Numer Algor (2016) 73:517–534
DOI 10.1007/s11075-016-0106-y

ORIGINAL PAPER

A discontinuous Galerkin method for time fractional diffusion equations with variable coefficients

K. Mustapha[1] · B. Abdallah[2] · K. M. Furati[1] · M. Nour[1]

Received: 17 September 2015 / Accepted: 1 February 2016 / Published online: 13 February 2016

Abstract We propose a piecewise-linear, time-stepping discontinuous Galerkin method to solve numerically a time fractional diffusion equation involving Caputo derivative of order $\mu \in (0, 1)$ with variable coefficients. For the spatial discretization, we apply the standard continuous Galerkin method of total degree ≤ 1 on each spatial mesh elements. Well-posedness of the fully discrete scheme and error analysis will be shown. For a time interval $(0, T)$ and a spatial domain Ω, our analysis suggest that the error in $L^2\big((0, T), L^2(\Omega)\big)$-norm is $O(k^{2-\frac{\mu}{2}} + h^2)$ (that is, short by order $\frac{\mu}{2}$ from being optimal in time) where k denotes the maximum time step, and h is the maximum diameter of the elements of the (quasi-uniform) spatial mesh. However, our numerical experiments indicate optimal $O(k^2 + h^2)$ error bound in the stronger $L^\infty\big((0, T), L^2(\Omega)\big)$-norm. Variable time steps are used to compensate the singularity of the continuous solution near $t = 0$.

Keywords Fractional diffusion · Variable coefficients · Discontinuous Galerkin method · Convergence analysis

The support of the Science Technology Unit at KFUPM through King Abdulaziz City for Science and Technology (KACST) under National Science, Technology and Innovation Plan (NSTIP) project No. 13-MAT1847-04 is gratefully acknowledged.

✉ K. Mustapha
kassem@kfupm.edu.sa

[1] Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia

[2] Palestine Technical University-Kadoorie, Tulkarm, Palestine

1 Introduction

In this paper, we investigate a numerical solution that allows a time discontinuity for solving time fractional diffusion equations with variable diffusivity. Let Ω be a bounded convex polygonal domain in $\mathbb{R}^d$ ($d = 1, 2, 3$), with a boundary $\partial\Omega$, and $T > 0$ be a fixed time. Then the fractional model problem is given by:

$$\begin{aligned} {}^{\mathrm{c}}\mathrm{D}^{\mu} u(x,t) - \nabla \cdot (\mathcal{A}(x,t)\nabla u(x,t)) &= f(x,t) \quad \text{on } \Omega \times (0,T], \\ u(x,0) &= u_0(x) \quad \text{on } \Omega, \\ u(x,t) &= 0 \quad \text{on } \partial\Omega \times (0,T], \end{aligned} \tag{1}$$

where we assume that $\mathcal{A} \in \mathcal{C}^1([0,T], L^\infty(\Omega))$ and satisfies

$$0 < a_{\min} < \mathcal{A}(x,t) < a_{\max} < \infty \quad \text{on } \overline{\Omega} \times [0,T]. \tag{2}$$

Here, ${}^{\mathrm{c}}\mathrm{D}^{\mu}$ is the Caputo's fractional derivative defined by

$${}^{\mathrm{c}}\mathrm{D}^{\mu} v(t) = I^{1-\mu} v'(t) := \int_0^t \omega_{1-\mu}(t-s)\, v'(s)\, ds \quad \text{with} \quad \omega_{1-\mu}(t) := \frac{t^{-\mu}}{\Gamma(1-\mu)},$$

where throughout the paper, $0 < \mu < 1$. Noting that, $I^{1-\mu}$ is the Riemman Liouville fractional integral operator, and v' denotes the time partial derivative of v.

Over the past few decades, researchers have observed numerous biological, physical and financial systems in which some key underlying random motion conform to a model where the diffusion is anomalously slow (subdiffusion) and not to the classical model of diffusion. For instance, the fractional diffusion model problem (1) is known to capture well the dynamics of subdiffusion processes, in which the mean square variance grows at a rate slower than that in a Gaussian process. Two distinct approaches can be used for modelling fractional sub-diffusion. One is based on fractional Brownian motion and Langevin equations, this leads to a diffusion equation with a varying diffusion coefficient exhibiting a fractional power law scaling in time [22]. However, the other one is based on continuous time random walks and master equations with power law waiting time densities which leads to a diffusion equation with fractional order temporal derivatives operating on the spatial Laplacian [12].

Having variable diffusivity in the model problem (1) is indeed very interesting and also practically important. However, due to the additional difficulty in this case, there are only few papers in the existing literature which considered the numerical solution of (1) and only for *one-dimensional* spatial domain Ω. Alikhanov [1] proposed a finite difference scheme and $O(h^q + k^2)$ convergence (with $q \in \{2, 4\}$) was proved, where k is the temporal grid size and h is the spatial grid size. For time independent diffusivity, Zhao and Xu [26] constructed a compact $O(h^4+k^{2-\mu})$ difference scheme and a box-type scheme of order $O(h^2 + k^{2-\mu})$. Stabilities of both schemes were proved. In relation, for time fractional convection-diffusion problems, Cui [4] studied a compact exponential scheme. An $O(h^4 + k^{2-\mu})$ convergence rate was showen assuming that the coefficients of the model problem are constants. Saadatmandi et al. [21] investigated the Sinc-Legendre collocation method, and the accuracy of the numerical method was tested numerically without providing any stability or convergence results. In all these papers, the imposed regularity assumptions on the solution

u of (1) are not practically valid. Furthermore, in the numerical tests, only smooth solutions were considered which is unlikely the case in the presence of the Caputo derivatives.

The innovation of this paper is to investigate a piecewise linear time-stepping discontinuous Galerkin (DG) method, combined with the standard finite elements (FEs) in space of total degree ≤ 1 on each mesh elements, for solving numerically time fractional models with variable diffusion coefficients of the form (1). The DG methods have found numerous applications, including for the time discretization of fractional diffusion and fractional wave equations, [14, 15]. Their advantages include excellent stability properties and suitability for adaptive refinement based on a posteriori error estimate to handle problems with low regularity. We use quasi-uniform spatial meshes, however, variable time-steps will be employed to compensate the singular behaviour of the continuous solution near $t = 0$. The convergence analysis will be carried under suitable assumptions on u. The stability of the numerical scheme remains valid for high-order DG-FE methods. Moreover, our convergence analysis can be easily extended for high-order DG-FE methods, where some ideas from [16] will be used. More precisely, we need to modify the DG error analysis, while the FE error analysis is applicable for high-order approximations in space under some compatibility assumptions on the boundary of Ω. In conclusion, the proposed numerical scheme is dynamic in terms of the meshes, the degree of the approximate solutions, the regularity of the exact solutions, and also allows bounded spatial domains in $\mathbb{R}^d$ for $d = 1, 2, 3$. These are considerable advantages over the existing numerical methods mentioned in the above paragraph.

The present work is motivated by an earlier paper [14]. There in, the first author and McLean considered a piecewise-linear DG method for a fractional diffusion problem with a constant diffusivity:

$$u'(x,t) - {}^R\mathrm{D}^{1-\mu}\nabla^2 u(x,t) = f(x,t) \quad \text{for } (x,t) \in \Omega \times (0,T], \tag{3}$$

where ${}^R\mathrm{D}^{1-\mu}u := \frac{\partial}{\partial t}(I^\mu u)$ (Riemann–Liouville fractional derivative). Recently, high order hp-DG methods with exponential rates of convergence for fractional diffusion (3) and also for fractional wave equations were studied in [16, 19]. Noting that, when $\mathcal{A}$ is constant and $f \equiv 0$ in (1), one may look at (3) as an alternative representation of (1).

