the solution operator.*

**Remark 3.1.4.** *An inspection of the proof of Lemma 3.1.9 indicates that for  $0 < q \leq 1$ , one can obtain a slightly improved estimate*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^2\ell_h\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

### 3.2 Spatial semidiscretization by lumped mass method

Now we consider the more practical lumped mass FEM (see, e.g. [73, Chapter 15, pp. 239–244]) and study the convergence rates for smooth and non-smooth initial data. First we introduce this approximation. Let  $z_j^\tau$ ,  $j = 1, \dots, d+1$  be the vertices of the  $d$ -simplex  $\tau \in \mathcal{T}_h$ . Consider the quadrature formula

$$Q_{\tau,h}(f) = \frac{|\tau|}{d+1} \sum_{j=1}^{d+1} f(z_j^\tau) \approx \int_\tau f dx. \quad (3.24)$$

We then define an approximation of the  $L_2$ -inner product in  $V_h$  by

$$(w, \chi)_h = \sum_{\tau \in \mathcal{T}_h} Q_{\tau,h}(w\chi). \quad (3.25)$$

Then the lumped mass Galerkin FEM is: find  $\bar{u}_h(t) \in V_h$  such that

$$({}^C\partial_t^\alpha \bar{u}_h, \chi)_h + a(\bar{u}_h, \chi) = (f, \chi) \quad \forall \chi \in V_h, \quad t > 0, \quad \bar{u}_h(0) = v_h. \quad (3.26)$$

The lumped mass method leads to a diagonal mass matrix, which in practice is important for preserving the qualitative properties of the semidiscrete and fully discrete approximations.

We now introduce the discrete Laplacian  $-\bar{\Delta}_h : V_h \rightarrow V_h$ , corresponding to the inner product  $(\cdot, \cdot)_h$ ,

by

$$-(\bar{\Delta}_h \psi, \chi)_h = (\nabla \psi, \nabla \chi) \quad \forall \psi, \chi \in V_h. \quad (3.27)$$

Also, we introduce a projection operator  $\bar{P}_h : L_2(\Omega) \rightarrow V_h$  by

$$(\bar{P}_h f, \chi)_h = (f, \chi), \quad \forall \chi \in V_h.$$

The lumped mass method can then be written with  $f_h = \bar{P}_h f$  in operator form as

$${}^C \partial_t^\alpha \bar{u}_h(t) - \bar{\Delta}_h \bar{u}_h(t) = f_h(t) \quad \text{for } t \geq 0 \quad \text{with } \bar{u}_h(0) = v_h.$$

Similarly as in Section 3.1, we define the discrete operator  $F_h$  by

$$F_h(t)v_h = \sum_{j=1}^N E_{\alpha,1}(-\bar{\lambda}_j^h t^\alpha)(v_h, \bar{\varphi}_j^h)_h \bar{\varphi}_j^h, \quad (3.28)$$

where  $\{\bar{\lambda}_j^h\}_{j=1}^N$  and  $\{\bar{\varphi}_j^h\}_{j=1}^N$  are respectively the eigenvalues and the orthonormal eigenfunctions of  $-\bar{\Delta}_h$  with respect to  $(\cdot, \cdot)_h$ .

Analogously to (3.5), we introduce the operator  $\bar{E}_h$  by

$$\bar{E}_h f_h(t) = \sum_{j=1}^N t^{\alpha-1} E_{\alpha,\alpha}(-\bar{\lambda}_j^h t^\alpha)(f_h, \bar{\varphi}_j^h)_h \bar{\varphi}_j^h. \quad (3.29)$$

Then the solution  $\bar{u}_h$  to problem (3.26) can be represented as follows

$$\bar{u}_h(t) = F_h(t)v_h + \int_0^t \bar{E}_h(t-s)f_h(s)ds.$$

For our analysis we shall need the following modification of the discrete norm (3.7),  $\|\cdot\|_p$ , on the space  $V_h$

$$\|\psi\|_{\dot{H}^p(\Omega)}^2 = \sum_{j=1}^N (\bar{\lambda}_j^h)^p (\psi, \bar{\varphi}_j^h)_h^2 \quad \forall p \in \mathbb{R}. \quad (3.30)$$

The following norm equivalence result is useful.

**Lemma 3.2.1.** *The norm  $\|\cdot\|_{\dot{H}^p(\Omega)}$  defined in (3.30) is equivalent to the norm  $\|\cdot\|_{\dot{H}^p(\Omega)}$  on the space  $V_h$  for  $p = 0, 1$ .*

*Proof.* The proof is a simple consequence of the definitions and is omitted.  $\square$

We shall also need the following inverse inequality, whose proof is identical with that of Lemma 3.1.5:

$$\|\psi\|_{\dot{H}^l(\Omega)} \leq Ch^{s-l} \|\psi\|_{\dot{H}^s(\Omega)} \quad l > s. \quad (3.31)$$

We show the following discrete analogue of Lemma 3.1.3:

**Lemma 3.2.2.** *Let  $\bar{E}_h$  be defined by (3.29). Then we have for  $\psi \in V_h$  and all  $t > 0$ ,*

$$\|\bar{E}_h(t)\psi\|_{\dot{H}^p(\Omega)} \leq \begin{cases} Ct^{-1+\alpha(1+\frac{q-p}{2})} \|\psi\|_{\dot{H}^q(\Omega)}, & p-2 \leq q \leq p, \\ Ct^{-1+\alpha} \|\psi\|_{\dot{H}^q(\Omega)}, & p < q. \end{cases}$$

We need the quadrature error operator  $Q_h : V_h \rightarrow V_h$  defined by

$$(\nabla Q_h \chi, \nabla \psi) = \epsilon_h(\chi, \psi) := (\chi, \psi)_h - (\chi, \psi) \quad \forall \chi, \psi \in V_h. \quad (3.32)$$

The operator  $Q_h$ , introduced in [8], represents the quadrature error (due to mass lumping) in a special way. It satisfies the following error estimate [8, Lemma 2.4].

**Lemma 3.2.3.** *Let  $\bar{\Delta}_h$  and  $Q_h$  be defined by (3.27) and (3.32), respectively. Then*

$$\|\nabla Q_h \chi\|_{L^2(\Omega)} + h \|\bar{\Delta}_h Q_h \chi\|_{L^2(\Omega)} \leq Ch^{p+1} \|\nabla^p \chi\|_{L^2(\Omega)} \quad \forall \chi \in V_h, \quad p = 0, 1.$$

### 3.2.1 Error estimates for the homogeneous problem

We now establish error estimates for the lumped mass FEM for smooth initial data, i.e.,  $v \in \dot{H}^2(\Omega)$ .

**Theorem 3.2.1.** *Let  $u$  and  $\bar{u}_h$  be the solutions of (2.7) and (3.26), respectively, with  $v_h = R_h v$ . Then*

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} + h \|\nabla(\bar{u}_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^2 \|v\|_{\dot{H}^2(\Omega)}.$$

*Proof.* We split the error into  $\bar{u}_h(t) - u(t) = u_h(t) - u(t) + \delta(t)$  with  $\delta(t) = \bar{u}_h(t) - u_h(t)$  and  $u_h(t)$  being the solution by the standard Galerkin FEM. Upon noting the estimate for  $u_h - u$ , it suffices to show

$$\|\delta(t)\|_{L^2(\Omega)} + h \|\nabla \delta(t)\|_{L^2(\Omega)} \leq Ch^2 \|v\|_{\dot{H}^2(\Omega)}. \quad (3.33)$$

It follows from the definitions of  $u_h(t)$ ,  $\bar{u}_h(t)$ , and  $Q_h$  that

$${}^C\partial_t^\alpha \delta(t) - \bar{\Delta}_h \delta(t) = \bar{\Delta}_h Q_h {}^C\partial_t^\alpha u_h(t) \quad \text{for } t > 0, \quad \delta(0) = 0$$

and by Duhamel's principle we have

$$\delta(t) = \int_0^t \bar{E}_h(t-s) \bar{\Delta}_h Q_h {}^C\partial_t^\alpha u_h(s) ds. \quad (3.34)$$

Using Lemmas 3.2.1, 3.2.2, and 3.2.3 we get for  $\chi \in V_h$ :

$$\|\nabla \bar{E}_h(t) \bar{\Delta}_h Q_h \chi\|_{L^2(\Omega)} \leq C t^{\frac{\alpha}{2}-1} \|\bar{\Delta}_h Q_h \chi\|_{L^2(\Omega)} \leq C t^{\frac{\alpha}{2}-1} h \|\nabla \chi\|_{L^2(\Omega)}.$$

Similarly, for  $\chi \in V_h$

$$\|\bar{E}_h(t) \bar{\Delta}_h Q_h \chi\|_{L^2(\Omega)} \leq C t^{\frac{\alpha}{2}-1} \|\bar{\Delta}_h Q_h \chi\|_{\dot{H}^{-1}(\Omega)} \leq C t^{\frac{\alpha}{2}-1} \|\nabla Q_h \chi\|_{L^2(\Omega)} \leq C t^{\frac{\alpha}{2}-1} h^2 \|\nabla \chi\|_{L^2(\Omega)}.$$

Consequently, using Lemma 3.1.2 with  $l = 1$ ,  $p = 1$  and  $q = 2$  we get

$$\begin{aligned} \|\delta(t)\|_{L^2(\Omega)} + h \|\nabla \delta(t)\|_{L^2(\Omega)} &\leq C h^2 \int_0^t (t-s)^{\frac{\alpha}{2}-1} \|{}^C\partial_t^\alpha u_h(s)\|_{\dot{H}^1(\Omega)} ds \\ &\leq C h^2 \int_0^t (t-s)^{\frac{\alpha}{2}-1} s^{-\frac{\alpha}{2}} ds \|u_h(0)\|_{\dot{H}^2(\Omega)}. \end{aligned}$$

Since  $\Delta_h R_h = P_h \Delta$ , we deduce

$$\|u_h(0)\|_{\dot{H}^2(\Omega)} = \|\Delta_h R_h u(0)\|_{L^2(\Omega)} = \|P_h \Delta u(0)\|_{L^2(\Omega)} \leq \|u(0)\|_{\dot{H}^2(\Omega)} \leq C \|v\|_{\dot{H}^2(\Omega)},$$

which yields (3.33) and concludes the proof.  $\square$

An improved bound for  $\|\nabla \delta(t)\|_{L^2(\Omega)}$  can be obtained as follows. In view of Lemmas 3.2.1 and 3.2.3 and (3.31), we observe that for any  $\epsilon > 0$  and  $\chi \in V_h$

$$\|\nabla \bar{E}_h(t) \bar{\Delta}_h Q_h \chi\|_{L^2(\Omega)} \leq C t^{\frac{\epsilon}{2}\alpha-1} \|\bar{\Delta}_h Q_h \chi\|_{\dot{H}^{-1+\epsilon}(\Omega)} \leq C t^{\frac{\epsilon}{2}\alpha-1} h^{2-\epsilon} \|\nabla \chi\|_{L^2(\Omega)}.$$

Consequently,

$$\|\nabla\delta(t)\|_{L^2(\Omega)} \leq Ch^{2-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} \|C\partial_t^\alpha u_h(s)\|_{\dot{H}^1(\Omega)} ds. \quad (3.35)$$

Now, to (3.35) we apply Lemma 3.1.2 with  $\ell = 1$ ,  $p = 1$  and  $q = 2$  to get

$$\|\nabla\delta(t)\|_{L^2(\Omega)} \leq Ch^{2-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} s^{-\frac{\alpha}{2}} ds \|u_h(0)\|_{\dot{H}^2(\Omega)} \leq C \frac{1}{\epsilon} h^{2-\epsilon} t^{-\alpha\frac{1-\epsilon}{2}} \|v\|_{\dot{H}^2(\Omega)}.$$

**Remark 3.2.1.** In the above estimate, by choosing  $\epsilon = 1/\ell_h$ ,  $\ell_h = |\ln h|$ , we get

$$\|\nabla\delta(t)\|_{L^2(\Omega)} \leq Ch^2 \ell_h t^{-\frac{\alpha}{2}} \|v\|_{\dot{H}^2(\Omega)}, \quad (3.36)$$

which improves the bound of  $\|\nabla\delta(t)\|_{L^2(\Omega)}$  for any fixed  $t > 0$  by almost one order.

**Remark 3.2.2.** Instead, if we apply to (3.35) Lemma 3.1.2 with  $\ell = 1$ ,  $p = 1$  and  $q = 1$  we get an improved estimate for  $\delta(t)$  in the case of initial data  $v \in \dot{H}^1(\Omega)$ :

$$\|\nabla\delta(t)\|_{L^2(\Omega)} \leq Ch^2 \ell_h t^{-\alpha} \|v\|_{\dot{H}^1(\Omega)}. \quad (3.37)$$

Now we consider the case of nonsmooth initial data  $v \in L^2(\Omega)$  as well as the intermediate case  $v \in \dot{H}^1$ . Due to the lower regularity, we take  $v_h = P_h v$ . Like before, the idea is to split the error into  $\bar{u}_h(t) - u(t) = u_h(t) - u(t) + \delta(t)$  with  $\delta(t) = \bar{u}_h(t) - u_h(t)$  and  $u_h(t)$  being the solution of (3.1). Thus, in view of estimate in Theorem 3.1.2 it suffices to establish proper bounds for  $\delta(t)$ .

**Theorem 3.2.2.** Let  $u$  and  $\bar{u}_h$  be the solutions of (2.7) and (3.26), respectively, with  $v_h = P_h v$ . Then with  $\ell_h = |\ln h|$ , the following estimates are valid for  $t > 0$ :

$$\|\nabla(\bar{u}_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch \ell_h t^{-\alpha(1-\frac{q}{2})} \|v\|_{\dot{H}^q(\Omega)} \quad q = 0, 1, \quad (3.38)$$

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} \leq Ch^{q+1} \ell_h t^{-\alpha(1-\frac{q}{2})} \|v\|_{\dot{H}^q(\Omega)} \quad q = 0, 1. \quad (3.39)$$

Furthermore, if the quadrature error operator  $Q_h$  defined by (3.32) satisfies

$$\|Q_h \chi\|_{L^2(\Omega)} \leq Ch^2 \|\chi\|_{L^2(\Omega)} \quad \forall \chi \in V_h, \quad (3.40)$$

then the following almost optimal error estimate is valid:

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} \leq Ch^2 \ell_h t^{-\alpha} \|v\|_{L^2(\Omega)}. \quad (3.41)$$

*Proof.* By Duhamel's principle  $\delta(t) = \int_0^t \bar{E}_h(t-s) \bar{\Delta}_h Q_h^C \partial_t^\alpha u_h(s) ds$ . Then by appealing to the smoothing property of the operator  $\bar{E}_h$  in Lemma 3.2.2 and the inverse inequality (3.31), we get for  $\chi \in V_h$ ,  $\epsilon > 0$ , and  $p = 0, 1$

$$\begin{aligned} \|\bar{E}_h(t) \bar{\Delta}_h Q_h \chi\|_{\dot{H}^p(\Omega)} &\leq Ct^{\frac{\epsilon}{2}\alpha-1} \|\bar{\Delta}_h Q_h \chi\|_{\dot{H}^{p-2+\epsilon}(\Omega)} = Ct^{\frac{\epsilon}{2}\alpha-1} \|Q_h \chi\|_{\dot{H}^{p+\epsilon}(\Omega)} \\ &\leq Ct^{\frac{\epsilon}{2}\alpha-1} h^{-\epsilon} \|Q_h \chi\|_{\dot{H}^p(\Omega)} \leq Ct^{\frac{\epsilon}{2}\alpha-1} h^{-\epsilon} \|Q_h \chi\|_{\dot{H}^p(\Omega)}. \end{aligned} \quad (3.42)$$

Consequently, by Lemmas 3.2.3, 3.1.2 and  $\dot{H}^1$ - and  $L^2$ -stability of the operator  $P_h$  from Lemma 3.1.1, we deduce for  $q = 0, 1$

$$\begin{aligned} \|\nabla \delta(t)\|_{L^2(\Omega)} &\leq Ch^{q+1-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} \|{}^C \partial_t^\alpha u_h(s)\|_{\dot{H}^q(\Omega)} ds \\ &\leq Ch^{q+1-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} s^{-\alpha} ds \|u_h(0)\|_{\dot{H}^q(\Omega)} \\ &= Ch^{q+1-\epsilon} t^{-\alpha(1-\frac{\epsilon}{2})} B\left(\frac{\epsilon}{2}\alpha, 1-\alpha\right) \|P_h v\|_{\dot{H}^q(\Omega)} \leq C\epsilon^{-1} h^{q+1-\epsilon} t^{-\alpha(1-\frac{\epsilon}{2})} \|v\|_{\dot{H}^q(\Omega)}. \end{aligned}$$

Now the estimate (3.38) follows by triangle inequality from this and Theorem 3.1.2 by taking  $\epsilon = 1$  and  $\epsilon = 1/\ell_h$  for the cases  $q = 1$  and  $0$ , respectively.

Next we derive an  $L^2$ - error estimate. First, note that for  $\chi \in V_h$  we have

$$\|\bar{E}_h(t) \bar{\Delta}_h Q_h \chi\|_{L^2(\Omega)} \leq Ct^{\frac{\alpha}{2}-1} \|\bar{\Delta}_h Q_h \chi\|_{\dot{H}^{-1}(\Omega)} \leq Ct^{\frac{\alpha}{2}-1} \|\nabla Q_h \chi\|_{L^2(\Omega)}.$$

This estimate together with Lemma 3.2.3 gives

$$\begin{aligned} \|\delta(t)\|_{L^2(\Omega)} &\leq Ch^{q+1} \int_0^t (t-s)^{\frac{\alpha}{2}-1} \|{}^C \partial_t^\alpha u_h(s)\|_{\dot{H}^q(\Omega)} ds \\ &\leq Ch^{q+1} \int_0^t (t-s)^{\frac{\alpha}{2}-1} s^{-\alpha} ds \|u_h(0)\|_{\dot{H}^q(\Omega)} \\ &\leq Ch^{q+1} t^{-\frac{\alpha}{2}} \|P_h v\|_{\dot{H}^q(\Omega)} \leq Ch^{q+1} t^{-\frac{\alpha}{2}} \|v\|_{\dot{H}^q(\Omega)}, \quad q = 0, 1, \end{aligned}$$

which shows the desired estimate (3.39).

Finally, if (3.40) holds, by applying (3.42) with  $p = 0$  and  $\epsilon \in (0, \frac{1}{2})$ , we get

$$\begin{aligned} \|\delta(t)\|_{L^2(\Omega)} &\leq Ch^{-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} \|Q_h^C \partial_t^\alpha u_h(s)\|_{L^2(\Omega)} ds \leq Ch^{2-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} \|{}^C \partial_t^\alpha u_h(s)\|_{L^2(\Omega)} ds \\ &\leq Ch^{2-\epsilon} \int_0^t (t-s)^{\frac{\epsilon}{2}\alpha-1} s^{-\alpha} ds \|u_h(0)\|_{L^2(\Omega)} \leq C\epsilon^{-1} h^{2-\epsilon} t^{-\alpha(1-\frac{\epsilon}{2})} \|v\|_{L^2(\Omega)}. \end{aligned}$$

Then (3.41) follows immediately by choosing  $\epsilon = 1/\ell_h$ .  $\square$

**Remark 3.2.3.** *By interpolation (3.41) is valid also for  $0 < q < 1$ .*

**Remark 3.2.4.** *The condition (3.40) on the quadrature error operator  $Q_h$  is satisfied for symmetric meshes [8, Section 5]. In one dimension, it can be relaxed to almost symmetry [8, Section 6]. In case (3.40) does not hold, we were able to show only a suboptimal  $O(h)$ -convergence rate for  $L^2$ -norm of the error, which is reminiscent of that in the classical parabolic equation [8, Theorem 4.4].*

**Remark 3.2.5.** *We note that we have used a globally quasi-uniform meshes, while the results in [8] are valid for meshes that satisfy the inverse inequality only locally.*

**Remark 3.2.6.** *If the initial data is very weak, i.e.,  $v \in \dot{H}^q(\Omega)$  with  $-1 < q < 0$  and the mesh is quasi-uniform. Then Theorem 3.2.2 and the argument of [25, Theorem 2] yield the following optimal error estimate for the lumped mass finite element method (3.26)*

$$\|u(t) - \bar{u}_h(t)\|_{L^2(\Omega)} + h\|\nabla(u(t) - \bar{u}_h(t))\|_{L^2(\Omega)} \leq Ch^{2+q} \ell_h t^{-\alpha} \|v\|_{\dot{H}^q(\Omega)}. \quad (3.43)$$

### 3.2.2 Error estimates for inhomogeneous problems

Now we first derive an  $L^2(0, T; \dot{H}^p(\Omega))$ -error estimate,  $p = 0, 1$ , for the lumped mass method.

**Theorem 3.2.3.** *Let  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ , and  $u$  and  $u_h$  be the solutions of (2.7) and (3.26) with  $f_h = \bar{P}_h f$ , respectively. Then there hold*

$$\begin{aligned} \|\nabla(\bar{u}_h - u)\|_{L^2(0, T; L^2(\Omega))} &\leq Ch^{1+\min(q, 0)} \|f\|_{L^2(0, T; \dot{H}^q(\Omega))}, \\ \|\bar{u}_h - u\|_{L^2(0, T; L^2(\Omega))} &\leq Ch^{1+q} \|f\|_{L^2(0, T; \dot{H}^q(\Omega))}. \end{aligned}$$



*Proof.* By repeating the proof of Lemma 3.1.4, we deduce from (3.34), Lemma 3.2.1, and (3.32) that

$$\begin{aligned} \int_0^T \|\nabla \delta(t)\|_{L^2(\Omega)}^2 dt &\leq C \int_0^T \|\bar{\Delta}_h Q_h^C \partial_t^\alpha u_h(t)\|_{\dot{H}^{-1}(\Omega)}^2 dt \leq C \int_0^T \|Q_h^C \partial_t^\alpha u_h(t)\|_{\dot{H}^1(\Omega)}^2 dt \\ &\leq C \int_0^T \|\nabla Q_h^C \partial_t^\alpha u_h(t)\|_{L^2(\Omega)}^2 dt \leq Ch^2 \int_0^T \|{}^C \partial_t^\alpha u_h(t)\|_{L^2(\Omega)}^2 dt. \end{aligned}$$

The desired assertion for the case  $q \geq 0$  now follows immediately from Lemma 3.1.4.

For  $-1 < q < 0$  we use Lemmas 3.1.5, 3.1.4 and 3.1.1 to get

$$\begin{aligned} \int_0^T \|\nabla \delta(t)\|_{L^2(\Omega)}^2 dt &\leq Ch^{2+2q} \int_0^T \|{}^C \partial_t^\alpha u_h(t)\|_{\dot{H}^q(\Omega)}^2 dt \\ &\leq Ch^{2+2q} \int_0^T \|f_h(t)\|_{\dot{H}^q(\Omega)}^2 dt \leq Ch^{2+2q} \|f\|_{L^2(0,T;\dot{H}^q(\Omega))}^2. \end{aligned}$$

Now we turn to the  $L^2$ -estimate. By repeating the preceding arguments, we arrive at

$$\begin{aligned} \int_0^T \|\delta(t)\|_{L^2(\Omega)}^2 dt &\leq C \int_0^T \|\bar{\Delta}_h Q_h^C \partial_t^\alpha u_h(t)\|_{\dot{H}^{-2}(\Omega)}^2 dt = C \int_0^T \|Q_h^C \partial_t^\alpha u_h(t)\|_{L^2(\Omega)}^2 dt \\ &\leq C \int_0^T \|\nabla Q_h^C \partial_t^\alpha u_h(t)\|_{L^2(\Omega)}^2 dt \leq Ch^4 \int_0^T \|{}^C \partial_t^\alpha u_h(t)\|_{\dot{H}^1(\Omega)}^2 dt \\ &\leq Ch^{2+2q} \int_0^T \|{}^C \partial_t^\alpha u_h(t)\|_{\dot{H}^q(\Omega)}^2 dt, \end{aligned}$$

where the second line follows from the trivial inequality  $\|\chi\|_{L^2(\Omega)} \leq C\|\nabla\chi\|_{L^2(\Omega)}$  for  $\chi \in V_h$  and the norm equivalence in Lemma 3.2.1. The rest of the proof follows identically, and hence it is omitted.  $\square$

The estimate in  $L^2(0,T;L^2(\Omega))$ -norm of Theorem 3.2.3 is suboptimal for any  $q < 1$ . An optimal estimate can be obtained under an additional condition on the mesh.

**Theorem 3.2.4.** *Let the assumptions in Theorem 3.2.3 be fulfilled and the operator  $Q_h$  satisfy (3.40).*

*Then*

$$\|\bar{u}_h - u\|_{L^2(0,T;L^2(\Omega))} \leq Ch^{2+\min(q,0)} \|f\|_{L^2(0,T;\dot{H}^q(\Omega))}.$$

*Proof.* It follows from the condition on the operator  $Q_h$  that

$$\begin{aligned} \int_0^T \|\delta(t)\|_{L^2(\Omega)}^2 dt &\leq C \int_0^T \|\bar{\Delta}_h Q_h^C \partial_t^\alpha u_h(t)\|_{\dot{H}^{-2}(\Omega)}^2 dt \leq C \int_0^T \|Q_h^C \partial_t^\alpha u_h(t)\|_{L^2(\Omega)}^2 dt \\ &\leq Ch^4 \int_0^T \|{}^C \partial_t^\alpha u_h(t)\|_{L^2(\Omega)}^2 dt \leq Ch^{4+2\min(q,0)} \int_0^T \|{}^C \partial_t^\alpha u_h(t)\|_{\dot{H}^q(\Omega)}^2 dt. \end{aligned}$$

The rest of the proof is identical with that of Theorem 3.2.3.  $\square$

**Remark 3.2.7.** *The condition (3.40) on the operator  $Q_h$  is satisfied for symmetric meshes [8, section 5]. In one dimension, the symmetry requirement can be relaxed to almost symmetry [8, section 6].*

Next we derive an estimate in  $L^\infty(0, T; \dot{H}^p(\Omega))$ -norm for the lumped mass approximation  $\bar{u}_h$ .

