Finite Element Analysis of Time-Fractional Integro-differential Equation of Kirchhoff type for Non-homogeneous Materials

Lalit Kumar [*] Sivaji Ganesh Sista[†] Konijeti Sreenadh[*‡]

Abstract

In this paper, we study a time-fractional initial-boundary value problem of Kirchhoff type involving memory term for non-homogeneous materials ($\mathcal{P}^{\alpha}$). As a consequence of energy argument, we derive $L^{\infty}(0, T; H_0^1(\Omega))$ bound as well as $L^2(0, T; H^2(\Omega))$ bound on the solution of the problem ($\mathcal{P}^{\alpha}$) by defining two new discrete Laplacian operators. Using these a priori bounds, existence and uniqueness of the weak solution to the considered problem is established. Further, we study semi discrete formulation of the problem ($\mathcal{P}^{\alpha}$) by discretizing the space domain using a conforming FEM and keeping the time variable continuous. The semi discrete error analysis is carried out by modifying the standard Ritz-Volterra projection operator in such a way that it reduces the complexities arising from the Kichhoff type nonlinearity. Finally, we develop a new linearized L1 Galerkin FEM to obtain numerical solution of the problem ($\mathcal{P}^{\alpha}$) with a convergence rate of $O(h + k^{2-\alpha})$, where α $(0 < \alpha < 1)$ is the fractional derivative exponent, h and k are the discretization parameters in the space and time directions respectively. This convergence rate is improved to second order in the time direction by proposing a novel linearized L2-1_{σ} Galerkin FEM. We conduct a numerical experiment to validate our theoretical claims.

Keywords: Nonlocal, Finite element method (FEM), Fractional time derivative, Fractional Crank-Nicolson scheme, Integro-differential equation.

AMS subject classification. 34K30, 26A33, 65R10, 60K50.

[*]Department of Mathematics, Indian Institute of Technology Bombay, Mumbai-400076, India. lalitccc528@gmail.com

[†]Department of Mathematics, Indian Institute of Technology Bombay, Mumbai-400076, India. siva@math.iitb.ac.in

[‡]Department of Mathematics, Indian Institute of Technology Delhi, New Delhi-110016, India. sreenadh@math.iitd.ac.in

1 Introduction

Let Ω be a convex and bounded subset of $\mathbb{R}^d$ $(d \geq 1)$ with smooth boundary $\partial\Omega$ and $[0,T]$ is a fixed finite time interval. We consider the following integro-differential equation of Kirchhoff type involving fractional time derivative of order α $(0 < \alpha < 1)$ for non-homogeneous materials

$$\partial_t^\alpha u - \nabla \cdot \left(M\left(x, \int_\Omega |\nabla u(x,t)|^2 \, dx\right) \nabla u \right) = f(x,t) + \int_0^t b(x,t,s) u(s) \, ds \text{ in } \Omega \times (0,T], \tag{$\mathcal{P}^\alpha$}$$

with initial and boundary conditions

$$\begin{aligned} u(x,0) &= u_0(x) \quad \text{in } \Omega, \\ u(x,t) &= 0 \quad \text{on } \partial\Omega \times [0,T], \end{aligned}$$

where $u := u(x,t) : \Omega \times [0,T] \to \mathbb{R}$ is the unknown function and $\partial_t^\alpha u$ is called regularized Caputo fractional derivative of order $\alpha \in (0,1)$, which is defined in [16, 36] as

$$\partial_t^\alpha u = \frac{d}{dt} \int_0^t k(t-s)(u(s) - u(0)) \, ds, \tag{1.1}$$

with $k(t) = \frac{t^{-\alpha}}{\Gamma(1-\alpha)}$ and $\Gamma(\cdot)$ denotes the gamma function. Nonlocal diffusion coefficient M, initial data u_0, source term f, and memory operator $b(x,t,s)$ are known functions.

There are many physical and biological processes in which the mean-squared displacement of the particle motion grows only sublinearly with time t, instead of linear growth. For instance, acoustic wave propagation in viscoelastic materials [25], cancer invasion system [26], anomalous diffusion transport [28], which cannot be described accurately by classical models having integer order derivatives. Therefore, the study of fractional differential equations has evolved immensely in recent years.

Mathematical problems involving fractional time derivatives have been studied by many researchers, for instance, see [10, 14, 24, 32]. Analytical solutions of fractional differential equations are expressed in terms of Mittag-Leffler function, Fox H-functions, Green functions, and hypergeometric functions. Such special functions are more complex to compute, which restrict the applications of fractional calculus in applied sciences. This motivates the researchers to develop numerical algorithms for solving fractional differential equations. Lin and Xu in [22] studied the following linear time-fractional PDE in one space direction

$$\begin{aligned} {}^C D_t^\alpha u - \frac{\partial^2 u}{\partial x^2} &= f(x,t) \quad x \in (0,1), \ t \in (0,T], \\ u(x,0) &= g(x) \quad x \in (0,1), \\ u(0,t) = u(1,t) &= 0 \quad 0 \leq t \leq T, \end{aligned} \tag{1.2}$$

where ${}^C D_t^\alpha u$ is the Caputo fractional derivative defined as

$$ {}^C D_t^\alpha u = \int_0^t k(t-s) \frac{\partial u}{\partial s}(s)\, ds. \tag{1.3}$$

They have developed the L1 scheme based on piecewise linear interpolation for Caputo fractional derivative and Legendre spectral method in space and achieved the convergence estimates of $O(h^2 + k^{2-\alpha})$ for solutions in $C^2\left([0,T]; H^2(\Omega) \cap H_0^1(\Omega)\right)$. Recently, Alikhanov in [1] proposed a modification of the L1 type scheme in the time direction and difference scheme in the space direction for some linear extension to the problem (1.2). In his work, the author proved that the convergence rate is of $O(h^2 + k^2)$ for solutions belonging to $C^3\left([0,T]; C^4(\Omega)\right)$.

On a similar note, there has been considerable attention devoted to the nonlocal diffusion problems where diffusion coefficient depends on the entire domain rather than pointwise. Lions [23] studied the following problem

$$\frac{\partial^2 u}{\partial t^2} - M\left(x, \int_\Omega |\nabla u(x,t)|^2 dx\right) \Delta u = f(x,t) \quad \text{in} \quad \Omega \times [0,T],$$

which models transversal oscillations of an elastic string or membrane by considering the change in length during vibrations. The parabolic problems with Kirchhoff type principal operator have been studied by many researchers, for instance [9, 32, 35]. Also, the stationary case is investigated by many authors in [11, 31, 34] for example nonlocal perturbation of stationary Kirchhoff problem [11] and Kirchhoff equations with magnetic field [34].

The models discussed above behave accurately only for a perfectly homogeneous medium, but in real-life situations, a large number of heterogeneities are present, which cause some memory effect or feedback term [3, 8]. These phenomena cannot be described by classical PDEs, which motivate us to study time-fractional PDEs for non-homogeneous materials. We note that this class of equations has not been analyzed in the literature yet, and this is the first attempt to establish new results for the problem $(\mathcal{P}^\alpha)$.

Key difficulties and our approaches are given below:

1. Due to the appearance of Kirchhoff term, we cannot apply Laplace/Fourier transformation in the problem $(\mathcal{P}^\alpha)$, therefore explicit representation of its solution in terms of Fourier expansion is not possible. To resemble this issue, we use Galerkin method [32, 35] to show the well-posedness of the weak formulation of the problem $(\mathcal{P}^\alpha)$. We define two new discrete Laplacian operators (4.13), (4.14) to derive a priori bounds on the solution of the problem $(\mathcal{P}^\alpha)$.

2. To determine the semi discrete error estimates, we introduce a modified Ritz-Volterra projection operator (5.3) so that it reduces the complications caused by

the Kirchhoff term. This modified Ritz-Volterra projection operator follows the best approximation properties same as that of standard Ritz-Volterra projection operator [4]. These best approximation properties play a key role in deriving the semi discrete error estimates.

3. The fully discrete formulation of the considered problem produces a system of nonlinear algebraic equations. In general, numerical schemes based on the Newton method are adopted to solve this system [13]. The Kirchhoff term leads to the highly non-sparse Jacobian of this system [13]. As a result of which we require high computational cost as well as huge computer storage for solving this system. We reduce these costs by developing new linearization techniques (3.6), (3.15) for the nonlinearity.

4. The memory term incorporates the history of the phenomena under investigation by virtue of which we need to store the value of approximate solution at all previous time steps. This process demands large computer memory. We overcome this difficulty by discretizing the memory term using modified Simpson's rule (3.7), (3.17) [29].

To prove the well-posedness of the weak formulation of the problem $(\mathcal{P}^\alpha)$, we reduce the weak formulation onto a finite dimensional subspace of $H_0^1(\Omega)$. The theory of fractional differential equations [6] ensures the existence of Galerkin sequence of weak solutions. The a priori bounds on these Galerkin sequences are attained by employing the energy argument. We make use of these a priori bounds in Aubin-Lions type compactness lemma [18] to prove that the Galerkin sequence converges to the weak solution of the problem $(\mathcal{P}^\alpha)$.

To obtain the numerical solution, we construct two fully discrete formulations for the problem $(\mathcal{P}^\alpha)$ by discretizing the space domain using a conforming FEM [33] and the time direction by uniform mesh. First, we develop a new linearized L1 Galerkin FEM. This method comprises of L1 type approximation [22] of the Caputo fractional derivative, linearization technique for the Kirchhoff type nonlinearity, and modified Simpson's rule [29] for approximation of the memory term. We acquire the a priori bounds on the solution of this numerical scheme and show that this numerical scheme is accurate of $O(h + k^{2-\alpha})$.

Further, we increase the accuracy of this scheme in the time direction by replacing the L1 scheme with the L2-1_σ scheme [1] for the approximation of the Caputo fractional derivative. As a consequence, we propose a new linearized L2-1_σ Galerkin FEM which has a convergence rate of $O(h + k^2)$. These numerical results are supported by conducting a numerical experiment in MATLAB software.

Turning to the layout of this paper: In Section 2, we provide some notations, assumptions, and preliminaries results that will be used throughout this work. In

Section 3, we state main results of this article. Section 4 contains the proof of well-posedness of the weak formulation of the problem $(\mathcal{P}^\alpha)$. In Section 5, we define semi discrete formulation of the considered problem and derive a priori bounds as well as error estimates on semi discrete solutions. In Section 6, we develop a new linearized L1 Galerkin FEM. We derive a priori bounds on numerical solutions of the developed numerical scheme and prove its accuracy rate of $O(h+k^{2-\alpha})$. In Section 7, we achieve improved convergence rate of $O(h+k^2)$ by proposing a new linearized L2-1_σ Galerkin FEM. Section 8 includes a numerical experiment that confirms the sharpness of theoretical results. Finally, we conclude this work in Section 9.

2 Preliminaries

Let $L^1(\Omega)$ be the set of all equivalence classes of integrable functions on Ω with the norm

$$\|g\|_{L^1(\Omega)} = \int_\Omega |g(x)|\ dx \quad \text{for} \quad g \in L^1(\Omega). \tag{2.1}$$

The Sobolev space $W^{1,1}(\Omega)$ is the collection of all functions in $L^1(\Omega)$ such that its distributional derivative of order one is also in $L^1(\Omega)$, i.e.,

$$W^{1,1}(\Omega) = \left\{ g \in L^1(\Omega);\ Dg \in L^1(\Omega) \right\}. \tag{2.2}$$

The norm on the space $W^{1,1}(\Omega)$ is given by

$$\|g\|_{W^{1,1}(\Omega)} = \|g\|_{L^1(\Omega)} + \|Dg\|_{L^1(\Omega)} \quad \text{for} \quad g \in W^{1,1}(\Omega). \tag{2.3}$$

Let $L^2(\Omega)$ be the set of all equivalence classes of square integrable functions on Ω with the norm

$$\|g\|^2 = \int_\Omega |g(x)|^2\ dx \quad \text{for} \quad g \in L^2(\Omega). \tag{2.4}$$

The norm defined in (2.4) is induced by the inner product $(\cdot,\cdot)$ as follows

$$(g,h) = \int_\Omega g(x)h(x)\ dx \quad \text{for} \quad g,h \in L^2(\Omega). \tag{2.5}$$

The sobolev space $H^m(\Omega), (m \in \{1,2\})$ is the set of all functions in $L^2(\Omega)$ such that its distributional derivatives upto order m are also in $L^2(\Omega)$, i.e.,

$$H^m(\Omega) = \left\{ g \in L^2(\Omega);\ D^\beta g \in L^2(\Omega),\ |\beta| \le m \right\}, \tag{2.6}$$

where β is multiindex. The norm on the space $H^m(\Omega)$ is induced by the following inner product $(\cdot,\cdot)_m$ as follows

$$(g,h)_m = \sum_{|\beta| \le m} (D^\beta g, D^\beta h) \quad \text{for} \quad g,h \in H^m(\Omega). \tag{2.7}$$

We denote $H_0^m(\Omega), (m \in \{1,2\})$ be the closure of $C_C^\infty(\Omega)$ in $H^m(\Omega)$. The space $H_0^m(\Omega)$ can be characterised by the functions in $H^m(\Omega)$ having zero trace on the boundary $\partial\Omega$ [15, Section 2.7]. The dual space of $H_0^m(\Omega)$ is denoted by $H^{-m}(\Omega)$.

