



Research article

A high-order stabilizer-free weak Galerkin finite element method on nonuniform time meshes for subdiffusion problems

Şuayip Toprakseven^{1,*} and Seza Dinibutun²

¹ Artvin Vocational School, Accounting and Taxation, Artvin Çoruh University, Artvin, 08100, Turkey

² International University of Kuwait, Ardiya, Kuwait Department of Mathematics and Natural Sciences, Kuwait

* **Correspondence:** Email: topraksp@artvin.edu.tr; Tel: +904662151000.

Abstract: We present a stabilizer-free weak Galerkin finite element method (SFWG-FEM) with polynomial reduction on a quasi-uniform mesh in space and Alikhanov’s higher order L^2-1_σ scheme for discretization of the Caputo fractional derivative in time on suitable graded meshes for solving time-fractional subdiffusion equations. Typical solutions of such problems have a singularity at the starting point since the integer-order temporal derivatives of the solution blow up at the initial point. Optimal error bounds in H^1 norm and L^2 norm are proven for the semi-discrete numerical scheme. Furthermore, we have obtained the values of user-chosen mesh grading constant r , which gives the optimal convergence rate in time for the fully discrete scheme. The optimal rate of convergence of order $O(h^{k+1} + M^{-2})$ in the $L^\infty(L^2)$ -norm has been established. We give several numerical examples to confirm the theory presented in this work.

Keywords: sub-diffusion problems; graded temporal meshes; stabilizer-free weak Galerkin finite element method; optimal rate of convergence

Mathematics Subject Classification: 65M06, 65M12, 65M60, 65M70

1. Introduction

In this work, we consider the following time fractional diffusion equation (TFDE) with zero Dirichlet boundary value:

$$\begin{cases} {}_0^C D_T^\alpha u(x, t) - \nabla \cdot (A(x)\nabla u(x, t)) = f(x, t), & \forall (x, t) \in Q_T := \Omega \times J, \\ u(x, t)|_{\partial\Omega} = 0, & \forall t \in [0, T], \\ u(x, 0) = g(x), & \forall x \in \Omega, \end{cases} \quad (1.1)$$

where $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain with its boundary $\partial\Omega$, $J = (0, T]$ is the time interval with the final time T , $A(x) = (a_{ij}(x))_{2 \times 2}$ is a symmetric positive definite matrix-valued function in Ω , g and f are given sufficiently smooth and ${}^C_0D_T^\alpha$ denotes the Caputo fractional derivative of order α [1],

$${}^C_0D_T^\alpha u(x, t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \frac{\partial u(x, s)}{\partial s} ds, \quad \alpha \in (0, 1). \quad (1.2)$$

In particular, we assume that the function $A(x)$ satisfies

$$k_1 v^T v \leq v^T A v \leq k_2 v^T v, \quad \forall v \in \mathbb{R}^2, \quad (1.3)$$

where k_1 and k_2 are positive constants.

We introduce the left-hand sided and right-hand sided Riemann-Liouville fractional integrals and the left-hand sided and right-hand sided Riemann-Liouville fractional derivatives, respectively, defined by [1]

$${}_0D_t^{-\alpha} u(x, t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} u(x, s) ds, \quad \alpha \geq 0, \quad (1.4)$$

$${}_tD_T^{-\alpha} u(x, t) = \frac{1}{\Gamma(\alpha)} \int_t^T (s-t)^{\alpha-1} u(x, s) ds, \quad \alpha \geq 0, \quad (1.5)$$

$${}^R_0D_T^\alpha u(x, t) = \frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_0^t (t-s)^{-\alpha} u(x, s) ds, \quad \alpha \in (0, 1], \quad (1.6)$$

$${}^R_tD_T^\alpha u(x, t) = \frac{-1}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_t^T (s-t)^{-\alpha} u(x, s) ds, \quad \alpha \in (0, 1]. \quad (1.7)$$

We will use the following identity from [1] in our analysis

$${}^R_0D_T^\alpha u(x, t) = {}^C_0D_T^\alpha u(x, t) + \frac{u(x, 0)}{\Gamma(1-\alpha)} t^{-\alpha}. \quad (1.8)$$

In recent decades, many physical phenomenon or processes in science and engineering have been modeled using fractional partial differential equations since they describe better memory effect and hereditary properties (see, [2–5] and references therein). The TFDE considered in this paper is obtained from the classical diffusion problem by taking a fractional derivative of order α in the place of the first-order time derivative. TFDEs are obtained from a fractional Fick law describing transport equations with long memory [6]. The TFDE is formulated by a linear integro-differential equation. Several analytical and numerical techniques have been considered for solving TFDE in the literature. The Fourier transform method, the Laplace transform method, the Mellin transform method and the Green function technique are some examples of the analytical approach (see [1, 7] and references therein). In general, the analytical solutions of fractional differential or partial differential equations are not available; thus one must construct a numerical method for solving this equation. Many papers have investigated robust numerical approximations to the time fractional diffusion equations or sub-diffusion problems [8–18]. The TFDEs with the Caputo derivatives have been solved by the widely used $L1$ method in [8–11, 19, 20]. The compact difference methods [12, 14, 16, 21] have been considered to improve the order of time approximation, and spectral methods [9, 17, 18] have been proposed for

improving the accuracy in the space discretization. High order numerical methods [22–27] have been presented for solving fractional partial differential equations. However, typical solutions of (1.1) have an initial layer when $t \rightarrow 0$ and the first order time derivative of the solution blows up at $t = 0$. Since the solution is singular at $t = 0$, one should construct a reliable and efficient numerical scheme in time [1]. Furthermore, the forcing function f in our model problem may blow up at the starting point $t = 0$ even though $u(x, t)$ is a continuous function in time. Thus, one needs to pay attention to the case $f(x, 0)$ when proposing a numerical method for solving the problem (1.1). As a remedy, we use the Alikhanov's $L2 - 1_\sigma$ method on graded meshes to approximate the time derivative in (1.1) for higher order convergence.

Our aim of this paper is to combine the $L2-1_\sigma$ method [28] due to Alikhanov for time discretization of the Caputo fractional derivative on graded temporal meshes with the SFWG-FEM in space on quasi-uniform meshes to develop a robust scheme to solve the TFDE (1.1). The weak Galerkin finite element method first appeared in [29] for solving the second order elliptic problems. The novelty of this method lies on the definitions of weak function space and weak derivatives such as weak gradient operator on this weak function space. This renders the method more applicable in defining finite element space and mesh constructions; for example, one uses completely discontinuous function spaces in the approximation and polygonal meshes. In the last decade, this method has been applied to solve a variety of differential equations, e.g., parabolic problems [30, 31], second order elliptic problems [32–39], Stokes equations [31], Maxwell equations [40] and fractional differential equations [41, 42]. The stabilized-free weak Galerkin method is introduced in [43] for solving second-order elliptic partial differential equations. This new method does not have a stabilization component, which is essential in discontinuous approximations to enforce the jump across the element boundaries. Also, the stabilization term is needed to ensure the coercivity of the WG-FEM. The absence of stabilization terms in numerical methods makes the formulation much simpler and provides much flexibility in implementation. The main idea of the SFWG-FEM is to use a higher degree of polynomials in computing weak differential operators for the strong connectivity of weak functions on element boundaries. In [44] and [45], the authors applied the SFWG-FEM to second-order elliptic equations and proved that the method has supercloseness convergence in an energy norm and L^2 -norm on triangular meshes. For more discussions of the method, we refer the interested readers to [43, 46, 47]. Further, in [48] the authors derived the semi-discrete SFWG-FEM formulation in space and presented the stability results and error estimates. Then, they established the fully discrete SFWG-FEM by discretization of the time using the $L2-1_\sigma$ formula to approximate the fractional time Caputo derivative in uniform meshes. However, Alikhanov's higher order $L2-1_\sigma$ formula has a rate of convergence $3 - \alpha$ on a uniform mesh if the solution is smooth enough, while this rate of convergence reduces to first order for the case where the typical solutions with a weak singularity at the starting point. As pointed out in [49, Theorem 2.1], the smooth solutions exist only under acceptably restrictive data. These observations and facts lead to using the $L2-1_\sigma$ formula on graded meshes in time and SFWG-FEM in space in this paper. We show the order of convergence result that gives second-order convergence in time and an optimal rate of convergence in space.

The remaining parts of the paper are organized as follows. Section 2 introduces a semi-discrete SFWG-FEM. Stability and error analysis of the semi-discrete method are presented. The fully-discrete SFWG-FEM and its stability and error analyses are given in Section 3. We present several numerical examples to support the theory in Section 4. Finally, we close the paper with some concluding remarks

in Section 5.

Throughout the paper, we use the following notations. $\|u\|$ represents the L^2 -norm of a function u in Ω and $(u, v) := \int_{\Omega} uv \, dx$ denotes the L^2 inner product. For $D \subset \Omega$, the classical Sobolev spaces are denoted by $H^s(\Omega)$ with $s \geq 0$ an integer and the corresponding norm $\|u\|_{s,D}$ and semi-norms $|u|_s$. Further, the space $L^\infty(J; H^k(\Omega))$ with the norm

$$\|v\|_{k,\infty} = \text{ess sup} \|v(\cdot, t)\|_k \quad (1.9)$$

is used. When $D = \Omega$, we do not write D in the subscript.

2. Semi-discrete scheme

We start with discretization of (1.1) only in space using a SFWG finite element method. Let \mathcal{T}_h be a partition of the domain Ω consisting of polygons with a set of shape-regular requirements given in [29]. Let \mathcal{E}_h be the set of all edges in \mathcal{T}_h , and $\mathcal{E}_h^0 = \mathcal{E} \setminus \partial\Omega$ be the set of all interior edges. For any element $K \in \mathcal{T}_h$, h_K is the diameter of K and the mesh size $h = \max_{K \in \mathcal{T}_h} h_K$.

We first formulate the weak form of (1.1): For any $t \in (0, T]$, we seek a solution $u(x, t) \in H_0^1(\Omega)$ satisfying

$$\begin{cases} ({}^C D_T^\alpha u, w) + (K \nabla u, \nabla w) = (f, w), & \forall w \in H_0^1(\Omega), \\ u_h(0) = \pi_h g, \end{cases} \quad (2.1)$$

where $\pi_h g$ is the L^2 projection of the function g on the weak Galerkin finite element space defined below.

For given integers $k \geq 1$, the weak Galerkin finite element space S_h associated with \mathcal{T}_h is defined by

$$S_h = \{\omega = \{\omega_0, \omega_b\} : \omega_0|_K \in \mathbb{P}_k(K), \quad \omega_b|_e \in \mathbb{P}_{k-1}(e), \quad e \subset \partial K, K \in \mathcal{T}_h\}, \quad (2.2)$$

and its subspace S_h^0 is given by

$$S_h^0 = \{\omega = \{\omega_0, \omega_b\} \in S_h : \omega_b = 0 \quad \text{on } \partial K \cap \partial\Omega, \forall K \in \mathcal{T}_h\}, \quad (2.3)$$

where $\mathbb{P}_k(K)$ is the space of polynomials with degree of at most k on an element K and $\mathbb{P}_{k-1}(e)$ is the space of polynomials with degree of at most $k - 1$ on an edge $e \subset \partial K$. Observe that ω_b of $v = \{\omega_0, \omega_b\}$ is allowed to have different value from the trace of ω_0 on the boundary of each element. The configuration of the SFWG-FEM scheme consists of $(\mathbb{P}_k(K), \mathbb{P}_{k-1}(e), \mathbb{P}_j(K)^d)$ with $j = d + k - 1$. Using the definition of the classical definition of weak gradient given by (2.4), the configuration of the SFWG-FEM $(\mathbb{P}_k(K), \mathbb{P}_{k-1}(e), \mathbb{P}_j(K)^d)$ delivers only sub-optimal spatial accuracy in both energy norm and the L^2 norm, presented in Table 1 ([50]).

Table 1. Weak gradient calculated by (2.4), $\|\cdot\|_E = \mathcal{O}(h^{r_1})$ and $\|\cdot\| = \mathcal{O}(h^{r_2})$.

$\mathbb{P}_k(K)$	$\mathbb{P}_{k-1}(e)$	$[\mathbb{P}_{k+1}(K)]^d$	r_1	r_2
$k = 1$	$k = 1$	$k = 1$	0	0
$k = 2$	$k = 2$	$k = 2$	1	2
$k = 3$	$k = 3$	$k = 3$	2	3

The standard definition for a weak gradient $\nabla_w u \in [\mathbb{P}_j(K)]^2$ of a weak function $u = \{u_0, u_b\}$ is a unique polynomial on each $K \in \mathcal{T}_h$ satisfying [29, 43, 51]

$$(\nabla_w u, \mathbf{r})_K = -(u_0, \nabla \cdot \mathbf{r})_K + \langle u_b, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} \quad \forall \mathbf{r} \in [\mathbb{P}_j(K)]^2. \quad (2.4)$$

Following the ideas in [52], we shall use a different definition of weak gradient given by (2.5) for the SFWG-FEM using the configuration $(P_k(K), P_{k-1}(e), [P_j(K)]^2)$. With this new definition, our proposed SFWG-FEM has optimal order spatial convergence rates in the energy and L^2 norms, presented in Table 2.