Numerical solutions for model problems of the form (1) with constant diffusion parameter $\mathcal{A}$ have attracted considerable interest in recent years. For *one-dimensional* spatial domains, Zhang et al. [24] studied a class of finite difference (FD) methods. Stability properties were provided. Another FD scheme in time (with $L1$ approximation for the Caputo fractional derivative) combined with the spatial fourth order compact difference approach was studied by Ren et al. [20]. Convergence rates of order $k^{1+\mu}+h^4$ were proved. Murillo and Yuste [13] presented an implicit FD method over non-uniform time steps. An adaptive procedure was described to choose the size of the time meshes. Lin and Xu [8] combined a FD scheme in time and a spectral method in space, and $O(k^{1+\mu} + r^{-m})$ accuracy was proved, where r is the spatial polynomial degree, and m is related to the regularity of the exact solution u. Later, Li and Xu [7] developed and analyzed a time-space spectral method. Zhao and Sun [25] combined an order reduction approach and $L1$ discretization of the fractional derivative. A box-type scheme was constructed and $O(k^{1+\mu} + h^2)$ convergence had been proved. Finite central differences in time combined with the FE method in space was studied by Li and Xu [6]. For a smooth u, $O(k^2 + h^{\ell+1})$ convergence was achieved where ℓ is the degree of the FE solutions in space. For a high-order local DG method for space discretization, we refer to the work by Xu and Zheng [23].

For *two-* or *three-dimensional* spatial domains with $\mathcal{A} = 1$ in (1), Brunner et al. [2] used an algorithm that couples an adaptive time stepping and adaptive spatial basis selection approach for the numerical solution of (1). A semi-discrete piecewise linear Galerkin FE and lumped mass Galerkin methods were studied by Jin et al. [5]. An optimal error with respect to the regularity error estimates was established for $f \equiv 0$ and non-smooth initial data u_0. For *three-dimensional* spatial domains, a fractional alternating direction implicit scheme was proposed and analyzed by Chen et al. [3]. Mustapha et al. [17] proposed *low-high* order time stepping discontinuous Petrov-Galerkin methods combined with FEs in space. Using variable time meshes, $O(k^{m+(1-\mu)/2} + h^{r+1})$ convergence rates were shown, where m and r are the degrees of approximate solutions in the time and spatial variables, respectively. Optimal convergence rates in both variables were demonstrated numerically. In [18], a hybridizable DG method in space was extensively studied by Mustapha et al..

The outline of the paper is as follows. Section 2 introduces a fully discrete DG-FE scheme. In Section 3, we prove the stability of the discrete solution and provide a remark about the existence and uniqueness of the numerical solution. Section 4 is devoted to introduce time and space projection operators that will be used later to show the convergence of the numerical scheme. The error analysis is given in Section 5. Using suitable refined time-steps (towards $t = 0$) and quasi-uniform spatial meshes, in the $L^2((0, T), L^2(\Omega))$-norm, $O(k^{2-\frac{\mu}{2}} + h^2)$ convergence is achieved. Section 6 is dedicated to present a sample of numerical test which illustrate that our error bounds are pessimistic. For a strongly graded time mesh, in the stronger $L^\infty((0, T), L^2(\Omega))$ $(L^\infty(L^2))$-norm, we observe optimal $O(k^2 + h^2)$ convergence rates. We also tested the performance of our scheme using FE solutions of total degree ≤ 2 (quadratic) and $O(h^3)$ convergence was observed for $k^2 \leq h^3$.

2 The numerical method

To describe our fully discrete DG-FE method, we introduce a time partition of the interval $[0, T]$ given by the points: $0 = t_0 < t_1 < \cdots < t_N = T$. We set $I_n = (t_{n-1}, t_n)$ and $k_n = t_n - t_{n-1}$ for $1 \leq n \leq N$ with $k := \max_{1\leq n\leq N} k_n$. Let $S_h \subseteq H_0^1(\Omega)$ denotes the space of continuous, piecewise polynomials of total degree ≤ 1 with respect to a quasi-uniform partition of Ω into conforming triangular finite elements, with maximum diameter h. Next, we introduce our time-space finite dimensional DG-FE space:

$$\mathcal{W} = \{w \in L^2((0, T), S_h) : \ \ w|_{I_n} \in \mathcal{P}_1(S_h) \text{ for } 1 \leq n \leq N\}$$

where $\mathcal{P}_1(S_h)$ denotes the space of linear polynomials in the time variable t, with coefficients in S_h. We denote the left-hand limit, right-hand limit and jump at t_n by

$$w^n := w(t_n) = w(t_n^-), \quad w_+^n := w(t_n^+), \quad [w]^n := w_+^n - w^n,$$

respectively. The weak form of the fractional diffusion equation in (1) is

$$\int_{I_n} \big[\langle {}^{\mathrm{c}}\mathrm{D}^{\mu} u, v\rangle + a\big(t, u, v\big)\big]\, dt = \int_{I_n} \langle f, v\rangle\, dt, \quad \forall\, v \in L^2\big(I_n, H^1(\Omega)\big). \tag{4}$$

Throughout the paper, $\langle\cdot,\cdot\rangle$ denotes the L^2-inner product and $\|\cdot\|$ is the associated norm, and $\|\cdot\|_m$ ($m \ge 1$) denotes the norm on the Sobolev space $H^m(\Omega)$.

For each fixed $t \in (0, T]$, $a(t, \cdot, \cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is the bilinear form

$$a(t, v, w) = \langle \mathcal{A}(\cdot, t)\nabla v, \nabla w\rangle = \int_{\Omega} \mathcal{A}(x, t)\nabla v(x) \cdot \nabla w(x)\, dx$$

associated with the operator $\nabla \cdot (\mathcal{A}(\cdot, t)\nabla)$ which is symmetric and positive definite (by (2)), that is, there exist positive constants c_0 and c_1 such that

$$c_0\|v(t)\|_1^2 \le |v(t)|_1^2 := a(t, v, v) \le c_1\|v(t)\|_1^2 \quad \forall\ \ v(t) \in H_0^1(\Omega)\,. \tag{5}$$

The DG-FE approximation $U \in \mathcal{W}$ is defined as follows: Given $U(t)$ for $0 \le t \le t_{n-1}$, the solution $U \in \mathcal{P}_1(S_h)$ on I_n is determined by requesting that for $1 \le n \le N$,

$$\int_{I_n} \Big[\langle {}^{\mathrm{c}}\mathrm{D}_{dg}^{\mu} U + \sum_{j=0}^{n-1} \omega_{1-\mu}(t - t_j)\, [U]^j, X\rangle + a\big(t, U, X\big)\Big]\, dt = \int_{I_n} \langle f, X\rangle\, dt,\ \forall\, X \in \mathcal{P}_1(S_h),$$

with $U_+^0 = U^0 \in S_h$ is a suitable approximation of the initial data u_0, where

$${}^{\mathrm{c}}\mathrm{D}_{dg}^{\mu} U(t) := \sum_{j=1}^{n} \int_{t_{j-1}}^{\min\{t_j, t\}} \omega_{1-\mu}(t - s)\, U'(s)\, ds \quad \text{for}\ \ t \in I_n\,.$$

Since

$$\begin{aligned} {}^{\mathrm{R}}\mathrm{D}^{\mu} U(t) \ &:= \ \frac{\partial}{\partial t}\int_0^t \omega_{1-\mu}(t - s) U(s)\, ds \\ &= \ {}^{\mathrm{c}}\mathrm{D}_{dg}^{\mu} U(t) + \omega_{1-\mu}(t) U^0 + \sum_{j=1}^{n-1} \omega_{1-\mu}(t - t_j)\, [U]^j \quad \text{for}\ \ t \in I_n, \end{aligned} \tag{6}$$

our scheme can be rewritten in a compact form as follows: for $1 \le n \le N$,

$$\int_{I_n} \big[\big\langle {}^{\mathrm{R}}\mathrm{D}^{\mu} U, X\big\rangle + a\big(t, U, X\big)\big]\, dt = \int_{I_n} \langle f + \omega_{1-\mu}(t) U^0, X\rangle\, dt \quad \forall\, X \in \mathcal{P}_1(S_h). \tag{7}$$

Noting that, since the DG-FE scheme (7) amounts to a square linear system, the existence of the numerical solution U follows from its uniqueness. The uniqueness follows immediately from the above stability property in Theorem 1.