For any Hilbert space X, we denote $L^2(0,T;X)$ be the set of all measurable functions $g:[0,T]\to X$ such that

$$\int_0^T \|g(s)\|_X^2 \, ds < \infty. \tag{2.8}$$

The norm on the space $L^2(0,T;X)$ is given by

$$\|g\|_{L^2(0,T;X)}^2 = \int_0^T \|g(s)\|_X^2 \, ds \quad \text{for} \quad g \in L^2(0,T;X). \tag{2.9}$$

We also define a $L^2_\alpha(0,T;X)$ space consisting of all measurable functions $g:[0,T]\to X$ such that

$$\sup_{t\in(0,T)} \left(\frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \|g(s)\|_X^2 \, ds \right) < \infty. \tag{2.10}$$

The norm on the space $L^2_\alpha(0,T;X)$ is given by [19, (4.5)]

$$\|g\|_{L^2_\alpha(0,T;X)}^2 = \sup_{t\in(0,T)} \left(\frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \|g(s)\|_X^2 \, ds \right) \quad \text{for} \quad g \in L^2_\alpha(0,T;X). \tag{2.11}$$

One can observe that $L^2_\alpha(0,T;X) \subset L^2(0,T;X)$. The set of all measurable functions $g:[0,T]\to X$ such that

$$\operatorname{ess}\sup_{t\in(0,T)} \|g(t)\|_X < \infty \tag{2.12}$$

is denoted by $L^\infty(0,T;X)$. The norm on this space is given by

$$\|g\|_{L^\infty(0,T;X)} = \operatorname{ess}\sup_{t\in(0,T)} \|g(t)\|_X \quad \text{for} \quad g \in L^\infty(0,T;X). \tag{2.13}$$

For any two quantities a and b, the notation $a \lesssim b$ means that there exists a generic positive constant C such that $a \le Cb$, where C depends on data but independent of discretization parameters and may vary at different occurrences.

Throughout the paper, we assume the following hypotheses on data:

(H1) Initial data $u_0 \in H^2(\Omega)\cap H_0^1(\Omega)$ and source term $f \in L^\infty(0,T;L^2(\Omega))$.

(H2) Diffusion coefficient $M : \bar{\Omega}\times(0,\infty)\to(0,\infty)$ is a Lipschitz continuous function such that there exists a positive constant m_0 which satisfies

$$M(x,s) \ge m_0 > 0 \text{ for all } \ (x,s)\in\bar{\Omega}\times(0,\infty) \text{ and } \left(m_0 - 4L_M K^2\right) > 0,$$

where $K = \left(\|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))}\right)$ and L_M is a Lipschitz constant.

(H3) Memory operator $b(x,t,s)$ is a second order partial differential operator of the form

$$b(x,t,s)u(s) := -\nabla\cdot(b_2(x,t,s)\nabla u(s)) + \nabla\cdot(b_1(x,t,s)u(s)) + b_0(x,t,s)u(s),$$

with $b_2 : \bar{\Omega} \times [0,T] \times [0,T] \to \mathbb{R}^{d\times d}$ is a symmetric and positive definite matrix with entries $[b_2^{ij}(x,t,s)]$, $b_1 : \bar{\Omega} \times [0,T] \times [0,T] \to \mathbb{R}^d$ is a vector with entries $[b_1^j(x,t,s)]$ and $b_0 : \bar{\Omega} \times [0,T] \times [0,T] \to \mathbb{R}$ is a scalar function. We assume that b_2^{ij}, b_1^j, b_0 are smooth functions in all variables $(x,t,s) \in \bar{\Omega} \times [0,T] \times [0,T]$ for $i,j = 1,2,\ldots,d$.

We define a function $B(t,s,u(s),v)$ for all t,s in $[0,T]$ and for all $u(s), v$ in $H_0^1(\Omega)$ as

$$B(t,s,u(s),v) := (b_2(x,t,s)\nabla u(s), \nabla v) + (\nabla \cdot (b_1(x,t,s)u(s)), v) + (b_0(x,t,s)u(s), v). \tag{2.14}$$

Using (H3) and Poincaré inequality one can prove that there exists a positive constant B_0 such that

$$|B(t,s,u(s),v)| \leq B_0 \|\nabla u(s)\| \, \|\nabla v\| \quad \forall\, t,s \in [0,T], \text{ and } \forall\, u(s), v \in H_0^1(\Omega). \tag{2.15}$$

We indicate $*$ as the convolution of two integrable functions g and h on $[0,T]$ such that

$$(g*h)(t) = \int_0^t g(t-s)h(s)\, ds \quad \forall\, t \ \in [0,T]. \tag{2.16}$$

Remark 1. *Note that $l(t)$ defined by $l(t) := \frac{t^{\alpha-1}}{\Gamma(\alpha)}$ satisfies $k*l = 1$.*

Lemma 2.1. *[12, Lemma 18.4.1] Let H be a real Hilbert space and $T > 0$. Then for any $\tilde{k} \in W^{1,1}(0,T)$ and $v \in L^2(0,T;H)$ we have*

$$\begin{aligned}\left(\frac{d}{dt}\left(\tilde{k} * v\right)(t), v(t)\right)_H &= \frac{1}{2}\frac{d}{dt}\left(\tilde{k} * \|v\|_H^2\right)(t) + \frac{1}{2}\tilde{k}(t)\|v\|_H^2 \\ &\quad + \frac{1}{2}\int_0^t \left[-\tilde{k}'(s)\right] \|v(t) - v(t-s)\|_H^2 \, ds \quad a.e. \ \ t \in (0,T).\end{aligned} \tag{2.17}$$

Lemma 2.2. *[2, Theorem 8] Let u, v be two nonnegative integrable functions on $[a,b]$ and g a continuous function in $[a,b]$. Assume that v is nondecreasing in $[a,b]$ and g is nonnegative and nondecreasing in $[a,b]$. If*

$$u(t) \leq v(t) + g(t)\int_a^t (t-s)^{\alpha-1} u(s)\, ds \quad \textit{for } \alpha \in (0,1) \textit{ and } \forall\, t \in [a,b],$$

then

$$u(t) \leq v(t) E_\alpha \left[g(t)\Gamma(\alpha)(t-a)^\alpha\right] \quad \textit{for } \alpha \in (0,1) \textit{ and } \forall\, t \in [a,b],$$

where $E_\alpha(\cdot)$ is the one parameter Mittag-Leffler function [30, Section 1.2].

Lemma 2.3. *[6] Consider the following initial value problem*

$$\begin{aligned}\partial_t^\alpha y(t) &= g(t, y(t)), \ t \in (0,T], \ \alpha \in (0,1), \\ y(0) &= y_0.\end{aligned} \tag{2.18}$$

Let $y_0 \in \mathbb{R}, K^* > 0, t^* > 0$. Define $D = \{(t, y(t));\ t \in [0, t^*],\ |y - y_0| \leq K^*\}$. Let function $g : D \to \mathbb{R}$ be a continuous. Define $M^* = \sup_{(t,y(t))\in D} |g(t, y(t)|$. Then there exists a continuous function $y \in C[0, T^*]$ which solves the problem (2.18), where

$$T^* = \begin{cases} t^*; & M^* = 0, \\ \min\{t^*, \left(\frac{K^*\Gamma(1+\alpha)}{M^*}\right)^{\frac{1}{\alpha}}\}; & \textit{else}. \end{cases} \tag{2.19}$$

Lemma 2.4. *[19, Lemma 4.1] For $T > 0$ and $\alpha \in (0, 1)$. Let X, Y, and Z be the Banach spaces such that X is compactly embedded in Y and Y is continuously embedded in Z. Suppose that $W \subset L^1_{loc}(0, T; X)$ satisfies the following*

1. *There exist a constant $C_1 > 0$ such that for all $u \in W$*

$$\sup_{t\in(0,T)} \left(\frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha-1} \|u(s)\|_X^2 \, ds \right) \leq C_1. \tag{2.20}$$

2. *There exists a constant $C_2 > 0$ such that for all $u \in W$*

$$\|\partial_t^\alpha u\|_{L^2(0,T;Z)} \leq C_2. \tag{2.21}$$

Then W is relatively compact in $L^2(0, T; Y)$.

Lemma 2.5. *[36] Let k be the kernel defined in (1.1) then there exists a sequence of kernels k_n in $W^{1,1}(0, T)$ such that k_n is nonnegative and nonincreasing in $(0, \infty)$. Also*

$$k_n \to k \ \text{in} \ L^1(0, T) \ \text{as} \ n \to \infty, \tag{2.22}$$

and for $u \in L^2(0, T; L^2(\Omega))$

$$\frac{d}{dt}(k_n * u) \to \frac{d}{dt}(k * u) \ \text{in} \ L^2(0, T; L^2(\Omega)) \ \text{as} \ n \to \infty. \tag{2.23}$$

3 Main results

The weak formulation corresponding to the problem $(\mathcal{P}^\alpha)$ is to find $u \in \mathcal{Z}$ such that the following equations hold for all v in $H_0^1(\Omega)$ and *a.e.* t in $(0, T]$

$$\begin{gathered}(\partial_t^\alpha u, v) + \left(M\left(x, \|\nabla u\|^2\right) \nabla u, \nabla v\right) = (f, v) + \int_0^t B(t, s, u(s), v)\, ds, \ \text{in} \ \Omega \times (0, T], \\ u(x, 0) = u_0(x) \quad \text{in} \ \Omega, \end{gathered} \tag{$\mathcal{W}^\alpha$}$$

where the solution space $\mathcal{Z}$ is defined as

$$\mathcal{Z} := \left\{u \ ; \ u \in L^\infty\left(0, T; L^2(\Omega)\right) \cap L^2\left(0, T; H_0^1(\Omega)\right) \ \text{and} \ \partial_t^\alpha u \in L^2\left(0, T; L^2(\Omega)\right)\right\}. \tag{3.1}$$

Theorem 3.1. ***(Well-posedness of the weak formulation*** $(\mathcal{W}^\alpha)$***)*** *Under the hypotheses (H1), (H2), and (H3) the problem* $(\mathcal{W}^\alpha)$ *admits a unique solution that satisfies the following a priori bounds*

$$\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|u\|_{L^2_\alpha\left(0,T;H^1_0(\Omega)\right)} \lesssim \left(\|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))}\right), \tag{3.2}$$

$$\|u\|_{L^\infty(0,T;H^1_0(\Omega))} + \|u\|_{L^2_\alpha(0,T;H^2(\Omega))} \lesssim \left(\|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))}\right). \tag{3.3}$$

For the semi discrete formulation of the problem $(\mathcal{P}^\alpha)$, we discretize the domain in the space variable by a conforming FEM [33] and keep the time direction continuous. Let $\mathbb{T}_h$ be a shape regular (non overlapping), quasi-uniform triangulation of the domain Ω and h be the discretization parameter in the space direction. We define a finite dimensional subspace X_h of $H^1_0(\Omega)$ as

$$X_h := \{v_h \in C(\bar{\Omega}) \ : \ v_h|_\tau \text{ is a linear polynomial for all } \tau \in \mathbb{T}_h \text{ and } v_h = 0 \text{ on } \partial\Omega\}.$$

The semi discrete formulation for the problem $(\mathcal{P}^\alpha)$ is to seek u_h in X_h such that the following equations hold for all v_h in X_h and *a.e.* t in $(0,T]$

$$\begin{aligned}
&(\partial_t^\alpha u_h, v_h) + \left(M\left(x, \|\nabla u_h\|^2\right)\nabla u_h, \nabla v_h\right) \\
&\qquad\qquad = (f, v_h) + \int_0^t B(t,s,u_h(s),v_h)\, ds \text{ in } \mathbb{T}_h \times (0,T], \\
&u_h(x,0) = u_h^0 \text{ in } \mathbb{T}_h,
\end{aligned} \tag{$\mathcal{S}^\alpha$}$$

where initial condition u_h^0 is in X_h which will be chosen later in the proof of Theorem 3.2.