Table 2. Weak gradient calculated by (2.5), $\|\cdot\|_E = O(h^{r_1})$ and $\|\cdot\| = O(h^{r_2})$.

$\mathbb{P}_k(K)$	$\mathbb{P}_{k-1}(e)$	$[\mathbb{P}_{k+1}(K)]^d$	r_1	r_2
$k = 1$	$k = 1$	$k = 1$	1	2
$k = 2$	$k = 2$	$k = 2$	2	3
$k = 3$	$k = 3$	$k = 3$	3	4

For each $\omega = \{\omega_0, \omega_b\} \in S_h \cup H^1(\Omega)$, we define the *weak gradient* operator $\nabla_w \omega \in [\mathbb{P}_j(K)]^2$ of the weak function ω on K as a unique polynomial on K satisfying the following equation

$$\int_K \nabla_w \omega \phi \, dx = \int_K \nabla \omega_0 \phi \, dx + \int_{\partial K} \pi_b(\omega_0 - \omega_b) \phi \cdot \mathbf{n} \, ds, \quad \forall \phi \in [\mathbb{P}_j(K)]^2, \quad (2.5)$$

where \mathbf{n} is the unit outward normal vector to ∂K . In (2.5), we let $\omega_0 = \omega$ and $\omega_b = \omega$ if $\omega \in H^1(\Omega)$. Here, π_b is the L^2 -projection onto $\mathbb{P}_{k-1}(e)$ for any $e \in \mathcal{E}_h$. Let π_0 be the L^2 -projection onto the space $\mathbb{P}_k(K)$ for any $K \in \mathcal{T}_h$ and define $\pi_h u = \{\pi_0 u, \pi_b u\} \in S_h$ for any $u \in H^1(K)$. For any $K \in \mathcal{T}_h$, we denote by $\Pi_h : [L^2(K)]^2 \rightarrow [\mathbb{P}_j(K)]^2$ the L^2 -orthogonal projection defined by $\int_K (\Pi_h \tau - \tau) \sigma \, dx = 0$, $\forall \sigma \in [\mathbb{P}_j(K)]^2$. From now on we take $j = k + 1$.

We have the relation between the weak gradient and the projection Π_h stated in the following lemma.

Lemma 2.1. For any $v \in H^1(K)$ and $K \in \mathcal{T}_h$, one has

$$\nabla_w(\pi_h v) = \Pi_h \nabla v.$$

Proof. From (2.5) and integration by parts, one has for $\mathbf{r} \in [\mathbb{P}_{k+1}(K)]^2$

$$\begin{aligned} (\nabla_w \pi_h v, \mathbf{r})_K &= (\nabla \pi_0 v, \mathbf{r})_K + \langle \pi_b(\pi_b v - \pi_0 v), \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} \\ &= -(\pi_0 v, \nabla \cdot \mathbf{r})_K + \langle \pi_0 v, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} + \langle \pi_b v - \pi_0 v, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} \\ &= -(v, \nabla \cdot \mathbf{r})_K + \langle \pi_b v, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} \\ &= (\nabla v, \mathbf{r})_K - \langle v, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} + \langle \pi_b v, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial K} \\ &= (\Pi_h \nabla v, \mathbf{r})_K, \end{aligned}$$

where we used the definitions of projections π_0 , π_b and Π_h . The proof is completed. \square

For simplicity, we will use the following notations.

$$(z, w)_{\mathcal{T}_h} = \sum_{K \in \mathcal{T}_h} (z, w)_K = \sum_{K \in \mathcal{T}_h} \int_K zw \, dx, \quad \text{and} \quad \langle z, w \rangle = \sum_{K \in \mathcal{T}_h} \langle z, w \rangle_{\partial K} = \sum_{K \in \mathcal{T}_h} \int_{\partial K} zw \, ds.$$

For $u = \{u_0, u_b\} \in S_h$ and $\omega = \{\omega_0, \omega_b\} \in S_h$, we define the bilinear form as

$$A_h(u, \omega) := (A \nabla_w u, \nabla_w \omega)_{\mathcal{T}_h}. \quad (2.6)$$

Based on the weak form (2.1), the semi-discrete SFWG-FEM for the problem (1.1) is to find a numerical solution $u_h(t) = \{u_{0,h}(t), u_{b,h}(t)\} : (0, T] \rightarrow S_h^0$ such that

$$\begin{cases} ({}_0^C D_T^\alpha u_{0,h}, \omega_0) + A_h(u_h, \omega) = (f, \omega_0), \\ u_h(0) = \pi_h g, \end{cases} \quad (2.7)$$

for all $\omega = \{\omega_0, \omega_b\} \in S_h^0$.

We define two energy norms $\|\cdot\|_E$ and $\|\cdot\|_{1,h}$, respectively: For $\omega = \{\omega_0, \omega_b\} \in S_h^0$,

$$\|\omega\|_E := \left(\sum_{K \in \mathcal{T}_h} \|\nabla_w \omega\|_K^2 \right)^{1/2}, \quad (2.8)$$

$$\|\omega\|_{1,h} := \left(\sum_{K \in \mathcal{T}_h} \|\nabla \omega_0\|_K^2 + h_K^{-1} \|\pi_b(\omega_0 - \omega_b)\|_{\partial K}^2 \right)^{1/2}. \quad (2.9)$$

The following lemma shows the equivalence of the two norms defined above.

Lemma 2.2. For any $\omega = \{\omega_0, \omega_b\} \in S_h^0$, there are two constants $a_1, a_2 > 0$ such that

$$a_1 \|\omega\|_{1,h} \leq \|\omega\|_E \leq a_2 \|\omega\|_{1,h}. \quad (2.10)$$

Proof. To avoid repetition, the interested reader is referred to [52, Lemma 3.2] for the proof. \square

Lemma 2.3. There is a constant $C > 0$ such that, for any $\omega \in S_h^0$

$$A_h(\omega, \omega) \geq C \|\omega\|_E^2. \quad (2.11)$$

Proof. The proof follows from the definition the bilinear form $A_h(\cdot, \cdot)$ given by (2.6) and (1.3). \square

In order to analyze convergence and stability properties of the time-fractional partial differential equation, we introduce some function spaces.

Let $H^\alpha(J)$ denote the classical fractional Sobolev space with the norm $\|\cdot\|_{H^\alpha(J)}$. We denote the spaces of infinitely differentiable functions and compactly supported infinitely differentiable functions on J by $C^\infty(J)$ and $C_0^\infty(J)$, respectively. Let ${}_0C^\infty(J)$ be the the space of infinitely differentiable functions on the interior of J with compact support in J . Now, the closure of the space ${}_0C^\infty(J)$ in the norm $\|\cdot\|_{H^\alpha(J)}$ for $\alpha \in (0, 1)$ defines the Sobolev space ${}_0H^\alpha(J)$. The following Sobolev spaces can be found in [53] and [54]. The set \bar{A} denotes the closure of a set A .

Definition 2.1. [53] Define $H_r^\alpha(0, T) := \overline{{}_0C^\infty(0, T)}$ in the norm $\|\cdot\|_{H_r^\alpha(0, T)}$ defined by

$$\|v\|_{H_r^\alpha(0, T)} = \left(\|v\|_{L^2(0, T)}^2 + |v|_{H_r^\alpha(0, T)}^2 \right)^{1/2},$$

where $|v|_{H_r^\alpha(0, T)} = \|{}_t^R D_T^\alpha v\|_{L^2(0, T)}$.

Definition 2.2. [53] Define $H_c^\alpha(0, T) := \overline{C_0^\infty(0, T)}$ in the norm $\|\cdot\|_{H_c^\alpha(0, T)}$ defined by

$$\|v\|_{H_c^\alpha(0, T)} = \left(\|v\|_{L^2(0, T)}^2 + |v|_{H_c^\alpha(0, T)}^2 \right)^{1/2},$$

where $|v|_{H_c^\alpha(0, T)} = \left| \left({}_0^R D_T^\alpha v, {}_t^R D_T^\alpha v \right)_{L^2(0, T)} \right|^{1/2}$.

Lemma 2.4. [53] For $\alpha \in (0, 1)$, the spaces $H_r^\alpha(J)$, $H_c^\alpha(J)$, $H_0^\alpha(J)$ are equal and their norms $\|\cdot\|_{H_r^\alpha(J)}$, $\|\cdot\|_{H_c^\alpha(J)}$ and $\|\cdot\|_{H_0^\alpha(J)}$ are equivalent.

Lemma 2.5. [55] If $0 < \alpha < 2$, $\alpha \neq 1$, $u, v \in H^{\alpha/2}(J)$, then we have

$$\left({}_0^R D_T^\alpha u, v \right)_{L^2(J)} = \left({}_0^R D_T^{\alpha/2} u, {}_t^R D_T^{\alpha/2} v \right)_{L^2(J)}.$$

Lemma 2.6. [54] For $\alpha > 0$ and $u \in H^\alpha(J)$, we have

$$\left({}_\infty D_t^{-\alpha} \hat{u}, {}_t D_\infty^{-\alpha} \hat{u} \right)_{L^2(\mathbb{R})} = \cos(\pi\alpha) \| {}_\infty D_t^{-\alpha} \hat{u} \|_{L^2(\mathbb{R})}^2,$$

where \hat{u} is the extension of u by zero outside of J .

For a nonnegative real number α and the Sobolev space Y with the norm $\|\cdot\|_Y$, define the space

$$H^\alpha(J; Y) := \{u : \|u(\cdot, t)\|_Y \in H^\alpha(J)\},$$

with respect to the norm

$$\|u\|_{H^\alpha(J; Y)} := \left\| \|u(\cdot, t)\|_Y \right\|_{H^\alpha(J)}.$$

We also adapt the following notations:

$$\begin{aligned} (u, v)_{L^2(J \times \Omega)} &= \int_0^T (u, v)_{L^2(\Omega)} dt, & (u, v)_{L^2(J \times \mathcal{T}_h)} &= \int_0^T (u, v)_{\mathcal{T}_h} dt, \\ \|v\|_{L^2(J; H^\alpha(\Omega))}^2 &= \int_0^T \|v\|_{H^\alpha(\Omega)}^2 dt, & (\|v\|_E)_{L^2(J)}^2 &= \int_0^T \|v\|_E^2 dt. \end{aligned}$$

The broken Sobolev space $H^s(\mathcal{T}_h)$ is denoted by $H^s(\mathcal{T}_h) := \{v \in L^2(\Omega) : v|_K \in H^s(K), \forall K \in \mathcal{T}_h\}$.

The following theorem ensures the well-posedness of the semi-discrete SFWG-FEM (2.7) and the stability estimate for $t > 0$.

Theorem 2.1. For any $\alpha \in (0, 1)$ and $f \in L^q(J; L^2(\Omega))$, the solution u_h of the semi-discrete problem (2.7) satisfies the following stability estimate with $q = \frac{2}{1 + \alpha}$.

$$\|u_{0, h}\|_{H^{\alpha/2}(J; L^2(\Omega))} \leq C \|f\|_{L^q(J; L^2(\Omega))} + \|g\|_{L^2(\Omega)} \|t^{-\alpha}\|_{L^q(J)}. \quad (2.12)$$

Proof. Choosing $v = u_h = \{u_{0, h}, u_{b, h}\}$ in (2.7), then using (1.8) and the coercivity of the bilinear form $A_h(\cdot, \cdot)$, we get

$$\left({}_0^R D_T^\alpha u_{0, h}, u_{0, h} \right) + C \|u_h\|_E^2 \leq (f, u_{0, h}) + \left(\frac{g(x)t^{-\alpha}}{\Gamma(1 - \alpha)}, u_{0, h} \right), \quad (2.13)$$

where we have used that $u(x, 0) = g(x)$ and $A_h(u_h, u_h) \geq C\|u_h\|_E^2$. Employing Lemma 2.5 to the first term on LHS of (2.13), knowing that $\|u_h\|_E^2 \geq 0$, and using Cauchy-Schwarz inequality on the right side of (2.13), we obtain

$$({}_0^R D_T^{\alpha/2} u_{0,h}, {}_t^R D_T^{\alpha/2} u_{0,h}) \leq \|f\|_{L^2(\Omega)} \|u_{0,h}\|_{L^2(\Omega)} + \frac{t^{-\alpha}}{\Gamma(1-\alpha)} \|g\|_{L^2(\Omega)} \|u_{0,h}\|_{L^2(\Omega)}.$$

Integrating the above expression with respect to t , we get

$$({}_0^R D_T^{\alpha/2} u_{0,h}, {}_t^R D_T^{\alpha/2} u_{0,h})_{L^2(Q_T)} \leq \int_0^T (\|f\|_{L^2(\Omega)} \|u_{0,h}\|_{L^2(\Omega)} + \frac{t^{-\alpha}}{\Gamma(1-\alpha)} \|g\|_{L^2(\Omega)} \|u_{0,h}\|_{L^2(\Omega)}) dt. \quad (2.14)$$

By Lemma 2.4, we obtain

$$C({}_0^R D_T^{\alpha/2} u_{0,h}, {}_0^R D_T^{\alpha/2} u_{0,h})_{L^2(Q_T)} \leq \int_0^T (\|f\|_{L^2(\Omega)} \|u_{0,h}\|_{L^2(\Omega)} + \frac{t^{-\alpha}}{\Gamma(1-\alpha)} \|g\|_{L^2(\Omega)} \|u_{0,h}\|_{L^2(\Omega)}) dt.$$

Employing Hölder inequality on the left side of the above inequality gives

$$({}_0^R D_T^{\alpha/2} u_{0,h}, {}_0^R D_T^{\alpha/2} u_{0,h})_{L^2(Q_T)} \leq C(\|f\|_{L^q(J;L^2(\Omega))} + \|g\|_{L^2(\Omega)} \|t^{-\alpha}\|_{L^q(J)}) \|u_{0,h}\|_{L^p(J;L^2(\Omega))},$$

where $q = \frac{2}{1-\alpha}$ and $p = \frac{2}{1+\alpha}$.