3 Stability of the numerical solution

To show the stability of the DG-FE scheme (7), we claim first the identity: $v(t) = I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}v)(t)$ for any $v \in \mathcal{W}$.

Lemma 1 *If $v \in \mathcal{W}$, then*

$$v(t) = I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}v)(t) \quad \textit{for } t \in I_n \quad \textit{with } 1 \le n \le N.$$

Proof Since v has possible discontinuities at the time nodes $t_0, t_1, \cdots, t_{j-1}$, from (6),

$$ {}^{\mathrm{R}}\mathrm{D}^{\mu}v(s) = \omega_{1-\mu}(s)v_+^0 + \sum_{i=1}^{j-1}\omega_{1-\mu}(s-t_i)\,[v]^i + {}^{\mathrm{c}}\mathrm{D}^{\mu}_{dg}v(s) \;\text{ for } s \in I_j\,. \tag{8}$$

Applying the operator I^{μ} to both sides and using $I^{\mu}({}^{\mathrm{c}}\mathrm{D}^{\mu}_{dg}v)(t) = \int_0^t v'(s)\,ds$, we observe

$$\begin{aligned} I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}v)(t) &= v_+^0 + \sum_{j=2}^{n-1}\int_{I_j}\omega_{\mu}(t-s)\sum_{i=1}^{j-1}\omega_{1-\mu}(s-t_i)\,[v]^i\,ds \\ &\quad + \int_{t_{n-1}}^{t}\omega_{\mu}(t-s)\sum_{i=1}^{n-1}\omega_{1-\mu}(s-t_i)\,[v]^i\,ds + \int_0^t v'(s)ds \;\text{ for } t \in I_n\,. \end{aligned}$$

Now, changing the order of summations and rearranging the terms yield

$$\begin{aligned} I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}v)(t) &= v_+^0 + \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\int_{I_j}\omega_{\mu}(t-s)\omega_{1-\mu}(s-t_i)\,[v]^i\,ds \\ &\quad + \sum_{i=1}^{n-1}\int_{t_{n-1}}^{t}\omega_{\mu}(t-s)\omega_{1-\mu}(s-t_i)\,[v]^i\,ds + \sum_{j=1}^{n}\int_{t_{j-1}}^{\min\{t,t_j\}} v'(s)ds \\ &= v_+^0 + \sum_{i=1}^{n-2}\int_{t_i}^{t}\omega_{\mu}(t-s)\omega_{1-\mu}(s-t_i)\,[v]^i\,ds \\ &\quad + \int_{t_{n-1}}^{t}\omega_{\mu}(t-s)\omega_{1-\mu}(s-t_{n-1})\,[v]^{n-1}\,ds + \sum_{j=1}^{n}\int_{t_{j-1}}^{\min\{t,t_j\}} v'(s)ds\,. \end{aligned}$$

Integrating and simplifying, then we have

$$I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}v)(t) = v_+^0 + \sum_{i=1}^{n-1}[v]^i + \sum_{j=1}^{n-1}(v^j - v_+^{j-1}) + v(t) - v_+^{n-1} = v(t) \;\text{ for } t \in I_n\,.$$

□

We let $\mathcal{C}^0(J_n, L^2(\Omega))$ ($J_n := \cup_{j=1}^n I_j$) denote the space of functions $v : J_n \to L^2(\Omega)$ such that the restriction $v|_{I_j}$ extends to a continuous function on the closed interval $\overline{I}_j$ for $1 \le j \le n$. For later use, we let

$$\|v\|_{I_j} := \sup_{t\in I_j} \|v(t)\|, \quad \|v\|_{J_n} := \max_{j=1}^{n} \|v\|_{I_j} \quad \text{and} \quad \|v\|_J := \max_{j=1}^{N} \|v\|_{I_j} .$$

Next, we state some properties of the Riemman Liouville factional operators that will be used in our stability and convergence analysis of the numerical scheme.

Lemma 2 *For* $0 < \mu < 1$, *we have*

(i) *The operator* ${}^R\mathrm{D}^\mu$ *satisfies: for* $v \in \mathcal{W}$,

$$\int_0^T \langle {}^R\mathrm{D}^\mu v, v\rangle \, dt \ge \frac{2}{3} \cos\left(\frac{\mu\pi}{2}\right) T^{-\mu} \int_0^T \|v(t)\|^2 \, dt \, .$$

(ii) *The integral operator* I^μ *satisfies: for* $v, w \in \mathcal{C}^0(J_N, L^2(\Omega))$

$$\left|\int_0^T \langle I^\mu v, w\rangle \, dt\right|^2 \le \sec^2\left(\frac{\mu\pi}{2}\right) \int_0^T \langle I^\mu v, v\rangle \, dt \int_0^T \langle I^\mu w, w\rangle \, dt \, .$$

Proof The property (i) was proven in [11, Theorem A.1]. For the proof of the property (ii), see [19, Lemma 3.1]. □

The next theorem shows the stability of the DG-FE scheme (7).

Theorem 1 *Assume that* $U^0 \in L^2(\Omega)$ *and* $f \in L^2((0, T), L^2(\Omega))$. *Then,*

$$\int_0^T \|U\|_1^2 \, dt \le C T^{1-\mu} \|U^0\|^2 + C \int_0^T \|f\|^2 \, dt \, .$$

Proof Choosing $X = U$ in the DG-FE scheme (7), and then summing over n, we obtain

$$\int_0^T \left[\langle {}^\mathrm{R}\mathrm{D}^\mu U, U\rangle + a(t, U, U)\right] dt = \int_0^T \langle f + \omega_{1-\mu}(t) U^0, U\rangle \, dt.$$

Since $a(\cdot, U, U) \ge c_0 \|U\|_1^2$ by (5) and $\langle f, U\rangle \le \frac{1}{2c_0}\|f\|^2 + \frac{c_0}{2}\|U\|^2$, we have

$$\int_0^T \left[\langle {}^\mathrm{R}\mathrm{D}^\mu U, U\rangle + \frac{c_0}{2}\|U\|_1^2\right] dt \le \int_0^T \left(\langle \omega_{1-\mu}(t) U^0, U\rangle + \frac{1}{2c_0}\|f\|^2\right) dt \, .$$

Using the identity $U(t) = I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}U)(t)$ from Lemma 1, Lemma 2 (ii), the inequality $ab \leq \frac{a^2}{4} + b^2$, and the identity $I^{\mu}\omega_{1-\mu}(t) = 1$, yield

$$\begin{aligned}\int_0^T \langle \omega_{1-\mu}(t)U^0, U\rangle\, dt &= \int_0^T \langle \omega_{1-\mu}(t)U^0, I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}U)\rangle\, dt \\ &\leq \frac{1}{4}\int_0^T \left\langle {}^{\mathrm{R}}\mathrm{D}^{\mu}U, U\right\rangle dt + \sec^2(\mu\pi/2)\int_0^T \omega_{1-\mu}(t)(I^{\mu}\omega_{1-\mu})(t)\, dt\|U^0\|^2 \\ &\leq \frac{1}{4}\int_0^T \left\langle {}^{\mathrm{R}}\mathrm{D}^{\mu}U, U\right\rangle dt + C\,T^{1-\mu}\|U^0\|^2 . \end{aligned} \tag{9}$$

To complete the proof, we combine the above two equations and use the positivity property of the operator ${}^{\mathrm{R}}\mathrm{D}^{\mu}$ given by Lemma 2 (i). □

4 Space and time projections

In this section, we introduce space and time projections, and then derive some bounds and errors properties that will be used later in our convergence analysis.