**Theorem 3.2.5.** *Let  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 1$ , and  $u$  and  $\bar{u}_h$  be the solutions of (2.7) and (3.26), respectively, with  $\bar{f}_h = \bar{P}_h f$ . Then with  $\ell_h = |\ln h|$ , the following estimate is valid for  $t > 0$ :*

$$\|\nabla(\bar{u}_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^{1+q}\ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \quad \text{for } -1 < q \leq 0. \quad (3.44)$$

Moreover, for  $-1 < q \leq 1$  there holds

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} \leq Ch^{1+q}\ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}. \quad (3.45)$$

*Proof.* By Lemma 3.2.2 and (3.31), we have for  $\chi \in V_h$ ,  $\epsilon > 0$ , and  $p = 0, 1$

$$\begin{aligned} \|\|\bar{E}_h(t)\bar{\Delta}_h Q_h \chi\|\|_{\dot{H}^p(\Omega)} &\leq Ct^{\epsilon\alpha/2-1} \|\|\bar{\Delta}_h Q_h \chi\|\|_{\dot{H}^{p-2+\epsilon}(\Omega)} = Ct^{\epsilon\alpha/2-1} \|\|Q_h \chi\|\|_{\dot{H}^{p+\epsilon}(\Omega)} \\ &\leq Ct^{\epsilon\alpha/2-1} h^{-\epsilon} \|\|Q_h \chi\|\|_{\dot{H}^p(\Omega)}. \end{aligned} \quad (3.46)$$

We first prove estimate (3.44). Setting  $\chi = {}^C\partial_t^\alpha u_h(t)$  in (3.46) and (3.32) yields

$$\begin{aligned} \|\|\bar{E}_h(t-s)\bar{\Delta}_h Q_h {}^C\partial_t^\alpha u_h(s)\|\|_{\dot{H}^1(\Omega)} &\leq C(t-s)^{\epsilon\alpha/2-1} h^{-\epsilon} \|Q_h {}^C\partial_t^\alpha u_h(s)\|_{\dot{H}^1(\Omega)} \\ &\leq Ch^{1-\epsilon}(t-s)^{\epsilon\alpha/2-1} \|{}^C\partial_t^\alpha u_h(s)\|_{L^2(\Omega)}. \end{aligned} \quad (3.47)$$

Then it follows from (3.1) and the triangle and inverse inequalities that

$$\begin{aligned} \|\nabla\delta(t)\|_{L^2(\Omega)} &\leq Ch^{1-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} \|{}^C\partial_t^\alpha u_h(s)\|_{L^2(\Omega)} ds \\ &\leq Ch^{1-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (\|\Delta_h u_h(s)\|_{L^2(\Omega)} + \|f_h(s)\|_{L^2(\Omega)}) ds \\ &\leq Ch^{1-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (h^{-\epsilon} \|\|\Delta_h u_h(s)\|\|_{\dot{H}^{-\epsilon}(\Omega)} + \|f_h(s)\|_{L^2(\Omega)}) ds. \end{aligned}$$

Further, using Lemmas 3.1.4 and 3.1.1 we further get for  $-1 < q \leq 0$

$$\begin{aligned}
\|\nabla\delta(t)\|_{L^2(\Omega)} &\leq Ch^{1+q-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (h^{-\epsilon} \|\Delta_h u_h(s)\|_{\dot{H}^{q-\epsilon}(\Omega)} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq Ch^{1+q-2\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (\epsilon^{-1} s^{\epsilon\alpha/2} \|f_h\|_{L^\infty(0,s;\dot{H}^q(\Omega))} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq C\epsilon^{-1} h^{1+q-2\epsilon} \|f_h\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \int_0^t (t-s)^{\epsilon\alpha/2-1} s^{\epsilon\alpha/2} ds \\
&\leq C\epsilon^{-2} h^{1+q-2\epsilon} \|f_h\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \leq Ch^{1+q} \ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))},
\end{aligned}$$

where in the last inequality we have chosen  $\epsilon = 1/\ell_h$ . Now (3.44) follows from this and Theorem 3.1.5.

Next we derive the  $L^2$ -error estimate. Similar to the derivation of (3.46), we get

$$\begin{aligned}
\|\bar{E}_h(t-s)\bar{\Delta}_h Q_h^C \partial_t^\alpha u_h(s)\|_{L^2(\Omega)} &\leq C(t-s)^{\alpha/2-1} \|\nabla Q_h^C \partial_t^\alpha u_h(s)\|_{L^2(\Omega)} \\
&\leq Ch^2(t-s)^{\alpha/2-1} \|C \partial_t^\alpha u_h(s)\|_{\dot{H}^1(\Omega)}.
\end{aligned}$$

Consequently, by the triangle inequality, Lemmas 3.1.5 and 3.1.4, there holds

$$\begin{aligned}
\|\delta(t)\|_{L^2(\Omega)} &\leq Ch^2 \int_0^t (t-s)^{\alpha/2-1} (\|\Delta u_h(s)\|_{\dot{H}^1(\Omega)} + \|P_h f(s)\|_{\dot{H}^1(\Omega)}) ds \\
&\leq Ch^{1+q} \int_0^t (t-s)^{\alpha/2-1} (h^{-\epsilon} \|\Delta u_h(s)\|_{\dot{H}^{q-\epsilon}(\Omega)} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq Ch^{1+q} \int_0^t (t-s)^{\alpha/2-1} (\epsilon^{-1} h^{-\epsilon} s^{\epsilon\alpha/2} \|f_h\|_{L^\infty(0,s;\dot{H}^q(\Omega))} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds.
\end{aligned}$$

The  $L^2$ -estimate follows by setting  $\epsilon = 1/\ell_h$  in this inequality and Theorem 3.1.5.  $\square$

**Remark 3.2.8.** For  $q > 0$ , we have  $\|\nabla(\bar{u}_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch\ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}$  and it cannot be improved even if the function  $f$  is smoother. In view of Remark 3.1.4, for  $0 < q \leq 1$ , there holds

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} \leq Ch^{1+q} \ell_h \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

In the case of  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $0 < q \leq 1$ , we can obtain an improved estimate of  $\|\nabla\delta(t)\|_{L^2(\Omega)}$ :

$$\begin{aligned}
\|\nabla\delta(t)\|_{L^2(\Omega)} &\leq Ch^{2-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} \|{}^C\partial_t^\alpha u_h(s)\|_{\dot{H}^1(\Omega)} ds \\
&\leq Ch^{1+q-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (h^{-\epsilon} \|\Delta_h u_h(s)\|_{\dot{H}^{q-\epsilon}(\Omega)} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq Ch^{1+q-2\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (\epsilon^{-1} s^{\epsilon\alpha/2} \|f_h\|_{L^\infty(0,s;\dot{H}^q(\Omega))} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq C\epsilon^{-2} h^{1+q-2\epsilon} \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \leq Ch^{1+q} \ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.
\end{aligned}$$

We record this observation in a remark.

**Remark 3.2.9.** For  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $0 < q \leq 1$ , the estimate  $\|\nabla\delta(t)\|_{L^2(\Omega)}$  can be improved to  $(1+q)$ th-order at the expense of an additional factor  $\ell_h$ :

$$\|\nabla\delta(t)\|_{L^2(\Omega)} \leq Ch^{1+q} \ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

Like in the case of  $L^2(0, T; L^2(\Omega))$ -estimate, the  $L^\infty(0, T; L^2(\Omega))$  estimate is suboptimal for any  $q \in (-1, 1)$ , and can be improved to an almost optimal one by imposing condition (3.40).

**Theorem 3.2.6.** Let the conditions in Theorem 3.2.5 be fulfilled and (3.40) hold. Then with  $\ell_h = |\ln h|$ ,

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} \leq Ch^{2+\min(q,0)} \ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}, \quad -1 < q \leq 1.$$

*Proof.* If (3.40) holds, then applying (3.46) with  $p = 0$  we get

$$\begin{aligned}
\|\bar{E}_h(t-s) \bar{\Delta}_h Q_h {}^C\partial_t^\alpha u_h(s)\|_{L^2(\Omega)} &\leq C(t-s)^{\epsilon\alpha/2-1} h^{-\epsilon} \|Q_h {}^C\partial_t^\alpha u_h(s)\|_{L^2(\Omega)} \\
&\leq Ch^{2-\epsilon} (t-s)^{\epsilon\alpha/2-1} \|{}^C\partial_t^\alpha u_h(s)\|_{L^2(\Omega)}.
\end{aligned}$$

Consequently, this together with Lemmas 3.1.5, 3.1.2, and 3.1.1, yields

$$\begin{aligned}
\|\delta(t)\|_{L^2(\Omega)} &\leq Ch^{2-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (\|\Delta_h u_h(s)\|_{L^2(\Omega)} + \|f_h(s)\|_{L^2(\Omega)}) ds \\
&\leq Ch^{2+\min(q,0)-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (h^{-\epsilon} \|\Delta_h u_h(s)\|_{\dot{H}^{-\epsilon+q}(\Omega)} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq Ch^{2+\min(q,0)-\epsilon} \int_0^t (t-s)^{\epsilon\alpha/2-1} (h^{-\epsilon} \epsilon^{-1} s^{\epsilon\alpha/2} \|f_h\|_{L^\infty(0,s;\dot{H}^q(\Omega))} + \|f_h(s)\|_{\dot{H}^q(\Omega)}) ds \\
&\leq C\epsilon^{-2} h^{2+\min(q,0)-2\epsilon} \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} \leq Ch^{2+\min(q,0)} \ell_h^2 \|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))},
\end{aligned}$$

where the last line follows from the choice  $\epsilon = 1/\ell_h$ . This and Theorem 3.1.5 conclude the proof.  $\square$

**Remark 3.2.10.** *In case of an initial data  $u(0) \in \dot{H}^q(\Omega)$  and a right hand side  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 0$ , by combining Theorems 3.2.5 and 3.2.6 with Theorem 3.2.2, we deduce that with the choice  $\bar{u}_h(0) = P_h u(0)$  and  $f_h = \bar{P}_h f$ , the lumped mass approximation  $\bar{u}_h$  satisfies*

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} + \|\nabla(\bar{u}_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^{1+q}\ell_h^2 t^{-\alpha} (\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} + \|u(0)\|_{\dot{H}^q(\Omega)}).$$

Further, under condition (3.40), there holds

$$\|\bar{u}_h(t) - u(t)\|_{L^2(\Omega)} \leq Ch^{2+q}\ell_h^2 t^{-\alpha} (\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))} + \|u(0)\|_{\dot{H}^q(\Omega)}).$$

### 3.3 Numerical results

In this section, we present 2-D numerical results to verify the convergence theory in Sections 3.1 and 3.2. We present the errors  $\|(u - \bar{u}_h)(t)\|_{L^2(\Omega)}$  and  $\|\nabla(u - \bar{u}_h)(t)\|_{L^2(\Omega)}$  for the lumped mass method only, since the errors for the Galerkin FEM are almost identical. Below we use the following notation convention:  $\|(u - \bar{u}_h)(t)\|_{\dot{H}^p(\Omega)}$  for  $p = 0$  and  $p = 1$  is simply referred to as  $L^2$ -norm and  $H^1$ -norm error, respectively.

#### 3.3.1 Homogenous problems

We first consider the problem (2.7) on the unit square  $\Omega = (0, 1)^2$  for the homogeneous problem with the following data:

- (a) Smooth initial data:  $v(x, y) = x(1-x)y(1-y)$ ; in this case the initial data  $v$  is in  $H^2(\Omega) \cap H_0^1(\Omega)$ , and the exact solution  $u(x, t)$  can be represented by a rapidly converging Fourier series:

$$u(x, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{4c_n c_m}{m^3 n^3 \pi^6} E_{\alpha,1}(-\lambda_{n,m} t^\alpha) \sin(n\pi x) \sin(m\pi y),$$

where  $\lambda_{n,m} = (n^2 + m^2)\pi^2$ , and  $c_l = 4 \sin^2(l\pi/2) - l\pi \sin(l\pi)$ ,  $l = m, n$ .

- (b) Nonsmooth initial data:  $v(x) = \chi_{[\frac{1}{4}, \frac{3}{4}] \times [\frac{1}{4}, \frac{3}{4}]}$ .
- (c) Very weak data:  $v = \delta_\Gamma$  with  $\Gamma$  being the boundary of the square  $[\frac{1}{4}, \frac{3}{4}] \times [\frac{1}{4}, \frac{3}{4}]$  with  $\langle \delta_\Gamma, \phi \rangle = \int_\Gamma \phi(s) ds$ . One may view  $(v, \chi)$  for  $\chi \in X_h \subset \dot{H}^{\frac{1}{2}+\epsilon}(\Omega)$  as duality pairing between the spaces

$H^{-\frac{1}{2}-\epsilon}(\Omega)$  and  $\dot{H}^{\frac{1}{2}+\epsilon}(\Omega)$  for any  $\epsilon > 0$  so that  $\delta_\Gamma \in H^{-\frac{1}{2}-\epsilon}(\Omega)$ . Indeed, it follows from Hölder's inequality

$$\|\delta_\Gamma\|_{H^{-\frac{1}{2}-\epsilon}(\Omega)} = \sup_{\phi \in \dot{H}^{\frac{1}{2}+\epsilon}(\Omega)} \frac{|\int_\Gamma \phi(s) ds|}{\|\phi\|_{\frac{1}{2}+\epsilon, \Omega}} \leq |\Gamma|^{\frac{1}{2}} \sup_{\phi \in \dot{H}^{\frac{1}{2}+\epsilon}(\Omega)} \frac{\|\phi\|_{L^2(\Gamma)}}{\|\phi\|_{\frac{1}{2}+\epsilon, \Omega}},$$

and the continuity of the trace operator from  $\dot{H}^{\frac{1}{2}+\epsilon}(\Omega)$  to  $L^2(\Gamma)$ .

To discretize the problem, we divide  $(0, 1)$  into  $N = 2^k$  equally spaced subintervals with a mesh size  $h = 1/N$  so that unit square  $(0, 1)^2$  is divided into  $N^2$  small squares. We get a symmetric triangulation of the domain  $(0, 1)^2$  by connecting the diagonal of each small square. Therefore, the lumped mass method and standard Galerkin method have the same convergence rates. On these meshes,  $\bar{\lambda}_{n,m}^h$  and  $\bar{\varphi}_{n,m}^h$ ,  $1 \leq n, m \leq N - 1$ , i.e., eigenvalues and eigenvectors of the discrete Laplacian  $\bar{\Delta}_h$ , are explicitly given by:

$$\bar{\lambda}_{n,m}^h = \frac{4}{h^2} \left( \sin^2 \frac{n\pi h}{2} + \sin^2 \frac{m\pi h}{2} \right), \quad \bar{\varphi}_{n,m}^h(x_i, y_j) = 2 \sin(n\pi x_i) \sin(m\pi y_j).$$

respectively, where  $(x_i, y_j)$ ,  $i, j = 1, \dots, N - 1$ , is a mesh point. Then the semidiscrete approximation  $\bar{u}_h$  can be computed via the explicit representation (3.6). To accurately evaluate the Mittag-Leffler functions, we employ the algorithm developed in [70].

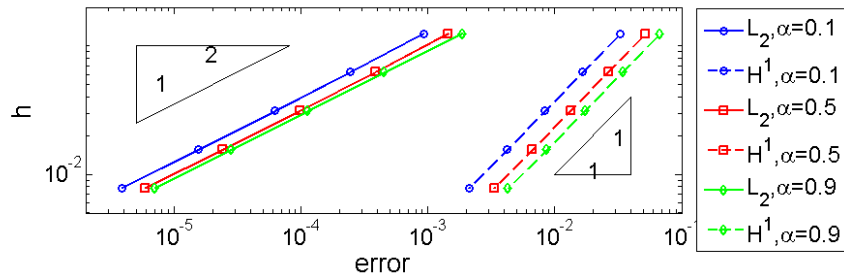


Figure 3.1: Error plots for smooth initial data, Example (a):  $\alpha = 0.1, 0.5, 0.9$  at  $t = 0.1$ .

**Smooth initial data: example (a).** In Table 3.1 we show the numerical results for  $t = 0.1$  and  $\alpha = 0.1, 0.5, 0.9$ , where **rate** refers to the empirical convergence rate as the mesh size  $h$  is halved. In

Figure 3.1, we plot the results from Table 3.1 in a log-log scale. The slopes of the error curves are 2 and 1, respectively, for  $L^2$ - and  $H^1$ -norm of the error. This confirms the theoretical result from Theorem 3.2.1.

Table 3.1: Numerical results for example (a) at  $t = 0.1$ , with  $\alpha = 0.1, 0.5$  and  $0.9$ , discretized on a uniform mesh,  $h = 2^{-k}$ .

$\alpha$	$k$	3	4	5	6	7	rate
0.1	$L^2$ -norm	9.25e-4	2.44e-4	6.25e-5	1.56e-5	3.85e-6	$\approx 2.01$ (2.00)
	$H^1$ -norm	3.27e-2	1.66e-2	8.40e-3	4.21e-3	2.11e-3	$\approx 1.00$ (1.00)
0.5	$L^2$ -norm	1.45e-3	3.84e-4	9.78e-5	2.41e-5	5.93e-6	$\approx 2.01$ (2.00)
	$H^1$ -norm	5.17e-2	2.64e-2	1.33e-2	6.67e-3	3.33e-3	$\approx 1.00$ (1.00)
0.9	$L^2$ -norm	1.88e-3	4.53e-4	1.13e-4	2.82e-5	7.06e-6	$\approx 2.00$ (2.00)
	$H^1$ -norm	6.79e-2	3.43e-2	1.73e-2	8.63e-3	4.31e-3	$\approx 1.00$ (1.00)

**Nonsmooth initial data: example (b).** In Table 3.2 we present the numerical results for problem (b). Here we are particularly interested in errors for  $t$  close to zero, and thus we also present the error at  $t = 0.001$  and  $t = 0.01$ . These numerical results fully confirm the theoretically predicted rates for nonsmooth data.

Table 3.2: Numerical results for example (b),  $\alpha = 0.5$ , at  $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh,  $h = 2^{-k}$ .

Time	$k$	3	4	5	6	7	rate
$t = 0.001$	$L^2$ -norm	1.55e-2	3.99e-3	1.00e-3	2.52e-4	6.26e-5	$\approx 2.01$ (2.00)
	$H^1$ -norm	6.05e-1	3.05e-1	1.48e-1	7.29e-2	3.61e-2	$\approx 1.00$ (1.00)
$t = 0.01$	$L^2$ -norm	8.27e-3	2.10e-3	5.28e-4	1.32e-4	3.29e-5	$\approx 2.01$ (2.00)
	$H^1$ -norm	3.32e-1	1.61e-1	7.90e-2	3.90e-2	1.93e-2	$\approx 1.01$ (1.00)
$t = 0.1$	$L^2$ -norm	2.12e-3	5.36e-4	1.34e-4	3.36e-5	8.43e-6	$\approx 2.00$ (2.00)
	$H^1$ -norm	8.23e-2	4.01e-2	1.96e-2	9.72e-3	4.84e-3	$\approx 1.00$ (1.00)

Very weak data: example (c). The empirical convergence rate for the weak data  $\delta_\Gamma$  agrees well with the theoretically predicted convergence rate in Remark 3.2.6, which gives a ratio of 2.82 and 1.41, respectively, for the  $L^2$ - and  $H^1$ -norm of the error; see Table 3.4. Interestingly, for the standard Galerkin scheme, the  $L^2$ -norm of the error exhibits super-convergence; see Table 3.3.

Table 3.3: Numerical results, i.e., errors  $\|u(t) - u_h(t)\|_{\dot{H}^p(\Omega)}$ ,  $p = 0, 1$ , for example (c),  $\alpha = 0.5$ , at  $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh,  $h = 2^{-k}$ .

Time	$k$	1/8	1/16	1/32	1/64	1/128	rate
$t = 0.001$	$L^2$ -norm	5.37e-2	1.56e-2	4.40e-3	1.23e-3	3.41e-4	$\approx 1.84$ (1.50)
	$H^1$ -norm	2.68e0	1.76e0	1.20e0	8.21e-1	5.68e-1	$\approx 0.53$ (0.5)
$t = 0.01$	$L^2$ -norm	2.26e-2	6.20e-3	1.67e-3	4.46e-4	1.19e-4	$\approx 1.90$ (1.50)
	$H^1$ -norm	9.36e-1	5.90e-1	3.92e-1	2.65e-1	1.84e-1	$\approx 0.52$ (0.5)
$t = 0.1$	$L^2$ -norm	8.33e-3	2.23e-3	5.90e-3	1.55e-3	4.10e-4	$\approx 1.91$ (1.50)
	$H^1$ -norm	3.08e-1	1.91e-1	1.26e-1	8.44e-2	5.83e-2	$\approx 0.53$ (0.5)



Table 3.4: Numerical results, i.e., errors  $\|u(t) - \bar{u}_h(t)\|_{\dot{H}^p(\Omega)}$ ,  $p = 0, 1$ , for example (c),  $\alpha = 0.5$ , at  $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh,  $h = 2^{-k}$

Time	$k$	3	4	5	6	7	rate
$t = 0.001$	$L^2$ -norm	1.98e-1	7.95e-2	3.00e-2	1.09e-2	3.95e-3	$\approx 1.51$ (1.50)
	$H^1$ -norm	5.56e0	4.06e0	2.83e0	2.02e0	1.41e0	$\approx 0.50$ (0.50)
$t = 0.01$	$L^2$ -norm	6.61e-2	2.56e-2	9.51e-3	3.47e-3	1.25e-3	$\approx 1.52$ (1.50)
	$H^1$ -norm	1.84e0	1.30e0	9.10e-1	6.40e-1	4.47e-1	$\approx 0.50$ (0.50)
$t = 0.1$	$L^2$ -norm	2.15e-2	8.13e-3	3.01e-3	1.09e-3	3.95e-4	$\approx 1.52$ (1.50)
	$H^1$ -norm	5.87e-1	4.14e-1	2.88e-1	2.03e-1	1.41e-1	$\approx 0.50$ (0.50)

### 3.3.2 Inhomogenous problems

Now we consider the problem (2.7) on the unit square  $\Omega = (0, 1)^2$  for with zero initial data and the following source data:

(d) Nonsmooth data:  $f(x, t) = (\chi_{[1/2, 1]}(t) + 1)\chi_{[1/4, 3/4] \times [1/4, 3/4]}$ .

(e) Very weak data:  $f(x, t) = (\chi_{[1/2, 1]}(t) + 1)\delta_\Gamma$  with  $\Gamma$  being the boundary of the square  $[1/4, 3/4] \times [1/4, 3/4]$  with  $\langle \delta_\Gamma, \phi \rangle = \int_\Gamma \phi(s) ds$ . One may view  $(v, \chi)$  for  $\chi \in X_h \subset \dot{H}^{1/2+\epsilon}(\Omega)$  as duality pairing between the spaces  $H^{-1/2-\epsilon}(\Omega)$  and  $\dot{H}^{1/2+\epsilon}(\Omega)$  for any small  $\epsilon > 0$  so that  $\delta_\Gamma \in H^{-1/2-\epsilon}(\Omega)$ .

**Numerical results for example (d)** In this example the right hand side  $f(x, t)$  is in the space  $L^\infty(0, 1; \dot{H}^{1/2-\epsilon}(\Omega))$  for any small  $\epsilon > 0$  and the numerical results were computed at  $t = 1$  for  $\alpha = 0.1, 0.5$  and  $0.95$ ; see Table 3.5. The slopes of the error curves in a log-log scale are 2 and 1 for  $L^2(\Omega)$ - and  $H^1(\Omega)$ -norm, respectively, which agrees well with the theoretical results for the nonsmooth case. We observe that the convergence rate is independent of  $\alpha$  value.

Table 3.5: Numerical results for example (d) at  $t = 1$ , with  $\alpha = 0.1, 0.5$  and  $0.9$ , discretized on a uniform mesh,  $h = 2^{-k}$ .

$\alpha$	$k$	3	4	5	6	7	rate
$\alpha = 0.1$	$L^2$ -norm	9.66e-4	2.48e-4	6.26e-5	1.57e-5	3.93e-6	$\approx 1.99$ (2.00)
	$H^1$ -norm	2.06e-2	1.04e-2	5.24e-3	2.63e-3	1.31e-3	$\approx 0.99$ (1.00)
$\alpha = 0.5$	$L^2$ -norm	9.82e-4	2.52e-4	6.36e-5	1.59e-5	3.99e-6	$\approx 1.99$ (2.00)
	$H^1$ -norm	2.10e-2	1.07e-2	5.35e-3	2.68e-3	1.34e-3	$\approx 0.99$ (1.00)
$\alpha = 0.95$	$L^2$ -norm	9.82e-4	2.52e-4	6.36e-5	1.61e-5	4.02e-6	$\approx 1.99$ (2.00)
	$H^1$ -norm	2.13e-2	1.08e-2	5.42e-3	2.71e-3	1.36e-3	$\approx 0.99$ (1.00)

**Numerical results for example (e)** In Table 3.6, we present the  $L^2(\Omega)$ - and  $H^1(\Omega)$ -norms of the error for this example. The  $H^1(\Omega)$ -norm of the error decays at the theoretical rate, however the  $L^2(\Omega)$ -norm of the error exhibits better convergence. This might be attributed to the fact that the boundary  $\Gamma$  is fully aligned with element edges. In contrast, if we choose  $P_h f$  as the discrete right hand side for the lumped mass method instead of  $\bar{P}_h f$ , then the  $L^2(\Omega)$ -norm of the error converges only at the standard order; see Table 3.7.

Table 3.6: Numerical results for example (e) at  $t = 1$ , discretized on a uniform mesh,  $h = 2^{-k}$ .

$\alpha$	$k$	3	4	5	6	7	rate
$\alpha = 0.1$	$L^2$ -norm	4.61e-3	1.25e-3	3.31e-4	8.68e-5	2.39e-5	$\approx 1.90$ (1.50)
	$H^1$ -norm	1.60e-1	9.92e-2	6.43e-2	4.43e-2	3.16e-2	$\approx 0.58$ (0.50)
$\alpha = 0.5$	$L^2$ -norm	4.67e-3	1.26e-3	3.34e-4	8.76e-5	2.40e-5	$\approx 1.91$ (1.50)
	$H^1$ -norm	1.60e-1	9.92e-2	6.44e-2	4.50e-2	3.17e-2	$\approx 0.58$ (0.50)
$\alpha = 0.95$	$L^2$ -norm	4.70e-3	1.27e-3	3.36e-4	8.81e-5	2.42e-5	$\approx 1.91$ (1.50)
	$H^1$ -norm	1.61e-1	9.98e-2	6.46e-2	4.50e-2	3.17e-2	$\approx 0.58$ (0.50)

Table 3.7: Numerical results for example (e) at  $t = 1$ , with  $f_h = P_h f$ , discretized on a uniform mesh,  $h = 2^{-k}$ .

$\alpha$	$k$	3	4	5	6	7	rate
$\alpha = 0.5$	$L^2$ -norm	1.19e-2	4.55e-3	1.69e-3	6.15e-4	2.22e-4	$\approx 1.44$ (1.50)
	$H^1$ -norm	3.28e-1	2.32e-1	1.67e-1	1.13e-1	8.21e-2	$\approx 0.50$ (0.50)

### 3.4 Extension to multi-term fractional diffusion

In this section, we develop a semidiscrete scheme and extend the error analysis in Section 3.1 to the multi-term counterpart (2.15). First we describe the semidiscrete scheme, and then derive almost optimal error estimates for the homogeneous and inhomogeneous problems separately. Analogous to (3.3), we may write the spatially discrete scheme as

$$P(\partial_t)u_h(t) - \Delta_h u_h(t) = f_h(t) \quad \text{for } t \geq 0 \quad \text{with } u_h(0) = v_h, \quad (3.48)$$

where  $f_h = P_h f$  and  $v_h$  is the Ritz or  $L^2$ -approximation of  $v$ . This represents a system of fractional ordinary differential equations.

