



# $\beta$ -Robust Superconvergent Analysis of a Finite Element Method for the Distributed Order Time-Fractional Diffusion Equation

Chaobao Huang<sup>1</sup> · Hu Chen<sup>2</sup> · Na An<sup>3</sup>

Received: 4 September 2020 / Revised: 27 October 2021 / Accepted: 3 November 2021 /

Published online: 12 December 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

A distributed order time fractional diffusion equation whose solution has a weak singularity near the initial time  $t = 0$  is considered. The numerical method of the paper uses the well-known L1 scheme on a graded mesh to discretize the time Caputo fractional derivative and a standard finite element method in space. A  $\beta$ -robust discrete fractional Grönwall inequality is investigated. By this inequality, the  $\beta$ -robust optimal-rate convergence and a superconvergence bound  $\|\nabla R_h u^n - \nabla u_h^n\|$  are proved. This superconvergence bound is also used to show that a simple postprocessing of the computed solution will yield a higher order of convergence in the spatial direction. The final convergence result reveals the optimal grading that one should use for the temporal graded mesh. Numerical results show that our analysis is sharp.

**Keywords** Distributed order derivative · Finite element method · Superconvergence · Weak singularity

**Mathematics Subject Classification** 65M60 · 65M12 · 35R11

---

✉ Hu Chen  
chenhu@ouc.edu.cn

Chaobao Huang  
huangcb@sdufe.edu.cn

Na An  
annabetter@sdnu.edu.cn

<sup>1</sup> School of Mathematics and Quantitative Economics, Shandong University of Finance and Economics, Jinan 250014, China

<sup>2</sup> School of Mathematical Sciences, Ocean University of China, Qingdao 266100, China

<sup>3</sup> School of Mathematics and Statistics, Shandong Normal University, Jinan 250014, China

## 1 Introduction

In the present paper, we consider the distributed order time-fractional diffusion equation with corresponding initial and boundary conditions as following:

$$\mathcal{D}_t^\omega u - \kappa \Delta u = f(x, t) \quad \forall (x, t) \in Q := \Omega \times (0, T], \quad (1.1a)$$

$$u|_{\partial\Omega} = 0 \quad \text{for } t \in (0, T], \quad (1.1b)$$

$$u(x, 0) = u_0(x) \quad \text{for } x \in \Omega, \quad (1.1c)$$

where  $\Omega \subset \mathbb{R}^d$  ( $d = 1, 2, 3$ ),  $\kappa$  is a positive constant, and  $f \in C(\overline{Q})$  with  $\overline{Q} := \Omega \times [0, T]$ . In (1.1a),  $\mathcal{D}_t^\omega u$  denotes the distributed order fractional derivative, which is defined by

$$\mathcal{D}_t^\omega u(x, t) = \int_0^\beta \omega(\alpha) D_t^\alpha u(x, t) d\alpha, \quad 0 < \beta \leq 1, \quad (1.2)$$

where  $\omega(\alpha) \geq 0$ ,  $\int_0^\beta \omega(\alpha) d\alpha = c_0 > 0$ ,  $D_t^\alpha u$  ( $0 < \alpha < 1$ ) is the fractional Caputo derivative of order  $\alpha$ , defined by

$$D_t^\alpha u(x, t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \frac{\partial u(x, s)}{\partial s} ds, \quad t > 0.$$

The analytic solutions of the distributed order time-fractional diffusion equation have been studied by many researchers [13,24,29,31]. However, only for a few problems the exact solutions can be displayed, and most of these solutions are consist of complex functions (Mittag-Leffler function, Wright function, etc.), which are not easy to compute. Thus it is very necessary to develop some efficient numerical methods to solve the distributed order time-fractional diffusion equation. Alikhanov [2] presented a priori estimates for the multi-term variable-distributed order diffusion equation by the method of the energy inequalities and investigated a difference scheme to solve it. Ye et al. [43] proposed a compact difference scheme for the problem (1.1) and got the stability and optimal convergent result for the proposed scheme. The numerical analysis of a finite difference method for the time distributed-order and Riesz space fractional diffusion equation was presented in [44]. Chen et al. [9] developed a fully discrete spectral method for the distributed order time-fractional reaction-diffusion equation, which will achieve the spectral accuracy. Bu et al. [8] investigated three efficient fully discrete finite element schemes to solve problem (1.1). Li et al. [23] developed two alternating direction implicit Galerkin-Legendre spectral methods for distributed-order differential equation in two-dimensional space. Samiee et al. [34,35] proposed a unified and fast Petrov-Galerkin spectral method for distributed-order partial differential equations, where Jacobi poly-fractionomials and Legendre polynomials were employed as temporal and spatial basis/test functions, respectively. Furthermore, some recent developments are given in [1,38,42].

It is worth noting that the analysis of the above schemes is based on the assumption that the solution is smooth enough in time direction. However, this assumption is unrealistic. The solution of the time-fractional partial differential equation typically exhibits a weak singularity near the initial time. Mclean [30] investigated the regularity result of solutions for time fractional diffusion equations and discovered these singular behaviors. Stynes et al. [39] investigated the L1 scheme on graded mesh to solve these weak singularities of the time-fractional reaction diffusion equation. Liao et al. [27] developed a discrete fractional Grönwall inequality on nonuniform mesh, which can be used to solve the time-fractional nonlinear problem with a weakly singular solution. Ren and Chen [32] investigated a finite

difference/spectral method to approximate a distributed order time fractional diffusion equation with initial singularity on two dimensional spatial domain, while the convergent result show that the bound will blow up as  $\beta \rightarrow 1^-$ . Bu et al. [7] proposed a space-time finite element method for the distributed order time fractional reaction diffusion equation with weakly singular solution. Moreover, there are some other relevant works about the weakly singular solution for the time-fractional partial differential equation, e.g. the finite difference method [22,26,36], the finite element method [3,20,21,41], the discontinuous Galerkin method [4,5,14,15,17,33], the collocation method [19,25].

Let  $p$  be a non-negative integer. Assume that  $u_0 \in D(\Delta^{p+2})$  and  $\partial_t^l f(\cdot, t) \in D(\Delta^p)$  for  $l = 1, 2$ , where the fractional power  $\Delta^p$  is defined in [16, p.3]. Imitating [16, Theorem 2.1], we obtained that the solution of the initial-boundary value problem (1.1) satisfies

$$\|u(\cdot, t)\|_p \leq C, \quad \|\partial_t^l u(\cdot, t)\|_p \leq C(1 + t^{\sigma-l}), \quad \|\mathcal{D}_t^\omega u(\cdot, t)\|_p \leq C \tag{1.3}$$

with  $l = 0, 1, 2$ , and  $0 < \sigma < 1$ .

In this paper, we will construct the finite difference/finite element method to solve the initial-boundary value problem (1.1), whose solutions behave a weak singularity as (1.3). In order to obtain the sharp  $H^1$ -norm convergent result, the fully discrete L1 finite element method with integral formula will be written as differential formula. By investigating a  $\beta$ -robust discrete Grönwall inequality, the  $\beta$ -robust  $H^1$ -norm stability and convergent results are obtained. Furthermore, the superconvergent result in space direction will achieve.

The rest of the paper will organized as follows. In Sect. 2, several operators will be introduced, which will be used in following error estimate. In Sect. 3, we will construct a fully discrete scheme, which is based on the L1 scheme in time direction and the finite element method in space direction. The sharp  $H^1$ -norm stability and convergent result will be presented in Sect. 4, while these bounds will blow up as  $\beta \rightarrow 1^-$ . To overcome this, a new analysis will be presented in Sect. 5, and the  $\beta$ -robust stability and convergent results are obtained by using a new  $\beta$ -robust Grönwall inequality. The superconvergent result in space direction is given in Sect. 6. Finally, in Sect. 7, the numerical experiments are presented to illustrate the sharpness of our theoretical analysis.

*Notation.*  $C$  denotes a generic constant, it is independent of mesh parameters  $N$  and  $h$  and can take different values in different places. We write  $\|\cdot\|_\infty$  and  $\|\cdot\|$  for the norms in  $L^\infty(\Omega)$  and  $L^2(\Omega)$  respectively. For each  $m \in \mathbb{N}$ , the notation  $H^m(\Omega)$  is used for the standard Sobolev space with its associated norm  $\|\cdot\|_m$  and seminorm  $|\cdot|_m$ . The  $L^2(\Omega)$  inner product is denoted by  $(\cdot, \cdot)$ .

## 2 Preliminaries

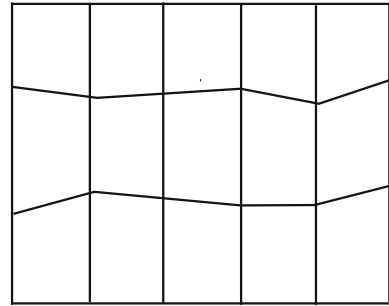
Recall that  $\Omega \subset \mathbb{R}^d$ . To construct standard finite element space, we write  $Q_k$  for the space of polynomials of degree  $k$  in  $d$  variables. Denote the reference element  $\widehat{K} := [0, 1]^d$ . Let  $\mathcal{T}_h$  be a quasiuniform partition of  $\Omega$  (see Figure 1) into elements  $K_m$  for  $m = 1, \dots, M$ , where each  $K_m = q_m(\widehat{K})$  for some  $q_m \in Q_k$ . Then the standard mapped  $Q_k$  functions are used on each element (see, e.g., [12, Section 3.7]), set

$$V_h := \left\{ v_h \in H^1(\Omega) : v_h|_{K_m} = \xi \circ q_m^{-1} \text{ with } \xi \in Q_k(\widehat{K}) \text{ and } q_m : \widehat{K} \rightarrow K_m \right\},$$

$$V_{0h} := \{v_h \in V_h : v_h|_{\partial\Omega} = 0\},$$

where  $h := \max_{1 \leq m \leq M} \{\text{diam}(K_m)\}$  is the mesh diameter.