With the aid of the facts that (see, e.g., [53])

$$\|{}_0^R D_T^{\alpha/2} v\|_{L^2(J)} \cong \|v\|_{H^{\alpha/2}(J)}, \quad \forall v \in_0 H^{\alpha/2}(J),$$

and the embedding theorem from [56]

$$H^{\alpha/2}(J) \hookrightarrow L^p(J), \quad \|u_{0,h}\|_{L^p(J;L^2(\Omega))} \leq C \|u_{0,h}\|_{H^{\alpha/2}(J;L^2(\Omega))},$$

one can conclude that

$$\|u_{0,h}\|_{H^{\alpha/2}(J;L^2(\Omega))}^2 \leq C(\|f\|_{L^q(J;L^2(\Omega))} + \|g\|_{L^2(\Omega)} \|t^{-\alpha}\|_{L^q(J)}) \|u_{0,h}\|_{H^{\alpha/2}(J;L^2(\Omega))},$$

from which the desired result follows after dividing both sides of the above inequality by $\|u_{0,h}\|_{H^{\alpha/2}(J;L^2(\Omega))}$. Thus, the proof is completed. \square

Lemma 2.7. *Let u be the solution of the problem (1.1). Then, for any $\omega = \{\omega_0, \omega_b\} \in S_h^0$, we have*

$$-(\nabla \cdot (A\nabla u), \omega_0)_{\mathcal{T}_h} = (A\nabla_w(\pi_h u), \nabla_w \omega)_{\mathcal{T}_h} - Z(u, \omega), \quad (2.15)$$

where $Z(u, \omega) := Z_1(u, \omega) + Z_2(u, \omega) + Z_3(u, \omega)$ with

$$Z_1(u, \omega) = (A\nabla_w(\pi_h u) - A\nabla u, \nabla_w \omega)_{\mathcal{T}_h}, \quad (2.16)$$

$$Z_2(u, \omega) = \langle (A\nabla u - \Pi_h(A\nabla u)) \cdot \mathbf{n}, \pi_b \omega_0 - \omega_b \rangle, \quad (2.17)$$

$$Z_3(u, \omega) = \langle A\nabla u \cdot \mathbf{n}, \omega_0 - \pi_b \omega_0 \rangle. \quad (2.18)$$

Proof. Using integration by parts, we have for any $\omega = \{\omega_0, \omega_b\} \in S_h^0$

$$\begin{aligned} -(\nabla \cdot (A\nabla u), \omega_0)_{\mathcal{T}_h} &= (A\nabla u, \nabla \omega_0)_{\mathcal{T}_h} - \langle A\nabla u \cdot \mathbf{n}, \omega_0 - \omega_b \rangle \\ &= (A\nabla u, \nabla \omega_0)_{\mathcal{T}_h} - \langle A\nabla u \cdot \mathbf{n}, \pi_b \omega_0 - \omega_b \rangle - \langle A\nabla u \cdot \mathbf{n}, \omega_0 - \pi_b \omega_0 \rangle, \end{aligned} \quad (2.19)$$

where we have used that $\omega_b|_{\partial\Omega} = 0$ and $A\nabla u$ is single value on \mathcal{E}_h .

Using the definition of projection Π_h , one has

$$(A\nabla u, \nabla \omega_0)_{\mathcal{T}_h} = (\Pi_h(A\nabla u), \nabla \omega_0)_{\mathcal{T}_h}. \quad (2.20)$$

On the other hand, it follows from the definition of weak gradient (2.5) that

$$(\Pi_h(A\nabla u), \nabla_w \omega)_{\mathcal{T}_h} = (\Pi_h(A\nabla u), \nabla \omega_0)_{\mathcal{T}_h} + \langle \pi_b(\omega_b - \omega_0), \Pi_h(A\nabla u) \cdot \mathbf{n} \rangle. \quad (2.21)$$

From (2.20), (2.21) and using the definitions of Π_h and π_b , we have

$$\begin{aligned} (A\nabla u, \nabla \omega_0)_{\mathcal{T}_h} &= (\Pi_h(A\nabla u), \nabla_w \omega)_{\mathcal{T}_h} + \langle \pi_b(\omega_0 - \omega_b), \Pi_h(A\nabla u) \cdot \mathbf{n} \rangle \\ &= (A\nabla u, \nabla_w \omega)_{\mathcal{T}_h} + \langle \pi_b(\omega_0 - \omega_b), \Pi_h(A\nabla u) \cdot \mathbf{n} \rangle \\ &= (A\nabla u, \nabla_w \omega)_{\mathcal{T}_h} + \langle \pi_b \omega_0 - \omega_b, \Pi_h(A\nabla u) \cdot \mathbf{n} \rangle. \end{aligned} \quad (2.22)$$

Plugging the above equation (2.22) into (2.19) yields

$$-(\nabla \cdot (A\nabla u), \omega_0)_{\mathcal{T}_h} = (A\nabla u, \nabla_w \omega)_{\mathcal{T}_h} + \langle \pi_b \omega_0 - \omega_b, (\Pi_h(A\nabla u) - A\nabla u) \cdot \mathbf{n} \rangle - \langle A\nabla u \cdot \mathbf{n}, \omega_0 - \pi_b \omega_0 \rangle,$$

which gives (2.15). Thus, we complete the proof. \square

Next, we present an error equation for the discretization error $e_h(t) = \{e_{0,h}(t), e_{b,h}(t)\} := \{\pi_0 u(t) - u_{0,h}(t), \pi_b u(t) - u_{b,h}(t)\}$. This error equation plays an important role in our error analysis.

Lemma 2.8. *Assume that $u(t)$ is the solution of the problem (1.1) and $u_h(t)$ is the solution of the semi-discrete problem (2.7) for $t \in (0, T]$. Then, we have for any $\omega = \{\omega_0, \omega_b\} \in S_h^0$*

$$({}_0^C D_T^\alpha e_{0,h}, \omega_0) + A_h(e_h, \omega) = Z(u, \omega). \quad (2.23)$$

Proof. Multiplying the first equation in (1.1) by a test function ω_0 of $\omega = \{\omega_0, \omega_b\} \in S_h^0$ yields

$$({}_0^C D_T^\alpha u, \omega_0)_{\mathcal{T}_h} - (\nabla \cdot (A\nabla u), \omega_0)_{\mathcal{T}_h} = (f, \omega_0)_{\mathcal{T}_h}.$$

From (2.15) in Lemma 2.7, we obtain

$$({}_0^C D_T^\alpha \pi_0 u, \omega_0)_{\mathcal{T}_h} - (A\nabla_w(\pi_h u), \nabla_w \omega)_{\mathcal{T}_h} = (f, \omega_0)_{\mathcal{T}_h} + Z(u, \omega), \quad (2.24)$$

where we have used that $({}_0^C D_T^\alpha u, \omega_0)_{\mathcal{T}_h} = ({}_0^C D_T^\alpha \pi_0 u, \omega_0)_{\mathcal{T}_h}$ which follows from the definition of projection π_0 .

Subtracting the first equation in (2.7) from (2.24), we get

$$({}_0^C D_T^\alpha e_{0,h}, \omega_0)_{\mathcal{T}_h} - (A\nabla_w e_h, \nabla_w \omega)_{\mathcal{T}_h} = Z(u, \omega), \quad \forall \omega \in S_h^0,$$

which is the desired result (2.23). Thus, we complete the proof. \square

For $\omega \in H^1(T)$, we have the trace inequality (see, e.g., [29])

$$\|\omega\|_e^2 \leq C \left(h_K^{-1} \|\omega\|_K^2 + h_K \|\nabla \omega\|_K^2 \right). \quad (2.25)$$

The following approximation results will be used in the sequel.

Lemma 2.9. [29] *Let u be the solution of the problem (1.1) and assume that \mathcal{T}_h is shape regular. Then, the L^2 projections π_0 and Π_h satisfy*

$$\sum_{K \in \mathcal{T}_h} \left(\|u - \pi_0 u\|_K^2 + h_K^2 \|\nabla(u - \pi_0 u)\|_K^2 \right) \leq Ch^{2(s+1)} \|u\|_{s+1}^2, \quad 0 \leq s \leq k, \quad (2.26)$$

$$\sum_{K \in \mathcal{T}_h} \left(\|\nabla u - \Pi_h \nabla u\|_K^2 + h_K^2 \|\nabla u - \Pi_h \nabla u\|_{1,K}^2 \right) \leq Ch^{2s} \|u\|_{s+1}^2, \quad 0 \leq s \leq k. \quad (2.27)$$

Lemma 2.10. *Assume that $w \in H^{k+1}(\Omega)$. Then, for any $\omega = \{\omega_0, \omega_b\} \in V_h^0$, one has*

$$|Z(w, \omega)| \leq Ch^k \|w\|_{k+1} \|\omega\|_E.$$

Proof. Using Cauchy-Schwarz inequality, (1.3), Lemmas 2.1 and Lemma 2.9 we obtain

$$\begin{aligned} |Z_1(w, \omega)| &= \left| \sum_{K \in \mathcal{T}_h} (A(\nabla_w(\pi_h u) - \nabla u), \omega_0 - \pi_b \omega_0)_T \right| \\ &\leq C \sum_{K \in \mathcal{T}_h} \|A\|_{L^\infty(T)} \|\Pi_h \nabla u - \nabla u\|_{L^2(T)} \|\nabla_w \omega\|_{L^2(T)} \\ &\leq Ch^k \|w\|_{H^{k+1}(\Omega)} \|\omega\|_E. \end{aligned} \quad (2.28)$$

It follows from the Cauchy-Schwarz inequality, the trace inequality (2.25), (1.3), (2.10) and Lemma 2.9 that

$$\begin{aligned} |Z_2(w, \omega)| &= \left| \sum_{K \in \mathcal{T}_h} \langle (A \nabla w - \Pi_h(A \nabla w)) \cdot \mathbf{n}, \pi_b \omega_0 - \omega_b \rangle_{\partial K} \right| \\ &\leq C \sum_{K \in \mathcal{T}_h} \|A \nabla w - \Pi_h(A \nabla w)\|_{L^2(\partial K)} \|\pi_b \omega_0 - \omega_b\|_{L^2(\partial K)} \\ &\leq C \left(\sum_{K \in \mathcal{T}_h} h_K \|A \nabla w - \Pi_h(A \nabla w)\|_{L^2(\partial K)}^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}_h} h_K^{-1} \|\pi_b \omega_0 - \omega_b\|_{L^2(\partial K)}^2 \right)^{1/2} \\ &\leq Ch^k \|w\|_{H^{k+1}(\Omega)} \|\omega\|_{1,h} \\ &\leq Ch^k \|w\|_{H^{k+1}(\Omega)} \|\omega\|_E, \end{aligned} \quad (2.29)$$

where we have used that

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} h_K \|A \nabla w - \Pi_h(A \nabla w)\|_{L^2(\partial K)}^2 &\leq \sum_{K \in \mathcal{T}_h} (\|A \nabla w - \Pi_h(A \nabla w)\|_{L^2(K)}^2 + h_K^2 \|A \nabla w - \Pi_h(A \nabla w)\|_{1,K}^2) \\ &\leq Ch^k \|w\|_{H^{k+1}(\Omega)}. \end{aligned} \quad (2.30)$$

From Cauchy-Schwarz inequality, the trace inequality (2.25), (1.3), (2.10) and Lemma 2.9 that

$$\begin{aligned}
 |Z_3(w, \omega)| &= \left| \sum_{K \in \mathcal{T}_h} \langle A \nabla w \cdot \mathbf{n}, \omega_0 - \pi_b \omega_0 \rangle_{\partial K} \right| \\
 &= \left| \sum_{K \in \mathcal{T}_h} \langle (A \nabla w - \Pi_{k-1}(A \nabla w)) \cdot \mathbf{n}, \omega_0 - \pi_b \omega_0 \rangle_{\partial K} \right| \\
 &\leq C \sum_{K \in \mathcal{T}_h} \|A \nabla w - \Pi_{k-1}(A \nabla w)\|_{L^2(\partial K)} \|\pi_b \omega_0 - \omega_0\|_{L^2(\partial K)} \\
 &\leq C \left(\sum_{K \in \mathcal{T}_h} h_K \|A \nabla w - \Pi_{k-1}(A \nabla w)\|_{L^2(\partial K)}^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}_h} h_K^{-1} \|\pi_0 \omega_0 - \omega_0\|_{L^2(\partial K)}^2 \right)^{1/2} \\
 &\leq C \left(\sum_{K \in \mathcal{T}_h} h_K \|K \nabla w - \Pi_{k-1}(A \nabla w)\|_{L^2(\partial K)}^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}_h} h_K^{-2} \|\pi_0 \omega_0 - \omega_0\|_{L^2(T)}^2 + \|\nabla(\pi_0 \omega_0 - \omega_0)\|_{L^2(T)}^2 \right)^{1/2} \\
 &\leq Ch^k \|w\|_{H^{k+1}(\Omega)} \|v\|_1 \\
 &\leq Ch^k \|w\|_{H^{k+1}(\Omega)} \|\omega\|_E,
 \end{aligned} \tag{2.31}$$

where Π_{k-1} is the L^2 projection operator onto the space $[\mathbb{P}_{k-1}(K)]^2$ on each $K \in \mathcal{T}_h$ and we have used the fact $\|\pi_b \omega_0 - \omega_0\|_{L^2(\partial K)} \leq \|\pi_0 \omega_0 - \omega_0\|_{L^2(\partial K)}$ and above estimate (2.30). Combining the estimates (2.28), (2.29) and (2.31) finishes the proof. \square

We now study the estimate for the error $e_h(t)$ of the numerical scheme (2.7) in the $\|\cdot\|_E$ -norm.