Projection in space For each $t \in [0, T]$, the elliptic projection operator $R_h : H_0^1(\Omega) \to S_h$ is defined by

$$a(t, R_h v - v, \chi) = 0 \quad \forall\ \chi \in S_h . \tag{10}$$

Assume that the diffusivity coefficient function $\mathcal{A} \in \mathcal{C}^1([0, T], L^{\infty}(\Omega))$, then the projection error $\xi := R_h u - u$ has the approximation property [9, (3.2)]: for each $t \in [0, T]$,

$$\|\xi(t)\| + h\|\nabla\xi(t)\| \leq C\,h^2\|u(t)\|_2 \quad \text{for}\ \ u(t) \in H^2(\Omega) \cap H_0^1(\Omega) . \tag{11}$$

Moreover, by [9, (3.3)], we have

$$\|\xi'(t)\| \leq C\,h^2(\|u(t)\|_2 + \|u'(t)\|_2) \quad \text{for}\ \ u(t),\ u'(t) \in H^2(\Omega) \cap H_0^1(\Omega) . \tag{12}$$

Projection in time The local L^2-projection operator $\Pi_k : L^2(I_n, L^2(\Omega)) \to \mathcal{C}(I_n, \mathcal{P}_1(L^2(\Omega))$ defined by:

$$\int_{I_n} \langle \Pi_k v - v,\ w\rangle\, dt = 0\ \forall\ \ w \in \mathcal{P}_1(L^2(\Omega)) \quad \text{for}\ \ 1 \leq n \leq N,$$

where $\mathcal{P}_1(L^2(\Omega))$ is the space of linear polynomials in the time variable t, with coefficients in $L^2(\Omega)$. Explicitly,

$$\Pi_k v(t) = \frac{12}{k_n^3}(t - t_{n-\frac{1}{2}}) \int_{I_n} (s - t_{n-\frac{1}{2}})\, v(s)\, ds + \frac{1}{k_n}\int_{I_n} v(s)\, ds \quad \text{for}\ t \in I_n\, , \tag{13}$$

where $t_{n-\frac{1}{2}} := (t_{n-1}+t_n)/2$. Hence, for $v' \in L^1(I_n, L^2(\Omega))$,

$$(\Pi_k v)'(t) = \frac{12}{k_n^3}\int_{I_n}(s-t_{n-\frac{1}{2}})\,v(s)\,ds = \frac{6}{k_n^3}\int_{I_n}(t_n-t)(s-t_{n-1})\,v'(s)\,ds\,.$$

Thus, for $t \in I_n$ with $1 \le n \le N$, we have

$$\|\Pi_k v(t)\| \le \frac{4}{k_n}\int_{I_n}\|v(s)\|\,ds \quad \text{and} \quad \|(\Pi_k v)'(t)\| \le \frac{3}{2k_n}\int_{I_n}\|v'(s)\|\,ds\,. \tag{14}$$

Setting $\eta_v = \Pi_k v - v$, we have the well-known projection error bound: for $t \in I_n$

$$\|\eta_v(t)\| + k_n\,\|\eta_v'(t)\| \le C\,k_n^{\ell-1}\int_{I_n}\|v^{(\ell)}(s)\|\,ds \quad \text{for } \ell = 1,\,2, \quad \text{with } v^{(\ell)} := \frac{\partial^\ell v}{\partial t^\ell}\,. \tag{15}$$

Next, we show an error bound property of Π_k that involves the operator ${}^{\mathrm{R}}\mathrm{D}^{\mu}$.

Lemma 3 *Let $v^{(\ell)}\big|_{I_n} \in L^1(I_n, L^2(\Omega))$ for $1 \le n \le N$ and for $\ell \in \{1,\,2\}$. We have*

$$\int_{I_n}\left\langle{}^{\mathrm{R}}\mathrm{D}^{\mu}\eta_v, \eta_v\right\rangle dt \le C\,k_n^{1-\mu}\max_{j=1}^{n} k_j^{2\ell-2}\Big(\int_{I_j}\|v^{(\ell)}\|\,dt\Big)^2 \quad \textit{for } 1 \le n \le N.$$

Proof We integrate by parts and notice that

$$\begin{aligned}\int_{I_n}\left\langle{}^{\mathrm{R}}\mathrm{D}^{\mu}\eta_v, \eta_v\right\rangle dt &= \langle I^{1-\mu}\eta_v(t), \eta_v(t)\rangle\Big|_{t_{n-1}^+}^{t_n^-} - \int_{I_n}\langle I^{1-\mu}\eta_v, \eta_v'\rangle\,dt\\ &= \langle \mathcal{I}_n(t_n), \eta_v(t_n)\rangle - \int_{I_n}\langle \mathcal{I}_n(t), \eta_v'(t)\rangle\,dt,\end{aligned} \tag{16}$$

where for $t \in I_n$,

$$\begin{aligned}\mathcal{I}_n(t) \;&:= I^{1-\mu}\eta_v(t) - I^{1-\mu}\eta_v(t_{n-1})\\ &= \int_0^{t_{n-1}}[\omega_{1-\mu}(t-s)-\omega_{1-\mu}(t_{n-1}-s)]\eta_v(s)\,ds + \int_{t_{n-1}}^{t}\omega_{1-\mu}(t-s)\eta_v(s)\,ds\,.\end{aligned}$$

Simplifying then integrating, we observe

$$\begin{aligned}\|\mathcal{I}_n(t)\| &\le \Big(\int_0^{t_{n-1}}[\omega_{1-\mu}(t_{n-1}-s) - \omega_{1-\mu}(t-s)]\,ds + \int_{t_{n-1}}^{t}\omega_{1-\mu}(t-s)\,ds\Big)\|\eta_v\|_{J_n}\\ &\le 2\,\omega_{2-\mu}(k_n)\,\|\eta_v\|_{J_n} \quad \text{for } t \in I_n\,.\end{aligned}$$

Therefore, an application of the Cauchy-Schwarz inequality gives

$$\int_{I_n}|\left\langle{}^{\mathrm{R}}\mathrm{D}^{\mu}\eta_v, \eta_v\right\rangle|\,dt \le 2\,\omega_{2-\mu}(k_n)\,\|\eta_v\|_{J_n}\left(\|\eta_v(t_n)\| + \int_{I_n}\|\eta_v'\|\,dt\right),$$

and hence, using the error projection in (15), we obtain the desired bound. □

The next estimate will be used to show the convergence of our scheme

Lemma 4 *We have*

$$\left|\int_0^T \left\langle {}^R\mathrm{D}^\mu \xi, \eta_\xi \right\rangle dt\right| + \left|\int_0^T \left\langle {}^R\mathrm{D}^\mu \Pi_k \xi, \xi \right\rangle dt\right| \le C\, T^{1-\mu} \left(\|\xi(0)\| + \int_0^T \|\xi'\|\, dt\right)^2 .$$

Proof Since ${}^R\mathrm{D}^\mu \xi(t) = \omega_{1-\mu}(t)\xi(0) + I^{1-\mu}\xi'(t)$ and since $\|\eta_\xi(t)\| \le C \int_{I_n} \|\xi'\|\, dt$ for t_n (by the time projection estimate in (15)), we have

$$\int_{I_n} |\langle {}^R\mathrm{D}^\mu \xi, \eta_\xi \rangle|\, dt \le C \int_{I_n} [\omega_{1-\mu}(t)\|\xi(0)\| + I^{1-\mu}(\|\xi'\|)]\, dt \int_{I_n} \|\xi'\|\, ds.$$