**Fig. 1** Quasiuniform partition of  $\Omega$



Next we will introduce three operators. First, define the  $L^2$  projector  $P_h : L^2(\Omega) \rightarrow V_{0h}$  by

$$(P_h w, v_h) = (w, v_h) \quad \forall v_h \in V_{0h}. \tag{2.1}$$

One can show [6] that

$$\|\nabla P_h v\| \leq C_p \|\nabla v\| \quad \forall v \in H^1(\Omega), \tag{2.2}$$

where  $C_p$  is a positive constant independent of the mesh size  $h$ .

The Ritz projector  $R_h : H_0^1(\Omega) \rightarrow V_{0h}$  is defined by  $(\nabla R_h w, \nabla v_h) = (\nabla w, \nabla v_h)$  for all  $v_h \in V_{0h}$ . From [40, Lemma 1.1] we get

$$\|w - R_h w\| + h\|w - R_h w\|_1 \leq Ch^{k+1}|w|_{k+1} \quad \forall w \in H^{k+1}(\Omega) \cap H_0^1(\Omega). \tag{2.3}$$

Next, we define the *discrete Laplacian*  $\Delta_h : V_{0h} \rightarrow V_{0h}$  [40, (1.33)] by

$$(\Delta_h v, w) = -(\nabla v, \nabla w) \quad \forall v, w \in V_{0h}. \tag{2.4}$$

Imitating [40, (1.34)], we get that these three operators are related by

$$\Delta_h R_h v = P_h \Delta v \quad \forall v \in H^2(\Omega). \tag{2.5}$$

### 3 The L1 FEM Method

In this section, we will approximate the problem (1.1) using the finite element method in space and the well-known L1 difference scheme in time, on a mesh that is uniform in space and graded in time.

#### 3.1 Temporal Discretisation

Let  $q$  be a positive integer. Divide the interval  $[0, \beta]$  into  $q$ -subintervals with  $h_\alpha = \beta/q$ . Denote  $\alpha := (\alpha_1, \dots, \alpha_q)$ , and  $\mathbb{D}_t^\alpha u := h_\alpha \sum_{s=1}^q \omega(\alpha_s) D_t^{\alpha_s} u$  with  $\alpha_s = (s - 1/2)h_\alpha$  for  $s = 1, \dots, q$ . Note that  $\alpha_s$  is the center of each cell  $[(s - 1)h_\alpha, sh_\alpha]$ . Thus the distributed order fractional derivative can be approximated by the multi-term fractional derivative

$$\mathcal{D}_t^\omega u = \mathbb{D}_t^\alpha u + R(t), \tag{3.1}$$

where  $R(t) := \mathcal{D}_t^\omega u - \mathbb{D}_t^\alpha u$  denotes the approximation error. Under the conditions  $\omega(\alpha) \in C^2[0, \beta]$  and  $D_t^\alpha u(\cdot, \cdot) \in C^2[0, \beta]$ , one has  $\|R(t)\| \leq Ch_\alpha^2$  by composite midpoint formula for numerical integration [11, (5.1.19)].

Let  $N$  be a positive integer. Set the temporal mesh  $t_n = T(n/N)^r$  for  $n = 0, 1, \dots, N$ , where the constant  $r$  satisfies  $r \geq 1$ . Denote the time step  $\tau_n := t_n - t_{n-1}$  for  $n = 1, \dots, N$ .

The Caputo fractional derivative  $D_t^\alpha u(\cdot, t_n)$  with  $0 < \alpha < 1$  is approximated by the well-known L1 approximation of [39, (3.1)]

$$D_N^\alpha u^n := d_{n,1}^{(\alpha)} u^n - d_{n,n}^{(\alpha)} u^0 - \sum_{i=1}^{n-1} \left( d_{n,i}^{(\alpha)} - d_{n,i+1}^{(\alpha)} \right) u^{n-i}, \tag{3.2}$$

where

$$d_{n,i}^{(\alpha)} := \frac{\tau_{n-i+1}^{-1}}{\Gamma(1-\alpha)} \int_{t_{n-i}}^{t_{n-i+1}} (t_n - \eta)^{-\alpha} d\eta \text{ for } i = 1, \dots, n.$$

Denote

$$d_{n,i} := h_\alpha \sum_{s=1}^q \omega(\alpha_s) d_{n,i}^{(\alpha_s)}. \tag{3.3}$$

Thus the distributed order fractional derivative (1.2) can be approximated by

$$\mathbb{D}_N^\alpha u^n = d_{n,1} u^n - d_{n,n} u^0 - \sum_{i=1}^{n-1} (d_{n,i} - d_{n,i+1}) u^{n-i}. \tag{3.4}$$

It is easy to see that

$$0 < d_{n,i+1} < d_{n,i} \text{ for } i \geq 1. \tag{3.5}$$

Next, we will present three Lemmas, which will be used in our later analysis.

**Lemma 3.1** [32, Lemma 2.2] *For any grid function  $\{v^n\}_{n=0}^N$ , one has*

$$|v^n| \leq |v^0| + \Gamma(1-\alpha_q) \max_{j=1, \dots, n} \left\{ \left( h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s} \right)^{-1} \mathbb{D}_N^\alpha |v^j| \right\} \text{ for } n = 1, \dots, N.$$

**Lemma 3.2** [32, Lemma 2.3] *Suppose  $h_\alpha$  is sufficiently small, one has*

$$h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s} \geq \frac{c_0}{2} \min\{t_j^{-\alpha_1}, t_j^{-\alpha_q}\} \text{ for } j = 1, \dots, N.$$

One should bear in mind that in the rest of our paper, we always take  $h_\alpha$  sufficiently small to ensure Lemma 3.2 is valid.

**Lemma 3.3** *Let  $\sigma \in (0, 1)$  and  $\alpha \in (0, 1)$ . Assume that  $\|u^{(l)}(\cdot, t)\|_1 \leq C(1 + t^{\sigma-l})$  for  $l = 0, 1, 2$ , and  $t \in (0, T]$ . Then*

$$\|D_N^\alpha u(t_n) - D_t^\alpha u(t_n)\|_1 \leq C t_n^{-\alpha} N^{-\min\{r\sigma, 2-\alpha\}} \text{ for } n = 1, \dots, N,$$

where  $C$  is a constant.

**Proof** The result is followed from [32, Lemma 2.4] and [32, Lemma 2.6]. □

### 3.2 The Fully Discrete L1 FEM

Firstly, we use the standard finite element method (FEM) to discretize (1.1a) in spatial direction. A weak formulation of (1.1) is: Find  $u(\cdot, t) \in H_0^1(\Omega)$  for each  $t \in (0, T]$ , such that

$$(\mathcal{I}_t^\omega u, v) + \kappa(\nabla u, \nabla v) = (f, v) \quad \forall v \in H_0^1(\Omega) \tag{3.6}$$

with  $u(x, 0) = u_0$ .

Our semi-discrete FEM is: Find  $u_h(\cdot, t) \in V_{0h}$  for each  $t \in (0, T]$ , such that

$$(\mathcal{I}_t^\omega u_h, v_h) + \kappa(\nabla u_h, \nabla v_h) = (f, v_h) \quad \forall v_h \in V_{0h} \tag{3.7}$$

with  $u_h^0 = R_h u_0$ .

Applying the L1 scheme (3.4) to approximate (3.7) in temporal direction, the fully discrete L1 finite element method (L1 FEM) takes the form: find  $u_h^n \in V_{0h}$  for  $n = 0, 1, \dots, N$  such that

$$(\mathbb{D}_N^\alpha u_h^n, v_h) + \kappa(\nabla u_h^n, \nabla v_h) = (f^n, v_h) \quad \text{with } u_h^0 = R_h u_0, \tag{3.8}$$

where  $f^n(\cdot) := f(\cdot, t_n)$ .

By (2.4) and (2.1), the L1 FEM (3.8) can be written as: find  $u_h^n \in V_{0h}$  for  $n = 0, 1, \dots, N$  such that

$$(\mathbb{D}_N^\alpha u_h^n, v_h) - \kappa(\Delta_h u_h^n, v_h) = (P_h f^n, v_h) \quad \text{with } u_h^0 = R_h u_0$$

for  $n = 1, \dots, N$ . Owing to  $\mathbb{D}_N^\alpha u_h^n, \Delta_h u_h^n$  and  $P_h f^n$  all belong to  $V_{0h}$ , this integral formulation of L1 FEM takes the differential form: find  $u_h^n \in V_{0h}$  for  $n = 0, 1, \dots, N$  such that

$$\mathbb{D}_N^\alpha u_h^n - \kappa \Delta_h u_h^n = P_h f^n \quad \text{with } u_h^0 = R_h u_0 \tag{3.9}$$

for  $n = 1, \dots, N$ .

### 4 $H^1$ -norm Stability of L1 FEM

In this section, we will present the sharp  $H^1$ -norm stability and convergent results of the computed solution given in (3.9).

Next the following important property of the L1 scheme will be stated.