Theorem 3.2. ***(Error estimate for the semi discrete formulation*** $(\mathcal{S}^\alpha)$***)*** *Suppose that hypotheses (H1), (H2), and (H3) hold. Then we have the following error estimate for the solution u_h of the semi discrete scheme* $(\mathcal{S}^\alpha)$

$$\|u - u_h\|_{L^\infty(0,T;L^2(\Omega))} + \|u - u_h\|_{L^2_\alpha\left(0,T;H^1_0(\Omega)\right)} \lesssim h, \tag{3.4}$$

provided that $u(t)$ is in $H^2(\Omega) \cap H^1_0(\Omega)$ for a.e. t in $[0,T]$.

Further, we move to the fully discrete formulation of the problem $(\mathcal{P}^\alpha)$ for that we divide the interval $[0,T]$ into sub intervals of uniform step size k and $t_n = nk$ for $n = 0,1,2,3,\ldots,N$ with $t_N = T$. We approximate the Caputo fractional derivative by L1 scheme, Kirchhoff type nonlinearity by linearization, and memory term by modified Simpson's rule as follows

L1 approximation scheme [18]: In this scheme, Caputo fractional derivative is approximated at the point t_n using linear interpolation or backward Euler difference formula as follows

$$
\begin{aligned}
{}^{C}D^{\alpha}_{t_n}u &= \frac{1}{\Gamma(1-\alpha)}\int_0^{t_n}\frac{1}{(t_n-s)^{\alpha}}\frac{\partial u}{\partial s}ds \\
&= \frac{1}{\Gamma(1-\alpha)}\sum_{j=1}^{n}\frac{u^j-u^{j-1}}{k}\int_{t_{j-1}}^{t_j}\frac{1}{(t_n-s)^{\alpha}}\ ds+\mathbb{Q}^n \\
&= \frac{k^{-\alpha}}{\Gamma(2-\alpha)}\sum_{j=1}^{n}a_{n-j}\left(u^j-u^{j-1}\right)+\mathbb{Q}^n \\
&= \mathbb{D}^{\alpha}_t u^n+\mathbb{Q}^n
\end{aligned}
\tag{3.5}
$$

where $a_i=(i+1)^{1-\alpha}-i^{1-\alpha},\ i\geq 0,\ u^j=u(x,t_j)$, and $\mathbb{Q}^n$ is the truncation error.
Linearization: For nonlinear term we use the following linearized approximation of u at t_n given by

$$
\begin{aligned}
u^n &\approx 2u^{n-1}-u^{n-2},\ \text{for } n\ \geq\ 2 \\
&:= \bar{u}^{n-1}.
\end{aligned}
\tag{3.6}
$$

Modified Simpson's rule [29]: Let $m_1=[k^{-1/2}]$, where $[\cdot]$ denotes the greatest integer function. Set $k_1=m_1k$ and $\bar{t}_j=jk_1$. Let j_n be the largest even integer such that $\bar{t}_{j_n}<t_n$ and introduce quadrature points

$$
\bar{t}^n_j=\begin{cases} jk_1, & 0\leq j\leq j_n, \\ \bar{t}^n_j+(j-j_n)k, & j_n\leq j\leq J_n, \end{cases}
$$

where $\bar{t}^n_{J_n}=t_{n-1}$. Then quadrature rule for any function g is as follows

$$
\begin{aligned}
\int_0^{t_n}g(s)\ ds &= \sum_{j=0}^{n-1}w_{nj}g(t_j)+q^n(g) \\
&= \frac{k_1}{3}\sum_{j=1}^{j_n/2}\left[g(\bar{t}^n_{2j})+4g(\bar{t}^n_{2j-1})+g(\bar{t}^n_{2j-2})\right] \\
&\quad+\frac{k}{2}\sum_{j=j_n+1}^{J_n}\left[g(\bar{t}^n_j)+g(\bar{t}^n_{j-1})\right]+kg(\bar{t}^n_{J_n})+q^n(g),
\end{aligned}
\tag{3.7}
$$

where w_{nj} are called quadrature weights and $q^n(g)$ is the quadrature error associated with the function g at t_n.

On the basis of approximations (3.5), (3.6), and (3.7) we develop the following linearized L1 Galerkin FEM.
Linearized L1 Galerkin FEM: Find u^n_h $(n=1,2,3,\ldots,N)$ in X_h with $\bar{u}^{n-1}_h=2u^{n-1}_h-u^{n-2}_h$ such that the following equations hold for all v_h in X_h
For $n\geq 2$,

$$
(\mathbb{D}^{\alpha}_t u^n_h,v_h)+\left(M\left(x,\|\nabla\bar{u}^{n-1}_h\|^2\right)\nabla u^n_h,\nabla v_h\right)=(f^n,v_h)+\sum_{j=1}^{n-1}w_{nj}B\left(t_n,t_j,u^j_h,v_h\right).
\tag{$\mathcal{E}^{\alpha}$}
$$

For $n=1$,

$$\left(\mathbb{D}_t^\alpha u_h^1, v_h\right) + \left(M\left(x, \|\nabla u_h^1\|^2\right)\nabla u_h^1, \nabla v_h\right) = \left(f^1, v_h\right) + kB\left(t_1, t_0, u_h^0, v_h\right),$$

with initial condition u_h^0 that is to be chosen later in the proof of Theorem 3.2.

To access the convergence rate of the developed numerical scheme $(\mathcal{E}^\alpha)$, we need the following discrete kernel corresponding to the kernel (a_j)

Lemma 3.3. *[18] Let p_n be a sequence defined by*

$$p_0 = 1, \quad p_n = \sum_{j=1}^{n}(a_{j-1} - a_j)p_{n-j} \quad \text{for} \quad n \geq 1.$$

Then p_n satisfies

$$0 < p_n < 1, \tag{3.8}$$

$$\sum_{j=k}^{n} p_{n-j}a_{j-k} = 1, \quad 1 \leq k \leq n, \tag{3.9}$$

$$\Gamma(2-\alpha)\sum_{j=1}^{n} p_{n-j} \leq \frac{n^\alpha}{\Gamma(1+\alpha)}. \tag{3.10}$$

Theorem 3.4. ***(Convergence estimate for the numerical scheme** $(\mathcal{E}^\alpha)$**)*** *Under the hypotheses (H1), (H2), and (H3) the fully discrete solution u_h^n $(1 \leq n \leq N)$ of the scheme $(\mathcal{E}^\alpha)$ converges to the solution u of the problem $(\mathcal{P}^\alpha)$ with the following rate of accuracy*

$$\max_{1\leq n\leq N} \|u(t_n) - u_h^n\| + \left(k^\alpha \sum_{n=1}^{N} p_{N-n}\|\nabla u(t_n) - \nabla u_h^n\|^2\right)^{1/2} \lesssim (h + k^{2-\alpha}). \tag{3.11}$$

At this point one can see that convergence rate is of $O(k^{2-\alpha})$ in the temporal direction. To improve this convergence rate a new linearized fractional Crank-Nicolson-Galerkin FEM is proposed. In this scheme we replace the L1 approximation of the Caputo fractional derivative with L2-1_σ ($\sigma = \frac{\alpha}{2}$) scheme [1] at $t_{n-\sigma}$ ($t_{n-\sigma} = (1-\sigma)t_n + \sigma t_{n-1}$), linearization technique for nonlinearity at $t_{n-\sigma}$, and modified Simpson's rule for the memory term at $t_{n-\sigma}$.

L2-1_σ approximation scheme [1]: In this scheme, Caputo fractional derivative is approximated at the point $t_{n-\sigma}$ as follows

$$^{C}D^\alpha_{t_{n-\sigma}} = \tilde{\mathbb{D}}^\alpha_{t_{n-\sigma}} u + \tilde{\mathbb{Q}}^{n-\sigma}, \tag{3.12}$$

where

$$\tilde{\mathbb{D}}^\alpha_{t_{n-\sigma}} u = \frac{k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{n} \tilde{c}^{(n)}_{n-j}\left(u^j - u^{j-1}\right), \tag{3.13}$$

with weights $\tilde{c}_{n-j}^{(n)}$ satisfying $\tilde{c}_0^{(1)} = \tilde{a}_0$ for $n = 1$ and for $n \geq 2$

$$\tilde{c}_j^{(n)} = \begin{cases} \tilde{a}_0 + \tilde{b}_1, & j = 0, \\ \tilde{a}_j + \tilde{b}_{j+1} - \tilde{b}_j, & 1 \leq j \leq n-2, \\ \tilde{a}_j - \tilde{b}_j, & j = n-1, \end{cases} \tag{3.14}$$

where

$$\tilde{a}_0 = (1-\sigma)^{1-\alpha} \text{ and } \tilde{a}_l = (l+1-\sigma)^{1-\alpha} - (1-\sigma)^{1-\alpha} \quad l \geq 1,$$

$$\tilde{b}_l = \frac{1}{(2-\alpha)}\left[(l+1-\sigma)^{2-\alpha} - (l-\sigma)^{2-\alpha}\right] - \frac{1}{2}\left[(l+1-\sigma)^{1-\alpha} + (l-\sigma)^{1-\alpha}\right] \; l \geq 1,$$

with $\tilde{\mathbb{Q}}^{n-\sigma}$ is the truncation error.

Linearization: Linearized approximation of the nonlinearity and diffusion at $t_{n-\sigma}$ is given below. For Kirchhoff term

$$\begin{aligned} u^{n-\sigma} &\approx (2-\sigma)u^{n-1} - (1-\sigma)u^{n-2}, \text{ for } n \geq 2 \\ &:= \bar{u}^{n-1,\sigma}, \end{aligned} \tag{3.15}$$

and for diffusion term

$$\begin{aligned} u^{n-\sigma} &\approx (1-\sigma)u^{n} + (\sigma)u^{n-1}, \text{ for } n \geq 1 \\ &:= \hat{u}^{n,\sigma}. \end{aligned} \tag{3.16}$$

Modified Simpson's rule: With a small modification in (3.7) we obtain the following approximation of memory term on $[0, t_{n-\sigma}]$

$$\begin{aligned} \int_0^{t_{n-\sigma}} g(s)\, ds &= \sum_{j=0}^{n-1} \tilde{w}_{nj} g(t_j) + \tilde{q}^{n-\sigma}(g) \\ &= \frac{k_1}{3}\sum_{j=1}^{j_n/2}\left[g(\bar{t}_{2j}^n) + 4g(\bar{t}_{2j-1}^n) + g(\bar{t}_{2j-2}^n)\right] \\ &+ \frac{k}{2}\sum_{j=j_n+1}^{J_n}\left[g(\bar{t}_j^n) + g(\bar{t}_{j-1}^n)\right] + (1-\sigma)\, k g(\bar{t}_{J_n}^n) + \tilde{q}^{n-\sigma}(g), \end{aligned} \tag{3.17}$$

where $\tilde{q}^{n-\sigma}(g)$ is the quadrature error associated with the function g at $t_{n-\sigma}$.
By combining all approximations (3.12)-(3.17), we construct the following linearized L2-1_σ Galerkin FEM.

Linearized L2-1_σ Galerkin FEM: Find u_h^n $(n = 1, 2, 3, \ldots, N)$ in X_h with $\bar{u}_h^{n-1,\sigma} = (2-\sigma)\, u_h^{n-1} - (1-\sigma)\, u_h^{n-2}$ and $\hat{u}_h^{n,\sigma} = (1-\sigma)\, u_h^n + (\sigma)\, u_h^{n-1}$ such that the following equations hold for all v_h in X_h

For $n \geq 2$,

$$\begin{aligned} \left(\tilde{\mathbb{D}}_{t_{n-\sigma}}^\alpha u_h^n, v_h\right) + \left(M\left(x, \|\nabla \bar{u}_h^{n-1,\sigma}\|^2\right)\nabla \hat{u}_h^{n,\sigma}, \nabla v_h\right) &= \sum_{j=1}^{n-1} \tilde{w}_{nj} B\left(t_{n-\sigma}, t_j, u_h^j, v_h\right) \\ &+ \left(f^{n-\sigma}, v_h\right). \end{aligned} \tag{$\mathcal{F}^\alpha$}$$

For $n = 1$,

$$\left(\tilde{\mathbb{D}}^{\alpha}_{t_{1-\sigma}} u_h^1, v_h\right) + \left(M\left(x, \|\nabla \hat{u}_h^{1,\sigma}\|^2\right) \nabla \hat{u}_h^{1,\sigma}, \nabla v_h\right) = (1-\sigma)\, kB\left(t_{1-\sigma}, t_0, u_h^0, v_h\right) + \left(f^{1-\sigma}, v_h\right),$$

with initial condition u_h^0 which is to be prescribed later in the proof of the Theorem 3.2.

Similar to the Lemma 3.3, we have the following discrete kernel corresponding to the kernel $(\tilde{c}_j^n)$.