Theorem 2.2. *Let u be the solution of the problem (1.1) and u_h be the solution of the semi-discrete problem (2.7). Suppose that $u, \frac{\partial u}{\partial t} \in L^2(0, t; H^{k+1}(\mathcal{T}_h))$ for any fixed $t \in J$. Then, we have the error estimate*

$$\|e_h(t)\|_E \leq C(g, u, \frac{\partial u}{\partial t}) h^k, \tag{2.32}$$

where $C(g, u, \frac{\partial u}{\partial t})$ depends on the norms of g, u and $\frac{\partial u}{\partial t}$.

Proof. For a fixed $t \in J$, choosing $\omega = e_h$ in (2.23) yields

$$({}_0^C D_T^\alpha e_{0,h}, e_{0,h}) + A_h(e_h, e_h) = Z(u, e_h).$$

The coercivity property (2.11) of the bilinear form $A_h(\cdot, \cdot)$ and Lemma 2.10 imply that

$$({}_0^C D_T^\alpha e_{0,h}, e_{0,h}) + \|e_h\|_E^2 \leq Ch^k \|u\|_{H^{k+1}(\Omega)} \|e_h\|_E \leq Ch^{2h} \|u\|_{H^{k+1}(\Omega)}^2 + \frac{1}{2} \|e_h\|_E^2.$$

Hence, we have

$$({}_0^C D_T^\alpha e_{0,h}, e_{0,h}) + \frac{1}{2} \|e_h\|_E^2 \leq Ch^{2h} \|u\|_{H^{k+1}(\Omega)}^2. \tag{2.33}$$

Since $e_h(0) = 0$, one has ${}_0^C D_T^\alpha e_{0,h} = {}_0^R D_T^\alpha e_{0,h}$, thus integrating (2.33) with respect to t and using Lemma 2.5, we have

$$\|e_{0,h}\|_{H^{\alpha/2}(0,t;L^2(\Omega))} + \int_0^t \|e_h(s)\|_E^2 ds \leq Ch^{2h} \int_0^t \|u(s)\|_{k+1}^2 ds. \tag{2.34}$$

Now, taking $v = \frac{\partial e_h}{\partial t}$ in (2.23) gives

$$({}^R D_T^\alpha e_{0,h}, \frac{\partial e_{0,h}}{\partial t}) + A_h(e_h, \frac{\partial e_h}{\partial t}) = Z(u, \frac{\partial e_h}{\partial t}),$$

or, equivalently

$$({}^R D_T^{-\beta} \frac{\partial e_{0,h}}{\partial t}, \frac{\partial e_{0,h}}{\partial t}) + \frac{1}{2} \frac{d}{dt} A_h(e_h, e_h) = \frac{d}{dt} Z(u, e_h) - Z(\frac{\partial u}{\partial t}, e_h),$$

where we have used the fact that ${}^R D_T^\alpha e_{0,h} = {}^R D_T^{\alpha-1} \frac{\partial e_{0,h}}{\partial t} = {}^R D_T^{-\beta} \frac{\partial e_{0,h}}{\partial t}$ with $\beta := 1 - \alpha$ (see, e.g., [1, 57]).

Hence, integrating on the time domain, using the adjoint property of the Riemann-Liouville fractional integrals and then Lemma 2.6, we get

$$\int_{\Omega} \cos(\pi\beta/2) \| {}^R D_t^{-\beta/2} \widehat{e_{0,h}} \|_{L^2(\mathbb{R})}^2 dx + \frac{1}{2} \| e_h \|_E^2 \leq Z(u(t), e_h(t)) + \int_0^t |Z(\frac{\partial u}{\partial t}(s), e_h(s))| ds,$$

where $\widehat{e_{0,h}}$ is the zero extension of $e_{0,h}$ in $(0, T]$. Using the fact that $\cos(\pi\beta/2)$ is non-negative for $\beta \in (0, 1)$ and Lemma 2.10, we obtain

$$\begin{aligned} \| e_h \|_E^2 &\leq Ch^k \| u(t) \|_{k+1} \| e_h \|_E + Ch^k \int_0^t \| \frac{\partial u}{\partial t}(s) \|_{k+1} \| e_h(s) \|_E ds \\ &\leq Ch^{2k} (\| u(t) \|_{k+1}^2 + \int_0^t \| \frac{\partial u}{\partial t}(s) \|_{k+1}^2 ds) + \frac{1}{2} \| e_h \|_E^2 + \frac{1}{2} \int_0^t \| e_h \|_E^2 ds. \end{aligned}$$

Appealing (2.34), one has

$$\| e_h \|_E^2 \leq Ch^{2k} (\| u(t) \|_{k+1}^2 + \int_0^t \| u(s) \|_{k+1}^2 ds + \int_0^t \| \frac{\partial u}{\partial t}(s) \|_{k+1}^2 ds). \quad (2.35)$$

Note that $u(t) = u(0) + \int_0^t \frac{\partial u}{\partial t}(s) ds = g + \int_0^t \frac{\partial u}{\partial t}(s) ds$. Thus, we get

$$\| u(t) \|_{k+1} \leq \| g \|_{k+1} + \int_0^t \| \frac{\partial u}{\partial t}(s) \|_{k+1} ds,$$

which combined with (2.35) gives

$$\| e_h \|_E^2 \leq Ch^{2k} (\| g \|_{k+1}^2 + \int_0^t \| u(s) \|_{k+1}^2 ds + \int_0^t \| \frac{\partial u}{\partial t}(s) \|_{k+1}^2 ds).$$

We complete the proof. \square

The optimal error rate in the L^2 -norm can be derived by introducing the elliptic (Ritz) projection $E_h : H_0^1(\Omega) \rightarrow S_h^0$ defined as follows. For any $w \in H_0^1(\Omega)$ [58]

$$A_h(E_h w, \phi) = (-\nabla \cdot (A \nabla w), \phi_0), \quad \forall \phi = \{\phi_0, \phi_b\} \in S_h^0. \quad (2.36)$$

In fact, $E_h w$ is the WG-FEM solution of the corresponding elliptic equation that has exact solution w . Note that $E_h w = \{E_0 w, E_b w\}$ for any $w \in H_0^1(\Omega)$, where $E_0 w$ is the WG-FEM solution in the inside of elements and $E_b w$ is the trace on \mathcal{E}_h .

The following error bounds for the elliptic projection E_h follow from [52, Theorem 4.4] and [52, Theorem 4.5].

Lemma 2.11. [52] Suppose that $w \in H_0^1(\Omega) \cap H^{k+1}(\mathcal{T}_h)$. Then, the elliptic projection E_h defined by (2.36) has the following error estimates

$$\begin{aligned}\|E_h w - \pi_h w\|_E &\leq Ch^k |w|_{k+1}, \\ \|E_0 w - \pi_0 w\| &\leq Ch^{k+1} |w|_{k+1}.\end{aligned}$$

We now state and prove the error estimate of a semi-discrete scheme in L^2 -norm.

Theorem 2.3. Let $u \in H^{k+1}(\Omega)$ and u_h be the solution of the problem (1.1) and the semi-discrete problem (2.7), respectively. Then for any fixed $t \in (0, T]$, there holds

$$\|u_{0,h}(t) - \pi_0 u(t)\|_{H^{\alpha/2}(0,T;L^2(\Omega))} \leq Ch^{k+1} \|{}_0^C D_T^\alpha u\|_{L^2(J;H^{k+1}(\Omega))}. \quad (2.37)$$

Proof. We first split the error $e_h(t) := \xi_h(t) + \theta_h(t)$, where

$$\xi_h(t) = \{\xi_0(t), \xi_b(t)\}, \quad \xi_0 := u_{0,h} - E_0 u, \quad \xi_b := u_{b,h} - E_b u, \quad (2.38)$$

$$\theta_h(t) = \{\theta_0(t), \theta_b(t)\}, \quad \theta_0 := E_0 u - \pi_0 u, \quad \theta_b := E_b u - \pi_b u. \quad (2.39)$$

It follows from Lemma 2.11 that

$$\|\theta_0\|_{H^\alpha(0,T;L^2(\Omega))} \leq Ch^{k+1} \|{}_0^C D_T^\alpha u\|_{L^2(J;H^{k+1}(\Omega))}. \quad (2.40)$$

We shall bound ξ_h using the semi-discrete problem (2.7), the definitions of the elliptic projection (2.36) and the projection π_0 as follows. For any $\phi = \{\phi_0, \phi_b\} \in S_h^0$, we have

$$\begin{aligned}({}_0^C D_T^\alpha \xi_0, \phi_0) + A_h(\xi_h, \phi) &= ({}_0^C D_T^\alpha u_{0,h}, \phi_0) + A_h(u_h, \phi) - ({}_0^C D_T^\alpha E_0, \phi_0) - A_h(E_h, \phi) \\ &= ({}_0^C D_T^\alpha u, \phi_0) - ({}_0^C D_T^\alpha E_0, \phi_0) \\ &= (\pi_0({}_0^C D_T^\alpha u), \phi_0) - ({}_0^C D_T^\alpha E_0, \phi_0) \\ &= -({}_0^C D_T^\alpha \theta_0, \phi_0).\end{aligned}$$

Choosing $\phi = \xi_h$ in the above equation, we get

$$({}_0^C D_T^\alpha \xi_0, \xi_0) + A_h(\xi_h, \xi_h) = -({}_0^C D_T^\alpha \theta_0, \xi_0).$$

Since $\xi_0(0) = 0$, one has ${}_0^C D_T^\alpha \xi_0 = {}_0^R D_T^\alpha \xi_0$ and Lemma 2.5 implies that

$$({}_0^R D_T^{\alpha/2} \xi_0, {}_0^R D_T^{\alpha/2} \xi_0) + \|\xi_h\|_E^2 \leq \|{}_0^C D_T^\alpha \theta_0\| \|\xi_0\|.$$

Recalling that

$$\|{}_0^R D_T^{\alpha/2} v\|_{L^2(0,T)} \cong \|v\|_{H^{\alpha/2}(0,T)}, \quad \forall v \in {}_0H^{\alpha/2}(0, T),$$

we get

$$\|\xi_0\|_{H^{\alpha/2}(0,T;L^2(\Omega))}^2 + (\|\xi_h\|_E^2)_{L^2(J)} \leq C \|{}_0^C D_T^\alpha \theta_0\|_{L^2(Q_T)} \|\xi_0\|_{L^2(Q_T)}.$$

Applying the arithmetic mean inequality yields

$$\begin{aligned}
\|\xi_0\|_{H^{\alpha/2}(0,T;L^2(\Omega))}^2 + (\|\xi_h\|_E^2)_{L^2(J)} &\leq C\|_0^C D_T^\alpha \theta_0\|_{L^2(Q_T)} \|\xi_0(s)\|_{L^2(J;H^1(\mathcal{T}_h))} \\
&\leq C\|_0^C D_T^\alpha \theta_0\|_{L^2(Q_T)}^2 + \frac{1}{2} \int_0^T \|\xi_0(s)\|_{H^1(\mathcal{T}_h)}^2 ds \\
&\leq C\|_0^C D_T^\alpha \theta_0\|_{L^2(Q_T)}^2 + \frac{1}{2} \int_0^T (\|\xi_h(s)\|_{1,h})^2 ds \\
&\leq C\|_0^C D_T^\alpha \theta_0\|_{L^2(Q_T)}^2 + \frac{1}{2} (\|\xi_h(s)\|_E^2)_{L^2(J)},
\end{aligned}$$

where we have used the equivalent of the norms (2.10) in the last inequality. From (2.40), we obtain

$$\|\xi_0\|_{H^{\alpha/2}(0,T;L^2(\Omega))}^2 + \frac{1}{2} (\|\xi_h\|_E^2)_{L^2(J)} \leq Ch^{2(k+1)} \|_0^C D_T^\alpha u\|_{L^2(J;H^{k+1}(\Omega))}^2,$$

which proves the desired result. \square

3. Temporal discretization

In order to present semi-discrete numerical scheme, we discretize (1.1) in time direction. We investigate time discretization using the well known $L2 - 1_\sigma$ formula on graded meshes to deal with the singularity of solution at $t = 0$.