#### 3.4.1 Error analysis

Spatial semidiscrete schemes. Like before, we give a solution representation of (3.48) using the eigenvalues and eigenfunctions  $\{\lambda_j^h\}_{j=1}^N$  and  $\{\varphi_j^h\}_{j=1}^N$  of the discrete Laplacian  $-\Delta_h$ . Next we introduce the operators  $E_h$  and  $\bar{E}_h$ , the discrete analogues of (2.17) and (2.18), for  $t > 0$ , defined respectively by

$$\begin{aligned} E_h(t)v_h &= \sum_{j=1}^N E_{\bar{\alpha},1}(-\lambda_j^h t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})(v, \varphi_j^h) \varphi_j^h \\ &+ \sum_{i=1}^m b_i t^{\alpha-\alpha_i} \sum_{j=1}^N E_{\bar{\alpha},1+\alpha-\alpha_i}(-\lambda_j^h t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})(v, \varphi_j^h) \varphi_j^h, \end{aligned} \quad (3.49)$$

and

$$\bar{E}_h(t)f_h = \sum_{j=1}^N t^{\alpha-1} E_{\bar{\alpha},\alpha}(-\lambda_j^h t^\alpha, -b_1 t^{\alpha-\alpha_1}, \dots, -b_m t^{\alpha-\alpha_m})(f_h, \varphi_j^h) \varphi_j^h. \quad (3.50)$$

Then the solution  $u_h$  of the discrete problem (3.48) can be expressed by:

$$u_h(t) = E_h(t)v_h + \int_0^t \bar{E}_h(t-s)f_h(s) ds. \quad (3.51)$$

**Lemma 3.4.1.** *Assume that the mesh  $\mathcal{T}_h$  is quasi-uniform. Then for any  $v_h \in X_h$  the function  $u_h(t) = E_h(t)v_h$  satisfies*

$$\| (P(\partial_t))^\ell u_h(t) \|_{\dot{H}^p(\Omega)} \leq C t^{-\alpha(\ell+(p-q)/2)} \| v_h \|_{\dot{H}^q(\Omega)}, \quad t > 0,$$

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

*Proof.* Upon noting  $\| P(\partial_t)E_h(t)v_h \|_{\dot{H}^p(\Omega)} = \| E_h(t)v_h \|_{\dot{H}^{p+2}(\Omega)}$ , it suffices to show the case  $\ell = 0$ .

Using the representation (3.51) and Lemma 2.2.2, we have for  $0 \leq p - q \leq 2$

$$\begin{aligned} \| E_h(t)v_h \|_{\dot{H}^p(\Omega)}^2 &\leq C \sum_{j=1}^N \frac{(\lambda_j^h)^p}{(1 + \lambda_j^h t^\alpha)^2} |(v_h, \varphi_j^h)|^2 \\ &\leq C t^{-(p-q)\alpha} \sum_{j=1}^N \frac{(\lambda_j^h t^\alpha)^{p-q}}{(1 + \lambda_j^h t^\alpha)^2} (\lambda_j^h)^q |(v_h, \varphi_j^h)|^2 \\ &\leq C t^{-(p-q)\alpha} \| v_h \|_{\dot{H}^q(\Omega)}^2, \end{aligned}$$

where the last inequality follows from  $\sup_{1 \leq j \leq N} \frac{(\lambda_j^h t^\alpha)^{p-q}}{(1 + \lambda_j^h t^\alpha)^2} \leq C$  for  $0 \leq p - q \leq 2$ .  $\square$

The next result is a discrete analogue to Lemma 2.4.1.

**Lemma 3.4.2.** *Let  $\bar{E}_h$  be defined by (3.50) and  $\chi \in X_h$ . Then for all  $t > 0$*

$$\| \bar{E}_h(t)\chi \|_{\dot{H}^p(\Omega)} \leq \begin{cases} C t^{-1+\alpha(1+(q-p)/2)} \| \chi \|_{\dot{H}^q(\Omega)}, & 0 \leq p - q \leq 2, \\ C t^{-1+\alpha} \| \chi \|_{\dot{H}^q(\Omega)}, & p < q. \end{cases}$$

*Proof.* The proof for the case  $0 \leq p - q \leq 2$  is similar to Lemma 2.4.1. The other assertion follows from the fact that  $\{\lambda_j^h\}_{j=1}^N$  are bounded from zero independent of  $h$ .  $\square$

**3.4.1.1 Error estimates for homogeneous problems.** We first consider the case of smooth initial data, i.e.,  $v \in \dot{H}^2(\Omega)$ , and derive error estimates. To this end, we split the error  $u_h(t) - u(t)$  into two terms:

$$u_h - u = (u_h - R_h u) + (R_h u - u) := \vartheta + \varrho.$$

By Lemma 3.1.1 and Theorem 2.4.1, we have for any  $t > 0$

$$\|\varrho(t)\|_{L^2(\Omega)} + h\|\nabla\varrho(t)\|_{L^2(\Omega)} \leq Ch^2\|v\|_{\dot{H}^2(\Omega)}. \quad (3.52)$$

So it suffices to get proper estimates for  $\vartheta(t)$ , which is given below. The proof is identical to that of Lemma 3.1.6 and hence omitted.

**Lemma 3.4.3.** *The function  $\vartheta(t) := u_h(t) - R_h u(t)$  satisfies for  $p = 0, 1$*

$$\|\vartheta(t)\|_{\dot{H}^p(\Omega)} \leq Ch^{2-p}\|v\|_{\dot{H}^2(\Omega)}.$$

Using (3.52), Lemma 3.4.3 and the triangle inequality, we arrive at our first estimate, which is formulated in the following Theorem:

**Theorem 3.4.1.** *Let  $v \in \dot{H}^2(\Omega)$  and  $f \equiv 0$ , and  $u$  and  $u_h$  be the solutions of (2.15) with  $v \in \dot{H}^q(\Omega)$  and (3.48) with  $v_h = R_h v$ , respectively. Then*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^2\|v\|_{\dot{H}^2(\Omega)}.$$

Now we turn to the nonsmooth case, i.e.,  $v \in \dot{H}^q(\Omega)$  with  $-1 < q \leq 1$ . Since the Ritz projection  $R_h$  is not well-defined for nonsmooth data, we use instead the  $L^2(\Omega)$ -projection  $v_h = P_h v$  and split the error  $u_h - u$  into:

$$u_h - u = (u_h - P_h u) + (P_h u - u) := \tilde{\vartheta} + \tilde{\varrho}. \quad (3.53)$$

By Lemma 3.1.1 and Theorem 2.4.1 we have for  $-1 < q \leq 1$

$$\begin{aligned} \|\tilde{\varrho}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\varrho}(t)\|_{L^2(\Omega)} &\leq Ch^{2+\min(0,q)}\|u(t)\|_{\dot{H}^{2+\min(0,q)}(\Omega)} \\ &\leq Ch^{2+\min(0,q)}t^{-\alpha(1-\max(q/2,0))}\|v\|_{\dot{H}^q(\Omega)}. \end{aligned}$$

Thus, we only need to estimate the term  $\tilde{\vartheta}(t)$ , which is stated in the following lemma and is an analogue to Lemma 3.1.9.

**Lemma 3.4.4.** *Let  $\tilde{\vartheta}(t) = u_h(t) - P_h u(t)$ . Then for  $p = 0, 1$ ,  $-1 < q \leq 1$ , there holds (with  $\ell_h = |\ln h|$ )*

$$\|\tilde{\vartheta}(t)\|_{\dot{H}^p(\Omega)} \leq Ch^{\min(q,0)+2-p}\ell_h t^{-\alpha(1-\max(q/2,0))}\|v\|_{\dot{H}^q(\Omega)}.$$

Now the triangle inequality yields an error estimate for nonsmooth initial data.

**Theorem 3.4.2.** *Let  $f \equiv 0$ ,  $u$  and  $u_h$  be the solutions of (2.15) with  $v \in \dot{H}^q(\Omega)$ ,  $-1 < q \leq 1$ , and (3.48) with  $v_h = P_h v$ , respectively. Then with  $\ell_h = |\ln h|$ , there holds*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^{\min(q,0)+2} \ell_h t^{-\alpha(1-\max(q/2,0))} \|v\|_{\dot{H}^q(\Omega)}.$$

**Remark 3.4.1.** *The log factor can be removed by the Fujita and Suzuki's technique analogue to Theorem 3.1.3.*

**3.4.1.2 Error estimates for inhomogeneous problems.** Now we derive error estimates for the semidiscrete Galerkin approximation of the inhomogeneous problem with  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 0$ , and  $v \equiv 0$ , in  $L^\infty$ -norm in time. To this end, we appeal again to the splitting (3.53). By Theorem 2.4.2 and Lemma 3.1.1, the following estimate holds for  $\tilde{\varrho}$ :

$$\|\tilde{\varrho}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\varrho}(t)\|_{L^2(\Omega)} \leq Ch^{2+q-\epsilon}t^{-\epsilon\alpha}\|u(t)\|_{\dot{H}^{2+q-\epsilon}(\Omega)} \leq C\epsilon^{-1}h^{2+q-\epsilon}\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))},$$

where the last inequality follows from the fact  $t \leq T$ , and  $t^{\epsilon\alpha}$  is bounded. Now the choice  $\ell_h = |\ln h|$ ,  $\epsilon = 1/\ell_h$ , yields

$$\|\tilde{\varrho}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\varrho}(t)\|_{L^2(\Omega)} \leq C\ell_h h^{2+q}\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}. \quad (3.54)$$

Thus, it suffices to bound the term  $\tilde{\vartheta}$ ; see the lemma below.

**Lemma 3.4.5.** *Let  $\tilde{\vartheta}(t)$  be defined by (3.53), and  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 0$ . Then with  $\ell_h = |\ln h|$ , there holds*

$$\|\tilde{\vartheta}(t)\|_{L^2(\Omega)} + h\|\nabla\tilde{\vartheta}(t)\|_{L^2(\Omega)} \leq Ch^{2+q}\ell_h^2\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

An inspection of the proof of Lemma 3.4.5 indicates that for  $0 < q < 1$ , one can get rid of one factor  $\ell_h$ . Now we can state an error estimate in  $L^\infty$ -norm in time.

**Theorem 3.4.3.** *Let  $v \equiv 0$ ,  $f \in L^\infty(0, T; \dot{H}^q(\Omega))$ ,  $-1 < q \leq 0$ , and  $u$  and  $u_h$  be the solutions of (2.15) and (3.48) with  $f_h = P_h f$ , respectively. Then with  $\ell_h = |\ln h|$  and  $t > 0$ , there holds*

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} + h\|\nabla(u_h(t) - u(t))\|_{L^2(\Omega)} \leq Ch^{2+q}\ell_h^2\|f\|_{L^\infty(0,t;\dot{H}^q(\Omega))}.$$

### 3.4.2 Numerical results

In this part we present two-dimensional numerical experiments on the unit square  $\Omega = (0, 1)^2$  to verify the error estimates in Sections 3.4.1. We consider the following datas:

- (a) Nonsmooth initial data:  $v = \chi_{(0,1/2) \times (0,1)}$  and  $f \equiv 0$ .
- (b) Very weak initial data:  $v = \delta_\Gamma$  with  $\Gamma$  being the union of  $\{1/4\} \times [1/4, 3/4] \cup [1/4, 3/4] \times \{3/4\}$  clockwise and  $[1/4, 3/4] \times \{1/4\} \cup \{3/4\} \times [1/4, 3/4]$  counterclockwise. The duality is defined

by  $\langle \delta_\Gamma, \phi \rangle = \int_\Gamma \phi(s) ds$ . By Hölder's inequality and the continuity of the trace operator from  $\dot{H}^{1/2+\epsilon}(\Omega)$  to  $L^2(\Gamma)$  [3], we deduce  $\delta_\Gamma \in H^{-1/2-\epsilon}(\Omega)$ .

(c) Nonsmooth right hand side:  $f(x, t) = (\chi_{[1/20, 1/10]}(t) + 1)\chi_{(0, 1/2) \times (0, 1)}(x)$  and  $v \equiv 0$ .

To discretize the problem, we divide each direction into  $N = 2^k$  equally spaced subintervals, with a mesh size  $h = 1/N$  so that the domain  $(0, 1)^2$  is divided into  $N^2$  small squares. We get a symmetric mesh by connecting the diagonal of each small square. In order to check the convergence rate of the semidiscrete scheme, we discretize the fractional derivatives using the L1 scheme (4.8) with a small time step  $\tau$  so that the temporal discretization error is negligible.

The numerical results for example (a) are shown in Table 3.8, which agree well with Theorem 3.4.2, with a rate  $O(h^2)$  and  $O(h)$ , respectively, for the  $L^2$ - and  $H^1$ -norm of the error. Interestingly, for example (b), both the  $L^2$ -norm and  $H^1$ -norm of the error exhibit super-convergence, cf. Table 3.9. The numerical results for example (c) confirm the theoretical results; see Table 3.10. The solution profiles for examples (b) and (c) at  $t = 0.1$  are shown in Fig. 3.2, from which the nonsmooth region of the solution can be clearly observed.

Table 3.8: Numerical results for (a) with  $\alpha = 0.5$  and  $\beta = 0.2$  at  $t = 0.1, 0.01, 0.001$ , discretized on a uniform mesh,  $h = 2^{-k}$  and  $\tau = t/10^4$ .

$t$	$k$	3	4	5	6	7	rate
$t = 0.1$	$L^2$ -norm	5.25e-3	1.35e-3	3.38e-4	8.24e-5	1.98e-5	$\approx 2.06$ (2.00)
	$H^1$ -norm	9.10e-2	4.53e-2	2.25e-2	1.09e-2	4.99e-3	$\approx 1.04$ (1.00)
$t = 0.01$	$L^2$ -norm	1.25e-2	3.23e-3	8.09e-4	1.97e-4	4.65e-5	$\approx 2.05$ (2.00)
	$H^1$ -norm	2.18e-1	1.08e-1	5.35e-2	2.62e-2	1.27e-2	$\approx 1.05$ (1.00)
$t = 0.001$	$L^2$ -norm	3.02e-2	7.84e-3	1.97e-3	4.81e-4	1.16e-4	$\approx 2.03$ (2.00)
	$H^1$ -norm	5.30e-1	2.64e-1	1.31e-1	6.38e-2	3.14e-2	$\approx 1.04$ (1.00)



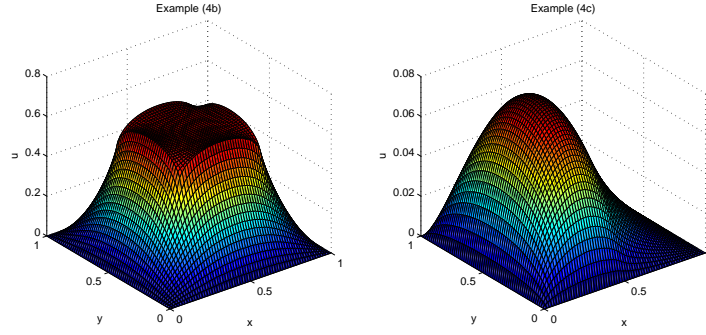


Figure 3.2: Numerical solutions of examples (b) and (c) with  $h = 2^{-6}$ ,  $\alpha = 0.5$ ,  $\beta = 0.2$  at  $t = 0.1$

Table 3.9: Numerical results for example (b) with  $\alpha = 0.5$  and  $\beta = 0.2$  at  $t = 0.1, 0.01, 0.001$  for a uniform mesh with  $h = 2^{-k}$  and  $\tau = t/10^4$ .

$t$	$k$	3	4	5	6	7	rate
$t = 0.1$	$L^2$ -norm	1.18e-2	3.18e-3	8.41e-4	2.18e-4	5.41e-5	$\approx 1.92$ (1.50)
	$H^1$ -norm	2.25e-1	1.13e-1	6.60e-2	3.40e-2	1.66e-2	$\approx 0.92$ (0.50)
$t = 0.01$	$L^2$ -norm	2.82e-2	7.62e-3	2.28e-3	5.26e-4	1.25e-4	$\approx 1.95$ (2.00)
	$H^1$ -norm	5.66e-1	3.09e-1	1.65e-1	8.52e-2	4.19e-2	$\approx 0.94$ (1.00)
$t = 0.001$	$L^2$ -norm	6.65e-2	1.83e-3	4.98e-3	1.33e-3	3.30e-4	$\approx 1.91$ (2.00)
	$H^1$ -norm	1.66e0	8.93e-1	4.75e-1	2.43e-1	1.21e-1	$\approx 0.95$ (1.00)

Table 3.10: Numerical results for example (c) with  $\alpha = 0.5$  and  $\beta = 0.2$  at  $t = 0.1, 0.01, 0.001$  for a uniform mesh with  $h = 2^{-k}$  and  $\tau = t/10^4$ .

$t$	$k$	3	4	5	6	7	rate
$t = 0.1$	$L^2$ -norm	2.28e-3	5.86e-4	1.47e-4	3.58e-5	7.91e-6	$\approx 2.07$ (2.00)
	$H^1$ -norm	3.97e-2	1.97e-2	9.77e-3	4.76e-3	2.13e-3	$\approx 1.06$ (1.00)
$t = 0.01$	$L^2$ -norm	1.06e-3	2.73e-4	6.86e-5	1.67e-6	3.70e-7	$\approx 2.06$ (2.00)
	$H^1$ -norm	1.85e-2	9.18e-3	4.56e-3	2.22e-3	9.94e-3	$\approx 1.06$ (1.00)
$t = 0.001$	$L^2$ -norm	8.66e-4	2.28e-4	5.75e-5	1.40e-6	3.11e-6	$\approx 2.04$ (2.00)
	$H^1$ -norm	1.56e-2	7.82e-3	3.88e-3	1.90e-3	8.47e-4	$\approx 1.05$ (1.00)

### 3.5 Extension to the diffusion-wave equation

Our analysis can be also extended to the diffusion-wave models (2.22) in the domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ . \* Like before, the semidiscrete Galerkin scheme for problem (2.22) reads: find  $u_h(t) \in X_h$  such that for  $t > 0$

$${}^C\partial_t^\alpha u_h(t) + A_h u_h(t) = f_h(t), \quad \text{with } u_h(0) = v_h \quad \text{and} \quad \partial_t u_h(0) = b_h, \quad (3.55)$$

where  $A_h = -\Delta_h$  and  $v_h \in X_h$  and  $b_h \in X_h$  are approximations to the initial data  $v$  and  $b$ , respectively. Following [73], we choose  $v_h \in X_h$  and  $b_h \in X_h$  depending on the smoothness of the data.

#### 3.5.1 Error analysis

Next we derive error estimates for the semidiscrete scheme (3.55). To this end we employ an operator technique developed in [16] due to the lack of smoothing properties and insufficiency of the spectral decomposition method in this case. First we derive an integral representation of the solution. Since the solution  $u : (0, T] \rightarrow L^2(\Omega)$  can be analytically extended to the sector  $\{z \in \mathbb{C}; z \neq 0, |\arg z| < \pi/2\}$  [68], we may apply the Laplace transform to (2.22) to deduce

$$z^\alpha \widehat{u}(z) + A\widehat{u}(z) = z^{\alpha-1}v + z^{\alpha-2}b, \quad (3.56)$$

---

\*The results for the diffusion-wave model given in this section were published in [29].

with  $A = -\Delta$ . Hence the solution  $u(t)$  can be represented by

$$u(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha I + A)^{-1} (z^{\alpha-1}v + z^{\alpha-2}b) dz, \quad (3.57)$$

where the contour  $\Gamma_{\theta,\delta}$  is given by (3.19). Throughout, we choose the angle  $\theta$  such that  $\pi/2 < \theta < \min(\pi, \pi/\alpha)$  and hence  $z^\alpha \in \Sigma_{\theta'}$  with  $\theta' = \alpha\theta < \pi$  for all  $z \in \Sigma_\theta := \{z \in \mathbb{C} : |\arg z| \leq \theta\}$ . Then there exists a constant  $C$  which depend only on  $\theta$  and  $\alpha$  such that

$$\|(z^\alpha I + A)^{-1}\| \leq Cz^{-\alpha}, \quad \forall z \in \Sigma'_\theta \quad \text{with} \quad \Sigma'_\theta = \Sigma_\theta \setminus \{0\}. \quad (3.58)$$

Similarly, the solution  $u_h$  to (3.55) can be represented by

$$u_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta}} e^{zt} (z^\alpha + A_h)^{-1} (z^{\alpha-1}v_h + z^{\alpha-2}b_h) dz. \quad (3.59)$$

The next lemma shows an important error estimate [16, 6].

**Lemma 3.5.1.** *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*

$$\|w_h - w\|_{L^2(\Omega)} + h\|\nabla(w_h - w)\|_{L^2(\Omega)} \leq Ch^2\|\varphi\|_{L^2(\Omega)}. \quad (3.60)$$

Now we can state an error estimate for the scheme (3.55) with nonsmooth initial data.

**Theorem 3.5.1.** *Let  $u$  and  $u_h$  be the solutions of problem (2.22) and (3.55) 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:*

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} + h\|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq Ch^2 (t^{-\alpha}\|v\|_{L^2(\Omega)} + t^{1-\alpha}\|b\|_{L^2(\Omega)}).$$

*Proof.* By (3.57) and (3.59), the error  $e(t) := u(t) - u_h(t)$  can be represented as

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

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 3.1.8 and choosing  $\delta = 1/t$  we have

$$\begin{aligned} \|\nabla e(t)\|_{L^2(\Omega)} &\leq Ch \left( \int_{-\theta}^{\theta} e^{\cos \psi} t^{-\alpha} d\psi + \int_{1/t}^{\infty} e^{rt \cos \theta} \rho^{\alpha-1} d\rho \right) \|v\|_{L^2(\Omega)} \\ &\quad + Ch \left( \int_{-\theta}^{\theta} e^{\cos \psi} t^{1-\alpha} d\psi + \int_{1/t}^{\infty} e^{rt \cos \theta} \rho^{\alpha-2} d\rho \right) \|b\|_{L^2(\Omega)} \\ &\leq Ch (t^{-\alpha} \|v\|_{L^2(\Omega)} + t^{1-\alpha} \|b\|_{L^2(\Omega)}). \end{aligned}$$

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 3.5.2.** *Let  $u$  and  $u_h$  be the solutions of problems (2.22) and (3.55) 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*

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} + h \|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \leq Ch^2 (\|v\|_{\dot{H}^2(\Omega)} + t \|b\|_{\dot{H}^2(\Omega)}). \quad (3.61)$$

*Proof.* Like before, the error  $e(t) := u(t) - u_h(t)$  can be represented as

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

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

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

where  $w^v(z) = (z^\alpha I + A)^{-1} A v$  and  $w_h^v(z) = (z^\alpha I + A_h)^{-1} A_h R_h v$ . Now Lemma 3.1.8 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,

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

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

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

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

**Remark 3.5.1.** For smooth initial data, we may also choose  $v_h = P_h v$  and  $b_h = P_h b$ . Then

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

where  $E$  and  $E_h$  are solution operators. The first term is already bounded in Theorem 3.5.2. By the same argument in Theorem 2.5.1, there holds the boundedness of  $E_h$  and then

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

An estimate on  $b_h$  follows analogously. Hence the error estimate (3.61) holds also for the choice  $v_h = P_h v$  and  $b_h = P_h b$ . By Theorem 3.1.3 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

$$\begin{aligned} \|u(t) - u_h(t)\|_{L^2(\Omega)} + h \|\nabla(u(t) - u_h(t))\|_{L^2(\Omega)} \\ \leq Ch^2 (t^{-\alpha(2-q)/2} \|v\|_{\dot{H}^q(\Omega)} + t^{1-\alpha(2-r)/2} \|b\|_{\dot{H}^r(\Omega)}). \end{aligned}$$

### 3.5.2 Numerical results

We consider the following five examples (with  $\epsilon \in (0, 1/2)$ ):

- (a)  $\Omega = (0, 1)$ ,  $b = 0$ , (a1)  $v = x(1-x) \in \dot{H}^{3/2-\epsilon}(\Omega)$  and (a2)  $v = 1 \in \dot{H}^{1/2-\epsilon}(\Omega)$ .
- (b)  $\Omega = (0, 1)$ ,  $v = 0$ , (b1)  $b = x\chi_{[0,1/2)} + (1-x)\chi_{[1/2,1]} \in \dot{H}^{3/2-\epsilon}(\Omega)$  and (b2)  $b = x^{-1/4} \in \dot{H}^{1/4-\epsilon}(\Omega)$ .

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

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. To observe the spatial error, we apply the SBD (second-order backward difference) time stepping method (from Chapter 5) to discretize the fractional derivative and let set the the time step  $\tau = 10^{-4}$  so that the temporal error is negligible. The fully discrete scheme will be analyzed later in Chapter 5.

**Numerical results for examples (a):** From Figure 3.3 for example (a1) we observe 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 3.5.1. These observations remain valid for nonsmooth data, cf. Table 3.11 and the spatial error deteriorates slightly as  $t \rightarrow 0$ .

It is widely accepted that as the fractional order  $\alpha$  increases from one to two, the model (2.22) transit from the classical diffusion equation to the wave equation [17]. 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 3.4. Further, the closer is the fractional order  $\alpha$  to 2, the slower is the decay of the solution (for  $t$  close to zero).

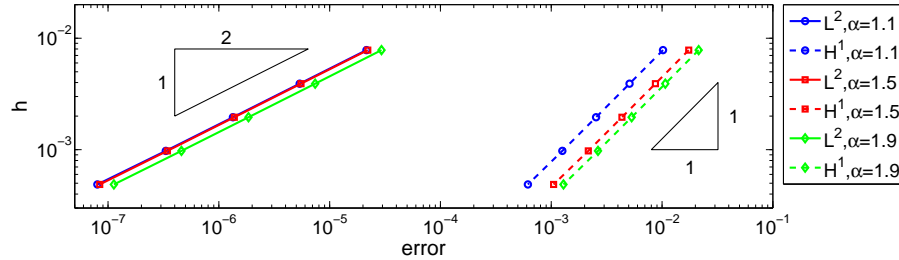


Figure 3.3: Errors of scheme SBD for example (a):  $N = 1000$  and  $t = 0.1$

Finally, in Table 3.12 we show the  $L^2$ -norm of the error for examples (a1) and (a2), for fixed  $h = 2^{-13}$  and  $t \rightarrow 0$ . 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  $t \rightarrow 0$ . This is in excellent

agreement with the theory: in view of Remark 3.5.1, the spatial error deteriorates as  $t \rightarrow 0$  like

$$\|u_h(t) - u(t)\|_{L^2(\Omega)} \leq Ct^{-\alpha(3+2\epsilon)/4} h^2 \|v\|_{\dot{H}^{1/2-\epsilon}(\Omega)} \quad \text{for any } \epsilon > 0.$$

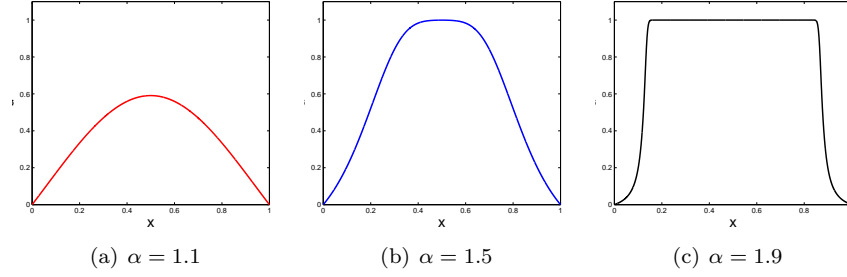


Figure 3.4: Numerical results for example (a2) using SBD method:  $t = 0.1$ ,  $h = 2^{-13}$ ,  $N = 160$ .