Summing over n and then integrating,

$$\begin{aligned}\int_0^T |\langle {}^R\mathrm{D}^\mu \xi, \eta_\xi \rangle|\, dt &\le C\Big(\omega_{2-\mu}(T)\|\xi(0)\| + \int_0^T \omega_{2-\mu}(T-s)\|\xi'(s)\|\, ds\Big)\int_0^T \|\xi'\|\, dt\\ &\le C\, T^{1-\mu}\Big(\|\xi(0)\| + \int_0^T \|\xi'(s)\|\, ds\Big)\int_0^T \|\xi'\|\, dt\, .\end{aligned}$$

On the other hand, noting that

$$\int_0^T \left\langle {}^R\mathrm{D}^\mu \Pi_k \xi, \xi \right\rangle dt = \langle I^{1-\mu}\Pi_k\xi(T), \xi(T)\rangle - \int_0^T \langle I^{1-\mu}\Pi_k\xi, \xi'\rangle\, dt,$$

and hence, by the Cauchy-Schwarz inequality and the first inequality in (14),

$$\begin{aligned}\left|\int_0^T \left\langle {}^R\mathrm{D}^\mu \Pi_k \xi, \xi \right\rangle dt\right| &\le \|\Pi_k\xi\|_J \int_0^T \Big[\omega_{1-\mu}(T-t)\|\xi(T)\| + \int_0^t \omega_{1-\mu}(t-s)\, ds\|\xi'(t)\|\Big]dt\\ &\le 4\|\xi\|_J\Big(\|\xi(T)\|\omega_{2-\mu}(T) + \int_0^T \omega_{2-\mu}(t)\|\xi'(t)\|\, dt\Big)\\ &\le C\, T^{1-\mu}\Big(\|\xi(0)\| + \int_0^T \|\xi'(t)\|\, dt\Big)^2,\end{aligned}$$

where in the last inequality we used $\|\xi(s)\| \le \|\xi(0)\| + \int_0^T \|\xi'\|\, dt$ for any $s \in [0, T]$.

Finally, the desired result follows from the above two inequalities. □

5 Error estimates

This section is devoted to investigate the convergence of the DG-FE scheme, (7). We decompose the error as follows:

$$U - u = \zeta + \Pi_k \xi + \eta_u \quad \text{with } \zeta = U - \Pi_k R_h u\, . \tag{17}$$

Recall that $\xi = R_h u - u$ and $\eta_u = \Pi_k u - u$. The main task now is to estimate ζ.

Theorem 2 *Choose $U^0 = R_h u_0$. For $1 \le n \le N$, we have*

$$\int_0^T \|\zeta(t)\|_1^2\, dt \le C(h^4 C_1(k, u) + C_2(k, u)) + C h^2 k^2 \int_0^T \|u(t)\|_2^2\, dt,$$

where for $\ell \in \{1, 2\}$,

$$C_1(k,u) = \max_{n=1}^{N} \Big(k_n^{-\frac{\mu}{2}} \int_{I_n} \|u'\|_2\, dt\Big)^2 + \Big(\|u_0\|_2 + \int_0^T \|u'\|_2\, dt\Big)^2,$$

$$C_2(k,u) = \max_{n=1}^{N} k_n^{2\ell-2-\mu}\Big(k_n^{-\mu}\Big(\int_{I_n} \|u^{(\ell)}\|\, dt\Big)^2 + \Big(\int_{I_n} \|\nabla u^{(\ell)}\|\, dt\Big)^2\Big). \quad (18)$$

Proof We start our proof by taking the inner product of (1) with ζ, using the identity ${}^{c}\mathrm{D}^{\mu}u(t) = {}^{R}\mathrm{D}^{\mu}u(t) - \omega_{1-\mu}(t)u_0$, and then integrating over the time subinterval I_n,

$$\int_{I_n} \big[\big\langle {}^{\mathrm{R}}\mathrm{D}^{\mu}u, \zeta\big\rangle + a\big(t, u, \zeta\big)\big]\, dt = \int_{I_n} \langle f + \omega_{1-\mu}(t)u_0, \zeta\rangle\, dt\,.$$

The above equation, the DG-FE scheme (7) and the decomposition in (17) imply

$$\int_0^T \Big(\big\langle {}^{R}\mathrm{D}^{\mu}\zeta, \zeta\big\rangle + |\zeta|_1^2\Big)\, dt = \int_0^T \langle \omega_{1-\mu}(t)\,\xi(0), \zeta\rangle\, dt - \int_0^T \Big[\big\langle {}^{R}\mathrm{D}^{\mu}(\Pi_k\xi + \eta_u), \zeta\big\rangle + a(t, \Pi_k\xi + \eta_u, \zeta)\Big]\, dt\,. \quad (19)$$

Now, using the identity $\zeta = I^{\mu}({}^{\mathrm{R}}\mathrm{D}^{\mu}\zeta)$ by Lemma 1, and the continuity property in Lemma 2 (ii), we notice that

$$\Big|\int_0^T \big\langle {}^{R}\mathrm{D}^{\mu}\eta_u, \zeta\big\rangle dt\Big| \le C\int_0^T \big\langle {}^{R}\mathrm{D}^{\mu}\eta_u, \eta_u\big\rangle dt + \frac{1}{4}\int_0^T \big\langle {}^{R}\mathrm{D}^{\mu}\zeta, \zeta\big\rangle dt,$$

$$\Big|\int_0^T \big\langle {}^{R}\mathrm{D}^{\mu}\Pi_k\xi, \zeta\big\rangle dt\Big| \le C\int_0^T \big\langle {}^{R}\mathrm{D}^{\mu}\Pi_k\xi, \Pi_k\xi\big\rangle dt + \frac{1}{4}\int_0^T \big\langle {}^{R}\mathrm{D}^{\mu}\zeta, \zeta\big\rangle dt\,.$$

In addition, following the steps in (9), we observe

$$\int_0^T \langle \omega_{1-\mu}(t)\,\xi(0), \zeta\rangle\, dt \le \frac{1}{4}\int_0^T \big\langle {}^{\mathrm{R}}\mathrm{D}^{\mu}\zeta, \zeta\big\rangle dt + C\, T^{1-\mu}\|\xi(0)\|^2$$

Inserting the above three inequalities in (19), then simplifying, and using the positivity property of ${}^{\mathrm{R}}\mathrm{D}^{\mu}$, Lemma 2 (i), yield

$$\int_0^T |\zeta|_1^2 dt \le C\, T^{1-\mu}\|\xi(0)\|^2 + C\int_0^T \Big(\big\langle {}^{R}\mathrm{D}^{\mu}\eta_u, \eta_u\big\rangle + \big\langle {}^{R}\mathrm{D}^{\mu}\Pi_k\xi, \Pi_k\xi\big\rangle\Big)\, dt + \sum_{n=1}^{N}\Big|\int_{I_n} a(t, \Pi_k\xi + \eta_u, \zeta) dt\Big|. \quad (20)$$