Let $M > 0$ be a natural number. We define temporal graded meshes by setting $t_m = T(\frac{m}{M})^r$ for $m = 0, 1, \dots, M$, where $r \geq 1$ is a user-chosen grading constant. Similar temporal graded meshes have been investigated in the literature, e.g, [59, 60]. The graded constant r will influence the convergence rate; thus we take it such a way that the rate is optimal as presented below.

One can easily show that

$$t_m \geq 2^{-r} t_{m+1} \quad \text{for } m = 1, \dots, M-1, \quad (3.1)$$

and

$$\tau_m := t_m - t_{m-1} \simeq T^{1/r} M^{-1} t_m^{1-1/r} \quad \text{for } m = 1, \dots, M. \quad (3.2)$$

For $m = 0, 1, \dots, M-1$ and $\sigma \in [0, 1]$, let $t_{m+\sigma} = t_m + \sigma\tau_{m+1}$. Following the ideas of [28], the Caputo derivative ${}_0^C D_T^\alpha u$ in (1.1) at $t = t_{m+\sigma}$ is approximated using the $L2 - 1_\sigma$ formula.

$${}_0^C D_T^\alpha u(t_{m+\sigma}) \approx \delta_{t_{m+\sigma}}^\alpha u = c_{m,m} u^{m+1} - \sum_{p=0}^m (c_{m,p} - c_{m,p-1}) u^p \quad \text{for } m = 0, \dots, M-1. \quad (3.3)$$

Here, $c_{0,0} = \tau_1^{-1} d_{0,0}$, $c_{m,-1} = 0$ and for $m \geq 1$ one has

$$c_{m,p} = \begin{cases} \tau_{p+1}^{-1} (d_{m,0} - f_{m,0}) & \text{if } p = 0, \\ \tau_{p+1}^{-1} (d_{m,p} + f_{m,p-1} - f_{m,p}) & \text{if } 1 \leq p \leq m-1, \\ \tau_{p+1}^{-1} (d_{m,m} + f_{m,m-1}) & \text{if } p = m, \end{cases}$$

where

$$\begin{aligned}
d_{m,m} &= \frac{1}{\Gamma(1-\alpha)} \int_{t_m}^{t_{m+\sigma}} (t_{m+\sigma} - \theta)^{-\alpha} d\theta = \frac{\sigma^{1-\alpha}}{\Gamma(2-\alpha)} \tau_{m+1}^{1-\alpha} \quad \text{for } m \geq 0, \\
d_{m,p} &= \frac{1}{\Gamma(1-\alpha)} \int_{t_p}^{t_{p+1}} (t_{m+\sigma} - \theta)^{-\alpha} d\theta \quad \text{for } m \geq 1, \text{ and } 0 \leq p \leq n-1, \\
f_{m,p} &= \frac{1}{\Gamma(1-\alpha)} \frac{2}{t_{p+2} - t_p} \int_{t_p}^{t_{p+1}} (t_{m+\sigma} - \theta)^{-\alpha} (\eta - t_{p+1/2}) d\theta \quad \text{for } m \geq 1, \text{ and } 0 \leq p \leq m-1.
\end{aligned}$$

The following technical lemmas will be useful in the later analysis.

Lemma 3.1. [28, Corollary 1] Assume that $0 \leq \sigma \leq 1$. we have the following coercivity property of a function $Z(\cdot, t)$ on the graded temporal mesh $\{t_m\}_{m=0}^M$,

$$(\sigma Z^{m+1} + (1-\sigma)Z^m, \delta_{t_{m+\sigma}}^\alpha Z) \geq \frac{1}{2} \delta_{t_{m+\sigma}}^\alpha \|Z\|^2 \quad \text{for } m = 0, \dots, M-1.$$

Lemma 3.2. [61, Lemma 7] Let $w \in C[0, T] \cap C^3(0, T)$. Suppose that $|w^{(d)}(t)| \leq C(1 + t^{\alpha-d})$ for $d = 0, 1, 2, 3$ and $t \in (0, T]$. Then

$$\begin{aligned}
\psi_w^{j+\sigma} &\leq CM^{-\min\{r\alpha, 3-\alpha\}} \quad \text{for } j = 0, \dots, M-1, \\
\psi_w^{j,s} &\leq CM^{-\min\{r\alpha, 3-\alpha\}} \quad \text{for } s = 1, \dots, j, \text{ when } j \geq 1,
\end{aligned}$$

where

$$\psi_v^{j+\sigma} = \tau_{j+1}^{3-\alpha} t_{j+\sigma}^\alpha \sup_{\eta \in (t_j, t_{j+1})} |v'''(\eta)| \quad \text{for } j = 1, \dots, M-1, \quad (3.4)$$

and

$$\psi_v^{j,s} = \tau_{j+1}^{-\alpha} \tau_s^2 (\tau_s + \tau_{s+1}) t_s^\alpha \sup_{\eta \in (t_{s-1}, t_{s+1})} |v'''(\eta)| \quad \text{for } 2 \leq s \leq j \leq M-1. \quad (3.5)$$

Lemma 3.3. [61, Lemma 5] Assume that $1 \geq \sigma \geq 1 - \alpha/2$. Then for any function $\{V^j\}_{j=0}^M$, one has

$$|V^{k+1}| \leq |V^0| + \Gamma(1-\alpha) \max_{m=0, \dots, k} \{t_{m+\sigma}^\alpha \delta_{t_{m+\sigma}}^\alpha |V|\} \quad \text{for } k = 0, \dots, M-1.$$

Lemma 3.4. [61, Lemma 1] Suppose $\sigma = 1 - \alpha/2$. Assume that $\tau_{j+1} \leq C\tau_j$ for $j \geq 2$ and $\tau_1/\tau_2 \leq \rho$, where ρ is any fixed positive constant. Then for any function $v(t) \in C^3(0, T]$, one has

$$\left| \delta_{t_{j+\sigma}}^\alpha v - D_t^\alpha v(t_{j+\sigma}) \right| \leq C t_{j+\sigma}^{-\alpha} \left(\psi_v^{j+\sigma} + \max_{s=1, \dots, j} \{ \psi_v^{j,s} \} \right) \quad \text{for } j = 0, \dots, M-1,$$

where $\psi_v^{j+\sigma}$ and $\psi_v^{j,s}$ are given by (3.4) and (3.5), respectively.

4. Fully discrete numerical scheme

We now formulate the fully discrete L2-1 $_{\sigma}$ SFWG-FEM for the problem (2.1) as follows: find $u_h^{m+1} = \{u_{0,h}^{m+1}, u_{b,h}^{m+1}\} \in S_h^0$ such that

$$\begin{cases} (\delta_{t_{m+\sigma}}^{\alpha} u_{0,h}, v_{0,h}) + (A \nabla_w u_h^{m,\sigma}, \nabla_w v_h) = (f^{m+\sigma}, v_{0,h}), & \forall v_h = \{v_{0,h}, v_{b,h}\} \in S_h^0, \\ u_h^0 = \pi_h g, \end{cases} \quad (4.1)$$

where $u_h^{m,\sigma} = \sigma u_h^{m+1} + (1 - \sigma)u_h^m$ and $f^{m+\sigma} = f(\cdot, t_{m+\sigma})$ for $m = 0, 1, \dots, M - 1$.

We will prove that the L^2 stability of the fully discrete L2-1 $_{\sigma}$ SFWG-FEM (4.1). First, we give a Poincare-type inequality in the WG finite element space S_h^0 .

Lemma 4.1. [62] *There is a positive constant C independent of the mesh such that*

$$\|\omega_0\| \leq C \|v\|_E, \quad \forall v = \{\omega_0, \omega_b\} \in S_h^0.$$

Lemma 4.2. *Let $\{u_h^j\}_{j=0}^M$ be the solution of the fully discrete problem (4.1). Then*

$$\|u_{0,h}^{m+1}\|^2 \leq \|u_{0,h}^0\|^2 + \frac{\Gamma(1 - \alpha)T^{\alpha}}{2} \max_{0 \leq m \leq M-1} \|f^{m+\sigma}\|^2 \text{ for } m = 0, \dots, M - 1.$$

Proof. For $m = 0, \dots, M - 1$, choosing $v_h = u_h^{m,\sigma}$ in (4.1) yields

$$(\delta_{t_{m+\sigma}}^{\alpha} u_{0,h}, u_{0,h}^{m,\sigma}) + \|\nabla_w u_h^{m,\sigma}\|^2 = (f^{m+\sigma}, u_{0,h}^{m,\sigma}).$$

Invoking Lemma 3.1, we obtain

$$\frac{1}{2} \delta_{t_{m+\sigma}}^{\alpha} \|u_{0,h}\|^2 + \|\nabla_w u_h^{m,\sigma}\|^2 \leq (f^{m+\sigma}, u_{0,h}^{m,\sigma}). \quad (4.2)$$

Using Cauchy-Schwarz inequality, the Young's inequality and Lemma 4.1, we obtain

$$(f^{m+\sigma}, u_{0,h}^{m,\sigma}) \leq \|f^{m+\sigma}\| \|u_{0,h}^{m,\sigma}\| \leq \|\nabla_w u_h^{m,\sigma}\|^2 + \frac{1}{4} \|f^{m+\sigma}\|^2.$$

Using this inequality in (4.2) gives

$$\delta_{t_{m+\sigma}}^{\alpha} \|u_{0,h}\|^2 \leq \frac{1}{2} \|f^{m+\sigma}\|^2.$$

Thus, from Lemma 3.3, one has

$$\begin{aligned} \|u_{0,h}^{m+1}\|^2 &\leq \|u_{0,h}^0\|^2 + \Gamma(1 - \alpha) \max_{m=0, \dots, k} \{t_{m+\sigma}^{\alpha} \delta_{t_{m+\sigma}}^{\alpha} \|u_{0,h}\|^2\} \\ &\leq \|u_{0,h}^0\|^2 + \frac{\Gamma(1 - \alpha)T^{\alpha}}{2} \max_{0 \leq m \leq M-1} \|f^{m+\sigma}\|^2, \end{aligned}$$

which completes the proof. \square

We now prove an error estimate for the fully discrete L2-1 $_{\sigma}$ SFWG-FEM (4.1). For the error analysis, similar to (2.38), we split the error

$$u^m - u_h^m = \xi_h^m + \theta_h^m, \quad (4.3)$$

where $\xi_h^m = u^m - E_h u^m$ and $\theta_h^m = E_h u^m - u_h^m$. The estimation of ξ_h^m follows from Lemma 2.9, and thus we shall estimate θ_h^m as follows. From (1.1) and (4.1), for $m = 0, \dots, M-1$ we have $\forall v_h = \{v_{0,h}, v_{b,h}\} \in S_h^0$,

$$\begin{aligned} & (\delta_{t_{m+\sigma}}^{\alpha} \theta_{0,h}, v_{0,h}) + (A \nabla_w \theta_h^{m,\sigma}, \nabla_w v_h) \\ &= (\delta_{t_{m+\sigma}}^{\alpha} E_0 u, v_{0,h}) + (A \nabla_w E_h u^{m,\sigma}, \nabla_w v_h) - (\delta_{t_{m+\sigma}}^{\alpha} u_{0,h}, v_{0,h}) - (A \nabla_w u_h^{m,\sigma}, \nabla_w v_h) \\ &= (E_0 \delta_{t_{m+\sigma}}^{\alpha} u, v_{0,h}) + (A \nabla_w E_h u^{m,\sigma}, \nabla_w v_h) - (f^{m+\sigma}, v_{0,h}) \\ &= ((E_0 - \pi_0) \delta_{t_{m+\sigma}}^{\alpha} u, v_{0,h}) + (\pi_0 (\delta_{t_{m+\sigma}}^{\alpha} u + \nabla \cdot (A \nabla u^{m,\sigma})), v_{0,h}) - (\pi_0 f^{m+\sigma}, v_{0,h}) \\ &= (\pi_0 (E_0 - I) \delta_{t_{m+\sigma}}^{\alpha} u, v_{0,h}) + (\pi_0 (\delta_{t_{m+\sigma}}^{\alpha} u + \nabla \cdot (A \nabla u^{m,\sigma})), v_{0,h}) \\ &\quad - (\pi_0 ({}^C D_T^{\alpha} u^{m+\sigma} - \nabla \cdot (A \nabla u^{m+\sigma})), v_{0,h}) \\ &= (-\pi_0 (\delta_{t_{m+\sigma}}^{\alpha} \xi_0 + \mathcal{R}^{m+\sigma} + \Phi^{m+\sigma}), v_{0,h}), \end{aligned} \quad (4.4)$$

where $\Phi^{m+\sigma} = {}^C D_T^{\alpha} u^{m+\sigma} - \delta_{t_{m+\sigma}}^{\alpha} u^m$ and $\mathcal{R}^{m+\sigma} = \nabla \cdot (A \nabla u^{m+\sigma}) - \nabla \cdot (A \nabla u^{m,\sigma})$.