Table 3.11: Numerical results example (e):  $\alpha = 1.5$ ,  $h = 2^{-k}$ .

$t$	$k$	7	8	9	10	11	rate
0.1	$L^2$ -norm	1.23e-4	3.08e-5	7.67e-6	1.90e-6	4.53e-7	$\approx 2.01$ (2.00)
	$H^1$ -norm	2.84e-2	1.42e-2	7.07e-3	3.51e-3	1.71e-3	$\approx 1.01$ (1.00)
0.01	$L^2$ -norm	1.58e-3	4.05e-4	1.01e-4	2.51e-5	6.00e-6	$\approx 2.01$ (2.00)
	$H^1$ -norm	3.98e-1	1.92e-1	9.46e-2	4.67e-2	2.29e-2	$\approx 1.02$ (1.00)
0.001	$L^2$ -norm	1.32e-2	4.28e-3	1.28e-3	3.30e-4	7.97e-5	$\approx 1.92$ (2.00)
	$H^1$ -norm	5.72e0	2.84e0	1.37e0	6.42e-1	3.07e-1	$\approx 1.00$ (1.00)

Table 3.12: The  $L^2$ -norm of the error for examples (a1), (a2), (b1) and (b2) with  $\alpha = 1.5$ :  $t \rightarrow 0$ ,  $h = 2^{-13}$ , and  $N = 10^5$ .

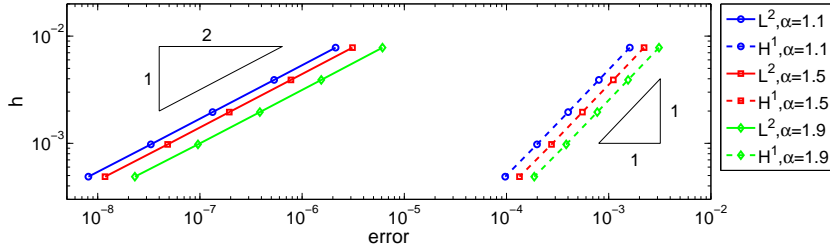
$t$	1	1e-1	1e-2	1e-3	1e-4	1e-5	rate
(a1)	1.31e-9	4.85e-9	5.72e-9	5.86e-9	5.88e-9	5.89e-9	$\approx -0.02$ (0)
(a2)	1.19e-9	2.30e-8	3.04e-7	4.06e-6	5.37e-5	5.66e-4	$\approx -1.12$ (-1.13)
(b1)	3.91e-10	6.62e-10	1.57e-10	3.72e-11	8.81e-12	2.04e-12	$\approx 0.63$ (0.63)
(b2)	4.81e-10	1.84e-9	3.59e-9	7.16e-9	1.43e-8	2.75e-8	$\approx -0.30$ (-0.31)

**Numerical results for examples (b):** Similarly to the results for example (a), we observe a first- and second-order convergence for the  $H^1$ - and  $L^2$ -norm of the error, cf. Figure 3.5. All the convergence rates are independent of the fractional order  $\alpha$ . For the nonsmooth case, i.e., example (b2), we are particularly interested in the errors for  $t$  close to zero, thus we also plot the error at  $t = 0.1, 0.01$  and  $0.001$ . These results fully confirm the analysis in Section 3.5.1. Further, by Remark 3.5.1

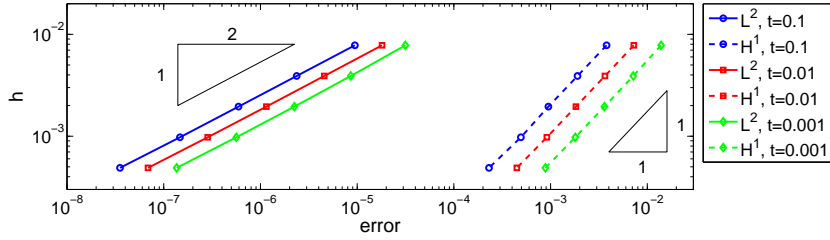
$$\|u(t) - u_h(t)\|_{L^2(\Omega)} \leq Ch^2 t_N^{1-\alpha(2-r)/2} \|b\|_{\dot{H}^r(\Omega)},$$

which is fully confirmed by 3.12.





(a) Error plots for Example (b1) at  $t = 0.1$



(b) Error plots for Example (b2) with  $\alpha = 1.5$

Figure 3.5: Numerical results for diffusion-wave equations using SBD with  $N = 1000$ .

**Numerical results for example (c):** Finally, we present numerical solutions of the two-dimensional example. The errors are showed in Table 3.13. The empirical results fully confirm our analysis. In Figure 3.6, we plot the solution profiles at  $t = 0.01$  and  $t = 0.1$ .

Table 3.13: Numerical results for example (h) with  $\alpha = 1.5$ , at  $t = 0.1$  with  $h = 2^{-k}$ .

$t$	$k$	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	$\approx 2.02$ (2.00)
	$H^1$ -norm	4.45e-1	2.19e-1	1.19e-1	5.39e-2	2.63e-2	$\approx 1.02$ (1.00)
0.01	$L^2$ -norm	1.65e-2	5.74e-3	1.47e-3	3.80e-4	9.26e-5	$\approx 1.99$ (2.00)
	$H^1$ -norm	5.14e-1	3.32e-1	1.59e-1	7.86e-2	3.81e-2	$\approx 1.02$ (1.00)

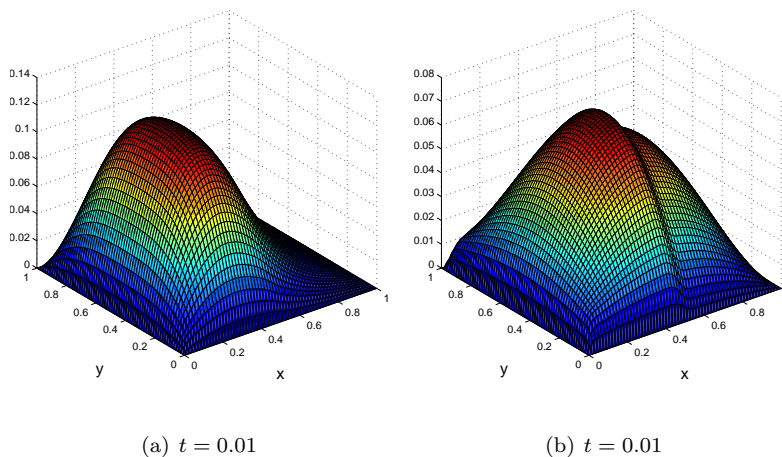


Figure 3.6: Numerical solutions of examples (h) with  $h = 2^{-6}$  and  $N = 1000$ ,  $\alpha = 1.5$  at  $t = 0.1$  and  $t = 0.01$ .

### 3.6 Conclusion and comments

In this chapter, we consider two spatial semidiscrete schemes for fractional diffusion model (2.7) using the standard Galerkin FEM and lumped mass FEM methods. The error estimate for homogeneous problems based on spectral representation is from [28], while the improved error bound in case of nonsmooth data relies on the argument by Fujita and Suzuki [16]; see also [26, 6]. The discussion on much weaker initial data, e.g.  $v \in \dot{H}^{-q}(\Omega)$  with  $q \in (0, 1)$  was proposed in [25]. Further, results for inhomogeneous problem were first discussed in [27]. The extension to the multi-term fractional diffusions and diffusion-wave equations were established in [23] and [29] respectively. For related works, we refer to [63, 55] for the discussion on fractional diffusions with a Riemann-Liouville type fractional derivative and [48, 56, 57] for evolution equations with a positive memory term.

#### 4. FULLY DISCRETE SCHEME BY L1 APPROXIMATION\*

In this chapter, we consider the L1 time-stepping scheme for the fractional diffusion model (2.7). It was shown in [45, equation (3.3)] (see also [72, Lemma 4.1]) that the local truncation error of the L1 approximation is bounded by  $c\tau^{2-\alpha}$  for some constant  $c$  depending only on  $u$ , provided that the solution  $u$  is twice continuously differentiable in time. Since its first appearance, the L1 scheme has been extensively used in practice, and currently it is one of the most popular and successful numerical methods for (2.7), including the case of nonsmooth data arising in inverse problems (see, e.g., [31]).

However, numerical experiments indicates that the  $O(\tau^{2-\alpha})$  convergence rate actually does not hold even for smooth initial data  $v$ , see Table 4.1. This is due to the lack of smoothness of the solution especially for  $t$  close to 0, cf. (1.6). The goal of this chapter is to fill the gap between the existing convergence theory and the numerical experiments, namely, establishing of optimal error bounds that are expressed directly in terms of the regularity of the problem data.

The rest of the chapter is organized as follows. In Section 4.1, we first recall preliminaries on the semidiscrete scheme in Chapter 3, and derive the solution representation for both semidiscrete and fully discrete schemes, which play an important role in the error analysis. The full technical details of the convergence analysis are presented in Subsections 4.1.1 and 4.1.2. The analysis can be applied to some generalized cases, such as replacing  $-\Delta$  by a general sectorial operator or a multi-term fractional model (2.15), see Sections 4.2 and 4.3. Numerical results are presented in Section 4.4 to confirm the convergence theory and the robustness of the scheme. Throughout, the notation  $c$ , with or without a subscript, denotes a generic constant, which may differ at different occurrences, but it is always independent of the spatial mesh-size  $h$  and the time step-size  $\tau$ .

##### 4.1 Fully discrete schemes by L1 time stepping

In order to establish an analysis for the time stepping, we first derive a proper solution representation for (2.7). Since the solution  $u : (0, T] \rightarrow L^2(\Omega)$  can be analytically extended to the sector  $\{z \in \mathbb{C}; z \neq 0, |\arg z| < \pi/2\}$  [68, Theorem 2.1], when  $f \equiv 0$ , we may apply the Laplace transform to equation (2.7) to deduce

$$z^\alpha \widehat{u}(z) + A\widehat{u}(z) = z^{\alpha-1}v, \tag{4.1}$$

---

\*The results in Sections 4.1, 4.2 and 4.4 of this chapter is reprinted with permission from "An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data" by Bangti Jin, Raytcho Lazarov and Zhi Zhou, IMA Journal of Numerical Analysis, in press, Copyright [2015] by Oxford University Press.

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

$$u(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} (z^\alpha I + A)^{-1} z^{\alpha-1} v dz, \quad (4.2)$$

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

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

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

$$\|(z^\alpha I + A)^{-1}\| \leq cz^{-\alpha}, \quad \forall z \in \Sigma'_\theta \quad \text{with} \quad \Sigma'_\theta = \Sigma_\theta \setminus \{0\}. \quad (4.3)$$

Now we briefly recall the spatial semidiscrete scheme based on the Galerkin finite element method. Let  $\mathcal{T}_h$  be a shape regular and quasi-uniform triangulation of the domain  $\Omega$  into  $d$ -simplexes, denoted by  $T$ . On the triangulation  $\mathcal{T}_h$  we define a continuous piecewise linear finite element space  $V_h$  by

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

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

$$\begin{aligned} (P_h \varphi, \chi) &= (\varphi, \chi) \quad \forall \chi \in V_h, \\ (\nabla R_h \varphi, \nabla \chi) &= (\nabla \varphi, \nabla \chi) \quad \forall \chi \in V_h, \end{aligned}$$

where  $(\cdot, \cdot)$  denotes the  $L^2(\Omega)$ -inner product. Then the semidiscrete Galerkin scheme for problem (2.7) reads: find  $u_h(t) \in V_h$  such that

$$({}^C \partial_t^\alpha u_h, \chi) + (\nabla u_h, \nabla \chi) = (f, \chi) \quad \forall \chi \in V_h, \quad (4.4)$$

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

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

the spatial semidiscrete scheme (4.4) can be rewritten into

$${}^C\partial_t^\alpha u_h(t) + A_h u_h(t) = f_h(t), \quad t > 0 \quad (4.5)$$

with  $u_h(0) = v_h \in V_h$ ,  $f_h = P_h f$  and  $A_h = -\Delta_h$ . Like before, the solution  $u_h$  to (4.5) with  $f_h \equiv 0$  can be represented by

$$u_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} (z^\alpha + A_h)^{-1} z^{\alpha-1} v_h dz. \quad (4.6)$$

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

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

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

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

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

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

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

$$w_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} K_1(z) v_h dz. \quad (4.7)$$

Now we describe the fully discrete scheme based on the L1 approximation. To this end, we divide the interval  $[0, T]$  into a uniform grid with a time step size  $\tau = T/N$ ,  $N \in \mathbb{N}$ , so that  $0 = t_0 < t_1 < \dots < t_N = T$ , and  $t_n = n\tau$ ,  $n = 0, \dots, N$ . The L1 approximation of the Caputo fractional derivative

${}^C\partial_t^\alpha u(x, t_n)$  is given by [45, Section 3]

$$\begin{aligned}
{}^C\partial_t^\alpha u(x, t_n) &= \frac{1}{\Gamma(1-\alpha)} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \frac{\partial u(x, s)}{\partial s} (t_n - s)^{-\alpha} ds \\
&\approx \frac{1}{\Gamma(1-\alpha)} \sum_{j=0}^{n-1} \frac{u(x, t_{j+1}) - u(x, t_j)}{\tau} \int_{t_j}^{t_{j+1}} (t_n - s)^{-\alpha} ds \\
&= \sum_{j=0}^{n-1} b_j \frac{u(x, t_{n-j}) - u(x, t_{n-j-1})}{\tau^\alpha} \\
&= \tau^{-\alpha} [b_0 u(x, t_n) - b_{n-1} u(x, t_0) + \sum_{j=1}^{n-1} (b_j - b_{j-1}) u(x, t_{n-j})] =: L_1^n(u),
\end{aligned} \tag{4.8}$$

where the weights  $b_j$  are given by

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

Then the fully discrete scheme reads: find  $U_h^n \in V_h$  for  $n = 1, 2, \dots, N$

$$(b_0 I + \tau^\alpha A_h) U_h^n = b_{n-1} U_h^0 + \sum_{j=1}^{n-1} (b_{j-1} - b_j) U_h^{n-j} + \tau^\alpha F_h^n, \tag{4.9}$$

with  $U_h^0 = v_h$  and  $F_h^n = P_h f(t_n)$ . We focus on the homogeneous case, i.e.,  $f \equiv 0$ . Throughout, we denote by

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

the generating function of a sequence  $\{\omega_j\}_{j=0}^{\infty}$ . To analyze the fully discrete scheme (4.9), we first derive a discrete analogue of the solution representation (4.7). The fully discrete solution  $W_h^n := U_h^n - U_h^0$  satisfies the following time-stepping scheme for  $n = 1, 2, \dots, N$

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

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

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

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

$$\begin{aligned}\sum_{n=1}^{\infty} L_1^n(W_h)\xi^n &= \tau^{-\alpha} \sum_{n=1}^{\infty} \left( b_0 W_h^n + \sum_{j=1}^{n-1} (b_j - b_{j-1}) W_h^{n-j} \right) \xi^n \\ &= \tau^{-\alpha} \sum_{n=1}^{\infty} \left( \sum_{j=0}^{n-1} b_j W_h^{n-j} \right) \xi^n - \tau^{-\alpha} \sum_{n=1}^{\infty} \left( \sum_{j=1}^{n-1} b_{j-1} W_h^{n-j} \right) \xi^n \\ &:= I - II.\end{aligned}$$

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

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

Similarly, the second term  $II$  can be written as

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

Hence, we arrive at

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

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

$$\begin{aligned}\tilde{b}(\xi) &= \frac{1}{\Gamma(2-\alpha)} \sum_{j=0}^{\infty} ((j+1)^{1-\alpha} - j^{1-\alpha}) \xi^j \\ &= \frac{1-\xi}{\xi \Gamma(2-\alpha)} \sum_{j=1}^{\infty} j^{1-\alpha} \xi^j = \frac{(1-\xi) \text{Li}_{\alpha-1}(\xi)}{\xi \Gamma(2-\alpha)},\end{aligned}$$

where  $\text{Li}_p(z)$  denotes the polylogarithm function defined by (see [37])

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

The polylogarithm function  $\text{Li}_p(z)$  is well defined for  $|z| < 1$ , and it can be analytically continued to the split complex plane  $\mathbb{C} \setminus [1, \infty)$ ; see [14]. With  $z = 1$ , it recovers the Riemann zeta function

$\zeta(p) = \text{Li}_p(1)$ . Therefore, the fully discrete solution  $\widetilde{W}_h(\xi)$  can be represented by

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

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

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

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

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

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

$$W_h^n = -\frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_{n-1}} \frac{\tau}{1-e^{-z\tau}} \left( \frac{(1-e^{-z\tau})^2}{e^{-z\tau} \tau^\alpha \Gamma(2-\alpha)} \text{Li}_{\alpha-1}(e^{-z\tau}) + A_h \right)^{-1} A_h v_h dz. \quad (4.10)$$

This representation is the basis of the error analysis.

#### 4.1.1 Error estimate for nonsmooth initial data

In this part, we derive optimal error estimates for the fully discrete scheme (4.9) in case of nonsmooth data, i.e.  $v \in L^2(\Omega)$ . The analysis is based on the representations of the semidiscrete and fully discrete solutions, i.e., (4.7) and (4.10). Upon subtracting them, we may write the difference between  $W_h^n$  and  $w_h(t_n)$  as

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

where the terms  $I$  and  $II$  are defined as

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



and

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

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

$$K_1(z) := -z^{-1}(z^\alpha + A_h)^{-1} A_h =: z^{-1} B_1(z) \quad (4.11)$$

and

$$K_2(z) := -\frac{\tau}{1 - e^{-z\tau}} \left( \frac{1 - e^{-z\tau}}{\tau^\alpha} \psi(z\tau) + A_h \right)^{-1} A_h =: \frac{\tau}{1 - e^{-z\tau}} B_2(z), \quad (4.12)$$

respectively, and the auxiliary function  $\psi$  is define by

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

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

$$\begin{aligned} \|K_1(z) - e^{-z\tau} K_2(z)\| &\leq |e^{-z\tau}| \|K_1(z) - K_2(z)\| + |1 - e^{-z\tau}| \|K_1(z)\| \\ &\leq c \|K_1(z) - K_2(z)\| + c|z|\tau \|K_1(z)\| \\ &\leq c \|K_1(z) - K_2(z)\| + c\tau, \end{aligned} \quad (4.13)$$

where the last line follows from the inequality, in view of (4.3):

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

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

**Lemma 4.1.1.** *Let  $\chi(z) = \tau^{-1}(1 - e^{-z\tau})$ . Then for all  $z \in \Gamma_\tau$ , there hold for some  $c_1, c_2 > 0$*

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

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

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

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

$$\begin{aligned} \left| \frac{1 - e^{-z\tau}}{z\tau} \right| &= \left| \frac{e^\xi - 1}{\xi} \right| = \frac{|e^{-\rho \cos\theta} \cos(\rho \sin\theta) - 1 - ie^{-\rho \cos\theta} \sin(\rho \sin\theta)|}{\rho} \\ &\geq \frac{(e^{-2\rho \cos\theta} + 1 - 2e^{-\rho \cos\theta})^{1/2}}{\rho} = \frac{e^{-\rho \cos\theta} - 1}{\rho} \geq -\cos\theta > 0 \end{aligned} \quad (4.14)$$

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

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

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

Then by the trivial inequality  $\|B_1(z)\| \leq c$  and Lemma 4.1.1, we deduce that

$$\begin{aligned} \|K_1(z) - K_2(z)\| &\leq |z^{-1} - \chi(z)^{-1}| \|B_1(z)\| + |\chi(z)^{-1}| \|B_1(z) - B_2(z)\| \\ &\leq \frac{|z - \chi(z)|}{|z\chi(z)|} + c|z|^{-1} \|B_1(z) - B_2(z)\| \\ &\leq c\tau + c|z|^{-1} \|B_1(z) - B_2(z)\|. \end{aligned} \quad (4.15)$$

Thus it suffices to establish a bound on  $\|B_1(z) - B_2(z)\|$ . This will be done using a series of lemmas. To this end, first we recall an important singular expansion of the function  $\text{Li}_p(e^{-z})$  [14, Theorem 1].

**Lemma 4.1.2.** *For  $p \neq 1, 2, \dots$ , the function  $\text{Li}_p(e^{-z})$  satisfies the singular expansion*

$$\text{Li}_p(e^{-z}) \sim \Gamma(1-p)z^{p-1} + \sum_{k=0}^{\infty} (-1)^k \zeta(p-k) \frac{z^k}{k!} \quad \text{as } z \rightarrow 0, \quad (4.16)$$

where  $\zeta$  is the Riemann zeta function.

**Remark 4.1.1.** *The singular expansion in Lemma 4.1.2 is stated only for  $z \rightarrow 0$ . However, we note that the expansion is valid in the sector  $\Sigma_{\theta, \delta}$  [14, pp. 377, proof of Lemma 2].*

For the subsequent analysis we shall need some additional results. The first result gives the absolute convergence of a special series involving the Riemann zeta function  $\zeta$ .

**Lemma 4.1.3.** *Let  $|z| \leq \pi/\sin \theta$  with  $\theta \in (\pi/2, 5\pi/6)$ , and  $p = \alpha - 1$ . Then the series (4.16) converges absolutely.*

*Proof.* Using the following well-known functional equation for the Riemann zeta function (see e.g., [34] for a short proof): for  $z \notin \mathbb{Z}$ , there holds

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

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

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

By Stirling's formula for the Gamma function  $\Gamma(x)$ ,  $x \rightarrow \infty$  [1, pp. 257]

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

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

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

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

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

**Lemma 4.1.4.** *Let  $\psi(z) = \frac{e^z-1}{\Gamma(2-\alpha)} \text{Li}_{\alpha-1}(e^{-z})$ . Then for the choice  $\theta \in (\pi/2, 5\pi/6)$ , there holds*

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

*Proof.* Upon noting the fact  $0 \leq |z\tau| \leq \pi/\sin \theta$  and using Taylor expansion and (4.16), we deduce that

for  $z \in \Gamma_\tau$ , there holds

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

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

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

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

and

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

In view of the choice  $\theta \in (\pi/2, 5\pi/6)$  and Lemma 4.1.3, the bound is uniform, since the series converges uniformly for  $z \in \Gamma_\tau$ . Consequently,

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

from which the desired assertion follows.  $\square$

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

**Lemma 4.1.5.** *Let  $\psi(z) = \frac{e^z - 1}{\Gamma(2-\alpha)} \text{Li}_{\alpha-1}(e^{-z})$ . Then for any  $\theta$  close to  $\pi/2$ , there holds for any  $\delta < \pi/2\tau$*

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

*Proof.* Since for  $z \in \Gamma_\tau$ ,  $|\Im z| \leq \pi/\tau$  and  $z \notin (-\infty, 0]$ , by [55, Lemma 1], there holds

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

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

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

It suffices to show that the real part

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

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

$$0 < 1 - e^{-r-s} \cos \phi \leq 1 - 2e^{-r-s} \cos \phi \leq 1 - 2e^{-r-s} \cos \phi + e^{-2r-2s}.$$

Consequently,

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

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

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

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

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

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

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

and

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

These two inequalities directly imply

$$\begin{aligned}\Re\psi(z\tau) &= c_\alpha \int_0^\infty \frac{s^{\alpha-2}(1-e^{-s})(1-e^{-r-s}\cos\phi)}{1-2e^{-r-s}\cos\phi+e^{-2r-2s}} ds \\ &\geq \frac{c_\alpha}{2} \int_0^\infty s^{\alpha-2}(1-e^{-s})^2 ds \geq c_3.\end{aligned}$$

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

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

**Lemma 4.1.6.** *Let  $\psi(z) = \frac{e^z-1}{\Gamma(2-\alpha)}\text{Li}_{\alpha-1}(e^{-z})$  and  $\chi_1(z) = \frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau)$ . Then there exists some  $\theta_0 \in (\pi/2, \pi)$  such that  $\chi_1(z) \in \Sigma_{\theta_0}$  for all  $z \in \Sigma_\theta$ .*

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

$$\Re\chi_1(z) = \frac{c_\alpha}{\tau^\alpha} \int_0^\infty \frac{s^{\alpha-2}(1-e^{-s})(1+e^{-2r-s}-e^{-r-s}\cos\phi-e^{-r}\cos\phi)}{1-2e^{-r-s}\cos\phi+e^{-2r-2s}} ds \quad (4.17)$$

and

$$\Im\chi_1(z) = \frac{c_\alpha}{\tau^\alpha} \int_0^\infty \frac{s^{\alpha-2}(1-e^{-s})^2 e^{-r} \sin \phi}{1-2e^{-r-s} \cos \phi + e^{-2r-2s}} ds, \quad (4.18)$$

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

$$\begin{aligned} 1 + e^{-2r-s} - e^{-r-s} \cos \phi - e^{-r} \cos \phi &\geq 1 + e^{-2r-s} - e^{-r-s} - e^{-r} \\ &= (1 - e^{-r-s})(1 - e^{-r}) \geq (1 - e^{-s})(1 - e^{-r}). \end{aligned}$$

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

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

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

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

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

and thus

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

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

$$\begin{aligned} 1 + e^{-2r-s} - e^{-r-s} \cos \phi - e^{-r} \cos \phi &\geq 1 + e^{-2r-s} \cos^2 \phi - e^{-r-s} \cos \phi - e^{-r} \cos \phi \\ &= (1 - e^{-r-s} \cos \phi)(1 - e^{-r} \cos \phi) \geq (1 - e^{-s})(1 - e^{-r} \cos \phi). \end{aligned}$$

Consequently,

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

In case (b), we may further assume  $\Re\chi_1(z) < 0$ , otherwise the statement follows directly. Then appealing to (4.17) and using the trivial inequality  $|\cos \phi| \leq 1$ , we deduce that  $e^{-r} > 1$  and

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

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

$$\begin{aligned} 0 &> 1 + e^{-2r-s} - e^{-r-s} \cos \phi - e^{-r} \cos \phi = (1 - e^{-r} \cos \phi) + e^{-s}(e^{-2r} - e^{-r} \cos \phi) \\ &\geq (\cos \phi - e^{-r}) + e^{-s}(e^{-2r} - e^{-r} \cos \phi) \geq e^{-r}(\cos \phi - e^{-r}) + e^{-s}(e^{-2r} - e^{-r} \cos \phi) \\ &= (1 - e^{-s})(e^{-r} \cos \phi - e^{-2r}), \end{aligned}$$

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

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

with the constant

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

Meanwhile, it follows directly from (4.18) that

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

Therefore,

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

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

$$\lim_{\rho \rightarrow 0} g(\rho) = -\tan \theta, \quad g_1(\rho), g_2(\rho) \geq 0, \quad g_1'(\rho) \leq 0 \quad \text{and} \quad g_2'(\rho) \geq 0,$$

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

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

This completes the proof of the lemma. □



**Remark 4.1.2.** By the proof of Lemma 4.1.6, the sector  $\Sigma_{\theta_0}$  depends on the choice of the angle  $\theta$ . For  $\theta \rightarrow \pi/2$ , it is contained in the sector  $\Sigma_{3\pi/4-\epsilon}$ , for any  $\epsilon > 0$ .