From the definitions of the time projection Π_k and the space projection R_h,

$$\int_{I_n} \langle \mathcal{A}(t_n)\nabla(\Pi_k\xi + \eta_u), \nabla\zeta\rangle\, dt = \int_{I_n} \langle \mathcal{A}(t_n)\nabla\xi, \nabla\zeta\rangle\, dt = \int_{I_n} \langle [\mathcal{A}(t_n) - \mathcal{A}(t)]\nabla\xi, \nabla\zeta\rangle\, dt$$

and so,

$$\begin{aligned}
&\Big|\int_{I_n} a(t, \Pi_k\xi + \eta_u, \zeta)dt\Big| \\
&= \Big|\int_{I_n} \langle \mathcal{A}(t_n)\nabla(\Pi_k\xi + \eta_u) + [\mathcal{A}(t) - \mathcal{A}(t_n)]\nabla(\Pi_k\xi + \eta_u), \nabla\zeta\rangle\, dt\Big| \\
&= \Big|\int_{I_n} \langle [\mathcal{A}(t) - \mathcal{A}(t_n)]\nabla(\eta_\xi + \eta_u), \nabla\zeta\rangle\, dt\Big| \\
&\le Ck_n \int_{I_n} \|\nabla(\eta_\xi + \eta_u)\|\, \|\nabla\zeta\|\, dt\,.
\end{aligned}$$

Thus, by the inequality $\|\nabla\eta_\xi(t)\| \le \|\nabla\xi(t)\| + 4k_n^{-1}\int_{I_n}\|\nabla\xi(s)\|\, ds$ (follows from the triangle inequality and the first property of Π_k in (14)) for $t \in I_n$, and property (5),

$$\Big|\int_{I_n} a(t, \Pi_k\xi + \eta_u, \zeta)dt\Big| \le C\, k_n^2 \int_{I_n} (\|\nabla\xi\|^2 + \|\nabla\eta_u\|^2)\, dt + \frac{1}{2}\int_{I_n} |\zeta|_1^2\, dt\,.$$

Inserting this in (20) and using (11) for $t = 0$, we get

$$\begin{aligned}
\int_0^T |\zeta|_1^2 dt \;&\le\; C\, h^4\|u_0\|_2^2 \\
&\quad + C\sum_{n=1}^N \int_{I_n} \Big(\big\langle {}^R\mathrm{D}^\mu \eta_u, \eta_u\big\rangle + \big\langle {}^R\mathrm{D}^\mu \Pi_k\xi, \Pi_k\xi\big\rangle + k_n^2(\|\nabla\xi\|^2 + \|\nabla\eta_u\|^2)\Big)\, dt\,.
\end{aligned}$$

But, for $t \in I_n$ and for $\ell \in \{1,\ 2\}$,

$$\begin{aligned}
\int_{I_n} \big\langle {}^R\mathrm{D}^\mu \eta_u, \eta_u\big\rangle\, dt \;&\le\; Ck_n \max_{j=1}^n k_j^{2\ell-2-\mu}\Big(\int_{I_j}\|u^{(\ell)}\|\, dt\Big)^2 \quad \text{by Lemma 3,} \\
\|\nabla\xi(t)\| \;&\le\; C\, h\|u(t)\|_2 \quad \text{by the elliptic projection error (11),} \\
\|\nabla\eta_u(t)\| \;&\le\; C\, k_n^{\ell-1}\int_{I_n}\|\nabla u^{(\ell)}\|\, ds \quad \text{by the time projection error (15),}
\end{aligned}$$

where in the first inequality we also used the non-increasing time step assumption. So,

$$\begin{aligned}
\int_0^T |\zeta|_1^2 dt \;&\le\; C\, h^4\|u_0\|_2^2 + C\int_0^T \big\langle {}^R\mathrm{D}^\mu \Pi_k\xi, \Pi_k\xi\big\rangle dt + C\, h^2k^2\int_0^T \|u\|_2^2 dt \\
&\quad + \; C\max_{n=1}^N k_n^{2\ell-2-\mu}\Big(\Big(\int_{I_n}\|u^{(\ell)}\|\, dt\Big)^2 + k_n^\mu\Big(\int_{I_n}\|\nabla u^{(\ell)}\|\, dt\Big)^2\Big)\,. \qquad (21)
\end{aligned}$$

It remains to estimate $\int_0^T \big\langle {}^R\mathrm{D}^\mu \Pi_k\xi, \Pi_k\xi\big\rangle\, dt$. From the decomposition:

$$\int_{I_n} \big\langle {}^R\mathrm{D}^\mu \Pi_k\xi, \Pi_k\xi\big\rangle\, dt = \int_{I_n} \Big[\big\langle {}^R\mathrm{D}^\mu \eta_\xi, \eta_\xi\big\rangle + \big\langle {}^R\mathrm{D}^\mu \xi, \eta_\xi\big\rangle + \big\langle {}^R\mathrm{D}^\mu \Pi_k\xi, \xi\big\rangle\Big]\, dt\,. \qquad (22)$$

By Lemma 3 with $\ell = 1$,

$$\int_{I_n} \big\langle {}^\mathrm{R}\mathrm{D}^\mu \eta_\xi, \eta_\xi\big\rangle\, dt \le C\, k_n^{1-\mu}\max_{j=1}^n\Big(\int_{I_j}\|\xi'\|\, dt\Big)^2 \le C\, k_n \max_{j=1}^n\Big(k_j^{-\frac{\mu}{2}}\int_{I_j}\|\xi'\|\, dt\Big)^2.$$

Inserting the above bound in (22), then summing over n and using the achieved bound in Lemma 4, we obtain

$$\int_0^T |\langle {}^R\mathrm{D}^\mu \Pi_k\xi, \Pi_k\xi\rangle|\, dt \le C \max_{n=1}^{N} \Big(k_n^{-\frac{\mu}{2}} \int_{I_n} \|\xi'\|\, dt\Big)^2 + C\Big(\|\xi(0)\| + \int_0^T \|\xi'\|\, dt\Big)^2.$$

Finally, to complete the proof, we combine (21) with the above bound, and use the Ritz projection error estimate in (11). □

In the next theorem we show our main convergence results of the DG-FE solution. We assume that the exact solution u of problem (1) satisfies the finite regularity assumptions:

$$\|u'(t)\|_2 + t\|u''(t)\|_1 \le \mathbf{M}\, t^{\sigma-1} \quad \text{for} \quad t > 0, \tag{23}$$

for some positive constants $\mathbf{M}$ and σ. To satisfy this inequality, some regularity and compatibility assumptions on the given data u_0, the source term f, and the variable diffusivity $\mathcal{A}$ are required. For instance, if $f \equiv 0$ and $\mathcal{A} \equiv 1$, we assume that $u_0 \in H^{2+\epsilon}(\Omega) \cap H_0^1(\Omega)$ for some $0 < \epsilon < 1/2$. In this case, $\sigma = \epsilon$ and $\mathbf{M}$ depends on $\|u_0\|_{2+\epsilon}$, see [10, Theorem 4.2] for more details.

Due to the singular behaviour u near $t = 0$, we employ a family of non-uniform meshes, where the time-steps are graded towards $t = 0$; see [14]. More precisely, for a fixed parameter $\gamma \ge 1$, we assume that

$$t_n = (n/N)^\gamma T \quad \text{for } 0 \le n \le N. \tag{24}$$

One can easily see that the sequence of time-step sizes $\{k_j\}_{j=1}^N$ is nondecreasing, that is, $k_i \le k_j$ for $1 \le i \le j \le N$. One can also show the following mesh property:

$$k_j \le \gamma k t_j^{1-1/\gamma}\,. \tag{25}$$

Theorem 3 *Let $u \in L^2((0,T), H^2(\Omega)$ be the solution of (1) satisfying the regularity property (23) with $\sigma > \mu/2$. Let U be the DG-FE solution defined by (7). Then, we have*

$$\int_0^T \|U - u\|^2\, dt \le C\,(h^4 + k^{\gamma(2\sigma-\mu)}) \quad \textit{for} \quad 1 \le \gamma \le \frac{4-\mu}{2\sigma-\mu}$$

where C is a constant that depends on T, μ, γ, σ, and on $\mathbf{M}$.