Lemma 3.5. *[20] Define*

$$\tilde{p}_0^{(n)} = \frac{1}{\tilde{c}_0^{(n)}}, \quad \tilde{p}_j^{(n)} = \frac{1}{\tilde{c}_0^{(n-j)}} \sum_{k=0}^{j-1} \left(\tilde{c}_{j-k-1}^{(n-k)} - \tilde{c}_{j-k}^{(n-k)}\right) \tilde{p}_k^{(n)} \quad \textit{for} \quad 1 \leq j \leq n-1.$$

Then $\tilde{p}_j^{(n)}$ satisfies

$$0 < \tilde{p}_{n-j}^{(n)} < 1, \tag{3.18}$$

$$\sum_{j=k}^{n} \tilde{p}_{n-j}^{(n)} \tilde{c}_{j-k}^{(j)} = 1, \quad 1 \leq k \leq n \leq N, \tag{3.19}$$

$$\Gamma(2-\alpha) \sum_{j=1}^{n} \tilde{p}_{n-j}^{(n)} \leq \frac{n^{\alpha}}{\Gamma(1+\alpha)}, \quad 1 \leq n \leq N. \tag{3.20}$$

Theorem 3.6. ***(Convergence estimate for the numerical scheme*** $(\mathcal{F}^{\alpha})$***)*** *Suppose that hypotheses (H1), (H2), and (H3) hold. Then the fully discrete solution u_h^n $(1 \leq n \leq N)$ of the scheme $(\mathcal{F}^{\alpha})$ satisfies the following convergence estimate*

$$\max_{1 \leq n \leq N} \|u(t_n) - u_h^n\| + \left(k^{\alpha} \sum_{n=1}^{N} \tilde{p}_{N-n}^{(N)} \|\nabla u(t_n) - \nabla u_h^n\|^2\right)^{1/2} \lesssim (h + k^2). \tag{3.21}$$

4 Well-posedness of the weak formulation

In this section, we prove the well-posedness of the weak formulation $(\mathcal{W}^{\alpha})$ using the Galerkin method. For this, we define two new discrete Laplacian operators and apply the energy argument to derive a priori bounds on every Galerkin sequence. As a consequence of compactness Lemma 2.4, these a priori bounds establish the convergence of the Galerkin sequence to the weak solution of the problem $(\mathcal{P}^{\alpha})$.

4.1 Proof of the Theorem 3.1

Proof. Let $\{(\lambda_i, \phi_i)\}_{i=1}^{\infty}$ be the eigenpair corresponding to the standard Laplacian operator with homogeneous Dirichlet boundary condition [7, Section 6.5]. For each fixed positive integer m, consider a finite dimensional subspace $\mathbb{V}_m$ of $H_0^1(\Omega)$ such that $\mathbb{V}_m = \text{span}\{\phi_i\}_{i=1}^m$. Further, reduce the problem $(\mathcal{W}^\alpha)$ onto this space $\mathbb{V}_m$ as to find $u_m \in \mathbb{V}_m$ with the identification

$$u_m(\cdot, t) = \sum_{j=1}^{m} \alpha_{mj}(t)\phi_j, \tag{4.1}$$

such that the following equations hold for all $v_m \in \mathbb{V}_m$ and *a.e.* $t \in (0, T]$

$$\begin{aligned}(\partial_t^\alpha u_m, v_m) + \left(M\left(x, \|\nabla u_m\|^2\right)\nabla u_m, \nabla v_m\right) &= (f, v_m) + \int_0^t B(t, s, u_m(s), v_m)\, ds,\\ u_m(\cdot, 0) &= \sum_{j=1}^{m}(u_0, \phi_j)\phi_j.\end{aligned} \tag{4.2}$$

Then by Riesz-Fischer theorem [21, Theorem 2.29]

$$u_m(\cdot, 0) \to u_0 \text{ in } L^2(\Omega) \text{ as } m \to \infty. \tag{4.3}$$

Put the values of u_m and $u_m(0)$ in (4.2), we obtain a coupled system of fractional order differential equations. Then by the theory of fractional order differential equations Lemma 2.3, the system (4.2) has a continuous solution $u_m(t)$ on some interval $[0, t_*), 0 < t_* < T$, with vanishing trace of $k*(u_m - u_m(0))$ at $t = 0$ [36, Theorem 3.1]. These local solutions $u_m(t)$ are extended to the whole interval by using the following a priori bounds [16, Lemma 3.1].

(A priori bounds) Take $v_m = u_m(t)$ in (4.2) to get

$$\begin{aligned}&\left(\frac{d}{dt}\left(k * u_m\right), u_m\right) + \left(M\left(x, \|\nabla u_m\|^2\right)\nabla u_m, \nabla u_m\right)\\ &\qquad = \left(\frac{d}{dt}\left(k * u_m(0)\right), u_m\right) + (f, u_m) + \int_0^t B(t, s, u_m(s), u_m(t))\, ds.\end{aligned} \tag{4.4}$$

Let $k_n, n \in \mathbb{N}$ be the sequence of kernels defined in Lemma 2.5, then equation (4.4) is rewritten as

$$\begin{aligned}&\left(\frac{d}{dt}(k_n * u_m), u_m\right) + \left(M\left(x, \|\nabla u_m\|^2\right)\nabla u_m, \nabla u_m\right)\\ &= (h_{mn}, u_m) + \left(\frac{d}{dt}\left(k_n * u_m(0)\right), u_m\right) + (f, u_m) + \int_0^t B(t, s, u_m(s), u_m(t))\, ds,\end{aligned} \tag{4.5}$$

with

$$h_{mn} = \frac{d}{dt}\left(k_n * (u_m - u_m(0))\right) - \frac{d}{dt}\left(k * (u_m - u_m(0))\right). \tag{4.6}$$

Use Lemma 2.1, positivity of diffusion coefficient (H2), continuity of $B(t, s, \cdot, \cdot)$ (2.15), Cauchy-Schwarz and Young's inequalities to obtain

$$\begin{aligned}&\frac{d}{dt}\left(k_n * \|u_m\|^2\right)(t) + \frac{\alpha}{\Gamma(1-\alpha)}\int_0^t \frac{\|u_m(t) - u_m(t-s)\|^2}{s^{1+\alpha}}\, ds + \|\nabla u_m\|^2\\ &\lesssim \|h_{mn}\|^2 + k_n(t)\|u_m(0)\|^2 + \|f\|^2 + \|u_m\|^2 + t\int_0^t \|\nabla u_m(s)\|^2\, ds.\end{aligned} \tag{4.7}$$

By convolving the equation (4.7) with the kernel $l(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$ and letting $n \to \infty$ in (4.7) then equation (4.7) reduces to

$$\begin{aligned} \|u_m(t)\|^2 + \left(l * \|\nabla u_m\|^2\right)(t) \lesssim \|u_m(0)\|^2 + \left(l * \|f\|^2\right)(t) \\ + l * \left(\|u_m\|^2 + t\int_0^t \|\nabla u_m(s)\|^2 \, ds\right). \end{aligned} \tag{4.8}$$

In (4.8) we have used the fact that

$$l * \frac{d}{dt}\left(k_n * \|u_m\|^2\right)(t) = \frac{d}{dt}\left(k_n * l * \|u_m\|^2\right)(t) \to \frac{d}{dt}\left(k * l * \|u_m\|^2\right)(t) = \|u_m\|^2, \tag{4.9}$$

with

$$(l * k_n)(t) \to (l * k)(t) = 1 \quad \text{and} \quad (l * h_{mn})(t) \to 0 \quad \text{as} \quad n \to \infty \quad \text{in} \quad L^1(0,T). \tag{4.10}$$

Denote $\tilde{u}_m(t) := \|u_m(t)\|^2 + \left(l * \|\nabla u_m\|^2\right)(t)$ and $\tilde{v}_m(t) := \|u_m(0)\|^2 + \left(l * \|f\|^2\right)(t)$. Then equation (4.8) is converted into

$$\begin{aligned} \tilde{u}_m(t) &\lesssim \tilde{v}_m(t) + l * \left(\|u_m\|^2 + t\int_0^t \|\nabla u_m(s)\|^2 \, ds\right) \\ &\lesssim \tilde{v}_m(t) + l * \left(\|u_m\|^2 + t\int_0^t (t-s)^{\alpha-1}(t-s)^{1-\alpha}\|\nabla u_m(s)\|^2 \, ds\right) \\ &\lesssim \tilde{v}_m(t) + \frac{t^{2-\alpha}}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\tilde{u}_m(s) \, ds. \end{aligned} \tag{4.11}$$

As a consequence of Lemma 2.2, (4.3), and Poincaré inequality, we deduce

$$\begin{aligned} \|u_m\|_{L^\infty(0,T;L^2(\Omega))} + \|u_m\|_{L^2_\alpha(0,T;H^1_0(\Omega))} &\lesssim \|u_m(0)\| + \|f\|_{L^2_\alpha(0,T;L^2(\Omega))} \\ &\lesssim \|u_0\| + \|f\|_{L^2_\alpha(0,T;L^2(\Omega))} \\ &\lesssim \|\nabla u_0\| + \|f\|_{L^2_\alpha(0,T;L^2(\Omega))}. \end{aligned} \tag{4.12}$$

Due to the presence of gradient type nonlinearity in $(\mathcal{P}^\alpha)$, these a priori bounds (4.12) are not sufficient to apply compactness Lemma 2.4. In order to use this lemma, we need to derive a priori bound on $L^\infty(0,T;H^1_0(\Omega))$ as well as $L^2_\alpha(0,T;H^2(\Omega))$. For these a priori bounds, we define two new discrete Laplacian operators Δ^M_m, $\Delta^{b_2}_m$: $\mathbb{V}_m \to \mathbb{V}_m$ such that

$$(-\Delta^M_m \, u_m, v_m) := (M(x, \|\nabla u_m\|^2)\nabla u_m, \nabla v_m) \quad \forall \; u_m, v_m \in \mathbb{V}_m, \; t \in (0,T], \tag{4.13}$$

and

$$(-\Delta^{b_2}_m \, u_m, v_m) := (b_2(x,t,s)\nabla u_m, \nabla v_m) \quad \forall \; u_m, v_m \in \mathbb{V}_m, \; t, s \; \in (0,T]. \tag{4.14}$$

Since diffusion coefficient is positive and b_2 is a symmetric positive definite matrix therefore Δ^M_m and $\Delta^{b_2}_m$ are well defined. We make use of these definitions (4.13) and

(4.14) to convert the equation (4.2) into

$$
\begin{aligned}
&(\partial_t^\alpha u_m, v_m) + \left(-\Delta_m^M\ u_m, v_m\right) \\
&= (f, v_m) + \int_0^t (-\Delta_m^{b_2}\ u_m(s), v_m)\ ds + \int_0^t (\nabla \cdot (b_1(x,t,s) u_m(s)), v_m)\ ds \\
&+ \int_0^t (b_0(x,t,s) u_m(s), v_m)\ ds.
\end{aligned} \tag{4.15}
$$

Put $v_m = -\Delta_m^M\ u_m(t)$ in (4.15) and apply Cauchy-Schwarz inequality together with Young's inequality to obtain

$$
\begin{aligned}
\left(\partial_t^\alpha \nabla u_m, M\left(x, \|\nabla u_m\|^2\right) \nabla u_m\right) + \|\Delta_m^M\ u_m\|^2 \\
\lesssim \int_0^t \|\Delta_m^{b_2}\ u_m(s)\|^2\ ds + \int_0^t \|\nabla u_m(s)\|^2\ ds \\
+ \int_0^t \|u_m(s)\|^2\ ds + \|f\|^2.
\end{aligned} \tag{4.16}
$$

Estimate $|(b_2(x,t,s)\nabla u_m(s), \nabla v_m)| \lesssim \|\nabla u_m(s)\|\ \|\nabla v_m\|$ implies

$$
\|\Delta_m^{b_2}\ u_m(s)\| = \sup_{v_m \in \mathbb{V}_m} \frac{|(b_2(x,t,s)\nabla u_m(s), \nabla v_m)|}{\|\nabla v_m\|} \lesssim \|\nabla u_m(s)\|. \tag{4.17}
$$

Hypothesis (H2) and estimate (4.17) yield

$$
\begin{aligned}
\left(\frac{d}{dt}\left[k * \nabla(u_m - u_m(0))\right](t), \nabla u_m\right) + \|\Delta_m^M\ u_m\|^2 \lesssim \int_0^t \left(\|\nabla u_m(s)\|^2 + \|u_m(s)\|^2\right)\ ds \\
+ \|f\|^2.
\end{aligned} \tag{4.18}
$$