**Lemma 4.1** *Let the functions  $v^j = v(\cdot, t_j)$  be in  $L^2(\Omega)$  for  $j = 0, 1, \dots, N$ . Then the discrete L1 scheme satisfies*

$$(\mathbb{D}_N^\alpha v^n, v^n) \geq (\mathbb{D}_N^\alpha \|v^n\|) \|v^n\| \quad \text{for } n = 1, 2, \dots, N.$$

**Proof** Let  $n \in \{1, 2, \dots, N\}$ . The definition of  $\mathbb{D}_N^\alpha v^n$  gives

$$\begin{aligned}
 (\mathbb{D}_N^\alpha v^n, v^n) &= d_{n,1}(v^n, v^n) - d_{n,n}(v^0, v^n) - \sum_{i=1}^{n-1} (d_{n,i} - d_{n,i+1})(v^{n-i}, v^n) \\
 &\geq d_{n,1}\|v^n\|^2 - d_{n,n}\|v^0\| \|v^n\| - \sum_{i=1}^{n-1} (d_{n,i} - d_{n,i+1})\|v^{n-i}\| \|v^n\| \\
 &= (\mathbb{D}_N^\alpha v^n) \|v^n\|,
 \end{aligned}$$

where we used Cauchy-Schwarz inequalities and  $0 < d_{n,i+1} < d_{n,i}$  given in (3.5). □

Now we give the stability of the L1 FEM (3.9) in following lemma.

**Lemma 4.2** *The solution  $u_h^n$  of the discrete problem (3.9) satisfies*

$$\|\nabla u_h^n\| \leq \|\nabla u_h^0\| + C_p \Gamma(1 - \beta) \frac{2T^\beta}{c_0} \max_{j=1,\dots,n} \|\nabla f^j\|. \tag{4.1}$$

**Proof** Fix  $n \in \{1, 2, \dots, N\}$ . Multiplying (3.9) by  $-\Delta_h u_h^n$  and integrating over  $\Omega$ , one has

$$(\mathbb{D}_N^\alpha u_h^n, -\Delta_h u_h^n) + \kappa \|\Delta_h u_h^n\|^2 = (P_h f^n, -\Delta_h u_h^n).$$

Applying  $\kappa > 0$ , then use the definition (2.4) of  $\Delta_h$  to get

$$(\mathbb{D}_N^\alpha (\nabla u_h^n), \nabla u_h^n) \leq (\nabla P_h f^n, \nabla u_h^n).$$

Now Lemma 4.1, a Cauchy-Schwarz inequality, and (2.2) yield

$$(\mathbb{D}_N^\alpha \|\nabla u_h^n\|) \|\nabla u_h^n\| \leq \|\nabla P_h f^n\| \|\nabla u_h^n\| \leq C_p \|\nabla f^n\| \|\nabla u_h^n\|. \tag{4.2}$$

Invoking (4.2) and Lemmas 3.1 and 3.2, one has

$$\begin{aligned}
 \|\nabla u_h^n\| &\leq \|\nabla u_h^0\| + \Gamma(1 - \alpha_q) \max_{j=1,\dots,n} \left\{ (h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s})^{-1} \mathbb{D}_N^\alpha \|\nabla u_h^j\| \right\} \\
 &\leq \|\nabla u_h^0\| + C_p \Gamma(1 - \alpha_q) \max_{j=1,\dots,n} \left\{ (h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s})^{-1} \|\nabla f^j\| \right\} \\
 &\leq \|\nabla u_h^0\| + C_p \Gamma(1 - \beta) \frac{2}{c_0} \max_{j=1,\dots,n} \left\{ \max\{t_j^{\alpha_1}, t_j^{\alpha_q}\} \|\nabla f^j\| \right\} \\
 &\leq \|\nabla u_h^0\| + C_p \Gamma(1 - \beta) \frac{2T^\beta}{c_0} \max_{j=1,\dots,n} \|\nabla f^j\|,
 \end{aligned}$$

where Lemma 3.2 is used for the penultimate inequality. □

Let  $u^n$  and  $u_h^n$  be the solutions of (1.1) and (3.9) respectively at time  $t = t_n$  for  $n = 0, 1, \dots, N$ . Denote  $R^n := \mathcal{D}_t^\omega u(t_n) - \mathbb{D}_t^\alpha u(t_n)$  for  $n = 0, 1, \dots, N$ . To facilitate the error analysis for the standard finite element method, we follow the writing

$$u^n - u_h^n = (R_h u^n - u_h^n) - (R_h u^n - u^n) = \zeta^n - \rho^n, \tag{4.3}$$

where  $\zeta^n := R_h u^n - u_h^n$  and  $\rho^n := R_h u^n - u^n$ . Now we consider the analysis of  $\zeta^n$ , because the bound of  $\rho^n$  can be approximated immediately using (2.3). From (1.1a), (3.9), and (2.5), one has

$$\begin{aligned}
 \mathbb{D}_N^\alpha \zeta^n - \kappa \Delta_h \zeta^n &= [R_h (\mathbb{D}_N^\alpha u^n) - \kappa \Delta_h R_h u^n] - (\mathbb{D}_N^\alpha u_h^n - \kappa \Delta_h u_h^n) \\
 &= (R_h - P_h) \mathbb{D}_N^\alpha u^n + P_h (\mathbb{D}_N^\alpha u^n - \kappa \Delta u^n) - P_h f^n \\
 &= P_h (R_h - I) \mathbb{D}_N^\alpha u^n + P_h (\mathcal{I}_t^\omega u^n - \kappa \Delta u^n + \mathbb{D}_N^\alpha u^n - \mathcal{I}_t^\omega u^n) - P_h f^n \\
 &= P_h (R_h - I) \mathbb{D}_N^\alpha u^n + P_h (f^n + \mathbb{D}_N^\alpha u^n - \mathbb{D}_t^\alpha u^n - R^n) - P_h f^n \\
 &= P_h (R_h - I) \mathbb{D}_N^\alpha u^n + P_h (f^n - \varphi^n - R^n) - P_h f^n \\
 &= P_h \mathbb{D}_N^\alpha \rho^n - P_h (\varphi^n + R^n),
 \end{aligned}$$

where  $\varphi^n(x) := \mathbb{D}_t^\alpha u^n - \mathbb{D}_N^\alpha u^n = \sum_{s=1}^q h_\alpha \omega(\alpha_s) [D_t^{\alpha_s} u(x, t_n) - D_N^{\alpha_s} u(x, t_n)]$ . Clearly

$$\begin{aligned}
 \mathbb{D}_N^\alpha \rho^n &= \mathbb{D}_N^\alpha \rho^n - \mathbb{D}_t^\alpha \rho^n + \mathbb{D}_t^\alpha \rho^n - \mathcal{I}_t^\omega \rho^n + \mathcal{I}_t^\omega \rho^n \\
 &= (\mathbb{D}_t^\alpha u^n - \mathbb{D}_N^\alpha u^n) - R_h (\mathbb{D}_t^\alpha u^n - \mathbb{D}_N^\alpha u^n) + (\mathcal{I}_t^\omega u^n - \mathbb{D}_t^\alpha u^n) \\
 &\quad - R_h (\mathcal{I}_t^\omega u^n - \mathbb{D}_t^\alpha u^n) + \mathcal{I}_t^\omega \rho^n \\
 &= \varphi^n + R^n - R_h (\varphi^n + R^n) + \mathcal{I}_t^\omega \rho^n.
 \end{aligned} \tag{4.4}$$

Thus

$$\mathbb{D}_N^\alpha \zeta^n - \kappa \Delta_h \zeta^n = P_h \mathcal{I}_t^\omega \rho^n - R_h (\varphi^n + R^n) \text{ for } n = 1, 2, \dots, N. \tag{4.5}$$

We can now prove the optimal-rate convergence of our L1 FEM (3.9) in  $L^\infty(H^1)$ .

**Theorem 4.1** (Error estimate for the L1 FEM) *Let  $u^n$  and  $u_h^n$  be the solutions of (1.1) and (3.9), respectively. Assume the hypotheses of (1.3) with  $k + 1 \leq p$ ,  $\omega(\alpha) \in C^2[0, \beta]$ , and  $D_t^\alpha u(\cdot, \cdot) \in C^2[0, \beta]$ . Then for  $n = 1, 2, \dots, N$ , there exists a constant  $C$  such that*

$$\|\nabla u^n - \nabla u_h^n\| \leq C \Gamma(1 - \beta) \left( h^k + h_\alpha^2 + N^{-\min\{r\sigma, 2-\alpha_q\}} \right). \tag{4.6}$$

**Proof** Fix  $n \in \{1, 2, \dots, N\}$ . Multiplying (4.5) by  $-\Delta_h \zeta^n$  and integrating over  $\Omega$ , one has

$$-(\mathbb{D}_N^\alpha \zeta^n, \Delta_h \zeta^n) + \|\Delta_h \zeta^n\|^2 = -(P_h \mathcal{I}_t^\omega \rho^n, \Delta_h \zeta^n) + (R_h (\varphi^n + R^n), \Delta_h \zeta^n). \tag{4.7}$$

Recalling the definition (2.4) of  $\Delta_h$  and the projection  $P_h$ , we get

$$(\mathbb{D}_N^\alpha (\nabla \zeta^n), \nabla \zeta^n) + \|\Delta_h \zeta^n\|^2 = (\nabla P_h \mathcal{I}_t^\omega \rho^n, \nabla \zeta^n) - (\nabla R_h (\varphi^n + R^n), \nabla \zeta^n).$$

By Lemma 4.1 and a Cauchy-Schwarz inequality, this gives

$$(\mathbb{D}_N^\alpha \|\nabla \zeta^n\|) \|\nabla \zeta^n\| \leq \|\nabla P_h \mathcal{I}_t^\omega \rho^n\| \|\nabla \zeta^n\| + \|\nabla R_h (\varphi^n + R^n)\| \|\nabla \zeta^n\|.$$