**Lemma 4.1.7.** Let  $\theta$  be close to  $\pi/2$ , and  $\delta < \pi/2\tau$ . Then for  $K_1(z)$  and  $K_2(z)$  defined in (4.11) and (4.12), respectively, there holds

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

*Proof.* For the operators  $B_1(z)$  and  $B_2(z)$  defined in (4.11) and (4.12), respectively, we have

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

where  $\chi_1(z) = \frac{1-e^{-z\tau}}{\tau^\alpha}\psi(z\tau)$ . Then by Lemma 4.1.4

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

Now we note that

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

By Lemma 4.1.6,  $\chi(z) \in \Sigma_{\theta_0}$  for some  $\theta_0 \in (\pi/2, \pi)$ . Thus by Lemma 4.1.5 and the resolvent estimate (4.3), we have

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

and the desired estimate follows immediately.  $\square$

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

**Theorem 4.1.1.** Let  $u_h$  and  $U_h^n$  be the solutions of problems (4.5) and (4.9) with  $v \in L^2(\Omega)$ ,  $U_h^0 = v_h = P_h v$  and  $f \equiv 0$ , respectively. Then there holds

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

*Proof.* It suffices to bound the terms  $I$  and  $II$ . With the choice  $\delta = t_n^{-1}$  and (4.3), we arrive at the following bound for the term  $I$

$$\begin{aligned} \|I\|_{L^2(\Omega)} &\leq c\tau \|v_h\|_{L^2(\Omega)} \left( \int_{1/t_n}^{\pi/(\tau \sin \theta)} e^{rt_n \cos \theta} dr + \int_{-\theta}^{\theta} e^{\cos \psi} t_n^{-1} d\psi \right) \\ &\leq ct_n^{-1} \tau \|v_h\|_{L^2(\Omega)}. \end{aligned} \quad (4.19)$$

By Lemma 4.1.7 and direct calculation, we bound the term  $II$  by

$$\begin{aligned} \|II\|_{L^2(\Omega)} &\leq c \int_{\pi/(\tau \sin \theta)}^{\infty} e^{rt_n \cos \theta} r^{-1} dr \|v_h\|_{L^2(\Omega)} \\ &\leq c\tau \|v_h\|_{L^2(\Omega)} \int_0^{\infty} e^{rt_n \cos \theta} dr \leq c\tau t_n^{-1} \|v_h\|_{L^2(\Omega)}. \end{aligned} \quad (4.20)$$

Combining estimates (5.17) and (5.16) yields  $\|w_h(t_n) - W_h^n\|_{L^2(\Omega)} \leq c\tau t_n^{-1} \|v_h\|_{L^2(\Omega)}$  and the desired result follows from the identity  $U_h^n - u_h(t_n) = W_h^n - w_h(t_n)$  and the stability of the projection  $P_h$  in  $L^2(\Omega)$ .  $\square$

Hence, the error estimates for the fully discrete scheme (4.9) in case of nonsmooth data follow from Theorems 3.1.3, 4.1.1 and the triangle inequality.

**Theorem 4.1.2.** *Assume that  $v \in L^2(\Omega)$  and  $f \equiv 0$ . Let  $u$  and  $U_h^n$  be the solutions of problems (2.7) and (4.9) with  $U_h^0 = P_h v$ , respectively. Then it holds that*

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

**Remark 4.1.3.** *The  $L^2(\Omega)$  stability of the  $L1$  scheme follows directly from Theorem 4.1.1 by*

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

#### 4.1.2 Error estimate for smooth initial data

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

**Lemma 4.1.8.** *Let  $\theta$  be close to  $\pi/2$ , and  $\delta < \pi/2\tau$ . Further, let  $K_1^s(z) = -z^{-1}(z^\alpha + A_h)^{-1}$  and*

$K_2^s(z) = -\chi(z)^{-1}(\chi_1(z) + A_h)^{-1}$ . Then for any  $z \in \Sigma_{\delta, \theta}$ , there hold

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

*Proof.* Like before, we set  $B_1^s(z) = -(z^\alpha + A_h)^{-1}$  and  $B_2^s(z) = -(\chi_1(z) + A_h)^{-1}$ . Then by Lemmas 4.1.4 and 4.1.5, we have

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

The rest follows analogously to the derivation (4.13).  $\square$

Now we can state an error estimate for  $v \in \dot{H}^2(\Omega)$ .

**Theorem 4.1.3.** *Let  $u_h$  and  $U_h^n$  be the solutions of problems (4.5) and (4.9) with  $v \in \dot{H}^2(\Omega)$ ,  $U_h^0 = v_h = R_h v$  and  $f \equiv 0$ , respectively. Then there holds*

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

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

$$\begin{aligned} w_h(t_n) - W_h^n &= \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta} \setminus \Gamma_\tau} e^{zt_n} K_1^s(z) A_h v_h dz \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_n} (K_1^s(z) - e^{-z\tau} K_2^s(z)) A_h v_h dz = I + II. \end{aligned} \tag{4.21}$$

By Lemma 4.1.8 we have for  $z \in \Gamma_\tau$

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

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

$$\begin{aligned} \|II\|_{L^2(\Omega)} &\leq c\tau \|A_h v_h\|_{L^2(\Omega)} \left( \int_{1/t_n}^{\pi/(\tau \sin \theta)} e^{rt_n \cos \theta} r^{-\alpha} dr + \int_{-\theta}^{\theta} e^{\cos \psi} t_n^{\alpha-1} d\psi \right) \\ &\leq ct_n^{\alpha-1} \tau \|A_h v_h\|_{L^2(\Omega)}. \end{aligned} \tag{4.22}$$

Now (4.3) implies that for all  $z \in \Gamma_{\delta, \theta}$

$$\begin{aligned} \|I\|_{L^2(\Omega)} &\leq c \|A_h v_h\|_{L^2(\Omega)} \int_{\pi/(\tau \sin \theta)}^{\infty} e^{r t_n \cos \theta} r^{-\alpha-1} dr \\ &\leq c \tau \|A_h v_h\|_{L^2(\Omega)} \int_0^{\infty} e^{r t_n \cos \theta} r^{-\alpha} dr \leq c \tau t_n^{\alpha-1} \|A_h v_h\|_{L^2(\Omega)}. \end{aligned} \quad (4.23)$$

Then the desired result follows directly from (4.22), (4.23) and the identities  $U_h^n - u_h(t_n) = W^n - w_h(t_n)$  and  $A_h R_h = P_h A$ .  $\square$

**Remark 4.1.4.** *The convergence behavior of the L1 scheme is identical with that for the convolution quadrature generated by the backward Euler method, which also converges at an  $O(\tau)$  rate, cf. [29]. In particular for smooth initial data  $v \in D(A)$ , the time discretization error by both schemes contains a singularity  $t_n^{\alpha-1}$ . This singularity reflects the limited smoothing property of the solution  $u$  [68, Theorem 2.1]*

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

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

**Example 4.1.4.** *To illustrate the convergence rate in Theorem 4.1.3, we give a trivial example. Consider the following initial value problem for the fractional ordinary differential equation:*

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

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

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

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

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

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

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

This confirms the convergence order in Theorem 4.1.3.

Last, the error estimates for the fully discrete scheme (4.9) in case of smooth data follow from Theorems 3.1.1, 4.1.3 and the triangle inequality.

**Theorem 4.1.5.** *Assume that  $v \in L^2(\Omega)$  and  $f \equiv 0$ . Let  $u$  and  $U_h^n$  be the solutions of problems (2.7) and (4.9) with  $U_h^0 = R_h v$ , respectively. Then it holds that*

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

**Remark 4.1.5.** *For  $v \in D(A)$ , we can also choose  $v_h = P_h v$  by the stability of the L1 scheme. Hence, by interpolation we deduce*

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

## 4.2 Extension to a more general sectorial operator

In this section, we show that our analysis maintains valid for a more general sectorial operator  $A$

- (a) The resolvent set  $\rho(A)$  contains the sector  $\{z : \theta \leq |\arg z| \leq \pi\}$  for some  $\theta \in (0, \pi/4)$ ;
- (b)  $\|(zI + A)^{-1}\| \leq M/|z|$  for  $z \in \Sigma_{\pi-\theta}$  and some constant  $M$ .

The technical restriction  $\theta \in (0, \pi/4)$  stems from Remark 4.1.2. This in particular covers the Riemann-Liouville fractional derivative of order  $\beta \in (3/2, 2)$ ; see Lemma 4.2.1 below. Specifically, we consider the following one-dimensional space-time fractional differential equation

$$\begin{aligned} {}^C\partial_t^\alpha u - {}^R D_x^\beta u &= f, & \text{in } \Omega = (0, 1) \quad T \geq t > 0, \\ u &= 0, & \text{on } \partial\Omega \quad T \geq t > 0, \\ u(0) &= v, & \text{in } \Omega, \end{aligned} \tag{4.24}$$

with  $\alpha \in (0, 1)$  and  $\beta \in (3/2, 2)$ . Here  ${}^R D_x^\beta$  with  $n - 1 < \beta < n$ ,  $n \in \mathbb{N}$ , denotes the left-sided Riemann-Liouville fractional derivative  ${}^R D_x^\beta u$  of order  $\beta$  defined by [33, pp. 70]:

$${}^R D_x^\beta u = \frac{1}{\Gamma(n-\beta)} \frac{d^n}{dx^n} \int_0^x (x-s)^{n-\beta-1} u(s) ds. \tag{4.25}$$

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

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

The model (4.24) is often adopted to describe anomalous diffusion process involving both long range interactions and history mechanism.

The variational formulation of (4.24) is to find  $u \in \tilde{H}^{\beta/2}(D) \equiv H_0^{\beta/2}(\Omega)$  such that (see [24])

$$({}^C \partial_t^\alpha u, \varphi) + A(u, \varphi) = (f, \varphi) \quad \forall \varphi \in \tilde{H}^{\beta/2}(D), \quad (4.26)$$

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

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

It is known (see [13, Lemma 3.1] and [24, Lemma 4.2]) that the sesquilinear form  $A(\cdot, \cdot)$  is coercive and bounded on the space  $\tilde{H}^{\beta/2}(D)$ . Then Riesz representation theorem implies that there exists a unique bounded linear operator  $\tilde{A} : \tilde{H}^{\beta/2}(D) \rightarrow H^{-\beta/2}(\Omega)$  such that

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

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

$$A(\varphi, \psi) = (A\varphi, \psi), \quad \varphi \in D(A), \quad \psi \in \tilde{H}^{\beta/2}(D). \quad (4.27)$$

We recall that the domain  $D(A)$  consists of functions of the form  ${}_0 I_x^\beta f - ({}_0 I_x^\beta f)(1)x^{\beta-1}$ , where  $f \in L^2(\Omega)$ , cf. [24]. The term  $x^{\beta-1} \in \tilde{H}_L^{\beta-1+\delta}(D)$ ,  $\delta \in [1-\beta/2, 1/2)$ , appears because it is in the kernel of the operator  ${}_0^R D_x^\beta$ . The presence of the term  $x^{\beta-1}$  indicates that the solution  $u$  usually can only have limited regularity.

**Lemma 4.2.1.** *For  $\beta \in (3/2, 2)$ , the resolvent set  $\rho(A)$  contains the sector  $\{z : \theta \leq |\arg(z)| \leq \pi\}$  for any  $\theta \in ((2-\beta)\pi/2, \pi/4)$ .*

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

$$\|u\|_{\tilde{H}^s(D)} = \|\xi|^s \mathcal{F}(\tilde{u})(\xi)\|_{L^2(\mathbb{R})},$$

where  $\mathcal{F}(\tilde{u})$  is the Fourier transform of  $\tilde{u}$ , is a consistent and well-defined norm on  $\tilde{H}^s(D)$ . Further, we note that for  $u \in \tilde{H}^{\beta/2}(D)$  (see [13] and [24])

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

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

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

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

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

Likewise, the right sided case follows:

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

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

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

Then by Lemma 2.1 of [26], we conclude that the resolvent set  $\rho(A)$  contains the sector  $\{z : \theta \leq |\arg z| \leq \pi\}$  for all  $\theta \in ((2 - \beta/2)\pi, \pi/2)$ . In particular, for  $\beta > 3/2$ , we may choose  $\theta \in ((2 - \beta/2)\pi, \pi/4)$ .  $\square$

By Lemma 4.2.1 and Remark 4.1.2, we can apply the theory in Section 4.1 to derive a fully discrete scheme based on the L1 scheme in time and the Galerkin finite element approximation in space. First we partition the unit interval  $\Omega$  into a uniform mesh with a mesh size  $h = 1/M$ . We then define  $V_h$  to be the set of continuous functions in  $V$  which are linear when restricted to the subintervals  $[x_i, x_{i+1}]$ ,

$i = 0, \dots, M - 1$ . Further, we define the discrete operator  $A_h : V_h \rightarrow V_h$  by

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

The corresponding Ritz projection  $R_h : \tilde{H}^{\beta/2}(D) \rightarrow V_h$  is defined by

$$A(R_h \psi, \chi) = A(\psi, \chi) \quad \forall \psi \in \tilde{H}^{\beta/2}(D), \chi \in V_h. \quad (4.28)$$

Then the fully discrete scheme for problem (4.24) based on the L1 approximation (4.8) reads: find  $U_h^n$  for  $n = 1, 2, \dots, N$

$$(I + \tau^\alpha A_h)U_h^n = b_n U_h^0 + \sum_{j=1}^n (b_{j-1} - b_j)U_h^{n-j} + \tau^\alpha F_h^n, \quad (4.29)$$

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

Last, we state the error estimates for the fully discrete scheme (4.29). These estimates follow from Theorems 4.1.1 and 4.1.3 and the error estimates for the semidiscrete Galerkin scheme, which can be proved using the operator trick in [26], and thus the proof is omitted.

**Theorem 4.2.1.** *Let  $u$  and  $U_h^n$  be the solutions of problems (4.24) and (4.29) with  $U_h^0 = v_h$  and  $f \equiv 0$ , respectively. Then for  $\delta \in [1 - \beta/2, 1/2)$  the following estimates hold.*

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

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

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

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

### 4.3 Extension to multi-term fractional diffusion

In this section, we extend our discussion to the following multi-term time fractional diffusion (2.15). The fully discrete scheme reads [23, (4.7)]: Find Now we arrive at the following fully discrete scheme: find  $U^{n+1} \in X_h$  such that

$$(P_\tau(\bar{\partial}_t)U^{n+1}, \chi) + (\nabla U^{n+1}, \nabla \chi) = (F^{n+1}, \chi) \quad \forall \chi \in X_h, \quad (4.30)$$



where  $F^{n+1} = f(x, t_{n+1})$  and the discrete differential operator  $P_\tau(\bar{\partial}_t)$  is defined by

$$P_\tau(\bar{\partial}_t)u(t_{n+1}) := \frac{1}{\Gamma(2-\alpha)} \sum_{j=0}^n P_j \frac{u(x, t_{n+1-j}) - u(x, t_{n-j})}{\tau^\alpha}, \quad (4.31)$$

where the coefficients  $\{P_j\}$  are defined by

$$P_j = d_{\alpha,j} + \sum_{i=1}^m \frac{\Gamma(2-\alpha)b_i d_{\alpha_i,j} \tau^{\alpha-\alpha_i}}{\Gamma(2-\alpha_i)}, \quad j \in \mathbb{N}.$$

In [23], the authors showed that the convergence is of order  $O(\tau^{2-\alpha})$  provided that  $u$  is sufficiently smooth. Unfortunately, this assumption is too restricted in general as the case of fractional diffusion, see Theorem 2.4.1. The goal of this section is to develop a robust error bound with respect to the data regularity.

We first recall that the solution regularity and analysis for the semidiscrete schemes were given in Section 2.4 and 3.4, respectively. Analogous to the argument in Section 4.1, we should first derive integral representations of  $w_h = u_h - v_h$  and  $W_h^n = U_h^n - v_h$ . Like before, the laplace transform gives formula for  $w_h$

$$w_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta,s}} e^{zt} K_1(z) v_h dz \quad \text{with} \quad K_1(z) = -z^{-1} (g(z) + A_h)^{-1} A_h, \quad (4.32)$$

where

$$g(z) = z^\alpha + \sum_{i=1}^m b_i z^{\alpha_i} \quad (4.33)$$

and the formula for  $W_h^n$

$$W_h^n = -\frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_{n-1}} \frac{\tau}{1 - e^{-z\tau}} K_2(z) v_h dz, \quad (4.34)$$

where the operator  $K_2(z)$  are defined as

$$K_2(z) = \left( \frac{1 - e^{-z\tau}}{\tau^\alpha} \left( \psi_\alpha(z\tau) + \sum_{i=1}^m b_i \tau^{\alpha-\alpha_i} \psi_{\alpha_i}(z\tau) \right) + A_h \right)^{-1} A_h, \quad (4.35)$$

with the function

$$\psi_\beta(z) = \frac{e^z - 1}{\Gamma(2-\beta)} \text{Li}_\beta(e^{-z}). \quad (4.36)$$

Next, analogue to Lemmas 4.1.4–4.1.6, we have the following results. Here we omit proofs since

they are the same as those of fractional diffusions.

**Lemma 4.3.1.** *Let  $g(z)$  and  $\psi_\beta(z)$  be respectively defined in (4.33) and (4.36). Then for the choice  $\theta \in (\pi/2, 5\pi/6)$ , there holds*

$$\left| \frac{1 - e^{-z\tau}}{\tau^\alpha} \left( \psi_\alpha(z\tau) + \sum_{i=1}^m b_i \tau^{\alpha - \alpha_i} \psi_{\alpha_i}(z\tau) \right) - g(z) \right| \leq c |z|^2 \tau^{2-\alpha} \quad \forall z \in \Gamma_\tau.$$

**Lemma 4.3.2.** *Let  $\psi_\beta(z)$  be defined in (4.33). Then for any  $\theta$  close to  $\pi/2$ , there holds for any  $\delta < \pi/2\tau$*

$$\left| \psi_\alpha(z\tau) + \sum_{i=1}^m b_i \tau^{\alpha - \alpha_i} \psi_{\alpha_i}(z\tau) \right| \geq c > 0 \quad \forall z \in \Gamma_\tau.$$

**Lemma 4.3.3.** *Let  $\psi_\beta(z)$  be defined in (4.33) and*

$$\chi_1(z) = \frac{1 - e^{-z\tau}}{\tau^\alpha} \left( \psi_\alpha(z\tau) + \sum_{i=1}^m b_i \tau^{\alpha - \alpha_i} \psi_{\alpha_i}(z\tau) \right).$$

*Then there exists some  $\theta_0 \in (\pi/2, \pi)$  such that  $\chi_1(z) \in \Sigma_{\theta_0}$  for all  $z \in \Sigma_\theta$ .*

As a result, we get the following bound for  $\|K_1(z) - K_2(z)\|$ .

**Lemma 4.3.4.** *Let  $\theta$  be close to  $\pi/2$ , and  $\delta < \pi/2\tau$ . Then for  $K_1(z)$  and  $K_2(z)$  defined in (4.32) and (4.35), respectively, there holds*

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

Then the estimate for  $\|u_h - U_h^n\|$  follows directly from Lemma 4.3.1–4.3.4.

**Theorem 4.3.1.** *Let  $u_h$  and  $U_h^n$  be the solutions of problems (3.48) and (4.30) with  $v \in L^2(\Omega)$ ,  $U_h^0 = v_h = P_h v$  and  $f \equiv 0$ , respectively. Then there holds*

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

Hence, the error estimate for the fully discrete scheme (4.9) in case of nonsmooth data follow from Theorems 3.4.2, 4.3.1 and the triangle inequality.

**Theorem 4.3.2.** *Assume that  $v \in L^2(\Omega)$  and  $f \equiv 0$ . Let  $u$  and  $U_h^n$  be the solutions of problems (2.15)*

and (4.30) with  $U_h^0 = P_h v$ , respectively. Then it holds that

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

Further the error estimate follows from the similar argument in Section 4.1.2.

**Theorem 4.3.3.** *Assume that  $v \in L^2(\Omega)$  and  $f \equiv 0$ . Let  $u$  and  $U_h^n$  be the solutions of problems (2.15) and (4.30) with  $U_h^0 = R_h v$ , respectively. Then it holds that*

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

#### 4.4 Numerical results

Now we present numerical results to verify the convergence theory, i.e., the  $O(\tau)$  convergence rate. We consider the fractional diffusion case (2.7), time-space fractional case (4.24) and multi-term fractional case (2.15) separately. For each model, we consider the following two initial data:

- (a)  $\Omega = (0, 1)$ , and  $v = \sin(2\pi x) \in H^2(\Omega) \cap H_0^1(\Omega)$ .
- (b)  $\Omega = (0, 1)$ , and  $v = x^{-1/4} \in H^{1/4-\epsilon}(\Omega)$ , with  $\epsilon \in (0, 1/4)$ ;

We measure the error  $e^n = u(t_n) - U_h^n$  by the normalized errors  $\|e^n\|_{L^2(\Omega)} / \|v\|_{L^2(\Omega)}$ . In the computations, we divide the unit interval  $\Omega = (0, 1)$  into  $M$  equally spaced subintervals with a mesh size  $h = 1/M$ . Likewise, we fix the time step size  $\tau$  at  $\tau = t/N$ , where  $t$  is the time of interest. In this work, we only examine the temporal convergence rate, since the space convergence rate has been examined earlier in Chapter 3. To this end, we take a small mesh size  $h = 2^{-13}$ , so that the spatial discretization error is negligible.

**Fractional diffusion.** The exact solution can be written explicitly by spectral decomposition as an infinite series involving the Mittag-Leffler function, which can be evaluated efficiently by the algorithm developed in [70]. The numerical results for cases (a) and (b) are shown in Table 4.1, where **rate** refers to the empirical convergence rate, and the number in the bracket refers to the theoretical rate. The results fully confirm the theoretical prediction: for both smooth and nonsmooth data, the fully discrete solution  $U_h^n$  converges at a rate  $O(\tau)$ , and the rate is independent of the fractional order  $\alpha$ . Further, for fixed  $t$ , the error increases with the fractional order  $\alpha$ . This might be attributed to the local decay behavior of the solution  $u$ : the larger is the fractional order  $\alpha$ , the slower is the solution

decay around  $t = 0$ . According to Remark 4.1.5, we have

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

Thus the temporal error deteriorates as the time  $t_n \rightarrow 0$  at a rate like  $t_n^\alpha$  and  $t_n^{\alpha/8-\alpha\epsilon}$  for  $Av \in L^2(\Omega)$  and  $v \in D(A^{1/8-\epsilon})$ , with  $\epsilon \in (0, 1/8)$ , respectively. In particular, for fixed  $N$ , the error behaves like  $t_n^{1/2}$  and  $t_n^{1/16}$  for cases (a) and (b) with  $\alpha = 0.5$ , respectively. This is clearly observed in Table 4.2, thereby showing the sharpness of the error estimates. Further, we observe that the L1 scheme fails to converge uniformly (in time) at a first order even though the initial data  $v$  in case (a) is very smooth, i.e.,  $v \in D(A^p)$  for any  $p \geq 0$ , cf. Table 4.2. This numerically confirms the observation in Remark 4.1.4.

Table 4.1: The  $L^2$ -norm of the error for the fractional diffusion equation with initial data (a) and (b) at  $t = 0.1$  with  $h = 2^{-13}$ ,  $\tau = 1/N$ .

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

Table 4.2: The  $L^2$ -norm of the error for the fractional diffusion equation with initial data (a) and (b) with  $\alpha = 0.5$ ,  $h = 2^{-13}$  and  $N = 10$ .

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

**Time-space fractional problem.** Now we present numerical results for the time-space fractional problem. Since the exact solution is not available in closed form, we compute the reference solution by the second-order backward difference scheme [29] on a much finer mesh, i.e.,  $N = 1000$ . The numerical results for case (a) with different  $\alpha$  and  $\beta$  values are presented in Table 4.3. A first-order convergence is observed, and the convergence rate is independent of the time- and space-fractional orders. Interestingly, the observation is valid also for the case  $\beta = 5/4$ , for which the theory in Section 4.2 does not cover, awaiting further study. For a fixed  $\alpha$  value, the error decreases with the increase of the fractional order  $\beta$ , which indicates that the solution decays faster as  $\beta$  approaches two. This is also consistent with the fact that the smaller is the  $\beta$  value, the more singular is the solution, and thus more challenging to approximate numerically [24]. For the nonsmooth case (b), we are particularly interested in the case of small  $t$ . Thus we present the numerical results for  $t = 0.1$ ,  $t = 0.01$  and  $t = 0.001$  in Table 4.4. The experiment shows the first order convergence is robust for nonsmooth data even if  $t$  is close to zero. Like before, for fixed  $n$ , let  $t_n \rightarrow 0$ , the error behaves like  $t_n^\alpha$  for case (a), which is fully confirmed by the numerical results, cf. Table 4.5, indicating the sharpness of the estimate in Theorem 4.2.1.

Table 4.3: The  $L^2$ -norm of the error for the time-space fractional problem with initial data (a) at  $t = 0.1$  with  $h = 2^{-13}$ ,  $\tau \rightarrow 0$ .

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

Table 4.4: The  $L^2$ -norm of the error for the time-space fractional problem with initial data (b) with  $\beta = 1.5$ ,  $h = 2^{-13}$ ,  $t = 0.1, 0.01$  and  $0.001$ .

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

Table 4.5: The  $L^2$ -norm of the error for the time-space fractional problem with initial data (a) and (b),  $\alpha = 0.5$  and  $\beta = 1.5$ , as  $t \rightarrow 0$  with  $h = 2^{-13}$  and  $N = 5$ .

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

**Multi-term fractional differential equations.** In this part, we perform numerical experiments for the multi-term time fractional diffusion equation (2.15) to verify theoretical convergence rates in Theorem 4.3.2 and 4.3.3. In order to find a reference solution, we compute the result on a much finer mesh, i.e.,  $N=2000$ , since the closed form of the solution is unavailable. For simplicity, we consider the case  $k = 1$  with  $\alpha_1 = 0.1$  and  $b_1 = 1$ . We present  $L^2$ -norm of the error and its behaviors when  $t_n \rightarrow 0$  for fixed  $n = 10$  in Tables 4.6 and 4.7, respectively. We observe that the convergence order and the error behaviors are the same as those of standard diffusion equations, which confirm our theoretical findings.