Following the similar lines of the proof of estimate (4.12), we reach at

$$
\begin{aligned}
\|u\|^2_{L^\infty(0,T;H_0^1(\Omega))} + \left(l * \|\Delta_m^M\ u_m\|^2\right)(t) &\lesssim \|\nabla u_m(0)\|^2 + \|f\|^2_{L^2_\alpha(0,T;L^2(\Omega))} \\
&\lesssim \|\nabla u_0\|^2 + \|f\|^2_{L^2_\alpha(0,T;L^2(\Omega))}.
\end{aligned} \tag{4.19}
$$

Finally, substitute $v_m = \partial_t^\alpha u_m$ in (4.15) to have

$$
\begin{aligned}
\|\partial_t^\alpha u_m\|^2 + \left(-\Delta_m^M\ u_m, \partial_t^\alpha u_m\right) \lesssim \|f\|^2 + \int_0^t (\|\Delta_m^{b_2}\ u_m(s)\|^2 + \|\nabla u_m(s)\|^2)\ ds. \\
+ \int_0^t \|u_m(s)\|^2\ ds.
\end{aligned} \tag{4.20}
$$

Proceeding further as estimate (4.19) is proved to conclude

$$
\begin{aligned}
\|\partial_t^\alpha u_m\|_{L^2(0,T;L^2(\Omega))} + \|u_m\|_{L^\infty(0,T;H_0^1(\Omega))} &\lesssim \|\nabla u_m(0)\| + \|f\|_{L^2_\alpha(0,T;L^2(\Omega))} \\
&\lesssim \|\nabla u_0\| + \|f\|_{L^2_\alpha(0,T;L^2(\Omega))}.
\end{aligned} \tag{4.21}
$$

Thus, estimates (4.12) and (4.21) provide a subsequence of (u_m) again denoted by (u_m) such that $u_m \rightharpoonup u$ in $L^2\left(0,T; H_0^1(\Omega)\right)$ and $\partial_t^\alpha u_m \rightharpoonup \partial_t^\alpha u$ in $L^2\left(0,T; L^2(\Omega)\right)$.

In the light of estimates (4.19) and (4.21), we apply compactness Lemma 2.4 to conclude $u_m \to u$ in $L^2(0,T;H_0^1(\Omega))$. Now using the fact that $M(x,\|\nabla u_m\|^2)$ and $B(t,s,u_m(s),v_m)$ are continuous and an application of Lebesgue dominated convergence theorem, we pass the limit inside (4.2) which establishes the existence of weak solutions to the problem $(\mathcal{P}^\alpha)$.

(Initial Condition) The weak solution u satisfies the following equation for all v in $H_0^1(\Omega)$

$$\left(\frac{d}{dt}\left[k*(u-u_0)\right](t),v\right)+\left(M\left(x,\|\nabla u\|^2\right)\nabla u,\nabla v\right)=(f,v)+\int_0^t B(t,s,u(s),v)ds. \tag{4.22}$$

Let ϕ in $C^1([0,T];H_0^1(\Omega))$ with $\phi(T)=0$, multiply (4.22) with ϕ and integrate by parts to get in time

$$\begin{aligned}-\int_0^T ((k*(u-u_0))(t),v)\,\phi'(t)dt+\int_0^T\left(M\left(x,\|\nabla u\|^2\right)\nabla u,\nabla v\right)\phi(t)dt\\ =\int_0^T (f,v)\,\phi(t)dt+\int_0^T\int_0^t B(t,s,u(s),v)\phi(t)dsdt\\ +((k*(u-u_0))(0),\phi(0)).\end{aligned} \tag{4.23}$$

Since $C^1([0,T];H_0^1(\Omega))$ is dense in $L^2(0,T;H_0^1(\Omega))$, thus using (4.2) and $(k*(u_m-u_m(0)))$ has vanishing trace at $t=0$, we have

$$\begin{aligned}-\int_0^T ((k*(u_m-u_m(0)))(t),v)\phi'(t)dt+\int_0^T\left(M\left(x,\|\nabla u_m\|^2\right)\nabla u_m,\nabla v\right)\phi(t)dt\\ =\int_0^T (f,v)\,\phi(t)dt+\int_0^T\int_0^t B(t,s,u_m(s),v)\phi(t)dsdt.\end{aligned} \tag{4.24}$$

Let m tend to infinity in (4.24) and comparing with (4.23) we obtain $((k*(u-u_0))(0),\phi(0))=0$. Since $\phi(0)$ is arbitrary, so we have $(k*(u-u_0))(0)=0$ which implies $u=u_0$ at $t=0$ for $\alpha\in\left(\frac{1}{2},1\right]$ [16, Proposition 6.7]. For the case $\alpha\in\left(0,\frac{1}{2}\right]$, we need to impose more compatibility conditions of data see [16, Theorem 1.3].

(Uniqueness) Suppose that u_1 and u_2 are solutions of the weak formulation $(\mathcal{W}^\alpha)$, then $z=u_1-u_2$ satisfies the following equation for all $v\in H_0^1(\Omega)$ and *a.e.* $t\in(0,T]$

$$\begin{aligned}\left(\frac{d}{dt}(k*z)(t),v\right)+\left(M\left(x,\|\nabla u_1\|^2\right)\nabla z,\nabla v\right)\\ =\left(\left[M\left(x,\|\nabla u_2\|^2\right)-M\left(x,\|\nabla u_1\|^2\right)\right]\nabla u_2,\nabla v\right)\\ +\int_0^t B(t,s,z(s),v)\;ds.\end{aligned} \tag{4.25}$$

Put $v=z(t)$ in (4.25) and using (H2), (H3), and a priori bound (3.3) on u_1,u_2 along with Cauchy-Schwarz and Young's inequality to obtain

$$\left(\frac{d}{dt}(k*z)(t),z(t)\right)+(m_0-4L_MK^2)\|\nabla z\|^2\lesssim\int_0^t\|\nabla z(s)\|^2ds. \tag{4.26}$$

Following the similar lines as in the proof of estimate (4.12) and using (H2) we conclude $\|z\|_{L^2\left(0,T;H_0^1(\Omega)\right)} = \|z\|_{L^\infty(0,T;L^2(\Omega))} = 0$. Thus uniqueness follows. $\square$

5 Semi discrete formulation and error estimate

In this section, we discuss the well-posedness of the semi discrete formulation $(\mathcal{S}^\alpha)$ and derive error estimate for the semi discrete solution by modifying Ritz-Volterra projection operator.

Theorem 5.1. *Suppose that hypotheses (H1), (H2), and (H3) hold. Then there exists a unique solution to the problem $(\mathcal{S}^\alpha)$ which satisfies the following a priori bounds*

$$\|u_h\|_{L^\infty(0,T;L^2(\Omega))} + \|u_h\|_{L^2_\alpha\left(0,T;H_0^1(\Omega)\right)} \lesssim \left(\|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))}\right), \tag{5.1}$$

$$\|\partial_t^\alpha u_h\|_{L^2(0,T;L^2(\Omega))} + \|u_h\|_{L^\infty(0,T;H_0^1(\Omega))} \lesssim \left(\|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))}\right). \tag{5.2}$$

Proof. This theorem is proved analogously to the proof of Theorem 3.1. $\square$

For the semi discrete error estimate, we define a new Ritz-Volterra type projection operator $W : [0,T] \to X_h$ by

$$\left(M\left(x, \|\nabla u\|^2\right)\nabla(u - W), \nabla v_h\right) = \int_0^t B(t, s, u(s) - W(s), v_h)\, ds \quad \forall\, v_h \in X_h. \tag{5.3}$$

This modified Ritz-Volterra projection operator W is well defined by the positivity of the Kirchhoff term M [5]. This projection operator satisfies the following stability and best approximation properties.

Lemma 5.2. *[17] Let W be the modified Ritz-Volterra projection operator defined in* (5.3)*, then $\|\nabla W\|$ is bounded for every t in $[0,T]$, i.e.,*

$$\|\nabla W\| \lesssim \|\nabla u\|.$$

To derive the best approximation properties of the modified Ritz-Volterra projection operator, we assume some additional regularity on the solution u of the problem $(\mathcal{P}^\alpha)$ such that [1, 22]

$$\|u(t)\|_2 \lesssim C \text{ and } \|u_t(t)\|_2 \lesssim C \;\forall\, t \in [0,T]. \tag{5.4}$$

Theorem 5.3. *Suppose that the solution u of the problem $(\mathcal{P}^\alpha)$ satisfies* (5.4)*. Then modified Ritz-Volterra projection operator has the following best approximation properties*

$$\begin{aligned} \|\rho(t)\| + h\|\nabla\rho(t)\| &\lesssim h^2 \quad \forall\, t \in [0,T], \\ \|\rho_t(t)\| + h\|\nabla\rho_t(t)\| &\lesssim h^2 \quad \forall\, t \in [0,T], \end{aligned} \tag{5.5}$$

where $\rho := u - W$.

Proof. For the proof of this theorem we refer the readers to [4, 5]. $\square$

Now error estimate for the semi discrete formulation $(\mathcal{S}^\alpha)$ is attained as stated in Theorem 3.2.

5.1 Proof of the Theorem 3.2

Proof. Denote $\theta := W - u_h$ such that $u - u_h = \rho + \theta$. Then put $u_h = W - \theta$ in the problem $(\mathcal{S}^\alpha)$ to have

$$\begin{aligned}(\partial_t^\alpha (W-\theta), v_h) &+ \left(M\left(x, \|\nabla u_h\|^2\right) \nabla (W-\theta), \nabla v_h\right) \\ &= (f, v_h) + \int_0^t B(t,s,W(s)-\theta(s), v_h)\; ds \quad \forall\; v_h \in X_h.\end{aligned}$$

Weak formulation $(\mathcal{W}^\alpha)$ and the definition (5.3) of the modified Ritz-Volterra projection operator W yield

$$\begin{aligned}&(\partial_t^\alpha \theta, v_h) + \left(M(x, \|\nabla u_h\|^2\right) \nabla\theta, \nabla v_h) \\ &= -(\partial_t^\alpha \rho, v_h) + \int_0^t B(t,s,\theta(s),v_h) ds + \left(\left(M(x,\|\nabla u_h\|^2 - M(x,\|\nabla u\|^2\right)\nabla W, \nabla v_h\right).\end{aligned} \tag{5.6}$$

Set $v_h = \theta(t)$ in (5.6) and employ (H2), (H3) to obtain

$$\begin{aligned}(\partial_t^\alpha \theta, \theta) + m_0\|\nabla\theta\|^2 &= \|\partial_t^\alpha \rho\|\|\theta(t)\| + \|\nabla\theta(t)\|\int_0^t \|\nabla\theta(s)\| ds \\ &\quad + L_M(\|\nabla u_h\| + \|\nabla u\|)(\|\nabla\rho\| + \|\nabla\theta\|)\|\nabla W\|\|\nabla\theta\|.\end{aligned} \tag{5.7}$$

By utilizing proved a priori bounds on $\|\nabla u\|$, $\|\nabla u_h\|$ and $\|\nabla W\|$ together with Cauchy-Schwarz and Young's inequality, we obtain

$$(\partial_t^\alpha \theta, \theta) + (m_0 - 4L_M K^2)\|\nabla\theta\|^2 \lesssim \|\partial_t^\alpha \rho\|^2 + \|\theta\|^2 + \int_0^t \|\nabla\theta(s)\|^2 ds + \|\nabla\rho\|^2. \tag{5.8}$$

Use (H2) and apply similar arguments as we prove estimate (4.12) to deduce

$$\|\theta\|^2_{L^\infty(0,T;L^2(\Omega))} + \left(l * \|\nabla\theta\|^2\right)(t) \lesssim \left[l * \left(\|\nabla\rho\|^2 + \|\partial_t^\alpha\rho\|^2\right)\right](t) + \|\nabla\theta(0)\|^2.$$

For absolutely continuous function ρ, we have $\partial_t^\alpha \rho = {}^C D_t^\alpha \rho$ [6]. Therefore,

$$\begin{aligned}\|\partial_t^\alpha \rho\| = \|{}^C D_t^\alpha \rho\| &= \left\|\frac{1}{\Gamma(1-\alpha)}\int_0^t (t-s)^{-\alpha}\frac{\partial\rho}{\partial s}(s)\; ds\right\| \\ &\lesssim \frac{1}{\Gamma(1-\alpha)}\int_0^t (t-s)^{-\alpha}\left\|\frac{\partial\rho}{\partial s}(s)\right\|\; ds \\ &\lesssim \frac{1}{\Gamma(1-\alpha)}\int_0^t (t-s)^{-\alpha} h^2\; ds \lesssim h^2.\end{aligned} \tag{5.9}$$

We choose $u_h^0 = W(0)$ such that $\theta(0) = 0$. Then apply approximation properties of modified Ritz-Volterra projection operator (5.5) and (5.9) to conclude

$$\|\theta\|^2_{L^\infty(0,T;L^2(\Omega))} + \left(l * \|\nabla\theta\|^2\right)(t) \lesssim h^2 + h^4 \lesssim h^2.$$

Finally, triangle inequality and estimate (5.5) finish the proof. □

6 Linearized L1 Galerkin FEM

In this section, we prove the well-posedness of the numerical scheme ($\mathcal{E}^\alpha$) and carry out its convergence analysis. The following two lemmas provide a priori bounds on the solution to the problem ($\mathcal{E}^\alpha$).