Proof From the decomposition of the error in (17), the triangle inequality, the bound in Theorem 2, the inequality $\int_0^T \|\Pi_k\xi\|^2 dt \le \int_0^T \|\xi\|^2 dt$ by (14), the elliptic projection error (11), the error from the time projection (15), we have

$$\int_0^T \|U - u\|^2\, dt \le C\Big(h^4 C_1(k,u) + C_2(k,u) + h^2(h^2 + k^2)\int_0^T \|u\|_2^2 dt\Big).$$

By the definitions of $C_1(k,u)$ and $C_2(k,u)$ in (18), the regularity assumption (23), and the inequality $h^2k^2 \le \frac{1}{2}(h^4+k^4)$, we observe

$$\begin{aligned}\int_0^T \|U-u\|^2\,dt &\le Ch^4 \max_{n=1}^{N}\Big(k_n^{-\frac{\mu}{2}}\int_{I_n} t^{\sigma-1}\,dt\Big)^2 + Ch^4\Big(1+\int_0^T t^{\sigma-1}\,dt\Big)^2\\ &\quad + C\,k_1^{-\mu}\Big(\int_{I_1} t^{\sigma-1}\,dt\Big)^2 + C\max_{n=2}^{N} k_n^{2-\mu}\Big(\int_{I_n} t^{\sigma-2}\,dt\Big)^2 + C\,h^2k^2\\ &\le C(h^4\max_{n=1}^{N} k_n^{2\sigma-\mu} + h^4 + k_1^{2\sigma-\mu} + \max_{n=2}^{N} k_n^{4-\mu}t_n^{2\sigma-4} + k^4)\\ &\le C\,(h^4 + k^{\min\{\gamma(2\sigma-\mu),4-\mu\}})\end{aligned}$$

where in the last inequality, by the mesh property (25), we used

$$k_n^{4-\mu}{t_n}^{2(\sigma-2)} \le C\,k^{4-\mu}t_n^{2(\sigma-2)+4-\mu-(4-\mu)/\gamma} \le C\,k^{\min\{\gamma(2\sigma-\mu),4-\mu\}}.$$

The proof is completed now. □

6 Numerical results

In this section, we present a sample of numerical tests to validate our theoretical convergence results.

Table 1 Errors and time convergence rates for various choices of γ

$\mu = 0.3$								
N	$\gamma = 1$		$\gamma = 2$		$\gamma = 3$			
10	5.899e-03		1.125e-03		9.933e-04			
20	3.598e-03	0.7134	4.116e-04	1.451	2.552e-04	1.960		
40	2.183e-03	0.7211	1.501e-04	1.456	6.453e-05	1.984		
80	1.321e-03	0.7247	5.470e-05	1.456	1.614e-05	1.999		
160	7.980e-04	0.7269	1.999e-05	1.452	4.008e-06	2.009		
320	4.817e-04	0.7284	7.348e-06	1.444	9.916e-07	2.015		
$\mu = 0.5$								
N	$\gamma = 1$		$\gamma = 2$		$\gamma = 3$		$\gamma = 4$	
10	1.149e-02		3.262e-03		1.560e-03		1.882e-03	
20	7.641e-03	0.589	1.619e-03	1.011	5.972e-04	1.385	4.869e-04	1.951
40	5.151e-03	0.569	8.037e-04	1.010	2.192e-04	1.446	1.209e-04	2.009
80	3.641e-03	0.500	3.997e-04	1.008	7.867e-05	1.478	2.933e-05	2.044
160	2.570e-03	0.503	1.992e-04	1.005	2.797e-05	1.492	7.011e-06	2.064
320	1.812e-03	0.504	9.940e-05	1.003	9.908e-06	1.497	1.774e-06	1.982

Table 2 Errors and time convergence rates for various choices of γ

$\mu = 2/3$								
N	$\gamma = 1$		$\gamma = 2$		$\gamma = 4$		$\gamma = 6$	
10	1.677e-02		7.579e-03		3.416e-03		3.261e-03	
20	1.327e-02	0.338	4.677e-03	0.696	1.393e-03	1.294	9.087e-04	1.843
40	1.044e-02	0.346	3.036e-03	0.623	5.553e-04	1.327	2.471e-04	1.879
80	8.191e-03	0.350	1.940e-03	0.646	2.205e-04	1.332	6.435e-05	1.941
160	6.427e-03	0.350	1.229e-03	0.658	8.753e-05	1.333	1.643e-05	1.970
$\mu = 0.7$								
N	$\gamma = 1$		$\gamma = 3$		$\gamma = 5$		$\gamma = 7$	
10	1.792e-02		5.149e-03		3.625e-03		3.991e-03	
20	1.446e-02	0.309	2.905e-03	0.826	1.318e-03	1.459	1.121e-03	1.832
40	1.160e-02	0.318	1.577e-03	0.881	4.673e-04	1.496	3.052e-04	1.877
80	9.290e-03	0.321	8.479e-04	0.895	1.652e-04	1.499	7.981e-05	1.935
160	7.447e-03	0.319	4.547e-04	0.899	5.843e-05	1.500		

Example 1 We consider a model problem in one space dimension, of the form (1) with $\Omega = (0, 1)$, $[0, T] = [0, 1]$, and $\mathcal{A}(x, t) = 1+t^{3/2}$. We choose $u_0(x) = \sin(\pi x)$ for the initial data and choose the source term f so that

$$u(x, t) = (1 + t^{1-\mu}) \sin(\pi x)\,. \tag{26}$$

One easily verifies that the regularity condition (23) holds for $\sigma = 1 - \mu$.

The numerical tests below reveal faster rates of convergence than those suggested by Theorem 3, and that our regularity assumptions are more restrictive than is needed in practice. More precisely, Theorem 3 shows suboptimal (in time) convergence of order $O(k^{2-\frac{\mu}{2}} + h^2)$ for sufficiently graded time meshes in the time-space L^2-norm. However, we demonstrate numerically optimal (in both time and space) rates of convergence in the stronger $L^\infty(L^2)$-norm. Let $\|v\|_{\mathcal{G}^m} := \max_{t\in\mathcal{G}^m} \|v(t)\|$ where

$$\mathcal{G}^m = \{\, t_{j-1} + \ell k_j/m : j = 1, 2, \ldots, N \text{ and } \ell = 0, 1, \ldots, m \,\},$$

Table 3 Errors and convergence rates in space with $\mu = 0.3$, 0.5 and 0.7

M	$\mu = 0.3$		$\mu = 0.5$		$\mu = 0.7$	
10	1.2156e-02		1.2780e-02		1.2563e-02	
20	3.1130e-03	1.9653	3.2743e-03	1.9646	3.1768e-03	1.9836
40	7.8803e-04	1.9820	8.2897e-04	1.9818	7.9873e-04	1.9918
80	1.9826e-04	1.9909	2.0864e-04	1.9903	2.0029e-04	1.9956
160	4.9724e-05	1.9954	5.2355e-05	1.9946	5.1065e-05	1.9717

Table 4 Errors and convergence rates from the spatial quadratic FE approximations

M	$\mu = 0.3$		$\mu = 0.5$	
10	2.6421e-04		2.5119e-04	
20	3.3554e-05	2.9771	3.1483e-05	2.9962
40	4.2194e-06	2.9913	3.9262e-06	3.0033
70	7.8874e-07	2.9967	7.5959e-07	2.9353
100	2.7028e-07	3.0026	2.6715e-07	2.9298

with $k_1, k_2, \cdots, k_N$ being the step sizes on the finest time meshes in each table. So that, for sufficiently large values of m, $\|U_h - u\|_{\mathcal{G}^m}$ approximates the uniform error $\|U_h - u\|_{L^\infty(L^2)}$. In all tables, we choose $m = 10$.