The following theorem shows convergence of the L2-1 $_{\sigma}$ SFWG-FEM in the norm $L^{\infty}(L^2)$.

Theorem 4.1. Assume that $\sigma = (2 - \alpha)/2$. Let u^m and u_h^m be the solutions of (1.1), (4.1), respectively. Assume that $u \in L^{\infty}(J; H_0^1(\Omega) \cap H^{k+1}(\Omega))$, ${}^C D_T^{\alpha} u \in L^{\infty}(J; H_0^1(\Omega) \cap H^{k+1}(\Omega))$ and $\left\| \frac{\partial^l u(\cdot, t)}{\partial t^l} \right\| \leq C(1 + t^{\alpha-l})$ for $l = 0, 1, 2, 3$. Then there exists a constant C such that

$$\max_{1 \leq m \leq M} \|u^m - u_{0,h}^m\| \leq C (M^{-\min\{r\alpha, 2\}} + h^{k+1}). \quad (4.5)$$

Proof. Let $m \in \{0, 1, \dots, M\}$ be a fixed number. Taking $v_h = \theta_h^{m,\sigma}$ in (4.4) yields

$$(\delta_{t_{m+\sigma}}^{\alpha} \theta_{0,h}, \theta_{0,h}^{m,\sigma}) + (A \nabla_w \theta_h^{m,\sigma}, \nabla_w \theta_h^{m,\sigma}) = -(\pi_0 (\delta_{t_{m+\sigma}}^{\alpha} \xi_0 + \mathcal{R}^{m+\sigma} + \Phi^{m+\sigma}), \theta_{0,h}^{m,\sigma}).$$

From Lemma 3.1 and Cauchy-Schwarz inequality, one has

$$\frac{1}{2} \delta_{t_{m+\sigma}}^{\alpha} \|\theta_{0,h}\|^2 \leq \|\pi_0 \beta\| \|\theta_{0,h}^{m,\sigma}\|,$$

where $\beta := \delta_{t_{m+\sigma}}^{\alpha} \xi_0 + \mathcal{R}^{m+\sigma} + \Phi^{m+\sigma}$. Now using the fact that the stability of the L^2 projection in L^2 -norm [63] and the triangle inequality, we have

$$\frac{1}{2} \delta_{t_{m+\sigma}}^{\alpha} \|\theta_{0,h}\|^2 \leq C (\|\delta_{t_{m+\sigma}}^{\alpha} \xi_0\| + \|\mathcal{R}^{m+\sigma}\| + \|\Phi^{m+\sigma}\|) \max_{1 \leq r \leq N} \|\theta_{0,h}^r\|.$$

Thus, Lemma 3.3 and $u_{0,h}^0 - E_0 u_0 = 0$ imply that

$$\|\theta_{0,h}^{m+1}\|^2 \leq C \max_{0 \leq i \leq M-1} \{t_{i+\sigma}^{\alpha} (\|\delta_{t_{i+\sigma}}^{\alpha} \xi_0\| + \|\mathcal{R}^{i+\sigma}\| + \|\Phi^{i+\sigma}\|)\} \max_{1 \leq r \leq M} \|\theta_{0,h}^r\|.$$

Therefore, one has

$$\max_{1 \leq m \leq M} \|\theta_{0,h}^m\| \leq C \max_{0 \leq i \leq M-1} \{t_{i+\sigma}^\alpha (\|\delta_{i+\sigma}^\alpha \xi_0\| + \|\mathcal{R}^{i+\sigma}\| + \|\Phi^{i+\sigma}\|)\}. \quad (4.6)$$

We shall estimate each term in (4.6) individually as follows. We first note that $\|E_0 \Phi^{i+\sigma}\| \leq \|\Phi^{i+\sigma}\|$ and Lemma 2.9 give that

$$\begin{aligned} \|\delta_{i+\sigma}^\alpha \xi_0\| &= \|\delta_{i+\sigma}^\alpha \xi_0 - {}^C D_T^\alpha \xi_0 + {}^C D_T^\alpha \xi_0\| \\ &= \|\delta_{i+\sigma}^\alpha u - {}^C D_T^\alpha u(t_{i+\sigma}) - E_0(\delta_{i+\sigma}^\alpha u - {}^C D_T^\alpha u(t_{i+\sigma})) - {}^C D_T^\alpha \xi_0\| \\ &\leq \|\Phi^{i+\sigma}\| + \|E_0 \Phi^{i+\sigma}\| + Ch^{k+1} \|{}^C D_T^\alpha u^{i+\sigma}\|_{k+1} \\ &\leq C \|\Phi^{i+\sigma}\| + Ch^{k+1} \|{}^C D_T^\alpha u\|_{L^\infty(H^{k+1}(\Omega))}. \end{aligned} \quad (4.7)$$

Now, from Lemma 3.4 and Lemma 3.2 we get

$$\begin{aligned} \max_{0 \leq i \leq M-1} \{t_{i+\sigma}^\alpha \|\Phi^{i+\sigma}\|\} &\leq C \max_{0 \leq i \leq M-1} \left\{ \|\psi_u^{i+\sigma}\| + \left\{ \max_{s=1, \dots, i} \|\psi_u^{i,s}\| \right\} \right\} \\ &\leq CM^{-\min\{r\alpha, 3-\alpha\}}. \end{aligned} \quad (4.8)$$

We next consider the second term $\max_{0 \leq i \leq M-1} \{t_{i+\sigma}^\alpha \|\mathcal{R}^{i+\sigma}\|\}$ in (4.6). When $i = 0$, the assumption $\|u(t)\| \leq C$ implies that

$$t_0^\alpha \|\mathcal{R}^\sigma\| \leq Ct_1^\alpha \simeq M^{-r\alpha}, \quad (4.9)$$

where we have used that $t_1 = M^{-r}$. Taylor's theorem [61, Lemma 9] and the assumption that $\left\| \frac{\partial^2 u(\cdot, t)}{\partial t^2} \right\| \leq C(1 + t^{\alpha-2}) \leq Ct^{\alpha-2}$, for each $i \geq 1$ yield

$$t_{i+\sigma}^\alpha \|\mathcal{R}^{i+\sigma}\| \leq Ct_{i+\sigma}^\alpha \tau_{i+1}^2 t_i^{\alpha-2} \leq Ct_{i+1}^\alpha M^{-2} t_{i+1}^{2-2/r} t_i^{\alpha-2} \leq Ct_{i+1}^{2\alpha-2/r} M^{-2},$$

where we have used (3.1) and (3.2). Therefore,

$$t_{i+\sigma}^\alpha \|\mathcal{R}^{i+\sigma}\| \leq C \begin{cases} M^{-2} & \text{for } i = 1, \dots, M-1 \text{ if } r \geq 1/\alpha, \\ M^{-2} t_1^{2\alpha-2/r} \simeq M^{-2r\alpha} & \text{for } i = 1, \dots, M-1 \text{ if } 1 \leq r \leq 1/\alpha. \end{cases}$$

This result, together with (4.9), gives

$$\max_{0 \leq i \leq M-1} \{t_{i+\sigma}^\alpha \|\mathcal{R}^{i+\sigma}\|\} \leq CM^{-\min\{r\alpha, 2\}}. \quad (4.10)$$

From (4.6)–(4.10), one can conclude that

$$\max_{1 \leq m \leq M} \|\theta_{0,h}^m\| \leq C (M^{-\min\{r\alpha, 2\}} + h^{k+1}). \quad (4.11)$$

Combining (4.11), Lemmas (2.9) and (4.3) gives the desired result (4.5). The proof is now completed. \square

Corollary 4.1. *Assume that $r\alpha \geq 2$. Then the error between the numerical solution u_h^m computed by $L2 - 1_\sigma$ SFWG-FEM scheme (4.1) satisfies*

$$\|u^m - u_{0,h}^m\| \leq C (h^{k+1} + M^{-2}) \text{ for } m = 0, 1, \dots, M.$$

Proof. From Theorem 4.1 and $r \geq 2/\alpha$, one has $\|u^m - u_{0,h}^m\| \leq C(h^{k+1} + M^{-2})$ for each m . The result now follows. The proof is now completed. \square

The convergence order $O(M^{-2})$ in time for the L2-1 $_{\sigma}$ SFWG-FEM proved in Corollary 4.1 is higher-order than the $O(M^{-(2-\alpha)})$ temporal rate obtained by the well-known L1-WG-FEM. The $O(h^{k+1})$ spatial convergence rate of the numerical methods is optimal in $L^2(\Omega)$.

5. Numerical results

In this section, we present some numerical experiments to show that our theoretical results are valid. For easy implementation and to avoid much algebra, we take $T = 1$ and $Q_T = [0, 1]^2 \times (0, 1]$. In the following examples, we have first divided the domain Ω into squares and then further divided these squares into triangles by connecting the diagonal lines (as illustrated in Figure 1). We use the space of piecewise linear polynomials on the triangles and constant polynomials on the edges. We aim to verify the spacial and temporal accuracy of the L2-1 $_{\sigma}$ SFWG-FEM scheme (4.1) on graded meshes. In the numerical solutions, the L2-1 $_{\sigma}$ SFWG-FEM solution $u_h^m = \{u_{0,h}^m, u_{b,h}^m\}$ is computed by the scheme (4.1). We compute the temporal errors $e_h = \{e_0, e_b\} = \{u^m - u_{0,h}^m, u^m - u_{b,h}^m\}$ in the $L^\infty(L^2)$ -norm using the following formula

$$\|e_0\|_{L^\infty(L^2),M} = \max_{0 \leq m \leq M} \|u(x, t_m) - u_{0,h}^m\|,$$

where we have used 5-point Gaussian quadrature rules on time mesh interval for each m to approximate $\|u(x, t_m) - u_{0,h}^m\|$. The order of convergence (OC) is computed using the following formula

$$OC = \log_2 \left(\frac{\|u - u_{0,h}^m\|_{L^\infty(L^2),M}}{\|u - u_{0,2h}^m\|_{L^\infty(L^2),2M}} \right).$$

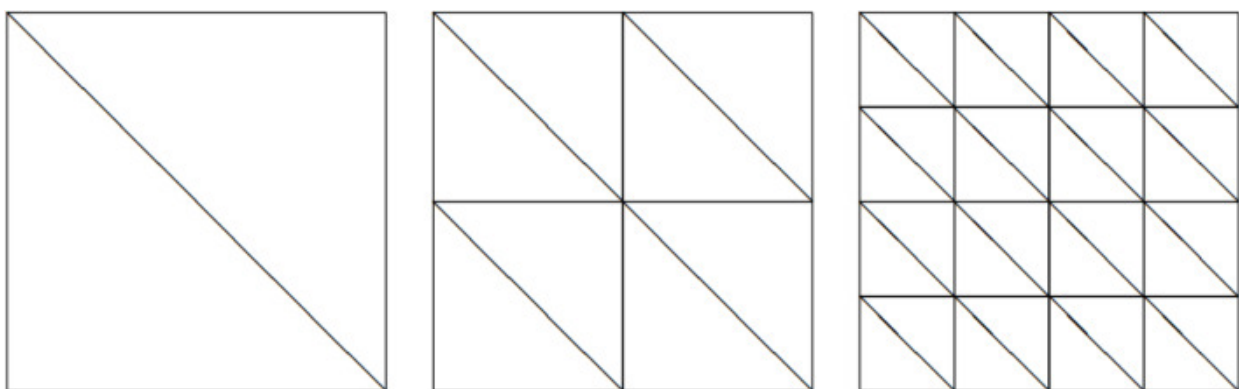


Figure 1. Uniform triangulation of the domain.

Example 5.1. Let $A = 1$ in (1.1). We take $f(x, y, t)$ so that the problem (1.1) has the following solution

$$u(x, y, t) = (t^\theta + t^2) \sin(\pi x) \sin(\pi y),$$

where θ is a fractional number.

If $\theta \in (0, 1)$, then the solution u will have a weak singularity at $t = 0$. That is, the integer-order derivatives of u with respect to the time will blow up at the starting point.