Lemma 6.1. *Under the hypotheses (H1), (H2), and (H3) the solution u_h^n $(n \geq 1)$ of the scheme ($\mathcal{E}^\alpha$) satisfy the following a priori bound*

$$\max_{1\leq m\leq N} \|u_h^m\|^2 + k^\alpha \sum_{n=1}^{N} p_{N-n} \|\nabla u_h^n\|^2 \lesssim \|\nabla u_0\|^2 + \max_{1\leq n\leq N} \|f^n\|^2. \tag{6.1}$$

Proof. Put $v_h = u_h^1$ for $n = 1$ in the formulation ($\mathcal{E}^\alpha$) to get

$$\left(\mathbb{D}_t^\alpha u_h^1, u_h^1\right) + \left(M\left(x, \|\nabla u_h^1\|^2\right)\nabla u_h^1, \nabla u_h^1\right) = \left(f^1, u_h^1\right) + kB\left(t_1, t_0, u_h^0, u_h^1\right).$$

Employing (H2) and (H3), we obtain

$$(1 - k^\alpha\Gamma(2-\alpha))\|u_h^1\|^2 + k^\alpha\|\nabla u_h^1\|^2 \lesssim k^\alpha\left(\|f^1\|^2 + k^2\|\nabla u_h^0\|^2\right) + \|u_h^0\|^2.$$

For sufficiently small k such that $k^\alpha < \frac{1}{\Gamma(2-\alpha)}$, we conclude

$$\|u_h^1\|^2 + k^\alpha\|\nabla u_h^1\|^2 \lesssim \|f^1\|^2 + \|\nabla u_h^0\|^2 \lesssim \|f^1\|^2 + \|\nabla u_0\|^2.$$

Further, set $v_h = u_h^n$ for $n \geq 2$ in the scheme ($\mathcal{E}^\alpha$) to have

$$(\mathbb{D}_t^\alpha u_h^n, u_h^n) + \left(M\left(x, \|\nabla \bar{u}_h^{n-1}\|^2\right)\nabla u_h^n, \nabla u_h^n\right) = (f^n, u_h^n) + \sum_{j=1}^{n-1} w_{nj} B\left(t_n, t_j, u_h^j, u_h^n\right). \tag{6.2}$$

Apply the identity $(\mathbb{D}_t^\alpha u_h^n, u_h^n) \geq \frac{1}{2}\mathbb{D}_t^\alpha\|u_h^n\|^2$ [18] and hypotheses (H2), (H3) along with Cauchy-Schwarz and Young's inequality to reach at

$$\mathbb{D}_t^\alpha\|u_h^n\|^2 + \|\nabla u_h^n\|^2 \lesssim \left(\|f^n\|^2 + \|u_h^n\|^2 + \sum_{j=1}^{n-1} w_{nj}\|\nabla u_h^j\|^2\right). \tag{6.3}$$

By the definition of $\mathbb{D}_t^\alpha\|u_h^n\|^2$ (3.5), the equation (6.3) reduces to

$$\frac{k^{-\alpha}}{\Gamma(2-\alpha)}\sum_{j=1}^{n} a_{n-j}\left(\|u_h^j\|^2 - \|u_h^{j-1}\|^2\right) + \|\nabla u_h^n\|^2 \lesssim \left(\|f^n\|^2 + \|u_h^n\|^2 + \sum_{j=1}^{n-1} w_{nj}\|\nabla u_h^j\|^2\right). \tag{6.4}$$

Multiply the equation (6.4) by discrete convolution P_{m-n} defined in Lemma 3.3 and take summation from $n = 1$ to m to obtain

$$\begin{aligned}
&\sum_{n=1}^{m} p_{m-n}\sum_{j=1}^{n} a_{n-j}\left(\|u_h^j\|^2 - \|u_h^{j-1}\|^2\right) + k^\alpha\Gamma(2-\alpha)\sum_{n=1}^{m} p_{m-n}\|\nabla u_h^n\|^2 \\
&\lesssim k^\alpha\Gamma(2-\alpha)\left(\sum_{n=1}^{m} p_{m-n}\|f^n\|^2 + \sum_{n=1}^{m} p_{m-n}\|u_h^n\|^2 + \sum_{n=1}^{m} p_{m-n}\sum_{j=1}^{n-1} w_{nj}\|\nabla u_h^j\|^2\right).
\end{aligned}$$

Interchanging the summation and property of discrete kernel (3.9) with $\alpha_0 = k^\alpha\Gamma(2-\alpha)$ yield

$$(1-\alpha_0)\|u_h^m\|^2 + k^\alpha\sum_{n=1}^{m} p_{m-n}\|\nabla u_h^n\|^2 \lesssim \sum_{n=1}^{m-1}\left(k^\alpha p_{m-n}\|u_h^n\|^2 + k_1 k^\alpha\sum_{j=1}^{n} p_{n-j}\|\nabla u_h^j\|^2\right) + k^\alpha\sum_{n=1}^{m} p_{m-n}\|f_h^n\|^2 + \|u_h^0\|^2.$$

Then for sufficiently small $k^\alpha < \frac{1}{\Gamma(2-\alpha)}$ one have

$$\|u_h^m\|^2 + k^\alpha\sum_{n=1}^{m} p_{m-n}\|\nabla u_h^n\|^2 \lesssim \sum_{n=1}^{m-1}\left(k^\alpha p_{m-n} + k_1\right)\left(\|u_h^n\|^2 + k^\alpha\sum_{j=1}^{n} p_{n-j}\|\nabla u_h^j\|^2\right) + k^\alpha\sum_{n=1}^{m} p_{m-n}\|f^n\|^2 + \|u_h^0\|^2.$$

Further, the discrete Grönwall's inequality provides

$$\begin{aligned}&\|u_h^m\|^2 + k^\alpha\sum_{n=1}^{m} p_{m-n}\|\nabla u_h^n\|^2 \\ &\quad\lesssim \left(k^\alpha\sum_{n=1}^{m} p_{m-n}\|f^n\|^2 + \|u_h^0\|^2\right)\exp\left(\sum_{n=1}^{m-1}\left(k^\alpha p_{m-n} + k_1\right)\right).\end{aligned} \tag{6.5}$$

Finally, using property of discrete kernel (3.10) one obtain

$$\begin{aligned}k^\alpha\sum_{n=1}^{m} p_{m-n}\|f^n\|^2 \lesssim \max_{1\le n\le N}\|f^n\|^2\left(k^\alpha\sum_{n=1}^{m} p_{m-n}\right) &\lesssim \max_{1\le n\le N}\|f^n\|^2 k^\alpha m^\alpha \\ &\lesssim T^\alpha\max_{1\le n\le N}\|f^n\|^2.\end{aligned} \tag{6.6}$$

Also,

$$\sum_{n=1}^{m-1}\left(k^\alpha p_{m-n} + k_1\right) \lesssim k^\alpha m^\alpha + mk_1 \lesssim T. \tag{6.7}$$

To conclude the result (6.1), put (6.6) and (6.7) in (6.5) as

$$\|u_h^m\|^2 + k^\alpha\sum_{n=1}^{m} p_{m-n}\|\nabla u_h^n\|^2 \lesssim \max_{1\le n\le N}\|f^n\|^2 + \|u_h^0\|^2 \lesssim \max_{1\le n\le N}\|f^n\|^2 + \|\nabla u_0\|^2. \tag{6.8}$$

□

Lemma 6.2. *Suppose that hypotheses (H1), (H2), and (H3) hold. Then the solution u_h^n $(n \ge 1)$ of the scheme $(\mathcal{E}^\alpha)$ satisfy the following a priori bound*

$$\max_{1\le m\le N}\|\nabla u_h^m\|^2 + k^\alpha\sum_{n=1}^{N} p_{N-n}\|\Delta_h^M u_h^n\|^2 \lesssim \|\nabla u_0\|^2 + \max_{1\le n\le N}\|f^n\|^2. \tag{6.9}$$

where $\Delta_h^M : X_h \to X_h$ is the discrete Laplacian operator defined in (4.13).

Proof. By making use of definitions of discrete Laplacian operator (4.13) and (4.14), the equation for $n = 1$ in the scheme $(\mathcal{E}^\alpha)$ is rewritten as

$$\begin{aligned}\left(\mathbb{D}_t^\alpha u_h^1, v_h\right) + (-\Delta_h^M u_h^1, v_h) = \left(f^1, v_h\right) + k(-\Delta_h^{b_2} u_h^0, v_h) + k(\nabla \cdot (b_1(x,t,s)u_h^0), v_h) \\ + k(b_0(x,t,s)u_h^0, v_h).\end{aligned} \tag{6.10}$$

Setting $v_h = -\Delta_{M_h} u_h^1$ in (6.10) one obtain

$$\left(\mathbb{D}_t^\alpha \nabla u_h^1, \nabla u_h^1\right) + \|\Delta_h^M u_h^1\|^2 \lesssim \|f^1\|^2 + k^2 \left(\|\Delta_h^{b_2} u_h^0\|^2 + \|\nabla u_h^0\|^2 + \|u_h^0\|^2\right). \tag{6.11}$$

Again identity $(\mathbb{D}_t^\alpha u_h^n, u_h^n) \geq \frac{1}{2}\mathbb{D}_t^\alpha \|u_h^n\|^2$ and estimate (4.17) simplify the equation (6.11) to

$$\|\nabla u_h^1\|^2 + k^\alpha \|\Delta_h^M u_h^1\|^2 \lesssim k^\alpha \|f^1\|^2 + k^\alpha k^2 \|\nabla u_h^0\|^2 + \|\nabla u_h^0\|^2. \tag{6.12}$$

For sufficiently small k as in Lemma 6.1, we deduce

$$\|\nabla u_h^1\|^2 + k^\alpha \|\Delta_h^M u_h^1\|^2 \lesssim \|f^1\|^2 + \|\nabla u_h^0\|^2 \lesssim \|f^1\|^2 + \|\nabla u_0\|^2. \tag{6.13}$$

Consider the scheme $(\mathcal{E}^\alpha)$ for $n \geq 2$ with definitions of discrete Laplacian operators (4.13) and (4.14)

$$\begin{aligned}(\mathbb{D}_t^\alpha u_h^n, v_h) + \left(-\Delta_h^M u_h^n, v_h\right) = (f^n, v_h) + \sum_{j=1}^{n-1} w_{nj} \left(-\Delta_h^{b_2} u_h^j, v_h\right) \\ + \sum_{j=1}^{n-1} w_{nj} \left(\nabla \cdot (b_1(x, t_n, t_j) u_h^j), v_h\right) \\ + \sum_{j=1}^{n-1} w_{nj} \left(b_0(x, t_n, t_j) u_h^j, v_h\right).\end{aligned} \tag{6.14}$$

Take $v_h = -\Delta_h^M u_h^n$ in (6.14) and apply (H2), (H3) along with Cauchy-Schwarz and Young's inequality to get

$$\begin{aligned}(\mathbb{D}_t^\alpha \nabla u_h^n, \nabla u_h^n) + \|\Delta_h^M u_h^n\|^2 \lesssim \|f^n\|^2 + \sum_{j=1}^{n-1} w_{nj} \|\Delta_h^{b_2} u_h^j\|^2 \\ + \sum_{j=1}^{n-1} w_{nj} \|\nabla u_h^j\|^2 + \sum_{j=1}^{n-1} w_{nj} \|u_h^j\|^2.\end{aligned} \tag{6.15}$$

By using the estimate (4.17) and the identity $(\mathbb{D}_t^\alpha u_h^n, u_h^n) \geq \frac{1}{2}\mathbb{D}_t^\alpha \|u_h^n\|^2$, the equation (6.15) is converted into

$$\mathbb{D}_t^\alpha \|\nabla u_h^n\|^2 + \|\Delta_h^M u_h^n\|^2 \lesssim \left(\|f^n\|^2 + \sum_{j=1}^{n-1} w_{nj} \|\nabla u_h^j\|^2\right). \tag{6.16}$$

Further, proceed as we prove estimate (6.1) to complete the proof. □

To show the existence of the fully discrete solution u_h^n $(n \geq 1)$ of the problem $(\mathcal{E}^\alpha)$, the following variant of Bröuwer fixed point theorem is used.