For the numerical illustration of the convergence rates in time, we choose M (the number of uniform spatial subintervals) to be sufficiently large such that the spatial error is negligible compared to the error from the time discretization. We employ a time mesh of the form (24). Tables 1 and 2 show the error (in the stronger $L^\infty(L^2)$ -norm) and the rates of convergence when $\mu = 0.3, 0.5, 2/3$ and 0.7 respectively, for various choices of N and γ. We observe optimal rates of order $O(k^{\gamma\sigma})$ for various choices of $1 \le \gamma \le \frac{2}{\sigma}$ which is faster than the rate $O(k^{\frac{\gamma}{2}(2\sigma-\mu)})$ for $1 \le \gamma \le \frac{4-\mu}{2\sigma-\mu}$ predicted by our theory in Theorem 3. Noting that, in Table 2, $\sigma \le \mu/2$ and thus the assumption $\sigma > \mu/2$ in this theorem is not sharp.

Next, we test the performance of the spatial piecewise linear FEs discretizaton of the scheme (7). A uniform spatial mesh that consists of M subintervals where each is of width h will be used. We refine the time mesh such that the spatial error is dominating. By Theorem 3, a convergence of order $O(h^2)$ is expected. We illustrate these results in Table 3. Moreover, the numerical results in Table 4 demonstrate an $O(h^3)$ convergence rate for the piecewise quadratic FEs, which is expected based on our theoretical results.

Example 2 In this example, we test the performance of the spatial FEs discretizaton of the scheme (7) on a *two-dimensional* fractional diffusion problem of the form (1).

Table 5 Errors and convergence rates in space with $\mu = 0.3$ and 0.5 for **Example 2**

M	$\mu = 0.3$		$\mu = 0.5$	
20	2.3247e-03		2.3232e-03	
40	5.7409e-04	2.0177	5.7340e-04	2.0185
70	1.8770e-04	1.9977	1.8774e-04	1.9951
100	9.1336e-05	2.0196	9.1384e-05	2.0186
140	4.6424e-05	2.0112	4.6476e-05	2.0094

We choose $\Omega = (0, 1) \times (0, 1)$, $[0, T] = [0, 1]$, and $\mathcal{A}(x, t) = 1 + t^{3/2}$. The initial data u_0 and the source term f will be chosen such that

$$u(x,t) = (1 + t^{1-\mu}) \sin(\pi x) \sin(\pi y), \quad \text{for } t \in [0, T] \text{ and } (x, y) \in \Omega. \tag{27}$$

As in **Example 1**, the regularity condition (23) holds for $\sigma = 1 - \mu$. We refine the time steps so that the FE errors are dominant. Hence, by Theorem 3, an error of order $O(h^2)$ is expected. To illustrate this, we let $\mathcal{T}_h$ be a family of uniform rectangular mesh of the domain Ω with diameter $h = \sqrt{2}/M$. The numerical results in Table 5 reveal optimal $O(h^2)$ rates of convergence.

References

1. Alikhanov, A.: A new difference scheme for the time fractional diffusion equation. J. Comput. Phys. **280**, 424–438 (2015)
2. Brunner, H., Ling, L., Yamamoto, M.: Numerical simulations of 2D fractional subdiffusion problems. J. Comput. Phys. **229**, 6613–6622 (2010)
3. Chen, J., Liu, F., Liu, Q., Chen, X., Anh, V., Turner, I., Burrage, K.: Numerical simulation for the three-dimension fractional sub-diffusion equation. Appl. Math. Model. **38**, 3695–3705 (2014)
4. Cui, M.: Compact exponential scheme for the time fractional convection-diffusion reaction equation with variable coefficients. J. Comput. Phys. **280**, 143–163 (2013)
5. Jin, B., Larzarov, R., Zhou, Z.: Error estimates for a semidiscrete FE method for fractional order parabolic equations. SIAM J. Numer. Anal. **51**, 445–466 (2013)
6. Li, W., Xu, D.: Finite central difference/FE approximations for parabolic integro-differential equations. Computing **90**, 89–111 (2010)
7. Li, X., Xu, C.: A Space-Time spectral method for the time fractional diffusion equation. SIAM J. Numer. Anal. **47**, 2108–2131 (2009)
8. Lin, Y., Xu, C.: Finite difference/spectral approximations for the time-fractional diffusion equation. J. Comput. Phys. **225**, 1533–1552 (2007)
9. Luskin, M., Rannacher, R.: On the smoothing property of the Galerkin method for parabolic equations. SIAM J. Numer. Anal. **19**, 93–113 (1982)
10. Mclean, W.: Regularity of solutions to a time-fractional diffusion equation. ANZIAM J. **52**, 123–138 (2010)
11. Mclean, W.: Fast summation by interval clustering for an evolution equation with memory. SIAM J. Sci. Comput. **34**, A3039–A3056 (2012)
12. Metzler, R., Klafter, J.: The random walk's guide to anomalous diffusion: a fractional dynamics approach. Physics Reports **339**, 1–77 (2000)
13. Murillo, J., Yuste, S.B.: A finite difference method with non-uniform timesteps for fractional diffusion and diffusion-wave equations. Eur. Phys. J. Special Topics **222**, 1987–1998 (2013)
14. Mustapha, K., McLean, W.: Piecewise-linear, discontinuous Galerkin method for a fractional diffusion equation. Numer. Algor. **56**, 159–184 (2011)
15. Mustapha, K., McLean, W.: Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations. SIAM J. Numer. Anal. **51**, 491–515 (2013)
16. Mustapha, K.: Time-stepping discontinuous Galerkin methods for fractional diffusion problems. Numer. Math. **130**, 497–516 (2015)
17. Mustapha, K., Abdallah, B., Furati, K.M.: A discontinuous Pertov-Galerkin method for time-fractional diffusion equations. SIAM J. Number. Anal. **52**, 2512–2529 (2014)
18. Mustapha, K., Nour, M., Cockburn, B.: Convergence and superconvergence analyses of HDG methods for time fractional diffusion problems. Adv. Comput. Math. (2015). doi:10.1007/s10444-015-9428-x
19. Mustapha, K., Schötzau, D.: Well-posedness of *hp*-version discontinuous Galerkin method for fractional wave equations. IMA J. Numer. Anal. **34**, 1426–1446 (2014)
20. Ren, J., Sun, Z., Zhao, X.: Compact difference scheme for the fractional sub-diffusion equation with Neumann boundary conditions. J. Comput. Phys. **232**, 456–467 (2013)

21. Saadatmandi, A., Dehghan, M., Azizi, M.R.: The Sinc-Legendre collocation method for a class of fractional convection diffusion equations with variable coefficients. Commun. Nonlinear Sci. Numer. Simul. **17**, 4125–4136 (2012). doi:10.1007/s10444-015-9428-x
22. Wang, K.G.: Long time correlation effects and biased anomalous diffusion. Phys. Rev. A **45**, 833–837 (1992)
23. Xu, Q., Zheng, Z.: Discontinuous Galerkin method for time fractional diffusion equation. J. Informat. Comput. Sci. **10**, 3253–3264 (2013)
24. Zhang, Y., Sun, Z., Liao, H.l.: Finite difference methods for the time fractional diffusion equation on non-uniform meshes. J. Comput. Phys. **265**, 195–210 (2014)
25. Zhao, X., Sun, Z.-Z.: A box-type scheme for fractional sub-diffusion equation with Neumann boundary conditions. J. Comput. Phys. **230**, 6061–6074 (2011)
26. Zhao, X., Xu, Q.: Efficient numerical schemes for fractional sub-diffusion equation with the spatially variable coefficient. Appl. Math. Model. **38**, 3848–3859 (2014)