We first fix $M = 2000$ which is small enough to avoid the time discretization errors and take the mesh size $h = \frac{1}{2^m}$, $m = 1, 2, 3, 4, 5, 6$. We present the spatial errors at $t = 1$ in the energy and L^2 norms and their orders of convergence is shown in Table 3 when $\alpha = 0.4$ and $\theta = 2$. The orders of convergence listed in this table show that the proposed method has the optimal convergence order of $\mathcal{O}(h^k)$ in the energy norm and $\mathcal{O}(h^{k+1})$ in the L^2 norm, as stated by Corollary 4.1.

Table 3. Spatial error histories and the rates of convergence at $t = 1$ of the SFWG-FEM using \mathbb{P}_2 element for Example 5.1 on triangular meshes.

Take $M = 2000, \theta = 2, \alpha = 0.4, r = 1$				
N	$\ e_h\ _E$	OC	$\ e_0\ $	OC
4	1.47E + 01	-	4.30E - 03	-
8	2.98E + 00	2.30	5.41E - 04	2.99
16	6.66E - 01	2.16	6.75E - 05	3.00
32	1.60E - 01	2.05	8.43E - 06	3.00
64	4.16E - 02	1.95	1.06E - 06	2.99
Expected OC		2.00		3.00

Then, we fix the mesh size h in space so that the errors in time dominate the errors in space. The results in Table 4, Table 5 and Table 6 show that the rate of convergence is of order $\mathcal{O}(M^{-2})$, which is in perfect agreement with Corollary 4.1. In these tables, when $r = r_{opt}$, we observe that the computed OC is slightly bigger than the expected ones. The reason for this is that the optimum value of the grading constant yields a better approximation in L^2 norm for the problem having a weak singularity at the initial point. Observe that if the regularity parameter $\theta \in (0, 1)$, then one cannot achieve the optimal convergence rate using the uniform mesh or non-optimal grading constant r due to the singularity of the solution at $t = 0$.

Table 4. Time error histories and the rates of convergence of the SFWG-FEM for Example 5.1 on triangular meshes.

Take $h = \pi/64, \theta = 0.8, \alpha = 0.4$						
M	$r = 1$		$r_{opt} = 5/2$		$r = 3$	
	$\ e_0\ $	OC	$\ e_0\ $	OC	$\ e_0\ $	OC
16	2.85E - 03	-	5.72E - 04	-	4.73E - 04	-
32	1.75E - 03	0.70	1.43E - 04	2.00	1.18E - 04	2.00
64	1.04E - 03	0.75	3.41E - 05	2.07	2.96E - 05	2.00
128	6.10E - 04	0.77	8.31E - 06	2.03	7.23E - 06	2.03
256	3.55E - 04	0.78	2.08E - 06	2.00	1.79E - 06	2.01
Expected OC	-	0.8	.	2.00	-	2.00

Table 5. Time error histories and the rate convergence of the SFWG-FEM for Example 5.1 on triangular meshes.

$h = \pi/64, \theta = 0.5, \alpha = 0.2$						
	$r = 2$		$r = 5/2$		$r_{opt} = 4$	
M	$\ e_0\ $	OC	$\ e_0\ $	OC	$\ e_0\ $	OC
16	3.06E – 03	-	4.03E – 03	-	4.07E – 03	-
32	1.74E – 03	0.81	1.84E – 03	1.13	1.92E – 03	1.08
64	9.03E – 04	0.95	8.53E – 04	1.11	6.88E – 04	1.48
128	4.54E – 04	0.99	3.67E – 04	1.21	1.90E – 04	1.86
256	2.21E – 04	1.04	1.43E – 04	1.36	5.01E – 05	1.92
Expected OC		1.00		1.25		2.00

Table 6. Time error histories and the rate convergence of the SFWG-FEM for Example 5.1 on triangular meshes.

$h = \pi/64, \theta = 0.5, r = 2/\theta$				
	$\alpha = 2/3$		$\alpha = 4/5$	
M	$\ e_0\ $	OC	$\ e_0\ $	OC
16	5.36E – 03	-	6.55E – 03	-
32	2.49E – 03	1.10	3.02E – 03	1.11
64	8.84E – 04	1.50	1.07E – 03	1.50
128	2.41E – 04	1.86	2.93E – 04	1.87
256	6.23E – 05	1.95	7.41E – 04	1.98
Expected OC		2.00		2.00

Example 5.2. Let $K = 0.1$ in (1.1). The function $f(x, y, t)$ is taken so that the exact solution of the problem (1.1) is

$$u(x, y, t) = (t^\theta + 1)(x(1-x)y(1-y))e^{x+y},$$

where θ is a regularity parameter.

In this example, we take $M = 2000$ so that the spatial error dominates the error in time. Table 7 lists the computed errors at $t = 1$ in the energy and L^2 norms and their orders of convergence. These rates of convergence displayed are in good agreement with the theory predicted by Corollary 4.1. These rates show that the optimal orders of convergence are obtained.

Next, we fix the spatial step size $h = 1/300$ to ensure that the temporal errors dominate the errors in space. We compute $\|u^m - u_{0,h}^m\|_{L^\infty(L^2), M}$ of L2-1 $_\sigma$ SFWG-FEM and their rates of convergence in Table 8 and in Table 9 for various values of α . These rates of convergence displayed are in good agreement with the theory predicted by Corollary 4.1. Further, Table 8 displays that the optimal rate of convergence $\mathcal{O}(M^{-2})$ is only obtained by the values of the grading constant $r \geq 2/\theta$.

Table 7. Spatial error histories and the rate convergence at $t = 1$ of the SFWG-FEM for Example 5.2 on triangular meshes.

Fix $M = 2000, \theta = 2, \alpha = 0.4, r = 1$				
N	$\ e_h\ _E$	OC	$\ e_0\ $	OC
4	2.62E + 00	-	7.51E - 04	-
8	6.56E - 01	2.00	9.56E - 05	2.97
16	1.65E - 01	1.99	1.20E - 05	2.99
32	4.18E - 02	1.98	1.50E - 06	3.00
64	1.06E - 02	1.97	1.88E - 07	3.00
Expected OC		2.00		3.00

Table 8. Time error histories and the rate convergence of the SFWG-FEM for Example 5.2 on triangular meshes.

$h = 1/300, \theta = 0.8, \alpha = 0.2$						
M	$r = 1$		$r_{opt} = 5/2$		$r = 3$	
	$\ e_0\ $	OC	$\ e_0\ $	OC	$\ e_0\ $	OC
16	1.25E - 04	-	2.07E - 05	-	3.95E - 05	-
32	7.65E - 05	0.70	5.61E - 06	1.88	1.10E - 05	1.84
64	4.57E - 05	0.74	1.61E - 06	1.80	2.96E - 06	1.90
128	2.68E - 05	0.77	4.35E - 07	1.89	8.11E - 07	1.87
256	1.54E - 05	0.80	1.12E - 07	1.96	2.14E - 07	1.92
Expected OC		0.80		2.00	-	2.00

Table 9. Time error histories and the rate convergence of the SFWG-FEM for Example 5.2 on triangular meshes.

$h = 1/300, \theta = 0.8, r = 2/\theta$				
M	$\alpha = 2/3$		$\alpha = 4/5$	
	$\ e_0\ $	OC	$\ e_0\ $	OC
16	4.27E - 05	-	5.37E - 05	-
32	1.99E - 05	1.10	2.42E - 05	1.14
64	6.98E - 06	1.50	8.58E - 06	1.50
128	1.92E - 06	1.86	2.32E - 06	1.89
256	4.93E - 07	1.96	5.79E - 07	2.00
Expected OC		2.00		2.00

In Example 5.3 we use triangular meshes with hanging nodes to illustrate the advantage of the SFWG method in treating irregular meshes shown in Figure 2.

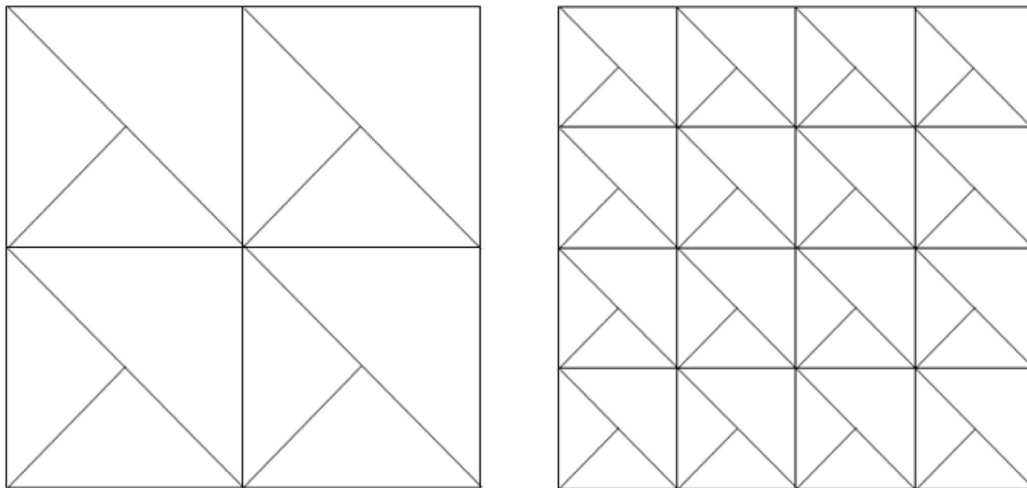


Figure 2. Triangulation of the domain with hanging nodes.

Example 5.3. Let $\Omega = (0, 1)^2$, $A = 1$, $T = 1$ in (1.1) and $f(x, y, t)$ be given such that the exact solution

$$u(x, y, t) = t^\alpha \sin(\pi x) \sin(\pi y).$$

In order to show the advantage of the SFWG method in dealing with the triangular meshes with hanging nodes, we report the measured errors in the energy and L^2 norms in Table 10. The results show the flexibility of the SFWG-FEM in treating hanging nodes in meshes, while the standard finite element method cannot be easily applied without any hp refinement, which makes the computations formidably expensive.

Table 10. Spatial error histories and the rates of convergence at $t = 1$ of the SFWG-FEM using \mathbb{P}_2 element for Example 5.3 on triangular meshes with hanging nodes.

Take $M = 2000$, $\alpha = 0.3$, $r = 2/\alpha$				
N	$\ e_h\ _E$	OC	$\ e_0\ $	OC
16	4.53E – 01	-	1.26E – 02	-
32	1.18E – 01	1.99	1.56E – 03	3.01
64	2.94E – 02	2.00	1.94E – 04	3.00
128	7.32E – 03	2.00	2.43E – 05	3.00
256	1.83E – 03	2.00	3.03E – 06	3.00
Expected OC		2.00		3.00

6. Conclusions

In this paper, we studied the SFWG-FEM in space and L^2 - 1_σ method in time for the fractional diffusion problems on graded meshes in time. We derived optimal error estimates of semi-discrete in the L^2 and H^1 norms and fully discrete numerical schemes in the L^2 norm. Because of the singularity at the initial time, graded meshes in time were used, and the optimal values of the grading constant gave the second order convergence in time. Various examples were carried out to verify the theory presented in this work. We will investigate the method of this paper to fractional diffusion problems with time-dependent diffusion coefficient in future work.