Theorem 6.3. *[15] Let H be finite dimensional Hilbert space. Let $G : H \to H$ be a continuous map such that $(G(w), w) > 0$ for all w in H with $\|w\| = r,\ r > 0$ then there exists a $\tilde{w}$ in H such that $G(\tilde{w}) = 0$ and $\|\tilde{w}\| \leq r$.*

Theorem 6.4. *Suppose that hypotheses (H1), (H2), and (H3) hold. Then there exists a unique solution u_h^n $(n \geq 1)$ to the problem $(\mathcal{E}^\alpha)$.*

Proof. **(Existence)** Take $n = 1$ in the scheme $(\mathcal{E}^\alpha)$ and apply the definition of $\mathbb{D}_t^\alpha u_h^1$ (3.5) with $\alpha_0 = k^\alpha \Gamma(2-\alpha)$ to obtain

$$\left(u_h^1 - u_h^0, v_h\right) + \alpha_0 \left(M\left(x, \|\nabla u_h^1\|^2\right) \nabla u_h^1, \nabla v_h\right) = \alpha_0 \left(f^1, v_h\right) + \alpha_0 k B\left(t_1, t_0, u_h^0, v_h\right). \tag{6.17}$$

In the view of (6.17) we define a map $G : X_h \to X_h$ by

$$\begin{aligned}\left(G\left(u_h^1\right), v_h\right) = \left(u_h^1, v_h\right) - \left(u_h^0, v_h\right) + \alpha_0 \left(M\left(x, \|\nabla u_h^1\|^2\right) \nabla u_h^1, \nabla v_h\right) \\ - \alpha_0 \left(f^1, v_h\right) - \alpha_0 k B\left(t_1, t_0, u_h^0, v_h\right).\end{aligned} \tag{6.18}$$

Then using (H2), (H3), and Cauchy-Schwarz inequality and Poincaré inequality with Poincaré constant C_p, we have

$$\begin{aligned}\left(G\left(u_h^1\right), u_h^1\right) &\geq \|u_h^1\|^2 - \|u_h^0\|\|u_h^1\| - \alpha_0\|f^1\|\|u_h^1\| + \alpha_0 m_0 \|\nabla u_h^1\|^2 \\ &\quad - \alpha_0 k B_0 \|\nabla u_h^0\|\|\nabla u_h^1\| \\ &\geq \|u_h^1\|\left(\|u_h^1\| - \|u_h^0\| - \alpha_0\|f^1\|\right) + \alpha_0\|\nabla u_h^1\|\left(m_0\|\nabla u_h^1\| - kB_0\|\nabla u_h^0\|\right) \\ &\geq \|u_h^1\|\left(\|u_h^1\| - \|u_h^0\| - \alpha_0\|f_h^1\|\right) \\ &\quad + \frac{\alpha_0 m_0}{C_p}\|\nabla u_h^1\|\left(\|u_h^1\| - \frac{kB_0C_p}{m_0}\|\nabla u_h^0\|\right).\end{aligned} \tag{6.19}$$

Thus, for $\|u_h^1\| > \|u_h^0\| + \alpha_0\|f^1\| + \frac{kB_0C_p}{m_0}\|\nabla u_h^0\|$ one have $\left(G\left(u_h^1\right), u_h^1\right) > 0$ and the map G defined by (6.18) is continuous as a consequence of continuity of M and B. Hence existence of u_h^1 follows by Theorem 6.3 immediately.

(Uniqueness) Suppose that X_h^1 and Y_h^1 are solutions of the scheme $(\mathcal{E}^\alpha)$ for $n = 1$, then $Z_h^1 = X_h^1 - Y_h^1$ satisfies the following equation for all v_h in X_h

$$\begin{aligned}\left(\mathbb{D}_t^\alpha Z_h^1, v_h\right) + \left(M\left(x, \|\nabla X_h^1\|^2\right) \nabla Z_h^1, \nabla v_h\right) \\ = \left(\left[M\left(x, \|\nabla Y_h^1\|^2\right) - M\left(x, \|\nabla X_h^1\|^2\right)\right] \nabla Y_h^1, \nabla v_h\right).\end{aligned} \tag{6.20}$$

Put $v_h = Z_h^1$ in (6.20) and using (H2) we get

$$\frac{1}{2}\mathbb{D}_t^\alpha \|Z_h^1\|^2 + m_0\|\nabla Z_h^1\|^2 \leq L_M \|\nabla Z_h^1\|\left(\|\nabla X_h^1\| + \|\nabla Y_h^1\|\right)\left(\nabla Y_h^1, \nabla Z_h^1\right)$$

Cauchy-Schwarz inequality and a priori bound (6.9) yield

$$\|Z_h^1\|^2 + k^\alpha \left(2m_0 - 4L_M K^2\right) \|\nabla Z_h^1\|^2 \leq 0.$$

At last, employ (H2) to obtain $\|Z_h^1\| = \|\nabla Z_h^1\| = 0$ that concludes the uniqueness of solution for $n = 1$ in the scheme $(\mathcal{E}^\alpha)$.

For $n \geq 2$, the numerical scheme $(\mathcal{E}^\alpha)$ is linear with a positive definite coefficient matrix as a result we get the existence and uniqueness of the solution u_h^n $(n \geq 2)$ for the problem $(\mathcal{E}^\alpha)$. □

To derive the convergence rate of developed numerical scheme $(\mathcal{E}^\alpha)$, first we discuss approximation properties of L1 scheme (3.5), linearization technique (3.6), and quadrature error (3.7).

Lemma 6.5. *[22] If $u \in C^2([0,T]; L^2(\Omega))$ then truncation error $\mathbb{Q}^n$ defined in* (3.5) *satisfies*

$$\|\mathbb{Q}^n\| \lesssim k^{2-\alpha} \text{ for } n \ \geq \ 1. \tag{6.21}$$

Lemma 6.6. *For any function $u \in C^2[0,T]$, the linearization error $(u^n - \bar{u}^{n-1}) := \bar{\mathbb{L}}^{n-1}$ defined in* (3.6) *undergoes*

$$|\bar{\mathbb{L}}^{n-1}| \lesssim k^2 \text{ for } n \geq 2. \tag{6.22}$$

Proof. Apply the Taylor's series expansion of u^n around u^{n-1} and u^{n-2} to obtain

$$|\bar{\mathbb{L}}^{n-1}| \lesssim k^2(u_{tt}(\xi_1) + u_{tt}(\xi_2)) \text{ for some } \xi_1 \ \in \ (t_{n-1}, t_n) \text{ and for some } \xi_2 \ \in \ (t_{n-2}, t_n). \tag{6.23}$$

As $u \in C^2[0,T]$ that implies the result (6.22). □

Lemma 6.7. *[29] If $u \in C^4[0,T]$ then quadrature error defined by* (3.7) *has the following error estimate*

$$|q^n(u)| \lesssim k^2 \text{ for } n \ \geq \ 1. \tag{6.24}$$

We prove the convergence estimate of the proposed numerical scheme $(\mathcal{E}^\alpha)$ by assuming that the solution u of the problem $(\mathcal{P}^\alpha)$ satisfies additional regularity used in Lemma 6.5 to Lemma 6.7, i.e., $u \in C^4([0,T]; H^2(\Omega) \cap H_0^1(\Omega))$ [1, 22].

6.1 Proof of the Theorem 3.4

Proof. First we prove the error estimate for the case $n = 1$. Substitute $u_h^1 = W^1 - \theta^1$ for $n = 1$ in the scheme $(\mathcal{E}^\alpha)$ and using weak formulation $(\mathcal{W}^\alpha)$ along with modified

Ritz-Volterra projection operator W at t_1 to get

$$\begin{aligned}
&\left(\mathbb{D}_t^\alpha \theta^1, v_h\right) + \left(M\left(x, \|\nabla u_h^1\|^2\right)\nabla\theta^1, \nabla v_h\right) \\
&\qquad = \left(\mathbb{D}_t^\alpha W^1 - {}^C D_{t_1}^\alpha u, v_h\right) - kB\left(t_1, t_0, W^0, v_h\right) + \int_0^{t_1} B(t_1, s, W(s), v_h) ds \\
&\qquad + \left(\left[M\left(x, \|\nabla u_h^1\|^2\right) - M\left(x, \|\nabla u^1\|^2\right)\right]\nabla W^1, \nabla v_h\right) \\
&\qquad + kB\left(t_1, t_0, \theta^0, v_h\right).
\end{aligned} \tag{6.25}$$

Set $v_h = \theta^1$ in (6.25) with $\theta^0 = 0$ and using (H2), (H3) together with Cauchy-Schwarz inequality and Young's inequality to obtain

$$\begin{aligned}
(1 - k^\alpha\Gamma(2-\alpha))\|\theta^1\|^2 &+ k^\alpha\left(m_0 - 4L_M K^2\right)\|\nabla\theta^1\|^2 \\
&\lesssim k^\alpha\|\mathbb{Q}^1\|^2 + k^\alpha\|\,{}^C D_{t_1}^\alpha \rho\|^2 + k^\alpha\|\nabla q^1(W)\|^2 + k^\alpha\|\nabla\rho^1\|^2.
\end{aligned} \tag{6.26}$$

For sufficiently small $k^\alpha < \frac{1}{\Gamma(2-\alpha)}$, we apply (H2) and the approximation properties (6.21), (5.5), (5.9), and (6.24) to conclude

$$\|\theta^1\|^2 + k^\alpha\|\nabla\theta^1\|^2 \lesssim \left(k^{2-\alpha} + h\right)^2. \tag{6.27}$$

Now we derive the error estimate for $n \geq 2$, for that take $u_h^n = W^n - \theta^n$ in the scheme $(\mathcal{E}^\alpha)$

$$\begin{aligned}
(\mathbb{D}_t^\alpha\theta^n, v_h) &+ \left(M\left(x, \|\nabla\bar{u}_h^{n-1}\|^2\right)\nabla\theta^n, \nabla v_h\right) \\
&= (\mathbb{Q}^n, v_h) - \left({}^C D_{t_n}^\alpha \rho, v_h\right) \\
&+ \left((M(x, \|\nabla\bar{u}_h^{n-1}\|^2) - M(x, \|\nabla u^n\|^2))\nabla W^n, \nabla v_h\right) \\
&+ (\nabla q^n(W), \nabla v_h) + \sum_{j=1}^{n-1} w_{nj} B(t_n, t_j, \theta^j, v_h).
\end{aligned} \tag{6.28}$$

Put $v_h = \theta^n$ in (6.28) it follows

$$\begin{aligned}
\mathbb{D}_t^\alpha\|\theta^n\|^2 + \|\nabla\theta^n\|^2 &\lesssim \|\mathbb{Q}^n\|^2 + \|\,{}^C D_{t_n}^\alpha\rho\|^2 + \|\theta^n\|^2 + \|\nabla\bar{\rho}^{n-1}\|^2 + \|\nabla\bar{\theta}^{n-1}\|^2 \\
&+ \|\nabla\bar{\mathbb{L}}^{n-1}\|^2 + \|\nabla q^n(W)\|^2 + \sum_{j=1}^{n-1} w_{nj}\|\nabla\theta^j\|^2.
\end{aligned} \tag{6.29}$$

Employ the approximation properties (6.21), (5.5), (5.9), (6.22), and (6.24) to deduce

$$\begin{aligned}
\mathbb{D}_t^\alpha\|\theta^n\|^2 + \|\nabla\theta^n\|^2 &\lesssim \left(k^{2-\alpha} + h^2 + h + k^2 + k^2\right)^2 + \|\theta^n\|^2 \\
&+ \left(\|\nabla\theta^{n-1}\|^2 + \|\nabla\theta^{n-2}\|^2 + \sum_{j=1}^{n-1} k_1\|\nabla\theta^j\|^2\right).
\end{aligned} \tag{6.30}$$

Now follow the similar arguments as we prove estimate (6.1) to conclude the result (3.11). $\square$

7 Linearized L2-1$_\sigma$ Galerkin scheme

In this section, we show that proposed numerical scheme ($\mathcal{F}^\alpha$) achieve the second order convergence in the time direction.