Thus

$$\mathbb{D}_N^\alpha \|\nabla \zeta^n\| \leq C_p \|\nabla \mathcal{I}_t^\omega \rho^n\| + \|\nabla (\varphi^n + R^n)\|. \tag{4.8}$$

Invoking Lemma 3.1 and (4.8), one has

$$\begin{aligned}
 &\|\nabla \zeta^n\| \\
 &\leq \|\nabla \zeta^0\| + \Gamma(1 - \alpha_q) \max_{j=1, \dots, n} \left\{ \left( h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s} \right)^{-1} \mathbb{D}_N^\alpha \|\nabla \zeta^j\| \right\} \\
 &\leq \Gamma(1 - \alpha_q) \max_{j=1, \dots, n} \left\{ \left( h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s} \right)^{-1} \left[ C_p \|\nabla \mathcal{I}_t^\omega \rho^j\| + \|\nabla (\varphi^j + R^j)\| \right] \right\}
 \end{aligned}$$



$$\begin{aligned} &\leq C\Gamma(1 - \alpha_q) \max_{j=1, \dots, n} \left\{ \left( h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s} \right)^{-1} \left[ h^k + h_\alpha^2 + h_\alpha \sum_{s=1}^q \omega(\alpha_s) t_j^{-\alpha_s} N^{-\min\{r\sigma, 2-\alpha_s\}} \right] \right\} \\ &\leq C\Gamma(1 - \beta) \max_{j=1, \dots, n} \left\{ \max\{t_j^{\alpha_1}, t_j^{\alpha_q}\} (h^k + h_\alpha^2) \right\} + C\Gamma(1 - \beta) N^{-\min\{r\sigma, 2-\alpha_q\}} \\ &\leq C\Gamma(1 - \beta) \left( h^k + h_\alpha^2 + N^{-\min\{r\sigma, 2-\alpha_q\}} \right), \end{aligned}$$

where Lemmas 3.2 and 3.3, (2.3), and  $\|\nabla \zeta^0\| = \|\nabla(R_h u^0 - u^0)\| = 0$  are used. Combining this bound and (2.3) with (4.3), we get (5.8).  $\square$

**Remark 4.1** The orders of convergence displayed in Theorem 4.1 indicate that the rates of convergence in space, distributed variable, and time are  $h^k$ ,  $h_\alpha^2$ , and  $N^{-\min\{r\sigma, 2-\alpha_q\}}$ , respectively.

It is obvious that the convergent result obtained in Theorem 4.1 will blow up as  $\beta \rightarrow 1^-$ . This phenomenon also appears in [32]. In the next part we will try to improve this convergent result by making it  $\beta$ -robust.

### 5 $\beta$ -robust $H^1$ -norm Error Analysis of the L1 FEM

In this section, we will present a  $\beta$ -robust discrete Grönwall inequality, which is an improve of [18, Lemma 8]. Applying this new discrete Grönwall inequality, a  $\beta$ -robust  $H^1$ -norm error estimate for the computed solution is obtained. Based on this result, a superconvergent result is achieved immediately.

As in [39, (4.6)], define the positive real numbers  $\theta_{n,j}$ , for  $n = 1, 2, \dots, N$  and  $j = 1, 2, \dots, n - 1$ , by

$$\theta_{n,n} = 1, \quad \theta_{n,j} = \sum_{k=1}^{n-j} \frac{1}{d_{n-k,1}} (d_{n,k} - d_{n,k+1}) \theta_{n-k,j}, \tag{5.1}$$

where  $d_{n,k}$  is defined in (3.3). Observe that (3.5) implies  $\theta_{n,j} > 0$  for all  $n, j$ .

**Lemma 5.1** [10, Lemma 5.1] For  $n = 1, 2, \dots, N$  and  $1 \leq k \leq n$ , one has

$$\sum_{j=k}^n d_{j,j+1-k} \theta_{n,j} = d_{n,1}. \tag{5.2}$$

**Lemma 5.2** Let  $\gamma \in (0, 1)$  be a constant. Then for  $n = 1, 2, \dots, N$ , one has

$$\sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) \frac{\Gamma(1 + \gamma)}{\Gamma(1 + \gamma - \alpha_s)} t_j^{\gamma - \alpha_s} \right) \theta_{n,j} \leq d_{n,1} t_n^\gamma. \tag{5.3}$$

**Proof** Imitating [10, Lemma 5.3], one has

$$\frac{\Gamma(1 + \gamma)}{\Gamma(1 + \gamma - \alpha_s)} t_j^{\gamma - \alpha_s} \leq \sum_{k=1}^j d_{j,j+1-k}^{(\alpha_s)} (t_k^\gamma - t_{k-1}^\gamma).$$

Thus

$$\sum_{s=1}^q h_\alpha \omega(\alpha_s) \frac{\Gamma(1 + \gamma)}{\Gamma(1 + \gamma - \alpha_s)} t_j^{\gamma - \alpha_s}$$

$$\begin{aligned} &\leq \sum_{s=1}^q h_\alpha \omega(\alpha_s) \sum_{k=1}^j d_{j,j+1-k}^{(\alpha_s)} (t_k^\gamma - t_{k-1}^\gamma) \\ &= \sum_{k=1}^j \sum_{s=1}^q h_\alpha \omega(\alpha_s) d_{j,j+1-k}^{(\alpha_s)} (t_k^\gamma - t_{k-1}^\gamma) \\ &= \sum_{k=1}^j d_{j,j+1-k} (t_k^\gamma - t_{k-1}^\gamma). \end{aligned}$$

Multiply this inequality by  $\theta_{n,j}$  then sum from  $j = 1$  to  $n$ . This yields

$$\begin{aligned} &\sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) \frac{\Gamma(1+\gamma)}{\Gamma(1+\gamma-\alpha_s)} t_j^{\gamma-\alpha_s} \right) \theta_{n,j} \\ &\leq \sum_{j=1}^n \theta_{n,j} \sum_{k=1}^j d_{j,j+1-k} (t_k^\gamma - t_{k-1}^\gamma) \\ &= \sum_{k=1}^n (t_k^\gamma - t_{k-1}^\gamma) \left( \sum_{j=k}^n \theta_{n,j} d_{j,j+1-k} \right) \\ &= d_{n,1} \sum_{k=1}^n (t_k^\gamma - t_{k-1}^\gamma) \\ &= d_{n,1} t_n^\gamma, \end{aligned}$$

by changing the order of summation then invoking Lemma 5.1. □

**Corollary 5.1** *Setting  $l_N = 1/\ln N$ , one has*

$$\frac{1}{d_{n,1}} \sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} \right) \theta_{n,j} \leq \frac{e^r \max_{1 \leq s \leq q} \Gamma(1+l_N-\alpha_s)}{\Gamma(1+l_N)}.$$

**Proof** Applying  $1 \leq j^{r l_N}$  and  $N^{r l_N} = e^r$ , one has

$$\begin{aligned} &\frac{1}{d_{n,1}} \sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} \right) \theta_{n,j} \\ &\leq \frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} j^{r l_N} \right) \\ &= \frac{N^{r l_N}}{T^{l_N}} \frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{l_N-\alpha_s} \right) \\ &\leq \frac{e^r \max_{1 \leq s \leq q} \Gamma(1+l_N-\alpha_s)}{T^{l_N} \Gamma(1+l_N)} \frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) \frac{\Gamma(1+l_N)}{\Gamma(1+l_N-\alpha_s)} t_j^{l_N-\alpha_s} \right). \end{aligned}$$

Choosing  $\gamma = l_N$  in Lemma 5.2 yields

$$\frac{1}{d_{n,1}} \sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} \right) \theta_{n,j}$$

$$\leq \frac{e^r \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{T^{l_N} \Gamma(1 + l_N)} t_n^{l_N} \leq \frac{e^r \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)},$$

where  $t_n^{l_N} \leq T^{l_N}$  is used. Thus the result follows. □

**Corollary 5.2** *Setting  $l_N = 1/\ln N$ , one has*

$$\frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \leq \frac{e^r \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \frac{2}{c_0} \max\{t_n^{\alpha_1}, t_n^{\alpha_q}\}.$$

**Proof** From Lemma 3.2 we have

$$\begin{aligned} & \frac{1}{d_{n,1}} \sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} \right) \theta_{n,j} \\ & \geq \frac{1}{d_{n,1}} \sum_{j=1}^n \frac{c_0}{2} \min\{t_j^{-\alpha_1}, t_j^{-\alpha_q}\} \theta_{n,j} \\ & \geq \frac{c_0}{2} \min\{t_n^{-\alpha_1}, t_n^{-\alpha_q}\} \frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j}. \end{aligned}$$

Thus the result follows from Corollary 5.1 and  $\frac{1}{\min\{t_n^{-\alpha_1}, t_n^{-\alpha_q}\}} \leq \max\{t_n^{\alpha_1}, t_n^{\alpha_q}\}$ . □

Next we will prove a new nonstandard Grönwall inequality, which is an improvement of [18, Lemma 8].