Table 4.6: The  $L^2$ -norm of the error for the multi-term fractional diffusion equation with initial data (a) and (b) with  $\alpha_1 = 0.1$  at  $t = 0.1$  with  $h = 2^{-12}$ ,  $\tau = 1/N$ .

$\alpha$	$N$	10	20	40	80	160	320	rate
$\alpha = 0.25$	Case (a)	9.19e-4	4.47e-4	2.19e-4	1.79e-4	5.27e-5	2.52e-5	$\approx 1.03$ (1.00)
	Case (b)	3.95e-4	1.93e-4	9.57e-5	4.76e-5	2.38e-5	1.19e-5	$\approx 1.03$ (1.00)
$\alpha = 0.5$	Case (a)	1.30e-3	6.26e-4	3.05e-4	1.49e-4	7.27e-5	3.47e-5	$\approx 1.04$ (1.00)
	Case (b)	2.26e-3	1.07e-3	5.19e-4	2.52e-4	1.22e-4	5.82e-5	$\approx 1.05$ (1.00)
$\alpha = 0.75$	Case (a)	2.76e-3	1.25e-3	5.91e-4	2.82e-4	1.35e-4	6.35e-4	$\approx 1.07$ (1.00)
	Case (b)	6.14e-3	2.86e-3	1.35e-3	6.44e-4	3.05e-4	1.43e-4	$\approx 1.08$ (1.00)

Table 4.7: The  $L^2$ -norm of the error for the the multi-term fractional diffusion equation with initial data (a) and (b) with  $\alpha = 0.5$ ,  $\alpha_1 = 0.1$ ,  $h = 2^{-13}$  and  $N = 10$ .

$t$	1e-5	1e-6	1e-7	1e-8	1e-9	1e-10	rate
(a)	2.88e-3	1.04e-3	3.43e-4	1.10e-4	3.49e-5	1.11e-5	$\approx 0.49$ (0.50)
(b)	3.44e-3	2.95e-3	2.53e-3	2.17e-3	1.87e-3	1.61e-3	$\approx 0.07$ (0.06)

#### 4.5 Conclusion and comments

In this chapter we discussed the popular L1 time-stepping scheme for discretizing the Caputo fractional derivative of order  $\alpha \in (0, 1)$ , arising in the modeling of fractional diffusion (2.7), and rigorously established the first order convergence for both smooth and nonsmooth initial data. The extensive numerical experiments fully verify the sharpness of the estimates and robustness of the scheme. The convergence analysis is valid for more general sectorial operators, and in particular, it covers also the space-time fractional differential equations. Last, we extend our discussion to the multi-term time fractional differential equations. We refer to our previous paper [30] for more details.

In view of the solution singularity for time  $t$  close to zero, it is natural to consider the L1 scheme on a nonuniform time mesh in order to arrive at a uniform first-order convergence. The generating function approach used in our analysis does not work directly in this case, and it is an interesting open question to rigorously establish error estimates directly in terms of the data regularity. The mesh graded time stepping based on discontinuous Petrov-Galerkin finite element method was discussed in [62]. In [55, 63, 61], the authors considered fractional diffusion involving a Riemann-Liouville type fractional derivative and discontinuous Galerkin finite element method in time with variable time steps. Some interesting results for an evolution equation with a positive-type memory term were provided in [58].



## 5. FULLY DISCRETE SCHEMES BY CONVOLUTION QUADRATURE\*

In this chapter, we develop two time stepping schemes based on convolution quadrature for the model (2.7) and to derive optimal error bounds for the approximate solution that are expressed in terms of the smoothness of the data, including the case of nonsmooth data, namely,  $v \in L^2(\Omega)$ . For example, this case is important in inverse problems and optimal control [31].

This chapter is organized as follows. In Section 5.1, 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 5.2. The analysis exploits the operational calculus framework [10] and recently also used in [6] for the Rayleigh-Stokes problem. A different approach for error analysis is given in Section 5.3 using the discrete Laplace transform. In Section 5.4, we present extensive one and two-dimensional numerical experiments to illustrate the convergence behavior of the schemes. A detailed comparison with several existing methods is also included. The numerical results indicate that the schemes can achieve first and second-order convergence, and they are robust with respect to data regularity. Finally, we extend the technique to the multi-term fractional diffusion and diffusion wave equations in Section 5.5. Throughout, the notation  $C$  denotes a generic constant, which may differ at different occurrences, but is always independent of the solution  $u$ , the spatial mesh-size  $h$  and the time step-size  $\tau$ .

### 5.1 Fractional convolution quadrature

In this part, we develop two fully discrete schemes for problem (2.7). This is achieved by applying convolution quadrature [47, 46] to the Riemann-Liouville derivative  ${}^R\partial_t^\alpha$  defined by (2.1). Specifically, we rewrite the semidiscrete problem (3.1) using the defining relation of the Caputo derivative (2.3). In particular, in the cases of subdiffusion,  $0 < \alpha < 1$ ,

$${}^C\partial_t^\alpha \varphi = {}^R\partial_t^\alpha (\varphi(t) - \varphi(0)).$$

Hence for  $t > 0$  the spatial semidiscrete scheme (3.1) can be rewritten as

$${}^R\partial_t^\alpha (u_h - v_h) + A_h u_h = f_h, \quad \text{for } 0 < \alpha < 1 \tag{5.1}$$

---

\*Sections 5.1.2, 5.2 and 5.4 is reprinted with permission from "On two schemes for fractional diffusion and diffusion wave equations" by Bangti Jin, Raytcho Lazarov and Zhi Zhou, submitted to SIAM Journal of Scientific Computing, Copyright [2015] by Society for Industrial and Applied Mathematics.

respectively, where  $f_h = P_h f(t)$ .

### 5.1.1 Approximating fractional derivatives by convolution quadrature

First, we briefly introduce the idea of convolution quadratures developed in [46, 47]. By Laplace transform and integration by parts, we may represent the Riemann-Liouville type fractional derivative (2.1) by

$$\begin{aligned} {}^R\partial_t^\alpha \varphi(t) &= \frac{d}{dt} \int_0^t \frac{s^{-\alpha}}{\Gamma(1-\alpha)} \varphi(t-s) ds = \frac{d}{dt} \int_0^t \left( \frac{1}{2\pi i} \int_\Gamma e^{zs} z^{\alpha-1} dz \right) \varphi(t-s) ds \\ &= \frac{1}{2\pi i} \int_\Gamma z^{\alpha-1} \left( \frac{d}{dt} \int_0^t e^{zs} \varphi(t-s) ds \right) dz. \end{aligned}$$

We set  $y = \int_0^t e^{zs} \varphi(t-s) ds$  and  $g(t) = y'$  and note that  $y$  is the solution of  $y' = zy + \varphi$  with  $y(0) = 0$ . Next we use a stable and consistent linear multistep method to approximate  $y_n = y(t_n)$ . We consider a uniform grid on the interval  $[0, T]$  with a time step size  $\tau = T/N$ ,  $N \in \mathbb{N}$ , so that,  $0 = t_0 < t_1 < \dots < t_N = T$ , and  $t_n = n\tau$ ,  $n = 0, \dots, N$ . Then by setting  $\varphi_n = \varphi(j\tau)$  for  $j \geq 0$ , the multistep method with generating functions (see, e.g. [22, p. 27, Lemma 2.3])

$$\sigma(\xi) = a_k \xi^k + \dots + a_0 \quad \text{and} \quad \rho(\xi) = b_k \xi^k + \dots + b_0$$

gives for  $n \geq 0$

$$\sum_{j=0}^k a_j y_{n+j-k} = \tau \sum_{j=0}^k b_j (z y_{n+j-k} + \varphi_{n+j-k}), \quad (5.2)$$

with starting values  $y_{-k} = \dots = y_{-1} = 0$  and  $\varphi_{-k} = \dots = \varphi_{-1} = 0$ . Multiplying (5.2) by  $\xi^n$  and summing over  $n$  from 0 to  $\infty$  we obtain

$$(a_0 \xi^k + \dots + a_k) \tilde{y}(\xi) = \tau (b_0 \xi^k + \dots + b_k) (z \tilde{y}(\xi) + \tilde{\varphi}(\xi)),$$

with the generating series  $\tilde{y}(\xi) = \sum_{n=0}^{\infty} y_n \xi^n$  and  $\tilde{\varphi}(\xi) = \sum_{n=0}^{\infty} \varphi_n \xi^n$ . Then rearranging the terms and solving this equation we get

$$\tilde{y}(\xi) = (\delta(\xi)/\tau - z)^{-1} \tilde{\varphi}(\xi),$$

where  $\delta(\xi) = \sigma(1/\xi)/\rho(1/\xi)$  is the quotient of the generating polynomials of the linear multistep method. Further, the similar argument yields

$$\tilde{y}(\xi) = (\delta(\xi)/\tau)\tilde{y}(\xi) = (\delta(\xi)/\tau)(\delta(\xi)/\tau - z)^{-1}\tilde{\varphi}(\xi).$$

Then by the Cauchy integral formula, the approximation for  ${}^R\partial_t^\alpha\varphi(t_n)$  is the  $n$ -th coefficient of

$$\frac{1}{2\pi i} \int_{\Gamma} z^{\alpha-1} (\delta(\xi)/\tau) (\delta(\xi)/\tau - z)^{-1} \tilde{\varphi}(\xi) dz = (\delta(\xi)/\tau)^\alpha \tilde{\varphi}(\xi),$$

which can be explicitly expressed as

$${}^R\partial_t^\alpha\varphi(t_n) \approx \sum_{0 \leq j\tau \leq t_n} \omega_j \varphi(t_n - j\tau) := \bar{\partial}_\tau^\alpha \varphi(t_n), \quad n \geq 0, \quad (5.3)$$

where the quadrature weights  $\{\omega_j\}_{j=0}^\infty$  are determined by the generating function:  $\sum_{j=0}^\infty \omega_j \xi^j = (\delta(\xi)/\tau)^\alpha$ . Then the fully discrete scheme follows from the semidiscrete scheme (5.1) and the approximation (5.3).

### 5.1.2 Discrete operational calculus framework

In order to derive an error estimate for the fully discrete schemes, we describe the framework developed in [10, Sections 2 and 3], initiated in [46, 47]. Let  $K$  be a complex valued or operator valued function which is analytic in a sector  $\Sigma_\theta := \{z \in \mathbb{C} : |\arg z| \leq \theta\}$ ,  $\theta \in (0, \pi/2)$  and is bounded by

$$\|K(z)\| \leq M|z|^{-\mu} \quad \forall z \in \Sigma_\theta, \quad (5.4)$$

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  $k$  :  $K(\partial_t)g = k * g$  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

$$K_1(\partial_t)K_2(\partial_t) = (K_1K_2)(\partial_t). \quad (5.5)$$

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  $0 = t_0 < t_1 < \dots < t_N = T$ , and  $t_n = n\tau$ ,  $n = 0, \dots, N$ . Then the convolution quadrature  $K(\bar{\partial}_\tau)g(t)$  of  $K(\partial_t)g(t)$  is defined by (see e.g. [47]):

$$K(\bar{\partial}_\tau)g(t) = \sum_{0 \leq j\tau \leq t} \omega_j g(t - j\tau), \quad t > 0, \quad (5.6)$$

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

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

We remark that the quadrature weights  $\{\omega_j\}$  can be computed efficiently via the fast Fourier transform [67]. The associativity property is also valid for convolution quadratures:

$$K_1(\bar{\partial}_\tau)K_2(\bar{\partial}_\tau) = (K_1K_2)(\bar{\partial}_\tau). \quad (5.7)$$

Next we derive the fully discrete schemes. First, we rewrite the semidiscrete scheme (5.1) in the form

$$u_h = (\partial_t^\alpha + A_h)^{-1} \partial_t^\alpha v_h + (\partial_t^\alpha + A_h)^{-1} f_h. \quad (5.8)$$

Then the associativity property (5.7) yields the BE scheme for the case  $0 < \alpha < 1$ :

$$U_h^n = (\bar{\partial}_\tau^\alpha + A_h)^{-1} \bar{\partial}_\tau^\alpha v_h + (\bar{\partial}_\tau^\alpha + A_h)^{-1} f_h. \quad (5.9)$$

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

$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha v_h + F_h^n, \quad \text{with } U_h^0 = v_h, \quad F_h^n = P_h f(t_n). \quad (5.10)$$

**Remark 5.1.1.** *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 [69, 48].*

Next we turn to the SBD scheme. It is known that the convolution quadrature (5.6) is only first-order accurate if  $g(0) \neq 0$ , e.g., for  $g \equiv 1$  [47, Theorem 5.1] [10, Section 3]. Hence, to get a second-order approximation one needs to introduce some correction. For this we follow the approach proposed in [48, 10]. Using the notation  $\partial_t^\beta u$ ,  $\beta < 0$  for the Riemann-Liouville integral

$$\partial_t^\beta u = \frac{1}{\Gamma(-\beta)} \int_0^t (t-s)^{-\beta-1} u(s) ds \quad \text{and} \quad (\partial_t^\alpha + A_h)^{-1} = \partial_t^{-\alpha} - (I + \partial_t^{-\alpha} A_h)^{-1} \partial_t^{-\alpha} A_h,$$

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 (5.8) as

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

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

$$U_h^n = v_h - (\bar{\partial}_\tau^\alpha + A_h)^{-1} \bar{\partial}_\tau \partial_t^{-1} A_h v_h + (\bar{\partial}_\tau^\alpha + A_h)^{-1} (\bar{\partial}_\tau \partial_t^{-1} f_{h,0} + \tilde{f}_h). \quad (5.11)$$

The purpose of keeping the operator  $\partial_t^{-1}$  is to achieve a second-order accuracy. Letting  $1_\tau = (0, 3/2, 1, \dots)$ , using the identity  $1_\tau = \bar{\partial}_\tau \partial_t^{-1} 1$  at grid points  $t_n$  [10], and the associativity (5.7), the scheme (5.11) 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 fractional diffusion case is: find  $U_h^n$ ,  $n \geq 1$  such that

$$\begin{aligned} \bar{\partial}_\tau^\alpha U_h^1 + A_h U_h^1 + \frac{1}{2} A_h U_h^0 &= \bar{\partial}_\tau^\alpha U_h^0 + F_h^1 + \frac{1}{2} F_h^0, \\ \bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n &= \bar{\partial}_\tau^\alpha U_h^0 + F_h^n, \quad 2 \leq n \leq N. \end{aligned} \quad (5.12)$$

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

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 [73].

**Remark 5.1.2.** *It is known that without a correction the SBD scheme in general is only first-order accurate. Lubich [46, 47] 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 [74, Section 3.2].*

Last we remark that the quadrature weights  $\{\omega_j\}$  can be computed efficiently via the fast Fourier transform [67]. Specifically, if  $f(z)$  is an analytic function in the closed unit disc, then the expansion coefficients  $\{f_j\}$  can be computed by Cauchy integrals, i.e.,

$$f(z) = \sum_{j=0}^{\infty} f_j z^j, \quad f_j = \frac{1}{2\pi i} \int_{\mathcal{C}} z^{-j-1} f(z) dz,$$

where the contour  $\mathcal{C}$  is the unit disc oriented counterclockwise. Upon introducing the variable  $z = e^{i\psi}$ , the coefficient  $f_j$  can be determined by

$$f_j = \frac{1}{2\pi} \int_0^{2\pi} e^{-ij\psi} f(e^{i\psi}) d\psi,$$

which can be evaluated efficiently via the fast Fourier transform.

**Remark 5.1.3.** *The SBD scheme for the Riemann-Liouville derivative was recently analyzed in [38], by means of a Fourier transform and uses substantially the zero extension  $\bar{\varphi}$  of  $\varphi$  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  $\alpha$  close to zero. These conditions are very restrictive, and do not hold even in the simplest case of a homogeneous problem [68]. In view of Remark 5.1.2, these conditions are not needed for the SBD scheme. To avoid these restrictions, we have employed the elegant technique developed in [10].*

## 5.2 Error analysis by discretized operational calculus

Now we analyze the fully discrete schemes derived in Section 5.2. 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. The semidiscrete method (3.1) and the appropriate error bounds for the error  $u(t) - u_h(t)$  have been discussed in Chapter 3. Next we analyze the error  $u_h(t_n) - U_h^n$  between the solutions of the semidiscrete and the fully discrete problems.

### 5.2.1 Error analysis for BE method

Now we derive  $L^2$  error estimates for the fully discrete schemes (5.10) using the technique developed in [47, 10]. Here we denote for  $z \in \Sigma_\theta$ ,  $\theta \in (\pi/2, \pi)$ ,  $G(z) = z^\alpha(z^\alpha + A_h)^{-1}$ . Then for the homogeneous problem ( $f \equiv 0$ ), by (5.8) and (5.9), the difference between  $U_h^n$  and  $u_h(t_n)$  can be represented by

$$U_h^n - u_h(t_n) = (G(\bar{\partial}_\tau) - G(\partial_t))v_h. \quad (5.13)$$

For the error analysis, we need the following estimate [47, Theorem 5.2].

**Lemma 5.2.1.** *Let  $K(z)$  be analytic in  $\Sigma_\theta$  and (5.4) hold. Then for  $g(t) = ct^{\beta-1}$ , the convolution quadrature based on the BE method satisfies*

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

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

**Lemma 5.2.2.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.10) with  $v \in L^2(\Omega)$ ,  $U_h^0 = v_h = P_h v$  and  $f \equiv 0$ , respectively. Then there holds*

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leq C\tau t_n^{-1} \|v\|_{L^2(\Omega)}, \quad n \geq 1.$$

*Proof.* By (4.3), there holds  $G(z) \leq C$  for  $z \in \Sigma_\theta$ . Hence (5.13) and Lemma 5.2.1 (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  $P_h$ .  $\square$

**Remark 5.2.1.** *The stability of the fully discrete scheme follows from Lemma 5.2.2*

$$\|U_h^n\|_{L^2(\Omega)} \leq \|U_h^n - u_h(t_n)\|_{L^2(\Omega)} + \|u_h(t_n)\|_{L^2(\Omega)} \leq C\|U_h^0\|_{L^2(\Omega)}.$$

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

**Lemma 5.2.3.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.10) with  $v \in \dot{H}^2(\Omega)$ ,  $U_h^0 = v_h = R_h v$  and  $f \equiv 0$ , respectively. Then there holds*

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leq C\tau t_n^{\alpha-1} \|Av\|_{L^2(\Omega)}, \quad n \geq 1.$$

*Proof.* Using the fact  $G(z) = I - (z^\alpha + A_h)^{-1} A_h$  and denoting  $G_s(z) = (z^\alpha + A_h)^{-1}$ , then  $U_h^n - u_h(t_n) = (G_s(\bar{\partial}_\tau) - G_s(\partial_t)) A_h v_h$ . Now using (4.3) and Lemma 5.2.1 (with  $\mu = \alpha$  and  $\beta = 1$ ) gives

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leq C\tau t_n^{\alpha-1} \|A_h v_h\|_{L^2(\Omega)}.$$

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 (5.10) follows from the triangle inequality and Theorems 3.1.1, 3.1.3.

**Theorem 5.2.1.** *Let  $u$  and  $U_h^n$  be the solutions of problem (2.7) and (5.10) 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  $n \geq 1$*

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

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

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

The next corollary gives the error estimate for the intermediate case, i.e.,  $v \in \dot{H}^1(\Omega)$ .

**Corollary 5.2.1.** *Using Remark 5.2.1 and interpolation we can get the following bound for data of intermediate smoothness, i.e., for  $v \in \dot{H}^1(\Omega)$  and  $v_h = P_h v$ :*

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



### 5.2.2 Error analysis for the SBD scheme

Now we turn to the analysis of the SBD scheme. Like Lemma 5.2.1, the following estimate holds [47, Theorem 5.2].

**Lemma 5.2.4.** *Let  $K(z)$  be analytic in  $\Sigma_\theta$  and (5.4) hold. Then for  $g(t) = ct^{\beta-1}$ , the convolution quadrature based on the SBD satisfies*

$$\|(K(\partial_t) - K(\bar{\partial}_\tau))g(t)\| \leq \begin{cases} Ct^{\mu-1}\tau^\beta, & 0 < \beta \leq 2, \\ Ct^{\mu+\beta-3}\tau^2, & \beta \geq 2. \end{cases}$$

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

**Lemma 5.2.5.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.12) with  $v \in L^2(\Omega)$ ,  $U_h^0 = v_h = P_h v$  and  $f \equiv 0$ , respectively. Then there holds*

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

*Proof.* Like before, the difference between  $u_h(t_n)$  and  $U_h^n$  can be represented 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  $G(z) = -z(z^\alpha + A_h)^{-1}A_h$ . By (4.3) and the identity

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

there holds  $\|G(z)\| \leq C|z|$ , for  $z \in \Sigma_\theta$ . Then Lemma 5.2.4 (with  $\mu = -1$  and  $\beta = 2$ ) gives

$$\|U_h^n - u_h(t_n)\|_{L^2(\Omega)} \leq C\tau^2 t_n^{-2} \|v_h\|_{L^2(\Omega)},$$

and the desired result follows directly from the  $L^2(\Omega)$  stability of  $P_h$ .  $\square$

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

**Lemma 5.2.6.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.12) with  $v \in \dot{H}^2(\Omega)$ ,  $U_h^0 =$*

$v_h = R_h v$  and  $f \equiv 0$ , respectively. Then there holds

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

*Proof.* By setting  $G_s(z) = -z(z^\alpha + A_h)^{-1}$ ,  $U_h^n - u_h(t_n)$  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 (4.3), we deduce

$$\|G_s(z)\| \leq M|z|^{1-\alpha} \quad \forall z \in \Sigma_\theta.$$

Now Lemma 5.2.4 (with  $\mu = \alpha - 1$  and  $\beta = 2$ ) gives

$$\|U_h^n - u_h(t_n)\|_{L^2(\Omega)} \leq C\tau^2 t_n^{\alpha-2} \|A_h v_h\|_{L^2(\Omega)},$$

and the desired estimate follows from the identity  $A_h R_h = P_h A$ . □

Then the error estimates for the fully discrete scheme (5.12) follows from the triangle inequality and Theorems 3.1.1, 3.1.3.

**Theorem 5.2.2.** *Let  $u$  and  $U_h^n$  be the solutions of problem (2.7) and (5.12) 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  $n \geq 1$*

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

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

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

### 5.3 Error analysis by discrete Laplace transform

We note that the error estimates in Theorems 5.2.1 and 5.2.2 can be derived using the technique discussed in Chapter 4. To this end, we consider the following splittings

$$w_h = u_h - v_h \quad \text{and} \quad W_h^n = U_h^n - v_h.$$

#### 5.3.1 Error analysis for the BE method

We first consider the BE scheme (5.10). Recall that the integral representation of  $w_h$  has been given by

$$w_h(t) = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \delta}} e^{zt} K(z) v_h dz \quad (5.14)$$

with  $K(z) = -z^{-1}(z^\alpha I + A_h)^{-1} A_h$ , while the representation of  $W_h$  is given in the following lemma.

**Lemma 5.3.1.** *Let  $K(z) = -z^{-1}(z^\alpha I + A_h)^{-1} A_h$ . Then  $W_h^n$  can be represented by*

$$W_h^n(t) = \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_{n-1}} K\left(\frac{1 - e^{-z\tau}}{\tau}\right) v_h dz,$$

with the contour  $\Gamma_\tau = \{z \in \Gamma_{\delta, \theta} : |\Im(z)| \leq \pi/\tau\}$ .

*Proof.* By the definition,  $W_h^n$  satisfies the time-stepping scheme

$$\bar{\partial}_\tau^\alpha W_h^n + A W_h^n = -A_h v_h, \quad (5.15)$$

with  $W_h^0 = 0$  and  $\bar{\partial}_\tau^\alpha W_h^n$  is defined in Section 5.1. Now multiplying both sides of the equation by  $\xi^n$  and summing from 1 to  $\infty$  yields

$$\sum_{n=1}^{\infty} (\bar{\partial}_\tau^\alpha W_h^n) \xi^n + A_h \widetilde{W}_h(\xi) = -\widetilde{1}_s(\xi) A_h v_h,$$

where  $\widetilde{1}_s$  denotes the right-sided shifted identity vector, i.e.,  $\widetilde{1}_s = (0, 1, 1, \dots)$  and hence  $\widetilde{1}_s(\xi) = \xi/(1-\xi)$ .

Using the fact  $W_h^0 = 0$ , we have

$$\sum_{n=1}^{\infty} (\bar{\partial}_\tau^\alpha W_h^n) \xi^n = \sum_{n=0}^{\infty} \tau^{-\alpha} \sum_{j=0}^n (\omega_{n-j} \xi^{n-j}) (W_h^j \xi^j) = ((1-\xi)/\tau)^\alpha \widetilde{W}_h(\xi).$$

Thus by simple calculation we have

$$\widetilde{W}_h(\xi) = (\xi/\tau)K((1-\xi)/\tau)v_h.$$

Since  $\widetilde{W}_h(\xi)$  is analytic at  $\xi = 0$ , Cauchy theorem implies that for  $\varrho$  small enough

$$W_h^n = \frac{1}{2\tau\pi i} \int_{|\xi|=\varrho} \xi^{-n} K((1-\xi)/\tau)v_h d\xi.$$

Now by changing variable  $\xi = e^{-z\tau}$  we obtain

$$W_h^n = \frac{1}{2\pi i} \int_{\Gamma^0} e^{zt_n-1} K((1-e^{-z\tau})/\tau)v_h dz,$$

where the contour  $\Gamma^0 = \{z = -\ln(\varrho)/\tau + iy : |y| \leq \pi/\tau\}$  is oriented counterclockwise. Finally, we obtain the desired representation by deforming the contour  $\Gamma^0$  to  $\Gamma_\tau = \{z \in \Gamma_{\theta,\delta} : |\Im(z)| \leq \pi/\tau\}$  and using the periodicity of the exponential function.  $\square$

The next lemma provides basic estimates on the kernel  $K$ . The estimates follow analogously to that in Chapter 4, and hence the proof is omitted.