Use of AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgements

The authors would like to express their deep gratitude to the editor and anonymous referees for their valuable comments and suggestions that improve the paper.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. I. Podlubny, *Fractional differential equations*, San Diego: Academic Press, 1998.
2. J. P. Bouchaud, A. Georges, Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications, *Phys. Rep.*, **195** (1990), 127–293. [https://doi.org/10.1016/0370-1573\(90\)90099-N](https://doi.org/10.1016/0370-1573(90)90099-N)
3. T. H. Solomon, E. R. Weeks, H. L. Swinney, Observation of anomalous diffusion and Levy flights in a two-dimensional rotating flow, *Phys. Rev. Lett.*, **71** (1993), 3975. <https://doi.org/10.1103/PhysRevLett.71.3975>
4. W. Wyss, The fractional diffusion equation, *J. Math. Phys.*, **27** (1986), 2782–2785. <https://doi.org/10.1063/1.527251>
5. R. Metzler, J. Klafter, The random walk's guide to anomalous diffusion: A fractional dynamics approach, *Phys. Rep.*, **339** (2000), 1–77. [https://doi.org/10.1016/S0370-1573\(00\)00070-3](https://doi.org/10.1016/S0370-1573(00)00070-3)
6. R. Gorenflo, F. Mainardi, D. Moretti, P. Paradisi, Time fractional diffusion: A discrete random walk approach, *Nonlinear Dynam.*, **29** (2002), 129–143 <https://doi.org/10.1023/A:1016547232119>
7. N. N. Leonenko, M. M. Meerschaert, A. Sikorskii, Fractional pearson diffusions, *J. Math. Anal. Appl.*, **403** (2013), 532–546. <https://doi.org/10.1016/j.jmaa.2013.02.046>
8. Z. Z. Sun, X. N. Wu, A fully discrete difference scheme for a diffusion-wave system, *Appl. Numer. Math.*, **56** (2006), 193–209. <https://doi.org/10.1016/j.apnum.2005.03.003>
9. Y. Lin, C. Xu, Finite difference/spectral approximations for the time-fractional diffusion equation, *J. Comput. Phys.*, **225** (2007), 1533–1552. <https://doi.org/10.1016/j.jcp.2007.02.001>
10. A. A. Alikhanov, Numerical methods of solutions of boundary value problems for the multi-term variable-distributed order diffusion equation, *Appl. Math. Comput.*, **268** (2015), 12–22. <https://doi.org/10.1016/j.amc.2015.06.045>
11. F. I. Taukenova, M. Kh. Shkhanukov-Lafishev, Difference methods for solving boundary value problems for fractional differential equations, *Comput. Math. Math. Phys.*, **46** (2006), 1785–1795. <https://doi.org/10.1134/S0965542506100149>

12. C. M. Chen, F. Liu, V. Anh, I. Turner, Numerical schemes with high spatial accuracy for a variable-order anomalous subdiffusion equation, *SIAM J. Sci. Comput.*, **32** (2010), 1740–1760. <https://doi.org/10.1137/090771715>
13. A. A. Alikhanov, Boundary value problems for the diffusion equation of the variable order in differential and difference settings, *Appl. Math. Comput.*, **219** (2012), 3938–3946. <https://doi.org/10.1016/j.amc.2012.10.029>
14. R. Du, W. R. Cao, Z. Z. Sun, A compact difference scheme for the fractional diffusion-wave equation, *Appl. Math. Model.*, **34** (2010), 2998–3007. <https://doi.org/10.1016/j.apm.2010.01.008>
15. G. H. Gao, Z. Z. Sun, H. W. Zhang, A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications, *J. Comput. Phys.*, **259** (2014), 33–50. <https://doi.org/10.1016/j.jcp.2013.11.017>
16. Y. N. Zhang, Z. Z. Sun, H. W. Wu, Error estimates of Crank-Nicolson-type difference schemes for the subdiffusion equation, *SIAM J. Numer. Anal.*, **49** (2011), 2302–2322. <https://doi.org/10.1137/100812707>
17. Y. Lin, X. Li, C. Xu, Finite difference/spectral approximations for the fractional cable equation, *Math. Comp.*, **80** (2011), 1369–1396. <https://doi.org/10.1090/S0025-5718-2010-02438-X>
18. X. J. Li, C. J. Xu, A space-time spectral method for the time fractional diffusion equation, *SIAM J. Numer. Anal.*, **47** (2009), 2108–2131. <https://doi.org/10.1137/080718942>
19. W. Qiu, D. Xu, J. Guo, J. Zhou, A time two-grid algorithm based on finite difference method for the two-dimensional nonlinear time-fractional mobile/immobile transport model, *Numer. Algorithms*, **85** (2020), 39–58. <https://doi.org/10.1007/s11075-019-00801-y>
20. X. Peng, D. Xu, W. Qiu, Pointwise error estimates of compact difference scheme for mixed-type time-fractional Burgers' equation, *Math. Comput. Simulation*, **208** (2023), 702–726. <https://doi.org/10.1016/j.matcom.2023.02.004>
21. G. H. Gao, Z. H. Sun, A compact finite difference scheme for the fractional sub-diffusion equations, *J. Comput. Phys.*, **230** (2011), 586–595. <https://doi.org/10.1016/j.jcp.2010.10.007>
22. X. Yang, L. Wu, H. Zhang, A space-time spectral order sinc-collocation method for the fourth-order nonlocal heat model arising in viscoelasticity, *Appl. Math. Comput.*, **457** (2023), 128192. <https://doi.org/10.1016/j.amc.2023.128192>
23. W. Wang, H. Zhang, X. Jiang, X. Yang, A high-order and efficient numerical technique for the nonlocal neutron diffusion equation representing neutron transport in a nuclear reactor, *Ann. Nucl. Energy*, **195** (2024), 110163. <https://doi.org/10.1016/j.anucene.2023.110163>
24. Q. Tian, X. Yang, H. Zhang, D. Xu, An implicit robust numerical scheme with graded meshes for the modified Burgers model with nonlocal dynamic properties, *Comput. Appl. Math.*, **42** (2023), 246. <https://doi.org/10.1007/s40314-023-02373-z>
25. Z. Zhou, H. Zhang, X. Yang, H^1 -norm error analysis of a robust ADI method on graded mesh for three-dimensional subdiffusion problems, *Numer. Algorithms*, 2023. <https://doi.org/10.1007/s11075-023-01676-w>

26. H. Zhang, X. Yang, Q. Tang, D. Xu, A robust error analysis of the OSC method for a multi-term fourth-order sub-diffusion equation, *Comput. Math. Appl.*, **109** (2022), 180–190. <https://doi.org/10.1016/j.camwa.2022.01.007>
27. H. Zhang, Y. Liu, X. Yang, An efficient ADI difference scheme for the nonlocal evolution problem in three-dimensional space, *J. Appl. Math. Comput.*, **69** (2023), 651–674. <https://doi.org/10.1007/s12190-022-01760-9>
28. A. A. Alikhanov, A new difference scheme for the time fractional diffusion equation, *J. Comput. Phys.*, **280** (2015), 424–438. <https://doi.org/10.1016/j.jcp.2014.09.031>
29. J. Wang, X. Ye, A weak Galerkin finite element method for second-order elliptic problems, *J. Comput. Appl. Math.*, **241** (2013), 103–115. <https://doi.org/10.1016/j.cam.2012.10.003>
30. Q. Li, J. Wang, Weak Galerkin finite element methods for parabolic equations, *Numer. Methods Partial Differ. Equ.*, **29** (2013), 2004–2024. <https://doi.org/10.1002/num.21786>
31. L. Mu, J. Wang, X. Ye, A modified weak Galerkin finite element method for the Stokes equations, *J. Comp. Appl. Math.*, **275** (2015), 79–90. <https://doi.org/10.1016/j.cam.2014.08.006>
32. T. Zhang, L. Tang, A weak finite element method for elliptic problems in one space dimension, *Appl. Math. Comput.*, **280** (2016), 1–10. <https://doi.org/10.1016/j.amc.2016.01.018>
33. J. Wang, X. Ye, A weak Galerkin mixed finite element method for second order elliptic problems, *Math. Comput.*, **83** (2014), 2101–2126.
34. S. Toprakseven, A weak Galerkin finite element method on temporal graded meshes for the multi-term time fractional diffusion equations, *Comput. Math. Appl.*, **128** (2022), 108–120. <https://doi.org/10.1016/j.camwa.2022.10.012>
35. S. Toprakseven, A weak Galerkin finite element method for time fractional reaction-diffusion-convection problems with variable coefficients, *Appl. Numer. Math.*, **168** (2021), 1–12. <https://doi.org/10.1016/j.apnum.2021.05.021>
36. S. Toprakseven, S. Dinibutun, Error estimations of a weak Galerkin finite element method for a linear system of $\ell \geq 2$ coupled singularly perturbed reaction-diffusion equations in the energy and balanced norms, *AIMS Mathematics*, **8** (2023), 15427–15465. <https://doi.org/10.3934/math.2023788>
37. S. Toprakseven, Optimal order uniform convergence of weak Galerkin finite element method on Bakhvalov-type meshes for singularly perturbed convection dominated problems, *Hacet. J. Math. Stat.*, **52** (2023), 850–875. <https://doi.org/10.15672/hujms.1117320>
38. S. Toprakseven, Optimal order uniform convergence in energy and balanced norms of weak Galerkin finite element method on Bakhvalov-type meshes for nonlinear singularly perturbed problems, *Comput. Appl. Math.*, **41** (2022), 377. <https://doi.org/10.1007/s40314-022-02090-z>
39. S. Toprakseven, P. Zhu, Error analysis of a weak Galerkin finite element method for two-parameter singularly perturbed differential equations in the energy and balanced norms, *Appl. Math. Comput.*, **441** (2023), 127683. <https://doi.org/10.1016/j.amc.2022.127683>
40. Y. Xie, M. Tang, C. Tang, A weak Galerkin finite element method for indefinite time-harmonic maxwell equations, *Appl. Math. Comput.*, **435** (2022), 127471. <https://doi.org/10.1016/j.amc.2022.127471>

41. X. Wang, F. Gao, Y. Liu, Z. Sun, A weak Galerkin finite element method for high dimensional time-fractional diffusion equation, *Appl. Math. Comput.*, **386** (2020), 125524. <https://doi.org/10.1016/j.amc.2020.125524>
42. H. Wang, D. Xu, J. Zhou, J. Guo, Weak Galerkin finite element method for a class of time fractional generalized Burgers' equation, *Numer. Methods Partial Differ. Equ.*, **37** (2021), 732–749. <https://doi.org/10.1002/num.22549>
43. X. Ye, S. Zhang, A stabilizer-free weak Galerkin finite element method on polytopal meshes, *J. Comput. Appl. Math.*, **371** (2020), 112699. <https://doi.org/10.1016/j.cam.2019.112699>
44. A. Al-Taweel, X. Wang, The lowest-order stabilizer free weak Galerkin finite element method, *Appl. Numer. Math.*, **157** (2020), 434–445. <https://doi.org/10.1016/j.apnum.2020.06.012>
45. A. Al-Taweel, X. Wang, X. Ye, S. Zhang, A stabilizer free weak Galerkin finite element method with supercloseness of order two, *Numer. Methods Partial Differ. Equ.*, **37** (2021), 1012–1029. <https://doi.org/10.1002/num.22564>
46. X. Ye, S. Zhang, A conforming discontinuous Galerkin finite element method, *Int. J. Numer. Anal. Model.*, **17** (2020), 110–117.
47. X. Ye, S. Zhang, Y. Zhu, Stabilizer-free weak Galerkin methods for monotone quasilinear elliptic PDEs, *Results Appl. Math.*, **8** (2020), 100097. <https://doi.org/10.1016/j.rinam.2020.100097>
48. J. Ma, F. Gao, N. Du, Stabilizer-free weak Galerkin finite element method with second-order accuracy in time for the time fractional diffusion equation, *J. Comput. Appl. Math.*, **414** (2022), 114407. <https://doi.org/10.1016/j.cam.2022.114407>
49. M. Stynes, Too much regularity may force too much uniqueness, *Fract. Calc. Appl. Anal.*, **19** (2016), 1554–1562. <https://doi.org/10.1515/fca-2016-0080>
50. X. Ye, S. Zhang, Numerical investigation on weak Galerkin finite elements, 2020, arXiv:2004.12483.255.
51. L. Mu, J. Wang, X. Ye, Weak Galerkin finite element methods on polytopal meshes, *Int. J. Numer. Anal. Model.*, **12** (2015), 31–53.
52. X. Ye, S. Zhang, A new weak gradient for the stabilizer free weak Galerkin method with polynomial reduction, 2020, arXiv:2004.13019.
53. N. J. Ford, J. Y. Xiao, Y. B. Yan, A finite element method for time fractional partial differential equations, *Fract. Calc. Appl. Anal.*, **14** (2011), 454–474. <https://doi.org/10.2478/s13540-011-0028-2>
54. V. J. Ervin, J. P. Roop, Variational formulation for the stationary fractional advection dispersion equation, *Numer. Methods Partial Differ. Equ.*, **22** (2006), 558–576. <https://doi.org/10.1002/num.20112>
55. X. Li, C. Xu, Existence and uniqueness of the weak solution of the space-time fractional diffusion equation and a spectral method approximation, *Commun. Comput. Phys.*, **8** (2010), 1016–1051. <https://doi.org/10.4208/cicp.020709.221209a>
56. R. A. Adams, J. J. F. Fournier, *Sobolev spaces*, Elsevier, 2003.
57. A. A. Kilbas, H. M. Srivastava, J. J. Trujillo, *Theory and applications of fractional differential equations*, Elsevier, 2006.

58. V. Thomee, *Galerkin finite element methods for parabolic problems*, Springer Science & Business Media, 2007.
59. M. Stynes, E. O’Riordan, J. L. Gracia, Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation, *SIAM J. Numer. Anal.*, **55** (2017), 1057–1079. <https://doi.org/10.1137/16M1082329>
60. M. Stynes, A survey of the L1 scheme in the discretisation of time-fractional problems, *Numer. Math. Theory Methods Appl.*, **15** (2022), 1173–1192. <https://doi.org/10.13140/RG.2.2.27671.60322>
61. H. Chen, M. Stynes, Error analysis of a second-order method on fitted meshes for a time-fractional diffusion problem, *J. Sci. Comput.*, **79** (2019), 624–647. <https://doi.org/10.13140/RG.2.2.34452.19846>
62. L. Mu, J. Wang, X. Ye, Weak Galerkin finite element methods on polytopal meshes, 2012, arXiv:1204.3655.
63. J. Douglas Jr., T. Dupont, L. Wahlbin, The stability in L^q of the L^2 -projection into finite element function spaces, *Numer. Math.*, **23** (1974), 193–197. <https://doi.org/10.1007/bf01400302>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)