**Lemma 5.3** *Assume that the sequences  $\{\xi^n\}_{n=1}^\infty, \{\eta^n\}_{n=1}^\infty$  are nonnegative and the grid function  $\{v^n : n = 0, 1, \dots, N\}$  satisfies  $v^0 \geq 0$  and*

$$(\mathbb{D}_N^\alpha v^n) v^n \leq \xi^n v^n + (\eta^n)^2 \text{ for } n = 1, 2, \dots, N. \tag{5.4}$$

Then

$$v^n \leq v^0 + \frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} (\xi^j + \eta^j) + \max_{1 \leq j \leq n} \{\eta^j\} \text{ for } n = 1, 2, \dots, N. \tag{5.5}$$

**Proof** For  $n = 1, \dots, N$ , set

$$\eta_*^n := \max_{1 \leq j \leq n} \{\eta^j\}, \quad g^n := \xi^n + \eta^n.$$

Our proof uses induction on  $n$ . First consider the case  $n = 1$ . If  $v^1 \leq \eta_*^1$ , then as  $v^0, \theta_{n,j}$  and  $\xi_j$  are all non-negative, the result for  $n = 1$  follows immediately. Otherwise  $v^1 > \eta_*^1$ , which implies  $v^1 > \eta^1 \geq 0$ . Hence the inequality (5.4) with  $n = 1$  gives us

$$\mathbb{D}_N^\alpha v^1 \leq \xi^1 + \eta^1 = g^1,$$

i.e.,

$$d_{1,1} v^1 - d_{1,1} v^0 \leq g^1$$

by (3.4). Rearranging this inequality then invoking (5.1), one has

$$v^1 \leq v^0 + \frac{1}{d_{1,1}} g^1 = v^0 + \frac{1}{d_{1,1}} \sum_{j=1}^1 \theta_{1,1} g^1$$

$$\leq v^0 + \frac{1}{d_{1,1}} \sum_{j=1}^1 \theta_{1,1} g^1 + \eta_*^1.$$

Thus the result is true for  $n = 1$ .

Fix  $k \in \{2, \dots, N\}$ . Assume that (5.5) is valid for  $n = 1, 2, \dots, k - 1$ . If  $v^k \leq \eta_*^k$ , then as  $v^0, \theta_{n,j}$  and  $\xi_j$  are all non-negative, the result for  $n = k$  follows immediately. Otherwise  $v^k > \eta_*^k$ , which implies  $v^k > \eta^k \geq 0$ . Hence the inequality (5.4) with  $n = k$  gives us

$$\mathbb{D}_N^\alpha v^k \leq \xi^k + \eta^k = g^k,$$

i.e.,

$$d_{k,1} v^k - d_{k,k} v^0 + \sum_{j=1}^{k-1} (d_{k,j+1} - d_{k,j}) v^{k-j} \leq g^k,$$

by (3.4). This is equivalent to

$$v^k \leq \frac{1}{d_{k,1}} \left[ g^k + d_{k,k} v^0 + \sum_{j=1}^{k-1} (d_{k,j} - d_{k,j+1}) v^{k-j} \right].$$

Combining this inequality with the inductive hypothesis yields

$$\begin{aligned} v^k &\leq \frac{1}{d_{k,1}} \left\{ g^k + d_{k,k} v^0 + \sum_{j=1}^{k-1} (d_{k,j} - d_{k,j+1}) \left[ v^0 + \frac{1}{d_{k-j,1}} \sum_{s=1}^{k-j} \theta_{k-j,s} g^s + \eta_*^{k-j} \right] \right\} \\ &\leq \frac{1}{d_{k,1}} \left\{ g^k + d_{k,k} v^0 + \sum_{j=1}^{k-1} (d_{k,j} - d_{k,j+1}) \left[ v^0 + \frac{1}{d_{k-j,1}} \sum_{s=1}^{k-j} \theta_{k-j,s} g^s \right] \right\} \\ &\quad + \frac{1}{d_{k,1}} \sum_{j=1}^{k-1} (d_{k,j} - d_{k,j+1}) \eta_*^{k-j} \\ &\leq \frac{1}{d_{k,1}} \left\{ g^k + d_{k,1} v^0 + \sum_{j=1}^{k-1} \left[ (d_{k,j} - d_{k,j+1}) \frac{1}{d_{k-j,1}} \sum_{s=1}^{k-j} \theta_{k-j,s} g^s \right] \right\} \\ &\quad + \frac{1}{d_{k,1}} (d_{k,1} - d_{k,k}) \eta_*^{k-1} \\ &\leq \frac{1}{d_{k,1}} \left\{ g^k + d_{k,1} v^0 + \sum_{s=1}^{k-1} g^s \left[ \sum_{j=1}^{k-s} \frac{1}{d_{k-j,1}} (d_{k,j} - d_{k,j+1}) \theta_{k-j,s} \right] \right\} \\ &\quad + \eta_*^k \\ &= v^0 + \frac{1}{d_{k,1}} \sum_{s=1}^k \theta_{k,s} g^s + \eta_*^k, \end{aligned}$$

where we used the relationship (5.1),  $d_{k,k} > 0$ , and  $\eta_*^n$  is nondecreasing for  $n$  increasing. We conclude the lemma is true by the principle of induction.  $\square$

Now we will achieve a  $\beta$ -robust stability result for the fully discrete L1 FEM by Lemma 5.3.

**Theorem 5.1** *The solution  $u_h^n$  of the discrete problem (3.9) satisfies*

$$\|\nabla u_h^n\| \leq \|\nabla u_h^0\| + \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \max_{1 \leq j \leq n} \{ \|f^j\| \}, \tag{5.6}$$

where  $l_N$  is defined in Corollary 5.2.

**Proof** Fix  $n \in \{1, 2, \dots, N\}$ . Multiplying (3.9) by  $-\Delta_h u_h^n$  and integrating over  $\Omega$ , one has

$$-(\mathbb{D}_N^\alpha u_h^n, \Delta_h u_h^n) + \kappa \|\Delta_h u_h^n\|^2 = -(P_h f^n, \Delta_h u_h^n) \leq \frac{1}{4\kappa} \|f^n\|^2 + \kappa \|\Delta_h u_h^n\|^2.$$

Discard the non-negative term  $\|\Delta_h u_h^n\|^2$ , then use the definition (2.4) of  $\Delta_h$  to get

$$(\mathbb{D}_N^\alpha (\nabla u_h^n), \nabla u_h^n) \leq \frac{1}{4\kappa} \|f^n\|^2.$$

Applying Lemma 4.1 yields

$$(\mathbb{D}_N^\alpha \|\nabla u_h^n\|) \|\nabla u_h^n\| \leq \frac{1}{4\kappa} \|f^n\|^2. \tag{5.7}$$

By Lemma 5.3, one has

$$\|\nabla u_h^n\| \leq \|\nabla u_h^0\| + \frac{1}{2\sqrt{\kappa}} \frac{1}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \|f^j\| + \frac{1}{2\sqrt{\kappa}} \max_{1 \leq j \leq n} \{ \|f^j\| \}.$$

Applied Corollary 5.2, the lemma is proved. □

We then prove the main result of the paper, which demonstrates convergence of our method in  $L^\infty(H^1)$  with an optimal and  $\beta$ -robust convergence rate.

**Theorem 5.2** (Error estimate) *Let  $u^n$  and  $u_h^n$  be the solutions of (1.1) and (3.9), respectively. Assume the hypotheses of (1.3) with  $k + 1 \leq p$ ,  $\omega(\alpha) \in C^2[0, \beta]$ , and  $D_t^\alpha u(\cdot, \cdot) \in C^2[0, \beta]$ . Then for  $n = 1, 2, \dots, N$ , one has*

$$\|\nabla u^n - \nabla u_h^n\| \leq \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^k + h_\alpha^2 \right), \tag{5.8}$$

$$\|\nabla R_h u^n - \nabla u_h^n\| \leq \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^{k+1} + h_\alpha^2 \right). \tag{5.9}$$

where  $l_N = 1/\ln N$  and  $C$  is a constant independent of  $h$  and  $N$ .

**Proof** Fix  $n \in \{1, 2, \dots, N\}$ . Multiplying (4.5) by  $-\Delta_h \zeta^n$  and integrating over  $\Omega$ , we arrive at

$$-(\mathbb{D}_N^\alpha \zeta^n, \Delta_h \zeta^n) + \kappa \|\Delta_h \zeta^n\|^2 = -(P_h \mathcal{D}_t^\omega \rho^n - R_h R^n, \Delta_h \zeta^n) + (R_h \varphi^n, \Delta_h \zeta^n). \tag{5.10}$$

Recalling the definition (2.4) of  $\Delta_h$  and the projection  $P_h$  yields

$$(\mathbb{D}_N^\alpha (\nabla \zeta^n), \nabla \zeta^n) + \kappa \|\Delta_h \zeta^n\|^2 = (-\mathcal{D}_t^\omega \rho^n + R_h R^n, \Delta_h \zeta^n) - (\nabla R_h \varphi^n, \nabla \zeta^n).$$

Applying the Lemma 4.1 and the Cauchy-Schwartz inequality, one has

$$(\mathbb{D}_N^\alpha \|\nabla \zeta^n\|) \|\nabla \zeta^n\| \leq \frac{1}{4\kappa} \|\mathcal{D}_t^\omega \rho^n - R_h R^n\|^2 + \|\nabla R_h \varphi^n\| \|\nabla \zeta^n\|.$$

Invoking (2.2) and (2.3), we get

$$\begin{aligned}
 (\mathbb{D}_N^\alpha \|\nabla \zeta^n\|) \|\nabla \zeta^n\| &\leq \frac{1}{4\kappa} (\|\mathcal{D}_t^\omega \rho^n\| + \|R_h R^n\|)^2 + \|\nabla R_h \varphi^n\| \|\nabla \zeta^n\| \\
 &\leq Ch^{2(k+1)} + C\|\nabla R_h R^n\|^2 + \|\nabla R_h \varphi^n\| \|\nabla \zeta^n\| \\
 &\leq C(h^{2(k+1)} + h_\alpha^4) + \|\nabla \varphi^n\| \|\nabla \zeta^n\|,
 \end{aligned}
 \tag{5.11}$$

where the inequality  $\|\nabla R_h w\| \leq \|\nabla w\| \forall w \in H_0^1(\Omega)$  is used. Observe that (5.11) is a particular case of (5.4). Thus we can invoke Lemma 5.3 to get

$$\|\nabla \zeta^n\| \leq \|\nabla \zeta^0\| + \frac{C}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \left( \|\nabla \varphi^j\| + h^{k+1} + h_\alpha^2 \right) + C(h^{k+1} + h_\alpha^2). \tag{5.12}$$