**Lemma 5.3.2.** *Let  $K(z) = -z^{-1}A_h(z^\alpha I + A_h)^{-1}$ . Then*

$$\begin{aligned} \|K(z)\| &\leq C|z|^{-1} \quad \text{and} \quad \|K'(z)\| \leq C|z|^{-2} \quad \forall z \in \Sigma'_\theta, \\ \|e^{-z\tau} K\left(\frac{1-e^{-z\tau}}{\tau}\right) - K(z)\| &\leq C\tau \quad \forall z \in \Gamma_\tau. \end{aligned}$$

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

**Theorem 5.3.1.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.10) with  $v \in L^2(\Omega)$ ,  $U_h^0 = v_h = P_h v$  and  $f \equiv 0$ , respectively. Then there holds*

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

*Proof.* By Lemma 5.3.1, the error can be split by

$$\begin{aligned} w_h(t_n) - W_h^n &= \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_n} \left( K(z) - e^{-z\tau} K\left(\frac{1-e^{-z\tau}}{\tau}\right) \right) v_h dz \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta} \setminus \Gamma_\tau} e^{zt_n} K(z) v_h dz = I + II. \end{aligned}$$

By Lemma 5.3.2, we arrive at the following bound on  $II$

$$\begin{aligned} \|II\|_{L^2(\Omega)} &\leq C \|v_h\|_{L^2} \int_{\pi/\tau}^{\infty} e^{rt_n \cos \theta} r^{-1} dr \\ &\leq C\tau \|v_h\|_{L^2} \int_{\pi/\tau}^{\infty} e^{rt_n \cos \theta} dr \leq C\tau t_n^{-1} \|v\|_{L^2}. \end{aligned} \tag{5.16}$$

With the choice  $\delta = t_n^{-1}$  and Lemma 5.3.2 we bound the term  $I$

$$\begin{aligned} \|I\|_{L^2(\Omega)} &\leq C\tau \|v_h\|_{L^2(\Omega)} \left( \int_{\pi/(t_n \sin \theta)}^{\pi/(\tau \sin \theta)} e^{-rt_n \cos \theta} dr + \int_{-\pi+\theta}^{\pi-\theta} e^{\cos \psi} t_n^{-1} d\psi \right) \\ &\leq C t_n^{-1} \tau \|v_h\|_{L^2(\Omega)}. \end{aligned} \tag{5.17}$$

Then (5.17), (5.16) and the  $L^2$ -stability of the projection  $P_h$  yields

$$\|w_h(t_n) - W_h^n\|_{L^2(\Omega)} \leq C\tau t_n^{-1} \|v\|_{L^2(\Omega)}.$$

Then the desired result follows directly from the identity  $U_h^n - u_h(t_n) = W_h^n - w_h(t_n)$ .  $\square$

**Remark 5.3.1.** *Note that high order estimates of the term  $II$  can be improved by*

$$\|II\|_{L^2(\Omega)} \leq C\tau^2 \|v_h\|_{L^2} \int_{\pi/\tau}^{\infty} e^{rt_n \cos \theta} r dr \leq C\tau^2 t_n^{-2} \|v\|_{L^2}$$

Next we turn to smooth initial data, i.e.,  $v \in \dot{H}^2(\Omega)$ . To this end, we first state some basic estimates on the solution kernel  $K_s(z) = -z^{-1}(z^\alpha I + A_h)^{-1}$ . The proof is analogous to that in Chapter 4, and hence omitted.

**Lemma 5.3.3.** *Let  $K_s(z) = -z^{-1}(z^\alpha I + A_h)^{-1}$ . Then*

$$\begin{aligned} \|K_s(z)\| &\leq C|z|^{-\alpha-1} \quad \text{and} \quad \|K'_s(z)\| \leq C|z|^{-\alpha-2} \quad \forall z \in \Sigma_{\pi-\theta}, \\ \|e^{-z\tau} K_s\left(\frac{1-e^{-z\tau}}{\tau}\right) - K_s(z)\| &\leq C\tau |z|^{-\alpha} \quad \forall z \in \Gamma_\tau. \end{aligned}$$

Now we state an estimate of the temporal error for smooth initial data.

**Theorem 5.3.2** ( $0 < \alpha < 1$ ). *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.10) with  $v \in \dot{H}^2(\Omega)$ ,  $U_h^0 = v_h = R_h v$  and  $f \equiv 0$ , respectively. Then there holds*

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

*Proof.* By Lemma 5.3.1, we may split the error into

$$\begin{aligned} u_h(t_n) - U_h^n &= \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_n} \left( K_s(z) - e^{-z\tau} K_s\left(\frac{1-e^{-z\tau}}{\tau}\right) \right) A_h v_h dz \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta} \setminus \Gamma_\tau} e^{zt_n} K_s(z) A_h v_h dz = I + II. \end{aligned}$$

By Lemma 5.3.3 and choosing  $\delta = 1/t_n$  we bound the term  $I$  by

$$\begin{aligned} \|I\|_{L^2(\Omega)} &\leq C_T \tau \|A_h v_h\|_{L^2(\Omega)} \left( \int_{\pi/(\tau \sin \theta)}^{\pi/(\tau \sin \theta)} e^{-rt_n \cos \theta} r^{-\alpha} dr + \int_{-\pi+\theta}^{\pi-\theta} e^{\cos \psi} t_n^{\alpha-1} d\psi \right) \\ &\leq C t_n^{\alpha-1} \tau \|A_h v_h\|_{L^2(\Omega)}. \end{aligned}$$

Appealing again to Lemma 5.3.3 yields

$$\begin{aligned} \|II\|_{L^2(\Omega)} &\leq C \|A_h v_h\|_{L^2(\Omega)} \int_{\pi/(\tau \sin \theta)}^{\infty} e^{-rt_n \cos \theta} r^{-\alpha-1} dr \\ &\leq C\tau \|A_h v_h\|_{L^2(\Omega)} \int_0^{\infty} e^{-rt_n \cos \theta} r^{-\alpha} dr \leq C\tau t_n^{\alpha-1} \|A_h v_h\|_{L^2(\Omega)}. \end{aligned}$$

The desired result follows from these two estimates and the identity  $A_h R_h = P_h A$ . □

**Remark 5.3.2.** *Note that high order estimates of the term  $II$  can be improved by*

$$\|II\|_{L^2(\Omega)} \leq C\tau^2 \|A_h v_h\|_{L^2(\Omega)} \int_0^{\infty} e^{-rt_n \cos \theta} r^{-\alpha+1} dr \leq C\tau^2 t_n^{\alpha-2} \|A_h v_h\|_{L^2(\Omega)}.$$

The error estimates in Theorem 5.2.1 follows from the triangle inequality and Section 3.1.

### 5.3.2 Error analysis for the SBD method

Now we derive the error estimates for the fully discrete scheme (5.12). Like before, we define  $W_h^n = U_h^n - v_h$ , which admits the following representation.

**Lemma 5.3.4.** *Let  $K(z) = -z^{-1}(z^\alpha I + A_h)^{-1}A_h$ . Then  $W_h^n$  can be represented by*

$$W_h^n = \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_n} \mu(e^{-\tau z}) K(\delta(e^{-\tau z})/\tau) v_h dz,$$

with  $\delta(\xi) = (1 - \xi) + (1 - \xi)^2/2$  and the contour  $\Gamma_\tau = \{z \in \Gamma_{\theta, \delta} : |\Im(z)| \leq \pi/\tau\}$ .

*Proof.* By the definition of scheme (5.12),  $W_h^n$  satisfies

$$\bar{\partial}_\tau^\alpha W_h^1 + A_h W_h^1 = -\frac{3}{2} A_h U_h^0 \quad \text{and} \quad \bar{\partial}_\tau^\alpha W_h^n + A_h W_h^n = -\frac{1}{2} A_h U_h^0 \quad (5.18)$$

with  $W_h^0 = 0$ . Multiplying by  $\xi^n$  (5.18), respectively, and summing from 1 to  $\infty$ , we obtain

$$(\delta(\xi)/\tau)^\alpha \widetilde{W}_h(\xi) + A_h \widetilde{W}_h(\xi) = -\frac{3\xi - \xi^2}{2(1 - \xi)} A_h v_h,$$

where  $\delta(\xi) = (1 - \xi) + (1 - \xi)^2/2$ . Consequently

$$\widetilde{W}_h(\xi) = \frac{3\xi - \xi^2}{2(1 - \xi)} B(\delta(\xi)/\tau) v_h,$$

with  $B(z) = -A_h((\delta(\xi)/\tau)^\alpha + A_h)^{-1}$ . Since  $\widetilde{W}_h(\xi)$  is analytic at  $\xi = 0$ , for  $\varrho$  small enough,  $W_h^n$ , the coefficient  $W_h^n$  of the Taylor series expansion of  $\widetilde{W}_h(\xi)$ , according to Cauchy's theorem, is given by

$$W_h^n = \frac{1}{2\pi i} \int_{|\xi|=\varrho} \frac{3\xi - \xi^2}{2(1 - \xi)\xi^{n+1}} B(p(\xi)/\tau) v_h d\xi = \frac{1}{2\tau\pi i} \int_{|\xi|=\varrho} \xi^{-n-1} \mu(\xi) K(p(\xi)/\tau) v_h d\xi,$$

where  $\mu(\xi) = \xi(3 - \xi)^2/4$ . Analogous to the proof of Lemma 5.3.1, we change the variable  $\xi = e^{-\tau z}$  and deform the integral contour to arrived at the desired representation.  $\square$

We state an analogue of Lemma 5.3.2 for the second-order scheme, which is key to the error analysis.

**Lemma 5.3.5.** *Let  $\mu(\xi) = \xi(3 - \xi)^2/4$ , and  $\theta \in (\pi/2, 3\pi/4)$ . Then for any  $z \in \Gamma_\tau$ , we have*

$$\|\mu(e^{-\tau z}) K(\delta(e^{-\tau z})/\tau) - K(z)\| \leq C\tau^2 |z|.$$

*Proof.* By the triangle inequality

$$\begin{aligned} \|\mu(e^{-\tau z})K(\delta(e^{-\tau z})/\tau) - K(z)\| &\leq |\mu(e^{-\tau z}) - 1| \|K(\delta(e^{-\tau z})/\tau)\| \\ &+ \|K(\delta(e^{-\tau z})/\tau) - K(z)\| = I + II. \end{aligned} \quad (5.19)$$

By Taylor expansion of the exponential function, there holds

$$\begin{aligned} \mu(e^{-\tau z}) &= 4^{-1}(9e^{-\tau z} - 6e^{-2\tau z} + e^{-3\tau z}) \\ &= 1 + (\tau z)^2 \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+2)!} (9(\tau z)^n - 6(2\tau z)^n + (3\tau z)^n). \end{aligned}$$

Now by observing  $|\tau z| \leq \pi/\sin\theta$ , the infinite series converges and can be bounded independent of  $z$ , and thus

$$|\mu(e^{-\tau z}) - 1| \leq C\tau^2|z|^2. \quad (5.20)$$

Next we let  $\chi(z) = p(e^{-\tau z})/\tau = (3/2 - 2e^{-\tau z} + e^{-2\tau z}/2)/\tau$ . Like before, we split the contour  $\Gamma_\tau$  into  $\Gamma_\tau = \Gamma_\tau^+ \cup \Gamma_\tau^c \cup \Gamma_\tau^-$ . Then

$$|\chi(z)| = |z| \left| \frac{3 - 4e^{-\tau z} + e^{-2\tau z}}{2\tau z} \right|.$$

Upon setting  $\zeta = -z\tau = re^{-i\theta}$  and  $r' = r \cos\theta$  and  $\phi = r \sin\theta$ , we obtain

$$\left| \frac{3 - 4e^{-z\tau} + e^{-2\tau z}}{2z\tau} \right| = \left| \frac{4e^\zeta - e^{2\zeta} - 3}{2\zeta} \right| = \frac{((e^{2r'} - 4e^{r'} \cos\phi + 3)^2 + 4e^{2r'} \sin^2(\phi))^{1/2}}{2r}.$$

This together with the inequality  $(a^2 + b^2)^{1/2} \geq (|a| + |b|)/\sqrt{2}$  for  $a, b \in \mathbb{R}$  yields

$$\left| \frac{3 - 4e^{-z\tau} + e^{-2\tau z}}{2z\tau} \right| \geq \frac{|e^{2r'} - 4e^{r'} \cos\phi + 3|}{2\sqrt{2}r} + \frac{e^{r'} \sin\phi}{\sqrt{2}r}.$$

The quantity on the right hand side is bounded from above and below by positive constants, by noting  $r \leq \pi/\sin\theta$ . Hence we obtain that

$$c|z| \leq |\chi(z)| \leq C|z|. \quad (5.21)$$

for some constants  $c, C > 0$  independent of  $z$ . This together with (5.20) and Lemma 5.3.2 implies

$$I = |\mu(e^{-\tau z}) - 1| \|K(\chi(z))\| \leq C\tau^2|z|.$$



Next we bound the second term  $II$ . By the mean value theorem, there exists some  $\eta \in [0, 1]$ ,  $z_\eta = \eta\chi(z) + (1 - \eta)z$ , such that

$$\|K(\chi(z)) - K(z)\| = \|K'(z_\eta)\| |\chi(z) - z| \leq C|z_\eta|^{-2}|z|^3\tau^2,$$

where the last inequality follows from Lemma 5.3.2 and the estimate

$$|\chi(z) - z| \leq \frac{1}{2}\tau^2|z|^3 \sum_{j=3}^{\infty} \frac{|\tau z|^{j-3}}{j!} (2^j - 4) \leq C\tau^2|z|^3.$$

Thus it suffices to show  $|z_\eta| > c|z|$ . We note that Taylor expansion yields

$$|z_\eta| = |z| \left| 1 + \eta \left( \sum_{j=2}^{\infty} (-1)^{j+1} \frac{(2^j - 2)(z\tau)^j}{(j+1)!} \right) \right|,$$

and  $|z\tau| \leq 1$  for  $z \in \Gamma_\tau^c$ . Hence  $|z_\eta| \geq c|z|$  for  $z \in \Gamma_\tau^c$ .

Next we show  $\Im(\chi(z)) \geq 0$  for  $z \in \Gamma_\tau^+$ . Let  $\tau z = re^{i(\pi-\theta)}$ . Then

$$\begin{aligned} \Im(\chi(z)) &= \tau^{-1}(4e^{r \cos \theta} \sin r \sin \theta - e^{2r \cos \theta} \sin 2r \sin \theta) \\ &= 2\tau^{-1}e^{r \cos \theta} \sin r \sin \theta(2 - e^{r \cos \theta} \cos r \sin \theta). \end{aligned}$$

We note that for any fixed  $\theta \in (0, \pi/2)$ , the function  $e^{r \cos \theta}$  is increasing in  $r$ , whereas  $\cos r \sin \theta$  is decreasing in  $r$  ( $r \sin \theta \leq \pi$ ). Obviously, for  $r \sin \theta \in [\pi/2, \pi]$ ,  $\Im(\chi) \geq 0$ . It suffices to consider the case  $r \sin \theta \in [0, \pi/2]$ . To this end, we consider the extreme of the function  $f(r) = 2 - e^{r \cos \theta} \cos r \sin \theta$ :

$$f'(r) = -e^{r \cos \theta}((\cos \theta)(\cos r \sin \theta) - (\sin \theta)(\sin r \sin \theta)) = -e^{r \cos \theta} \cos(\theta + r \sin \theta).$$

By noting the range of  $\theta$  and  $r \sin \theta$ , the only interior critical point is achieved at  $\theta + r \sin \theta = \pi/2$ , i.e.,  $r \sin \theta = \pi/2 - \theta$ , and at this point,  $\Im(\chi)$  achieves its minimum with a value

$$2\tau^{-1}e^{r \cos \theta} \sin r \sin \theta(2 - e^{(\pi/2-\theta) \cos \theta / \sin \theta} \sin \theta).$$

For  $\theta \in (\pi/4, \pi/2)$ , the term in the bracket is positive, and thus  $\Im(\chi) \geq 0$ . This and (5.21) yield for

$z \in \Gamma_\tau^+$

$$|z_\eta| \geq \min(|z|, |\chi(z)|) \cos \frac{\pi - \theta}{2} \geq c|z|.$$

Hence for  $z \in \Gamma_\tau$ , we derive a bound for the second term  $II$

$$II = \|K(\delta(e^{-\tau z})/\tau) - K(z)\| \leq C\tau^2|z|,$$

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

Now we can state the time discretization error for the scheme (5.12).

**Lemma 5.3.6.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.12) with  $v \in L^2(\Omega)$ ,  $U_h^0 = v_h = P_h v \in L^2(\Omega)$  and  $f \equiv 0$ , respectively. Then there holds*

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

*Proof.* Lemma 5.3.4 and the representation of  $w_h$  in Lemma 5.3.1 give the following representation of the error

$$\begin{aligned} W_h^n - w_h(t_n) &= \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_n} (\mu(e^{-\tau z})K(\delta(e^{-\tau z})/\tau) - K(z)) v_h dz \\ &\quad - \frac{1}{2\pi i} \int_{\Gamma_{\delta, \pi-\theta} \setminus \Gamma_\tau} e^{zt_n} K(z) v_h dz = I + II. \end{aligned}$$

For the term  $I$ , we appeal to Lemma 5.3.5 and the choice  $\delta = 1/t_n$  to deduce

$$\begin{aligned} \|I\|_{L^2(\Omega)} &\leq C_T \tau^2 \|v_h\|_{L^2(\Omega)} \left( \int_0^{\pi/(\tau \sin \theta)} e^{-rt_n \cos \theta} r dr + \int_{-\pi+\theta}^{\pi-\theta} e^{\cos \psi} t_n^{-2} d\psi \right) \\ &\leq C_T t_n^{-2} \tau^2 \|v_h\|_{L^2(\Omega)}. \end{aligned}$$

Then the error estimate for nonsmooth data follows by this, Remark 5.3.1 and Lemma 3.1.1.  $\square$

Next we turn to the case of smooth initial data, i.e.,  $v \in \dot{H}^2(\Omega)$ . We shall need the following estimate. The proof is similar to Lemma 5.3.5 and hence omitted.

**Lemma 5.3.7.** *Let  $K_s(z) = -z^{-1}(z^\alpha + A_h)^{-1}$ ,  $\mu(\xi) = \xi(3 - \xi)^2/4$ , and  $\theta \in (\pi/4, \pi/2)$ . Then there*

exists a constant  $C$  independent of  $z$  such that

$$\|\mu(e^{-z\tau})K_s(\delta(e^{-\tau z})/\tau) - K_s(z)\| \leq C\tau^2(1 + |z|^\alpha) \quad \forall z \in \Gamma_\tau.$$

**Lemma 5.3.8.** *Let  $u_h$  and  $U_h^n$  be the solutions of problem (3.1) and (5.12) with  $v \in \dot{H}^2(\Omega)$ ,  $U_h^0 = v_h = R_h v$  and  $f \equiv 0$ , respectively. Then there holds*

$$\|u_h(t_n) - U_h^n\|_{L^2(\Omega)} \leq C\tau^2 t_n^{\alpha-2} \|v\|_{\dot{H}^2(\Omega)}.$$

*Proof.* For  $v \in D(A) = \dot{H}^2(\Omega)$ , the error representation (4.21) can be rewritten as

$$\begin{aligned} W_h^n - w_h(t_n) &= \frac{1}{2\pi i} \int_{\Gamma_\tau} e^{zt_n} (\mu(e^{-\tau z})K_s(\delta(e^{-\tau z})/\tau) - K_s(z)) A_h v_h dz \\ &\quad - \frac{1}{2\pi i} \int_{\Gamma_{\theta,\delta} \setminus \Gamma_\tau} e^{zt_n} \widehat{K}_s(z) A_h v_h dz = I + II. \end{aligned} \tag{5.22}$$

By Lemma 5.3.7 and the choice  $\delta = t_n^{-1}$ , we derive the following bound:

$$\begin{aligned} \|I\|_{L^2(\Omega)} &\leq C_T \tau^2 \|A_h v_h\|_{L^2(\Omega)} \left( \int_0^{\pi/(\tau \sin \theta)} e^{-rt_n \cos \theta} r^\alpha dr + \int_{-\pi+\theta}^{\pi-\theta} e^{\cos \psi} t_n^{-1-\alpha} d\psi \right) \\ &\leq C_T t_n^{-1+\alpha} \tau^2 \|A_h v_h\|_{L^2(\Omega)}. \end{aligned}$$

This, Remark 5.3.2 and the identity  $A_h R_h = P_h A$  yield the desired estimate.  $\square$

The error estimates in Theorem 5.2.2 follow from Lemmas 5.3.6 and 5.3.8, the triangle inequality and Section 3.1.

## 5.4 Numerical results

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 5.2. 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 fractional diffusion equation was already studied in [28, 25], 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 5.1. For comparison we have also included the numerical results obtained by the three existing schemes described in [29, Section 2]. In all tables, **rate** refers to the empirical convergence rate of the errors when the time step size  $\tau$  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. We also present the numerical results for the L1 time stepping schemes discussed in Chapter 4. As to the two schemes due to Zeng et al [77], 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 indicates that existing time stepping schemes may not work well for nonsmooth data, whereas our fully discrete schemes are robust and accurate.

Table 5.1: The  $L^2$ -norm of the error for example (a) at  $t = 0.1$  with  $h = 2^{-12}$ .

$\alpha \backslash N$	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	6.57e-7	1.66e-7	4.44e-8	1.98 (2.00)
	L1	4.48e-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)
	L1	4.15e-3	1.96e-3	9.50e-4	4.65e-4	2.29e-5	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	1.65e-2	8.36e-3	4.21e-3	2.11e-3	1.06e-3	5.30e-4	1.00 (1.00)
	SBD	8.61e-4	2.20e-4	5.54e-5	1.39e-5	3.49e-6	8.82e-7	1.99 (2.00)
	L1	1.75e-2	8.72e-3	4.31e-3	2.13e-3	1.04e-3	5.18e-4	1.01
	Zeng I	9.94e-2	6.77e-2	4.60e-2	3.10e-2	2.07e-2	1.35e-2	0.55
	Zeng II	1.71e-2	1.75e-3	3.09e-4	1.60e-4	8.17e-5	4.12e-5	1.00

**Numerical results for examples (b) and (c):** In Table 5.2, we report the numerical results for examples (b) and (c) with  $\alpha = 0.5$ ; see also Figure 5.1 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 5.2. Further, if the spatial error is negligible,  $N$  is fixed and  $t_N \rightarrow 0$ , then we deduce from Theorem 5.2.1 and interpolation that

$$\|U_h^N - u(t_N)\|_{L^2(\Omega)} \leq C t_N^{q\alpha/2} N^{-1} \|v\|_{\dot{H}^q(\Omega)}. \quad (5.23)$$

In Table 5.3 and Figure 5.2 we show the  $L^2$ -norm of the error for examples (b) and (c), for fixed  $N = 10$  and  $t_N \rightarrow 0$  with  $\alpha = 0.5$ . It is observed that in the smooth case (b), the temporal error decreases like  $O(t^{1/2})$ , whereas in the nonsmooth case (c), it decays like  $O(t^{1/8})$ . Note that in example (c), the initial data  $v \in \dot{H}^{1/2-\epsilon}(\Omega)$  for any  $\epsilon > 0$ , (5.23) predicts an error decay rate  $O(t^{\alpha/4}) = O(t^{1/8})$ . Hence the empirical rates in Table 5.3 and Figure 5.2 agree well with the theoretical prediction.

Table 5.2: The  $L^2$ -norm of the error for examples (b) and (c) for  $t = 0.1$ ,  $\alpha = 0.5$ , and  $h = 2^{-9}$ .

$N$	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)

Table 5.3: The  $L^2$ -norm of the error for examples (b) and (c) as  $t \rightarrow 0$  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)

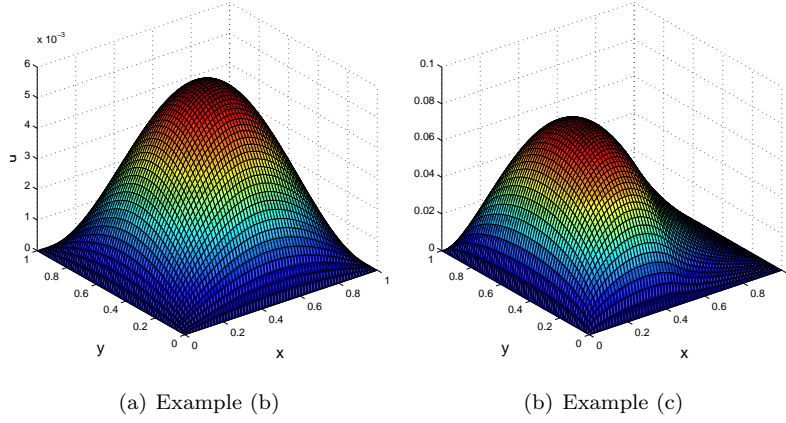


Figure 5.1: Numerical solutions of examples (b) and (c) by SBD with  $h = 2^{-7}$  and  $N = 1000$ ,  $\alpha = 0.5$  at  $t = 0.1$

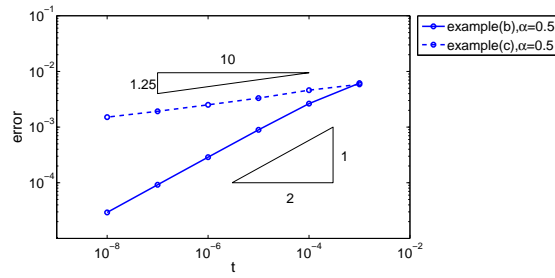


Figure 5.2: Numerical results for examples (b) and (c) using the BE method with  $h = 2^{-9}$  and  $N = 10$ ,  $\alpha = 0.5$  for  $t \rightarrow 0$ .

## 5.5 Extensions

In this section, we show that the framework discussed in Sections 5.1–5.4 can be extended to other fractional models, such as diffusion-wave equation (2.22) and multi-term model (2.15).

### 5.5.1 Extension to the diffusion-wave equation

Now we first establish the fully discrete schemes for diffusion-wave equations (2.22). To this end, we recall (2.3) and noting that

$${}^C \partial_t^\alpha \varphi = {}^R \partial_t^\alpha (\varphi(t) - \varphi(0) - t\varphi'(0)), \quad \text{for } \alpha \in (1, 2).$$

Hence we may rewrite the spatial semidiscrete scheme (3.55) by

$$\partial_t^\alpha(u_h - v_h - tb_h) + A_h u_h = f_h. \quad (5.24)$$

Using the same framework in Section 5.1 we get the approximation of the diffusion-wave equation. In this case the fully discrete scheme is: find  $U_h^n$  for  $n = 1, 2, \dots$ , such that

$$\bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n = \bar{\partial}_\tau^\alpha v_h + (\bar{\partial}_\tau^\alpha t)b_h + F_h^n, \quad \text{with } U_h^0 = v_h, \quad F_h^n = P_h f(t_n), \quad (5.25)$$

where  $\bar{\partial}_\tau^\alpha$  is defined in Section 5.1 based on  $\delta(\xi) = 1 - \xi$ .

Further, the second-order fully discrete scheme reads

$$\begin{aligned} \bar{\partial}_\tau^\alpha U_h^1 + A_h U_h^1 + \frac{1}{2} A_h U_h^0 &= \bar{\partial}_\tau^\alpha U_h^0 + \bar{\partial}_\tau^\alpha (tb_h) + F_h^1 + \frac{1}{2} F_h^0, \\ \bar{\partial}_\tau^\alpha U_h^n + A_h U_h^n &= \bar{\partial}_\tau^\alpha U_h^0 + \bar{\partial}_\tau^\alpha (tb_h) + F_h^n, \quad 2 \leq n \leq N, \end{aligned} \quad (5.26)$$

with  $U_h^0 = v_h$  and  $F_h^n = P_h f(t_n)$ . Here  $\bar{\partial}_\tau^\alpha$  is defined in Section 5.1 based on  $\delta(\xi) = (1 - \xi) + (1 - \xi)^2/2$ .