By Lemma 3.3, we get  $\|\nabla \varphi^j\| \leq C \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} N^{-\min\{2-\alpha_s, r\sigma\}}$ . Substituting this inequality into (5.12) and invoking Corollaries 5.1 and 5.2 yields

$$\begin{aligned}
 \|\nabla \zeta^n\| &\leq \|\nabla \zeta^0\| + \frac{C}{d_{n,1}} \sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} N^{-\min\{2-\alpha_s, r\sigma\}} \right) \theta_{n,j} \\
 &\quad + \frac{C}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \left( h^{k+1} + h_\alpha^2 \right) + C(h^{k+1} + h_\alpha^2) \\
 &\leq CN^{-\min\{2-\alpha_q, r\sigma\}} \frac{1}{d_{n,1}} \sum_{j=1}^n \left( \sum_{s=1}^q h_\alpha \omega(\alpha_s) t_j^{-\alpha_s} \right) \theta_{n,j} \\
 &\quad + \frac{C}{d_{n,1}} \sum_{j=1}^n \theta_{n,j} \left( h^{k+1} + h_\alpha^2 \right) + C(h^{k+1} + h_\alpha^2) \\
 &\leq \frac{e^r \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} C \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^{k+1} + h_\alpha^2 \right),
 \end{aligned}$$

where we used  $\|\nabla \zeta^0\| = \|\nabla(R_h u^0 - u_h^0)\| = 0$ . Combining this bound and (2.3) with (4.3), we get (5.8). □

**Remark 5.1** No blow up appear in the error estimate given in Theorem 5.2 as  $\beta \rightarrow 1^-$ , unlike the bound in Theorem 4.1 and [32, Theorem 3.1].

**Corollary 5.3** Assume the hypotheses of Theorem 5.2 are satisfied. Then

$$\|u^n - u_h^n\| \leq \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^{k+1} + h_\alpha^2 \right). \tag{5.13}$$

**Proof** Applying Poincare inequality, one has  $\|R_h u^n - u_h^n\| \leq C\|\nabla R_h u^n - \nabla u_h^n\|$ . Combining this result with (5.9) and (2.3), we get (5.13). □

## 6 Superconvergence Analysis

In this section, a superconvergent result for the distributed order time-fractional diffusion equation (1.1) in two-dimensions will be presented, where the bilinear element ( $k = 1$ ) is used in our finite element space.

To obtain the superconvergent result, we will introduce two operators as follows. Let  $\pi_h : H^2(\Omega) \rightarrow V_{0h}$  be the interpolation operator satisfying  $\pi_h v(a_i) = v(a_i)$ , where  $a_i$ , ( $i = 1, 2, 3, 4$ ) are the four vertices of  $K_m$ . By [37, (8)], we get the  $H^1$ -norm estimate

$$\|R_h w - \pi_h w\|_1 \leq Ch^2 \|w\|_3, \quad \forall w \in H_0^1(\Omega) \cap H^3(\Omega), \tag{6.1}$$

which will play an important role in the superclose and superconvergence analysis.

Now we adopt the interpolation postprocessing operator  $\pi_{2h}$  as the same in [28], which satisfies

$$\pi_{2h}\pi_h w = \pi_{2h} w, \quad \forall w \in H^2(\Omega), \tag{6.2a}$$

$$\|w - \pi_{2h} w\|_1 \leq Ch^2 |w|_3, \quad \forall w \in H^3(\Omega), \tag{6.2b}$$

$$\|\pi_{2h} w_h\|_1 \leq C \|w_h\|_1, \quad \forall w_h \in V_{0h}. \tag{6.2c}$$

Next we will state the global superconvergence result.

**Corollary 6.1** *Assume the hypotheses of Theorem 5.2 are satisfied with  $p = 3$ . Using bilinear element ( $k = 1$ ) in our finite element space. If the domain  $\Omega$  is rectangular with sides parallel to the coordinate axes, one has*

$$\|\nabla \pi_h u^n - \nabla u_h^n\| \leq \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^2 + h_\alpha^2 \right), \tag{6.3}$$

$$\|\nabla u^n - \nabla \pi_{2h} u_h^n\| \leq \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^2 + h_\alpha^2 \right). \tag{6.4}$$

**Proof** By triangle inequality, we arrive at  $\|\nabla \pi_h u^n - \nabla u_h^n\| \leq \|\nabla \pi_h u^n - \nabla R_h u^n\| + \|\nabla R_h u^n - \nabla u_h^n\|$ . Then the bound (6.3) follows from (6.1) and (5.9) immediately.

Furthermore, combining the bound (6.3) with (6.2) yields

$$\begin{aligned} \|\nabla u^n - \nabla \pi_{2h} u_h^n\| &\leq \|\nabla u^n - \nabla \pi_{2h} \pi_h u^n\| + \|\nabla \pi_{2h} \pi_h u^n - \nabla \pi_{2h} u_h^n\| \\ &= \|\nabla u^n - \nabla \pi_{2h} u^n\| + \|\nabla \pi_{2h} (\pi_h u^n - u_h^n)\| \\ &\leq Ch^2 + C \|\nabla \pi_h u^n - \nabla u_h^n\| \\ &\leq \frac{C \max_{1 \leq s \leq q} \Gamma(1 + l_N - \alpha_s)}{\Gamma(1 + l_N)} \left( N^{-\min\{2-\alpha_q, r\sigma\}} + h^2 + h_\alpha^2 \right). \end{aligned}$$

Thus the proof is complete. □

**Remark 6.1** Under the same condition of Corollary 6.1, the spatial  $H^1$ -norm error of exact solution and numerical solution given in (5.8) reaches  $O(h)$  convergence for the bilinear element in space (note that the degree of the polynomial is  $k = 1$ ). However, according to the superconvergent result (6.4), the  $H^1$ -norm error of the exact solution and numerical solution after postprocessing is improved to  $O(h^2)$  convergence.

## 7 Numerical Experiments

In this section we will present some numerical results for the initial-boundary value problem (1.1) whose solution mimic the behaviour described in (1.3) with  $\sigma = \beta$ .

Define the errors  $E_0^{M,N}$ ,  $E_1^{M,N}$ , and  $E_2^{M,N}$  for the computed solutions by

$$E_0^{M,N} := \max_{0 \leq n \leq N} \|\nabla u^n - \nabla u_h^n\|,$$

**Table 1** Example 7.1:  $E_0^{M,N}$  errors and rates of convergence in temporal direction

	N=10	N=20	N=30	N=40
$\beta = 0.3$	6.8409E-2	2.1408E-2	1.0737E-2	6.5847E-3
		<b>1.6759</b>	<b>1.7017</b>	<b>1.6998</b>
$\beta = 0.5$	1.0902E-1	3.8769E-2	2.1148E-2	1.3792E-2
		<b>1.5046</b>	<b>1.4947</b>	<b>1.4857</b>
$\beta = 0.7$	1.7449E-1	6.9783E-2	4.1538E-2	2.8837E-2
		<b>1.3222</b>	<b>1.2794</b>	<b>1.2685</b>

$$E_1^{M,N} := \max_{0 \leq n \leq N} \|\nabla \pi_h u^n - u_h^n\|, \quad E_2^{M,N} := \max_{0 \leq n \leq N} \|\nabla u^n - \nabla \pi_{2h} u_h^n\|.$$

**Example 7.1** Consider the problem (1.1) with  $\kappa = 0.1, T = 1, \omega(\alpha) = \Gamma(\beta + 1 - \alpha)/\Gamma(1 + \beta), \Omega = (0, 1) \times (0, 1)$ , and the function  $f$  is chosen such that the exact solution of this problem is  $u(x, y, t) = t^\beta \sin(\pi x) \sin(\pi y)$ .

To solve Example 7.1 numerically, a uniform rectangular partition of  $\Omega$  with  $M + 1$  nodes in each spatial direction and the bilinear polynomial in spatial are used. By taking  $r = (2 - \beta)/\beta$ , one obtains the optimal rates of convergence in Theorem 5.2 and Corollary 6.1, viz.,  $O(h + h_\alpha^2 + N^{-(2-\beta)})$  for  $\|\nabla u^n - \nabla u_h^n\|$  and  $O(h^2 + h_\alpha^2 + N^{-(2-\beta)})$  for  $\|\nabla \pi_h u^n - \nabla u_h^n\|$  and  $\|\nabla u^n - \nabla \pi_{2h} u_h^n\|$ .

Firstly, we verify the temporal accuracy of our fully discrete L1 FEM (3.9). Table 1 shows the  $E_0^{M,N}$  errors for  $\beta = 0.3, 0.5, 0.7$ . Here  $M = \lceil N^{2-\beta} \rceil$  and  $q = 100$  are taken so that the temporal error dominates the result. The orders of convergence displayed indicate that the rate of convergence is  $N^{-(2-\beta)}$ , as predicted by (5.8) of Theorem 5.2. Tables 2 and 3 display the  $E_1^{M,N}$  and  $E_2^{M,N}$  error and their associated orders of convergence for  $\beta = 0.3, 0.5, 0.7$ , with  $M = \lceil N^{1-\beta/2} \rceil$  and  $q = 100$  so that the temporal error dominates the distributed variable error and the spatial error. The orders of convergence displayed indicate that the rate of convergence is  $N^{-(2-\beta)}$ , as predicted by Corollary 6.1.