Then applying Lemma 5.2.1 and the same argument as that in Section 5.2.1 we arrive at the following error bounds for the fully discrete scheme (5.25) for the diffusion-wave equation.

**Theorem 5.5.1** ( $1 < \alpha < 2$ , backward Euler). *Let  $u$  and  $U_h^n$  be the solutions of problem (2.22) and (5.25) 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*

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leq C \left( (\tau t_n^{-1} + h^2 t_n^{-\alpha}) \|v\|_{L^2(\Omega)} + (\tau + h^2 t_n^{1-\alpha}) \|b\|_{L^2(\Omega)} \right).$$

Then for the fully discrete scheme (5.26), we may derive the following error estimate by Theorems 3.5.1, 3.5.1, Lemma 5.2.4 and the same argument as that in Section 5.2.2.

**Theorem 5.5.2** ( $1 < \alpha < 2$ , second-order backward scheme). *Let  $u$  and  $U_h^n$  be the solutions of problem (2.22) and (5.26) 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

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

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

$$\|u(t_n) - U_h^n\|_{L^2(\Omega)} \leq C \left( (\tau^2 t_n^{-2} + h^2 t_n^{-\alpha}) \|v\|_{L^2(\Omega)} + (\tau^2 t_n^{-1} + h^2 t_n^{1-\alpha}) \|b\|_{L^2(\Omega)} \right).$$

### 5.5.2 Extension to the multi-term fractional diffusion

In this part, we consider the multi-term fractional differential equation (2.15). The spatial semidiscrete scheme (3.48) can be written as

$$P(\bar{\partial}_\tau)(u_h - v_h) + A_h u_h = f_h. \quad (5.27)$$

Using the same framework in Section 5.1 we get the the fully discrete scheme. The backward Euler scheme is: find  $U_h^n$  for  $n = 1, 2, \dots$ , such that

$$P(\bar{\partial}_\tau)U_h^n + A_h U_h^n = P(\bar{\partial}_\tau)v_h + (P(\bar{\partial}_\tau)t)b_h + F_h^n, \quad \text{with } U_h^0 = v_h, \quad F_h^n = P_h f(t_n). \quad (5.28)$$

where the symbol  $P(\bar{\partial}_\tau)$  is defined by  $P(\bar{\partial}_\tau) = \bar{\partial}_\tau^\alpha + \sum_{k=1}^m b_k \bar{\partial}_\tau^{\alpha_k}$  and  $\bar{\partial}_\tau^\alpha$  is defined in Section 5.1 based on  $\delta(\xi) = 1 - \xi$ .

Further, the second-order backward finite difference scheme reads

$$\begin{aligned} P(\bar{\partial}_\tau)U_h^1 + A_h U_h^1 + \frac{1}{2}A_h U_h^0 &= P(\bar{\partial}_\tau)U_h^0 + \bar{\partial}_\tau^\alpha(t b_h) + F_h^1 + \frac{1}{2}F_h^0, \\ P(\bar{\partial}_\tau)U_h^n + A_h U_h^n &= P(\bar{\partial}_\tau)U_h^0 + P(\bar{\partial}_\tau)(t b_h) + F_h^n, \quad 2 \leq n \leq N, \end{aligned} \quad (5.29)$$

with  $U_h^0 = v_h$  and  $F_h^n = P_h f(t_n)$ . Here  $\bar{\partial}_\tau^\alpha$  is defined in Section 5.1 based on  $\delta(\xi) = (1 - \xi) + (1 - \xi)^2/2$ .

Then the following error estimates follows from the same argument as that in Section 5.1.

**Theorem 5.5.3** (Multi-term). *Let  $u$  and  $U_h^n$  be the solutions of problem (2.15). Then the following estimates hold for  $0 \leq q \leq 2$ .*



(a) If  $U_h^n$  is the solution of the fully discrete scheme (5.28) with  $v_h = P_h v$ , then

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

(b) If  $U_h^n$  is the solution of the fully discrete scheme (5.29) with  $v_h = P_h v$ , then

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

### 5.5.3 Numerical results

In this section, we present numerical results for diffusion-wave equation. We consider the following five examples (with  $\epsilon \in (0, 1/2)$ ):

(a)  $\Omega = (0, 1)$ ,  $b = 0$ , (a1)  $v = x(1-x) \in \dot{H}^2(\Omega)$  and (a2)  $v = 1 \in \dot{H}^{1/2-\epsilon}(\Omega)$ .

(b)  $\Omega = (0, 1)$ ,  $v = 0$ , (b1)  $b = x\chi_{[0,1/2)} + (1-x)\chi_{[1/2,1]} \in \dot{H}^{3/2-\epsilon}(\Omega)$  and  $b = x^{-1/4} \in \dot{H}^{1/4-\epsilon}(\Omega)$ .

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

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 (a):** From Table 5.4 and Figure 3.3 for example (a1) we observe a temporal convergence rate  $O(\tau)$  and  $O(\tau^2)$  for the BE and SBD scheme, respectively. These results are in full agreement with the analysis in Section 5.5.1. These observations remain valid for nonsmooth data (a2), cf. Tables 5.5. Further, we observe that in Table 5.5, the convergence of the SBD for  $\alpha = 1.9$  is only of order  $O(\tau^{1.68})$ . 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 5.6. In Tables 5.4 and 5.5, we present also the results by the popular Crank-Nicolson scheme [72, 79, 39], which converges at a rate  $O(\tau^{3-\alpha})$  for smooth solutions [72]. For smooth data, i.e., example (a1), 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.

Table 5.4: The  $L^2$ -norm of the error for example (a1) 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	$\approx 0.99$ (1.00)
	SBD	4.76e-4	1.05e-4	2.47e-5	5.97e-6	1.46e-6	$\approx 2.03$ (2.00)
	CN	9.00e-5	2.73e-5	8.95e-6	2.88e-6	8.94e-7	$\approx 1.66$
1.5	BE	4.95e-3	2.59e-3	1.32e-3	6.69e-4	3.36e-4	$\approx 0.98$ (1.00)
	SBD	4.25e-4	1.11e-4	2.79e-5	6.98e-6	1.76e-6	$\approx 1.99$ (2.00)
	CN	1.24e-3	4.59e-4	1.67e-4	6.00e-5	2.15e-5	$\approx 1.47$
1.9	BE	5.38e-3	2.72e-3	1.37e-3	6.89e-4	3.45e-4	$\approx 1.00$ (1.00)
	SBD	3.08e-4	8.81e-5	2.41e-5	6.28e-6	1.60e-6	$\approx 1.95$ (2.00)
	CN	4.33e-3	2.05e-3	9.67e-4	4.54e-5	2.12e-5	$\approx 1.09$

Like before, if the spatial error is negligible then for fixed  $N$  and  $t_N \rightarrow 0$ , the error decay estimate (5.23) holds. In Table 5.7, 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 (5.23).

Table 5.5: The  $L^2$ -norm of the error for example (a2) 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	$\approx 1.00$ (1.00)
	SBD	1.01e-3	1.77e-4	3.88e-5	9.13e-6	2.21e-6	5.35e-7	$\approx 2.07$ (2.00)
	CN	9.63e-2	6.67e-2	4.59e-2	3.14e-2	2.10e-2	1.35e-2	$\approx 0.63$
1.5	BE	2.58e-2	1.39e-2	7.30e-3	3.75e-3	1.90e-3	9.58e-4	$\approx 0.98$ (1.00)
	SBD	7.40e-3	1.93e-3	4.69e-4	1.16e-4	2.87e-5	7.12e-6	$\approx 2.02$ (2.00)
	CN	5.48e-2	3.49e-2	2.17e-2	1.28e-2	6.39e-3	1.91e-3	$\approx 0.97$
1.9	BE	7.11e-2	5.00e-2	3.37e-2	2.17e-2	1.31e-2	7.63e-3	$\approx 0.74$ (1.00)
	SBD	4.82e-2	2.78e-2	1.36e-2	5.28e-3	1.62e-3	4.25e-4	$\approx 1.68$ (2.00)
	CN	5.48e-2	3.49e-2	2.04e-2	1.10e-2	5.71e-3	2.94e-3	$\approx 0.92$

Table 5.6: The  $L^2$ -norm of the error for example (a2) 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	$\approx 0.91$ (1.00)
SBD	5.28e-3	1.62e-3	4.25e-4	1.06e-4	2.88e-5	$\approx 1.94$ (2.00)

Table 5.7: The  $L^2$ -norm of the error for examples (a1), (a2), (b1) and (b2) with  $\alpha = 1.5$ :  $t \rightarrow 0$ ,  $h = 2^{-13}$ , and  $N = 10$ .

$t$	1	1e-1	1e-2	1e-3	1e-4	1e-5	rate
(a1)	1.71e-2	4.95e-3	2.84e-4	9.52e-6	3.04e-7	1.13e-8	$\approx 1.47$ (1.50)
(a2)	1.54e-2	2.58e-2	1.06e-2	4.48e-3	1.87e-3	7.30e-4	$\approx 0.38$ (0.38)
(b1)	2.20e-2	1.56e-3	1.17e-5	8.79e-8	6.61e-10	5.73e-12	$\approx 2.11$ (2.13)
(b2)	1.75e-2	1.67e-3	1.02e-4	6.42e-5	4.14e-6	4.18e-7	$\approx 1.15$ (1.19)

**Numerical results for example (b):** Similarly to the results for examples (b1) and (b2), we observe a first-order and second-order convergence for the BE and SBD scheme, respectively, cf. Tables 5.8. All the convergence rates are independent of the fractional order  $\alpha$ . For the nonsmooth case, i.e., example (b2), we are particularly interested in the errors for  $t$  close to zero, thus we also plot the error at  $t = 0.1, 0.01$  and  $0.001$ . These results fully confirm the analysis in Section 5.5.1. Further, by Theorems 5.5.1 and 5.5.2

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

Hence, if we fix  $\alpha = 1.5$ ,  $N = 10$  and let  $t \rightarrow 0$ , this estimate predicts a behavior  $O(t_N^{17/8-3\epsilon/4})$  and  $O(t_N^{19/16-3\epsilon/4})$  for examples (b1) and (b2), respectively, if the spatial error is negligible, which agree with the results in Table 5.7.

Table 5.8: The  $L^2$ -norm of the error for examples (b1) and (b2) at  $t = 0.1$  with  $h = 2^{-13}$ .

Case	$\alpha$	$N$	10	20	40	80	160	320	rate
(f)	1.1	BE	1.99e-3	1.01e-3	5.12e-4	2.57e-4	1.29e-4	6.41e-5	$\approx 1.00$ (1.00)
		SBD	2.34e-5	4.15e-6	8.38e-7	1.83e-7	4.00e-8	7.15e-9	$\approx 2.31$ (2.00)
	1.5	BE	1.56e-3	7.83e-4	3.92e-4	1.96e-4	9.81e-5	4.91e-5	$\approx 1.00$ (1.00)
		SBD	1.12e-4	2.85e-5	7.17e-6	1.79e-6	4.43e-7	1.04e-7	$\approx 2.03$ (2.00)
	1.9	BE	8.78e-4	4.39e-4	2.20e-4	1.10e-4	5.52e-5	2.76e-5	$\approx 1.00$ (1.00)
		SBD	9.03e-5	2.56e-5	6.88e-6	1.76e-6	4.28e-7	9.65e-8	$\approx 2.02$ (2.00)
(g)	1.1	BE	1.58e-3	8.04e-4	4.06e-4	2.04e-4	1.02e-4	5.12e-5	$\approx 1.00$ (1.00)
		SBD	3.08e-5	6.54e-6	1.52e-6	3.67e-7	8.83e-8	2.00e-8	$\approx 2.08$ (2.00)
	1.5	BE	1.68e-3	8.71e-4	4.45e-4	2.25e-4	1.13e-4	5.68e-5	$\approx 0.99$ (1.00)
		SBD	2.18e-4	5.65e-5	1.40e-5	3.46e-6	8.39e-7	1.89e-7	$\approx 2.07$ (2.00)
	1.9	BE	2.92e-3	1.70e-3	9.70e-4	5.38e-4	2.91e-4	1.53e-4	$\approx 0.90$ (1.00)
		SBD	1.04e-3	4.64e-4	1.77e-4	5.63e-5	1.51e-5	3.54e-6	$\approx 1.92$ (2.00)

**Numerical results for example (c):** Finally, we present numerical solutions of the two-dimensional example. The temporal error is showed in Table 5.9. The empirical results fully confirm our analysis.

Table 5.9: The  $L^2$ -norm of the error for example (c) 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	$\approx 0.97$ (1.00)
SBD	1.60e-2	4.70e-3	1.29e-3	3.30e-4	8.23e-5	$\approx 1.95$ (2.00)

**Numerical results for multi-term time-fractional diffusion:** Now we present numerical results to confirm our theory for the multi-term counterpart, i.e.,

$$({}^C \partial_t^\alpha + \sum_{j=1}^m b_j {}^C \partial_t^{\beta_j})u(x, t) - \Delta u(x, t) = f(x, t) \quad (x, t) \in \Omega \times (0, T). \quad (5.30)$$

with  $1 > \alpha > \beta_1 > \beta_2 > \dots > \beta_m > 0$  and constant coefficients  $b_j > 0$ ,  $j = 1, 2, \dots, m$ . Error estimates for the space semidiscrete Galerkin FEM for problem (5.30) with nonsmooth initial data were established in [23]. Now we illustrate the scheme for  $\alpha = 0.5$ ,  $\beta = 0.1$  and  $b = 1$  with the nonsmooth initial data  $v = 1 \in \dot{H}^{1/2-\epsilon}(\Omega)$ . The results indicate that the BE and SBD schemes yield a first- and second-order convergence for (5.30), respectively, cf. Tables 5.10 and 5.11.

Table 5.10: 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	$\approx 1.00$ (1.00)
SBD	5.12e-4	1.19e-4	2.87e-5	7.06e-6	1.76e-6	$\approx 2.00$ (2.00)

Table 5.11: The  $L^2$ -norm of the error for multi-term time-fractional parabolic equation as  $t \rightarrow 0$  with  $h = 2^{-9}$  and  $N = 10$ .

$\tau$	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	$\approx 0.13$ (0.13)
SBD	3.00e-4	2.39e-4	1.80e-4	1.36e-4	1.05e-4	8.46e-5	$\approx 0.11$ (0.13)

## 5.6 Conclusions and comments

In this chapter we develop two simple fully discrete schemes for the fractional diffusion equations. The time stepping schemes employ the convolution quadrature generated by the backward Euler method and the second-order backward difference method. We provide a complete error analysis for both schemes, and derived optimal error estimates for both smooth and nonsmooth initial data. The analysis in Section 5.1 and the extension to the diffusion-wave equations were proposed in [29].

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 problems.

## 6. CONCLUSIONS AND FUTURE WORK

This dissertation has provided a rigorous analysis of various numerical schemes for fractional-order differential equations, which models anomalous diffusion phenomena in highly heterogeneous aquifers and complex viscoelastic materials. Our analysis focuses on the fractional diffusion model and then extends to diffusion-wave and multi-term counterparts.

Due to significant potentials of problem (2.7) in practical applications, a number of efficient schemes, notably based on finite difference in space and various discretizations in time, have been developed. The corresponding error analysis is often based on Taylor expansion and the error bounds are expressed in terms of certain smoothness of the solution. Many existing numerical methods for problem (1.1) require that the solution is a  $C^2$  or  $C^3$  function in time, cf. Table 1.1. Unfortunately, the high regularity requirement in the convergence analysis in these works is too restrictive. The goal of our research is to construct fully discrete schemes and establish optimal error bounds that are expressed directly in terms of the regularity of the problem data. We are especially interested in the case of nonsmooth data arising in many practical applications, e.g., inverse and control problems.

In Chapter 3 we have constructed two semidiscrete schemes by standard Galerkin finite element method and lumped mass method. Error estimates for homogeneous and inhomogeneous problems were established separately in terms of the smoothness of the data. In Chapter 4 we revisited the most popular fully discrete scheme based on L1-type approximation in time and Galerkin finite element method in space. The gap between the existing convergence theory and numerical results was filled. We showed the first order convergence by the discrete Laplace transform technique. Two fully discrete schemes based on convolution quadratures were developed in Chapter 5. The error bounds were given using two different techniques, i.e., discretized operational calculus and discrete Laplace transform. Our analysis can be easily extended to some other fractional-order differential models.

Next we list several perspectives for our future research:

1. In view of the solution singularity for time  $t$  close to zero, it is natural to consider the time stepping on a nonuniform time mesh in order to arrive at a uniform higher order convergence. The generating function approach used in our analysis does not work directly in this case, and it is still an open question to rigorously establish error estimates directly in terms of the data regularity.

2. A second interesting topic is to developing a fast algorithm for solving (2.7). Due to the nonlocal property of the fractional derivative, all previous solutions have to be saved in order to compute the solution at the current time level. Hence the fully discrete scheme based on finite difference/element method requires substantial storage. One possible choice is the proper orthogonal decomposition Galerkin finite element method [35]. We hope to derive error estimates depending on the number of POD basis functions and on the time discretization.
3. In Chapters 4 and 5, we only considered fully discrete schemes on homogeneous problem with initial data, i.e.,  $f \equiv 0$  since the generating function approach to inhomogeneous problems are still unclear. The technical difficulty is that the kernel  $\bar{E}(t)$  of solution representation (3.6) is weakly singular. That may requires higher regularity of the source data  $f$ . One open question is the development of robust error analysis for L1 time stepping on diffusion-wave equations even for the homogeneous case  $f \equiv 0$ .
4. Finally, it is of interest to analyze related control and inverse problems [32].



## REFERENCES

- [1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.
- [2] 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.
- [3] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- [4] D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo. *Fractional Calculus*. World Scientific, Hackensack, NJ, 2012. Models and numerical methods.
- [5] E. Bazhlekova. Properties of the fundamental and the impulse-response solutions of multi-term fractional differential equations. In *Complex Analysis and Applications '13, Sofia, 31 Oct.-2 Nov. 2013*, pages 55–64, 2013.
- [6] E. Bazhlekova, B. Jin, R. Lazarov, and Z. Zhou. An analysis of the Rayleigh-Stokes problem for the generalized second grade fluid. *Numer. Math.*, page in press. arXiv:1401.8049, 2014.
- [7] D. A. Benson, S. W. Wheatcraft, and M. M. Meerschaert. The fractional-order governing equation of Lévy motion. *Water Resour. Res.*, 36(6):1413–1423, 2000.
- [8] P. Chatzipantelidis, R. Lazarov, and V. Thomée. Some error estimates for the lumped mass finite element method for a parabolic problem. *Math. Comp.*, 81(277):1–20, 2012.
- [9] 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.
- [10] E. Cuesta, C. Lubich, and C. Palencia. Convolution quadrature time discretization of fractional diffusion-wave equations. *Math. Comp.*, 75(254):673–696, 2006.
- [11] M. Djrbashian. *Harmonic Analysis and Boundary Value Problems in the Complex Domain*. Birkhäuser, Basel, 1993.

- [12] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer-Verlag, New York, 2004.
- [13] V. J. Ervin and J. P. Roop. Variational formulation for the stationary fractional advection dispersion equation. *Numer. Methods Partial Diff. Eq.*, 22(3):558–576, 2006.
- [14] P. Flajolet. Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.*, 215(1-2):371–381, 1999.
- [15] N. J. Ford, J. Xiao, and Y. Yan. A finite element method for time fractional partial differential equations. *Fract. Calc. Appl. Anal.*, 14(3):454–474, 2011.
- [16] 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.
- [17] Y. Fujita. Integrodifferential equation which interpolates the heat and the wave equation. *Osaka J. Math.*, 27(2):309–321, 1990.
- [18] 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.
- [19] S. B. Hadid and Y. F. Luchko. An operational method for solving fractional differential equations of an arbitrary real order. *Panamer. Math. J.*, 6(1):57–73, 1996.
- [20] 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.
- [21] R. Hilfer and L. Anton. Fractional master equations and fractal time random walks. *Physical Review E*, 51(2):R848, 1995.
- [22] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge University Press, 2009.
- [23] B. Jin, R. Lazarov, Y. Liu, and Z. Zhou. The Galerkin finite element method for a multi-term time-fractional diffusion equation. *J. Comput. Phys.*, 281:825–843, 2015.
- [24] B. Jin, R. Lazarov, J. Pasciak, and W. Rundell. Variational formulation of problems involving fractional order differential operators. *Math. Comp.*, page in press. arXiv:1307.4795, 2015.

- [25] 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.
- [26] B. Jin, R. Lazarov, J. Pasciak, and Z. Zhou. Error analysis of a finite element method for a space-fractional parabolic equation. *SIAM J. Numer. Anal.*, 52(5):2272–2294, 2014.
- [27] B. Jin, R. Lazarov, J. Pasciak, and Z. Zhou. Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion. *IMA Numer. Anal.*, 35(2):561–582, 2015.
- [28] B. Jin, R. Lazarov, and Z. Zhou. Error estimates for a semidiscrete finite element method for fractional order parabolic equations. *SIAM J. Numer. Anal.*, 51(1):445–466, 2013.
- [29] B. Jin, R. Lazarov, and Z. Zhou. On two schemes for fractional diffusion and diffusion-wave equations. preprint, arXiv:1404.3800, 2014.
- [30] B. Jin, R. Lazarov, and Z. Zhou. An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data. *IMA J. Numer. Anal.*, page in press, 2015.
- [31] B. Jin and W. Rundell. An inverse problem for a one-dimensional time-fractional diffusion problem. *Inverse Problems*, 28(7):075010, 19, 2012.
- [32] B. Jin and W. Rundell. A tutorial on inverse problems for anomalous diffusion processes. *Inverse Problems*, 31(3):035003, 2015.
- [33] A. Kilbas, H. Srivastava, and J. Trujillo. *Theory and Applications of Fractional Differential Equations*. Elsevier, Amsterdam, 2006.
- [34] M. Knopp and S. Robins. Easy proofs of Riemann’s functional equation for  $\zeta(s)$  and of Lipschitz summation. *Proc. Amer. Math. Soc.*, 129(7):1915–1922, 2001.
- [35] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik*, 90(1):117–148, 2001.
- [36] 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.
- [37] L. Lewin. *Polylogarithms and associated functions*. North-Holland, 1981.

- [38] C. Li and H. Ding. Higher order finite difference method for the reaction and anomalous-diffusion equation. *Appl. Math. Model.*, page in press, 2014.
- [39] 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.
- [40] W. Li and D. Xu. Finite central difference/finite element approximations for parabolic integro-differential equations. *Computing*, 90(3-4):89–111, 2010.
- [41] 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.
- [42] Z. Li, Y. Liu, and M. Yamamoto. Initial-boundary value problems for multi-term time-fractional diffusion equations with positive constant coefficients. preprint, arXiv:1312.2112, 2013.
- [43] Z. Li and M. Yamamoto. Initial-boundary value problems for linear diffusion equations with multiple time-fractional derivatives. preprint, arXiv:1306.2778, 2013.
- [44] Y. Lin, X. Li, and C. Xu. Finite difference/spectral approximations for the fractional cable equation. *Math. Comp.*, 80(275):1369–1396, 2011.
- [45] Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.*, 225(2):1533–1552, 2007.
- [46] C. Lubich. Discretized fractional calculus. *SIAM J. Math. Anal.*, 17(3):704–719, 1986.
- [47] C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.*, 52(2):129–145, 1988.
- [48] 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.
- [49] Y. Luchko. Initial-boundary-value problems for the generalized multi-term time-fractional diffusion equations. *J. Math. Anal. Appl.*, 374(2):538–548, 2011.
- [50] Y. Luchko and R. Gorenflo. An operational method for solving fractional differential equations with the Caputo derivatives. *Acta Math. Vietnam.*, 24(2):207–233, 1999.

- [51] F. Mainardi. Fractional relaxation-oscillation and fractional diffusion-wave phenomena. *Chaos, Solitons & Fractals*, 7(9):1461–1477, 1996.
- [52] F. Mainardi. *Fractional Calculus and Waves in Linear Viscoelasticity*. Imperial College Press, London, 2010.
- [53] W. McLean. Regularity of solutions to a time-fractional diffusion equation. *ANZIAM J.*, 52(2):123–138, 2010.
- [54] W. McLean and K. Mustapha. Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. *Numer. Algor.*, 52(1):69–88, 2009.
- [55] W. McLean and K. Mustapha. Time-stepping error bounds for fractional diffusion problems with non-smooth initial data. *J. Comput. Phys.*, 293(15):201–217, 2015.
- [56] W. McLean and V. Thomée. Time discretization of an evolution equation via Laplace transforms. *IMA J. Numer. Anal.*, 24(3):439–463, 2004.
- [57] 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.
- [58] W. McLean, V. Thomée, and L. B. Wahlbin. Discretization with variable time steps of an evolution equation with a positive-type memory term. *J. Comput. Appl. Math.*, 69(1):49–69, 1996.
- [59] R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339(1):1–77, 2000.
- [60] R. Metzler and J. Klafter. The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A: Math. Gen.*, 37(31):R161–R208, 2004.
- [61] K. Mustapha. An implicit finite-difference time-stepping method for a sub-diffusion equation, with spatial discretization by finite elements. *IMA J. Numer. Anal.*, 31(2):719–739, 2011.
- [62] K. Mustapha, B. Abdallah, and K. Furati. A discontinuous petrov–galerkin method for time-fractional diffusion equations. *SIAM J. Numer. Anal.*, 52(5):2512–2529, 2014.

- [63] K. Mustapha and W. McLean. Piecewise-linear, discontinuous galerkin method for a fractional diffusion equation. *Numer. Algor.*, 56:159–184, 2011.
- [64] 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.
- [65] R. Nigmatulin. The realization of the generalized transfer equation in a medium with fractal geometry. *Phys. Stat. Sol. B*, 133:425–430, 1986.
- [66] K. B. Oldham and J. Spanier. *The Fractional Calculus*. Academic Press, New York, 1974.
- [67] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego, CA, 1999.
- [68] 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.
- [69] J. M. Sanz-Serna. A numerical method for a partial integro-differential equation. *SIAM J. Numer. Anal.*, 25(2):319–327, 1988.
- [70] H. Seybold and R. Hilfer. Numerical algorithm for calculating the generalized Mittag-Leffler function. *SIAM J. Numer. Anal.*, 47(1):69–88, 2008/09.
- [71] T. H. Solomon, E. R. Weeks, and H. L. Swinney. Observation of anomalous diffusion and lévy flights in a two-dimensional rotating flow. *Phys. Rev. Lett.*, 71(24):3975–3978, 1993.
- [72] Z.-Z. Sun and X. Wu. A fully discrete scheme for a diffusion wave system. *Appl. Numer. Math.*, 56(2):193–209, 2006.
- [73] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2006.
- [74] S. 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.
- [75] S. B. Yuste. Weighted average finite difference methods for fractional diffusion equations. *J. Comput. Phys.*, 216(1):264–274, 2006.

- [76] 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.
- [77] 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.
- [78] 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.
- [79] 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.