Next we test the accuracy in spatial direction. Table 4 shows the  $E_0^{M,N}, E_1^{M,N}$ , and  $E_2^{M,N}$  errors and their associated orders of convergence for  $\beta = 0.3, 0.5, 0.7$ . Here  $N = 200$  and  $q = 100$  are taken so that the spatial error dominates the results. We observe  $O(h)$  convergence for  $E_0^{M,N}$  and  $O(h^2)$  convergence for  $E_1^{M,N}$  and  $E_2^{M,N}$ , again as predicted by Theorem 5.2 and Corollary 6.1.

At last we check the convergence order for distributed variable. Table 5 shows the  $E_1^{M,N}$  error and the associated order of convergence for  $\beta = 0.3, 0.5, 0.7$ , where  $N = 1000$  and  $M = 200$  are taken to eliminate the temporal error and the spatial error.

These numerical results demonstrate the sharpness of our theoretical convergence bounds in Theorem 5.2 and Corollary 6.1.

**Example 7.2** Consider the problem (1.1) with  $\kappa = 0.1, T = 1, \omega(\alpha) = \Gamma(\beta + 1 - \alpha)/\Gamma(1 + \beta), \Omega = (0, 1) \times (0, 1)$ , and  $\phi(x, y) = x^{2.5}(x - 1)y^{2.5}(y - 1)$ . The function  $f$  is chosen such that the exact solution of this problem is  $u(x, y, t) = (1 + t^\beta)x^{2.5}(x - 1)y^{2.5}(y - 1)$ , which is nonsmooth in spatial direction.

In this example, we just test the convergent result of  $E_1^{M,N}$  and  $E_2^{M,N}$ , and the selection of  $M, N$ , and  $q$  is same as Example 7.1. Tables 6 and 7 show that  $E_1^{M,N}$  and  $E_2^{M,N}$  have the global truncation error  $O(N^{-(2-\beta)})$  in temporal direction. Table 8 shows that  $O(h^2)$  convergence for  $E_1^{M,N}$  and  $E_2^{M,N}$  in spatial direction is observed.



**Table 2** Example 7.1:  $E_1^{M,N}$  errors and rates of convergence in temporal direction

	N=40	N=80	N=160	N=320	N=640
$\beta = 0.3$	3.4599E-3	1.1331E-3 <b>1.6104</b>	3.5575E-4 <b>1.6703</b>	1.0994E-4 <b>1.6943</b>	3.4313E-5 <b>1.6812</b>
$\beta = 0.5$	8.3643E-3	2.9477E-3 <b>1.5046</b>	1.0639E-3 <b>1.4755</b>	3.7353E-4 <b>1.5029</b>	1.3185E-3 <b>1.5025</b>
$\beta = 0.7$	1.8759E-2	7.0388E-3 <b>1.4141</b>	2.9176E-3 <b>1.2704</b>	1.2397E-3 <b>1.2345</b>	5.1130E-4 <b>1.2789</b>

**Table 3** Example 7.1:  $E_2^{M,N}$  errors and rates of convergence in temporal direction

	N=40	N=80	N=160	N=320	N=640
$\beta = 0.3$	6.6453E-3	2.1717E-3 <b>1.6134</b>	6.7214E-4 <b>1.6923</b>	2.0878E-4 <b>1.6849</b>	6.4730E-5 <b>1.6916</b>
$\beta = 0.5$	1.5621E-2	5.1626E-3 <b>1.5637</b>	1.8896E-3 <b>1.4497</b>	6.7802E-4 <b>1.4790</b>	2.3912E-4 <b>1.5042</b>
$\beta = 0.7$	2.9737E-2	1.2300E-2 <b>1.1681</b>	5.0884E-3 <b>1.2740</b>	2.1082E-3 <b>1.3037</b>	8.7630E-4 <b>1.2561</b>

**Table 4** Example 7.1:  $E_0^{M,N}$ ,  $E_1^{M,N}$ , and  $E_2^{M,N}$  errors and convergence rates in spatial direction

	M	$E_0^{M,N}$	Order	$E_1^{M,N}$	Order	$E_2^{M,N}$	Order
$\beta = 0.3$	8	4.3761E-1	-	3.1868E-2	-	6.0159E-2	-
	16	2.1826E-1	<b>1.0035</b>	8.0456E-3	<b>1.9858</b>	1.5089E-2	<b>1.9953</b>
	32	1.0906E-1	<b>1.0008</b>	2.0105E-3	<b>2.0006</b>	3.7721E-3	<b>2.0000</b>
	64	5.4525E-2	<b>1.0002</b>	4.9683E-4	<b>2.0162</b>	9.3996E-4	<b>2.0046</b>
$\beta = 0.5$	8	4.3790E-1	-	3.3903E-2	-	6.1283E-2	-
	16	2.1830E-1	<b>1.0042</b>	8.5523E-3	<b>1.9870</b>	1.5366E-2	<b>1.9957</b>
	32	1.0907E-1	<b>1.0010</b>	2.1318E-3	<b>2.0042</b>	3.8381E-3	<b>2.0012</b>
	64	5.4525E-2	<b>1.0002</b>	5.2149E-4	<b>2.0327</b>	9.5323E-4	<b>2.0098</b>
$\beta = 0.7$	8	4.3817E-1	-	3.5708E-2	-	6.2322E-2	-
	16	2.1833E-1	<b>1.0049</b>	8.9989E-3	<b>1.9884</b>	1.5621E-2	<b>1.9962</b>
	32	1.0907E-1	<b>1.0012</b>	2.2359E-3	<b>2.0088</b>	3.8970E-3	<b>2.0029</b>
	64	5.4526E-2	<b>1.0003</b>	5.3981E-4	<b>2.0498</b>	9.6338E-4	<b>2.0167</b>

**Table 5** Example 7.1:  $E_1^{M,N}$  errors and convergence rates for the distributed variable

	N=2	N=4	N=8	N=16
$\beta = 0.3$	5.2242E-3	1.3312E-3 <b>1.9723</b>	3.3415E-4 <b>1.9942</b>	8.3631E-5 <b>1.9989</b>
$\beta = 0.5$	7.1872E-3	1.8175E-3 <b>1.9834</b>	4.4140E-4 <b>2.0431</b>	9.7649E-5 <b>2.1758</b>
$\beta = 0.7$	9.2441E-3	2.3636E-3 <b>1.9697</b>	5.8322E-4 <b>2.0166</b>	1.3611E-4 <b>2.0992</b>

**Table 6** Example 7.2:  $E_1^{M,N}$  errors and rates of convergence in temporal direction

	N=40	N=80	N=160	N=320	N=640
$\beta = 0.3$	2.1250E-4	6.9723E-5	2.1326E-5	6.6642E-6	2.0707E-6
		<b>1.6116</b>	<b>1.7103</b>	<b>1.6772</b>	<b>1.6858</b>
$\beta = 0.5$	4.8051E-4	1.5851E-4	5.8931E-5	2.1619E-5	7.6261E-6
		<b>1.5970</b>	<b>1.4326</b>	<b>1.4472</b>	<b>1.5012</b>
$\beta = 0.7$	8.5790E-4	3.8711E-4	1.6112E-4	6.5448E-5	2.7437E-5
		<b>1.1486</b>	<b>1.2652</b>	<b>1.2996</b>	<b>1.2551</b>

**Table 7** Example 7.2:  $E_2^{M,N}$  errors and rates of convergence in temporal direction

	N=40	N=80	N=160	N=320	N=640
$\beta = 0.3$	4.6671E-4	1.5443E-4	4.7852E-5	1.5154E-5	4.7729E-6
		<b>1.6004</b>	<b>1.6848</b>	<b>1.6559</b>	<b>1.6720</b>
$\beta = 0.5$	1.0445E-3	3.4508E-4	1.2938E-4	4.7984E-5	1.7129E-5
		<b>1.5981</b>	<b>1.4192</b>	<b>1.4262</b>	<b>1.4890</b>
$\beta = 0.7$	1.8559E-3	8.3009E-4	3.4631E-4	1.4175E-4	5.9976E-5
		<b>1.1609</b>	<b>1.2623</b>	<b>1.2848</b>	<b>1.2428</b>

**Table 8** Example 7.2:  $\beta = 0.3$ ;  $E_1^{M,N}$ ,  $E_2^{M,N}$  errors, and convergence rates in spatial direction

Polynomial	M	$E_1^{M,N}$	Order	$E_2^{M,N}$	Order
$Q_1$	8	1.8181E-3	-	4.1434E-3	-
	16	4.7576E-4	<b>1.9334</b>	1.0421E-3	<b>1.9930</b>
	32	1.2027E-4	<b>1.9879</b>	2.6426E-4	<b>1.9779</b>
	64	3.0118E-5	<b>1.9951</b>	6.7221E-5	<b>1.9967</b>

**Acknowledgements** The research of Chaobao Huang is supported in part by the National Natural Science Foundation of China under grants Nos. 12101360 and 12171278 and the Natural Science Foundation of Shandong Province under grant ZR2020QA031. The research of Hu Chen is supported in part by the National Natural Science Foundation of China under grant 11801026, sponsored by OUC Scientific Research Starting Fund of Introduced Talent. The research of Na An is supported in part by the National Natural Science Foundation of China under grant 11801332.

**Data Availability** The datasets generated during the current study are available from the corresponding author on reasonable request.

### Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Code availability** Codes of the current study are available from the authors on reasonable request.

## References

1. Abbaszadeh, Mostafa, Dehghan, Mehdi: An improved meshless method for solving two-dimensional distributed order time-fractional diffusion-wave equation with error estimate. *Numer. Algorithms* **75**(1), 173–211 (2017)
2. Alikhanov, Anatoly A.: Numerical methods of solutions of boundary value problems for the multi-term variable-distributed order diffusion equation. *Appl. Math. Comput.* **268**, 12–22 (2015)
3. An, N.: Superconvergence of a finite element method for the time-fractional diffusion equation with a time-space dependent diffusivity. *Adv. Differ. Equ.* **2020**(1), 1–11 (2020)
4. An, Na., Huang, Chaobao, Xijun, Yu.: Error analysis of direct discontinuous Galerkin method for two-dimensional fractional diffusion-wave equation. *Appl. Math. Comput.* **349**, 148–157 (2019)
5. An, Na., Huang, Chaobao, Xijun, Yu.: Error analysis of discontinuous Galerkin method for the time fractional KdV equation with weak singularity solution. *Discrete Contin. Dyn. Syst. Ser. B* **25**(1), 321–334 (2020)
6. Bramble, James H., Pasciak, Joseph E., Steinbach, Olaf: On the stability of the  $L^2$  projection in  $H^1(\Omega)$ . *Math. Comput.* **71**(237), 147–156 (2002)
7. Weiping, Bu., Ji, Lun, Tang, Yifa, Zhou, Jie: Space-time finite element method for the distributed-order time fractional reaction diffusion equations. *Appl. Numer. Math.* **152**, 446–465 (2020)
8. Weiping, Bu., Xiao, Aiguo, Zeng, Wei: Finite difference/finite element methods for distributed-order time fractional diffusion equations. *J. Sci. Comput.* **72**(1), 422–441 (2017)
9. Chen, Hu., Lü, Shujuan, Chen, Wenping: Finite difference/spectral approximations for the distributed order time fractional reaction-diffusion equation on an unbounded domain. *J. Comput. Phys.* **315**, 84–97 (2016)
10. Chen, Hu., Stynes, Martin: Blow-up of error estimates in time-fractional initial-boundary value problems. *IMA J. Numer. Anal.* **41**(2), 974–997 (2021)
11. Dahlquist, Germund, Björck, Åke.: *Numerical methods in scientific computing*, vol. 1. Society for Industrial and Applied Mathematics, USA (2008)
12. Ganesan, Sashikumar, Tobiska, Lutz: *Finite elements Theory and algorithms*. Cambridge University Press, Delhi (2017)
13. Gorenflo, Rudolf, Luchko, Yuri, Stojanović, Mirjana: Fundamental solution of a distributed order time-fractional diffusion-wave equation as probability density. *Fract. Calc. Appl. Anal.* **16**(2), 297–316 (2013)
14. Huang, Chaobao, An, Na., Xijun, Yu.: A local discontinuous Galerkin method for time-fractional diffusion equation with discontinuous coefficient. *Appl. Numer. Math.* **151**, 367–379 (2020)
15. Huang, Chaobao, An, Na., Xijun, Yu., Zhang, Huili: A direct discontinuous Galerkin method for time-fractional diffusion equation with discontinuous diffusive coefficient. *Complex Var. Elliptic Equ.* **65**(9), 1445–1461 (2020)
16. Chaobao Huang and Martin Stynes: Superconvergence of a finite element method for the multi-term time-fractional diffusion problem. *J. Sci. Comput.* **82**(1), 1–17 (2020)
17. Huang, Chaobao, Stynes, Martin, An, Na.: Optimal  $L^\infty(L^2)$  error analysis of a direct discontinuous Galerkin method for a time-fractional reaction-diffusion problem. *BIT* **58**(3), 661–690 (2018)
18. Huang, Chaobao, Stynes, Martin, Chen, Hu.: An  $\alpha$ -robust finite element method for a multi-term time-fractional diffusion problem. *J. Comput. Appl. Math.* **389**, 113334 (2021)
19. Jia, Jinhong, Wang, Hong, Zheng, Xiangcheng: A fast collocation approximation to a two-sided variable-order space-fractional diffusion equation and its analysis. *J. Comput. Appl. Math.* **388**, 113234 (2021)
20. Kopteva, Natalia, Meng, Xiangyun: Error analysis for a fractional-derivative parabolic problem on quasi-graded meshes using barrier functions. *SIAM J. Numer. Anal.* **58**(2), 1217–1238 (2020)
21. Li, Dongfang, Chengda, Wu., Zhang, Zhimin: Linearized Galerkin FEMs for nonlinear time fractional parabolic problems with non-smooth solutions in time direction. *J. Sci. Comput.* **80**(1), 403–419 (2019)
22. Hongwei Li and Yuchen Wu: Artificial boundary conditions for nonlinear time fractional Burgers' equation on unbounded domains. *Appl. Math. Lett.* **120**, 107277 (2021)
23. Li, Xiaoli, Rui, Hongxing, Liu, Zhengguang: Two alternating direction implicit spectral methods for two-dimensional distributed-order differential equation. *Numer. Algorithms* **82**(1), 321–347 (2019)
24. Li, Zhiyuan, Luchko, Yuri, Yamamoto, Masahiro: Asymptotic estimates of solutions to initial-boundary-value problems for distributed order time-fractional diffusion equations. *Fract. Calc. Appl. Anal.* **17**(4), 1114–1136 (2014)
25. Liang, Hui, Stynes, Martin: Collocation methods for general Caputo two-point boundary value problems. *J. Sci. Comput.* **76**(1), 390–425 (2018)
26. Liao, Hong-lin, Li, Dongfang, Zhang, Jiwei: Sharp error estimate of the nonuniform L1 formula for linear reaction-subdiffusion equations. *SIAM J. Numer. Anal.* **56**(2), 1112–1133 (2018)

27. Liao, Hong-lin, McLean, William, Zhang, Jiwei: A discrete Grönwall inequality with applications to numerical schemes for subdiffusion problems. *SIAM J. Numer. Anal.* **57**(1), 218–237 (2019)
28. Qun Lin and Jiafu Lin: *Finite element methods: accuracy and improvement*. Elsevier, Netherlands (2007)
29. Luchko, Yury: Boundary value problems for the generalized time-fractional diffusion equation of distributed order. *Fract. Calc. Appl. Anal.* **12**(4), 409–422 (2009)
30. McLean, William: Regularity of solutions to a time-fractional diffusion equation. *ANZIAM J.* **52**(2), 123–138 (2010)
31. Igor Podlubny. *Fractional differential equations*, volume 198 of *Mathematics in Science and Engineering*. Academic Press, Inc., San Diego, CA, 1999. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications
32. Ren, Jincheng, Chen, Hu.: A numerical method for distributed order time fractional diffusion equation with weakly singular solutions. *Appl. Math. Lett.* **96**, 159–165 (2019)
33. Ren, Jincheng, Huang, Chaobao, An, Na.: Direct discontinuous Galerkin method for solving nonlinear time fractional diffusion equation with weak singularity solution. *Appl. Math. Lett.* **102**, 106111 (2020)
34. Samiee, Mehdi, Kharazmi, Ehsan, Meerschaert, Mark M., Zayernouri, Mohsen: A unified Petrov-Galerkin spectral method and fast solver for distributed-order partial differential equations. *Commun. Appl. Math. Comput.* **3**(1), 61–90 (2021)
35. Mehdi Samiee, Ehsan Kharazmi, Mohsen Zayernouri, and Mark M Meerschaert. Petrov-galerkin method for fully distributed-order fractional partial differential equations, 2018
36. Shen, Jinye, Li, Changpin, Sun, Zhi-zhong: An H2N2 interpolation for Caputo derivative with order in (1, 2) and its application to time-fractional wave equations in more than one space dimension. *J. Sci. Comput.* **83**(2), 29 (2020)
37. Shi, Dong Yang, Wang, Fen Ling, Fan, Ming Zhi, Zhao, Yan Min: A new approach of the lowest-order anisotropic mixed finite element high-accuracy analysis for nonlinear sine-Gordon equations. *Math. Numer. Sin.* **37**(2), 148–161 (2015)
38. Shi, Y.H., Liu, F., Zhao, Y.M., Wang, F.L., Turner, I.: An unstructured mesh finite element method for solving the multi-term time fractional and Riesz space distributed-order wave equation on an irregular convex domain. *Appl. Math. Model.* **73**, 615–636 (2019)
39. Stynes, Martin, O’Riordan, Eugene, Gracia, José Luis.: Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. *SIAM J. Numer. Anal.* **55**(2), 1057–1079 (2017)
40. Thomée, Vidar: *Galerkin finite element methods for parabolic problems*. Springer, Berlin (2006)
41. Wang, Feng, Chen, Huanzhen, Wang, Hong: Finite element simulation and efficient algorithm for fractional Cahn-Hilliard equation. *J. Comput. Appl. Math.* **356**, 248–266 (2019)
42. Wei, Leilei, Liu, Lijie, Sun, Huixia: Stability and convergence of a local discontinuous Galerkin method for the fractional diffusion equation with distributed order. *J. Appl. Math. Comput.* **59**(1–2), 323–341 (2019)
43. Ye, H., Liu, F., Anh, V.: Compact difference scheme for distributed-order time-fractional diffusion-wave equation on bounded domains. *J. Comput. Phys.* **298**, 652–660 (2015)
44. Ye, H., Liu, F., Anh, V., Turner, I.: Numerical analysis for the time distributed-order and Riesz space fractional diffusions on bounded domains. *IMA J. Appl. Math.* **80**(3), 825–838 (2015)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.