Lemma 7.1. *Under the hypotheses (H1), (H2), and (H3) the solution u_h^n $(n \geq 1)$ of the scheme ($\mathcal{F}^\alpha$) satisfy the following a priori bound*

$$\max_{1\leq m\leq N} \|u_h^m\|^2 + k^\alpha \sum_{n=1}^{N} \tilde{p}_{N-n}^{(N)} \|\nabla u_h^n\|^2 \lesssim \max_{1\leq n\leq N} \|f^{n-\frac{\alpha}{2}}\|^2 + \|\nabla u_0\|^2. \tag{7.1}$$

Proof. For $n = 1$ the scheme ($\mathcal{F}^\alpha$) is

$$\begin{aligned}\left(\tilde{\mathbb{D}}_{t_{1-\sigma}}^\alpha u_h^1, v_h\right) + \left(M\left(x, \|\nabla \hat{u}_h^{1,\sigma}\|^2\right)\nabla \hat{u}_h^{1,\sigma}, \nabla v_h\right) &= (1-\sigma)\, kB\left(t_{1-\sigma}, t_0, u_h^0, v_h\right) \\ &\quad + \left(f^{1-\sigma}, v_h\right).\end{aligned} \tag{7.2}$$

Substitute $v_h = u_h^1$ in (7.2) to get

$$\begin{aligned}\frac{k^{-\alpha}(1-\sigma)^{1-\alpha}}{\Gamma(2-\alpha)}&(u_h^1 - u_h^0, u_h^1) + (1-\sigma)\left(M\left(x, \|\nabla \hat{u}_h^{1,\sigma}\|^2\right)\nabla u_h^1, \nabla u_h^1\right) \\ &= -\sigma\left(M\left(x, \|\nabla \hat{u}_h^{1,\sigma}\|^2\right)\nabla u_h^0, \nabla u_h^1\right) \\ &\quad + (1-\sigma)\, kB\left(t_{1-\sigma}, t_0, u_h^0, u_h^1\right) + \left(f^{1-\sigma}, u_h^1\right).\end{aligned} \tag{7.3}$$

Simplification of (7.3) using (H2) and (H3) with $\alpha_0 = k^\alpha \Gamma(2-\alpha)$ and $\tilde{a}_0 = (1-\sigma)^{1-\alpha}$ yields

$$\left(1 - \frac{\alpha_0}{\tilde{a}_0}\right)\|u_h^1\|^2 + k^\alpha \|\nabla u_h^1\|^2 \lesssim k^\alpha \|\nabla u_h^0\|^2 + \|u_h^0\|^2 + k^{2+\alpha}\|\nabla u_h^0\|^2 + k^\alpha \|f^{1-\sigma}\|^2. \tag{7.4}$$

Take sufficiently small k to conclude

$$\|u_h^1\|^2 + k^\alpha \|\nabla u_h^1\|^2 \lesssim \|\nabla u_h^0\|^2 + \|f^{1-\sigma}\|^2 \lesssim \|\nabla u_0\|^2 + \|f^{1-\sigma}\|^2. \tag{7.5}$$

For u_h^n $(n \geq 2)$ in the scheme ($\mathcal{F}^\alpha$) to have

$$\begin{aligned}\left(\tilde{\mathbb{D}}_{t_{n-\sigma}}^\alpha u_h^n, v_h\right) + \left(M\left(x, \|\nabla \bar{u}_h^{n-1,\sigma}\|^2\right)\nabla \hat{u}_h^{n,\sigma}, \nabla v_h\right) &= \sum_{j=1}^{n-1} \tilde{w}_{nj} B\left(t_{n-\sigma}, t_j, u_h^j, v_h\right) \\ &\quad + \left(f^{n-\sigma}, v_h\right).\end{aligned} \tag{7.6}$$

Put $v_h = u_h^n$ in (7.6) to obtain

$$\begin{aligned}\left(\tilde{\mathbb{D}}_{t_{n-\sigma}}^\alpha u_h^n, u_h^n\right) &+ (1-\sigma)\left(M\left(x, \|\nabla \bar{u}_h^{n-1,\sigma}\|^2\right)\nabla u_h^n, \nabla u_h^n\right) \\ &= \sum_{j=1}^{n-1} \tilde{w}_{nj} B\left(t_{n-\sigma}, t_j, u_h^j, u_h^n\right) + \left(f^{n-\sigma}, u_h^n\right) \\ &\quad - \sigma\left(M\left(x, \|\nabla \bar{u}_h^{n-1,\sigma}\|^2\right)\nabla u_h^{n-1}, \nabla u_h^n\right).\end{aligned} \tag{7.7}$$

We invoke the identity $\left(\tilde{\mathbb{D}}^{\alpha}_{t_{n-\sigma}} u_h^n, u_h^n\right) \geq \frac{1}{2}\tilde{\mathbb{D}}^{\alpha}_{t_{n-\sigma}}\|u_h^n\|^2$ [1] and apply (H2), (H3) along with Cauchy-Schwarz and Young's inequality to reach at

$$\tilde{\mathbb{D}}^{\alpha}_{t_{n-\sigma}}\|u_h^n\|^2 + \|\nabla u_h^n\|^2 \lesssim \|f^{n-\frac{\alpha}{2}}\|^2 + \sum_{j=1}^{n-1} \tilde{w}_{nj}\|\nabla u_h^j\|^2 + \|\nabla u_h^{n-1}\|^2 + \|u_h^n\|^2. \tag{7.8}$$

We follow the similar arguments as we prove estimate (6.1) to obtain (7.1). ☐

Lemma 7.2. *Under the assumptions (H1), (H2), and (H3) the solution u_h^n $(n \geq 1)$ of the scheme $(\mathcal{F}^{\alpha})$ satisfy the following a priori bound*

$$\max_{1\leq m\leq N}\|\nabla u_h^m\|^2 + k^{\alpha}\sum_{n=1}^{N}\tilde{p}^{(N)}_{N-n}\|\Delta_h^M u_h^n\|^2 \lesssim \max_{1\leq n\leq N}\|f^{n-\frac{\alpha}{2}}\|^2 + \|\nabla u_0\|^2. \tag{7.9}$$

Proof. We combine the idea of Lemma 6.2 and Lemma 7.1 to prove the result (7.9). ☐

Theorem 7.3. *Suppose that (H1), (H2), and (H3) hold. Then there exists a unique solution to the problem $(\mathcal{F}^{\alpha})$.*

Proof. For the case $n \geq 2$, the scheme $(\mathcal{F}^{\alpha})$ is linear having positive definite coefficient matrix. Thus existence and uniqueness in this case follow immediately. For the case $n = 1$ in the scheme $(\mathcal{F}^{\alpha})$, the equation is nonlinear so we again use Bröuwer fixed point Theorem 6.3. Consider the case for $n = 1$ in the problem $(\mathcal{F}^{\alpha})$ with $\alpha_0 = k^{\alpha}\Gamma(2-\alpha)$ and $\tilde{a}_0 = (1-\sigma)^{1-\alpha}$

$$\begin{aligned}\left(u_h^1 - u_h^0, v_h\right) &+ \frac{\alpha_0}{\tilde{a}_0}\left(M\left(x, \|\nabla\hat{u}_h^{1,\sigma}\|^2\right)\nabla\hat{u}_h^{1,\sigma}, \nabla v_h\right)\\ &= \frac{\alpha_0}{\tilde{a}_0}\left(f^{1-\sigma}, v_h\right) + \frac{\alpha_0}{\tilde{a}_0}kB\left(t_{1-\sigma}, t_0, u_h^0, v_h\right).\end{aligned} \tag{7.10}$$

Multiply the equation (7.10) by $(1-\sigma)$ to obtain

$$\begin{aligned}\left(\hat{u}_h^{1,\sigma}, v_h\right) - \left(u_h^0, v_h\right) &+ (1-\sigma)\frac{\alpha_0}{\tilde{a}_0}\left(M\left(x, \|\nabla\hat{u}_h^{1,\sigma}\|^2\right)\nabla\hat{u}_h^{1,\sigma}, \nabla v_h\right)\\ &= (1-\sigma)\frac{\alpha_0}{\tilde{a}_0}\left(f^{1-\sigma}, v_h\right) + (1-\sigma)\frac{\alpha_0}{\tilde{a}_0}kB\left(t_{1-\sigma}, t_0, u_h^0, v_h\right).\end{aligned} \tag{7.11}$$

Further, proceeding analogously to the proof of Theorem 6.4 we conclude the existence of $\hat{u}_h^{1,\sigma}$. Hence existence of u_h^1 follows. ☐

Now, we derive the convergence estimate for the numerical scheme $(\mathcal{F}^{\alpha})$. This estimate is proved with the help of the following lemmas.

Lemma 7.4. *[1] If $u \in C^3([0,T]; L^2(\Omega))$ then truncation error $\tilde{\mathbb{Q}}^{n-\sigma}$ defined in* (3.12) *satisfies*

$$\|\tilde{\mathbb{Q}}^{n-\sigma}\| \lesssim k^{3-\alpha} \text{ for } n \geq 1. \tag{7.12}$$

Lemma 7.5. *For any function $u \in C^2[0,T]$, the linearization error $(u^n - \bar{u}^{n-1,\sigma}) := \bar{\mathbb{L}}^{n-1,\sigma}$ defined in* (3.15) *and $(u^n - \hat{u}^{n,\sigma}) := \hat{\mathbb{L}}^{n,\sigma}$ defined in* (3.16) *undergo*

$$|\bar{\mathbb{L}}^{n-1,\sigma}| \lesssim k^2 \text{ for } n \geq 2. \tag{7.13}$$

and

$$|\hat{\mathbb{L}}^{n,\sigma}| \lesssim k^2 \text{ for } n \geq 1. \tag{7.14}$$

Proof. This lemma is proved by an application of Taylor's series expansion as we have proved Lemma (6.6). □

Lemma 7.6. *[29] If $u \in C^4[0,T]$ then quadrature error defined by* (3.17) *has the following error estimate*

$$|\tilde{q}^{n-\sigma}(u)| \lesssim k^2 \text{ for } n \geq 1. \tag{7.15}$$

7.1 Proof of the Theorem 3.6

Proof. Take $u_h^1 = W^1 - \theta^1$ with $\theta^0 = 0$ in the scheme $(\mathcal{F}^\alpha)$ for $n = 1$, we have the following error equation for θ^1

$$\begin{aligned} &\left(\tilde{\mathbb{D}}^\alpha_{t_{1-\sigma}}\theta^1, v_h\right) + \left(M\left(x, \|\nabla \hat{u}_h^{1,\sigma}\|^2\right)\nabla\hat{\theta}^{1,\sigma}, \nabla v_h\right) \\ &= \left(\tilde{\mathbb{Q}}^{1-\sigma}, v_h\right) - \left({}^C D^\alpha_{t_{1-\sigma}}\rho, v_h\right) + \left(M(x, \|\nabla u^{1-\sigma}\|^2)(\nabla\hat{W}^{1,\sigma} - \nabla W^{1-\sigma}), \nabla v_h\right) \\ &+ \left((M(x, \|\nabla\hat{u}_h^{1,\sigma}\|^2) - M(x, \|\nabla u^{1-\sigma}\|^2))\nabla\hat{W}^{1,\sigma}, \nabla v_h\right) + \left(\nabla\tilde{q}^{1-\sigma}(W), \nabla v_h\right). \end{aligned} \tag{7.16}$$

Set $v_h = \theta^1$ in (7.16) with $\alpha_0 = k^\alpha\Gamma(2-\alpha)$ and $\tilde{a}_0 = (1-\sigma)^{1-\alpha}$ to have

$$\begin{aligned} &\left(1 - \frac{\alpha_0}{\tilde{a}_0}\right)\|\theta^1\|^2 + k^\alpha\left(m_0 - 4L_M K^2\right)\|\nabla\theta^1\|^2 \\ &\qquad \lesssim k^\alpha\|\tilde{\mathbb{Q}}^{1-\sigma}\|^2 + k^\alpha\|\,{}^C D^\alpha_{t_{1-\sigma}}\rho\|^2 + k^\alpha\|\nabla\hat{\mathbb{L}}^{1,\sigma}\|^2 + k^\alpha\|\nabla\hat{\rho}^{1,\sigma}\|^2 \\ &\qquad + k^\alpha\|\nabla\tilde{q}^{1-\sigma}(W)\|^2. \end{aligned} \tag{7.17}$$

For small value of k with hypothesis (H2) and approximation properties (7.12), (5.5), (5.9), (7.14), and (7.15), we deduce

$$\|\theta^1\|^2 + k^\alpha\|\nabla\theta^1\|^2 \lesssim \left(k^{3-\alpha} + h^2 + k^2 + h + k^2\right)^2 \lesssim \left(k^2 + h\right)^2. \tag{7.18}$$

For $n \geq 2$, substitute $u_h^n = W^n - \theta^n$ in the scheme $(\mathcal{F}^\alpha)$, then θ^n satisfies

$$\begin{aligned} &\left(\tilde{\mathbb{D}}^\alpha_{t_{n-\sigma}}\theta^n, v_h\right) + \left(M\left(x, \|\nabla\bar{u}_h^{n-1,\sigma}\|^2\right)\nabla\hat{\theta}^{n,\sigma}, \nabla v_h\right) \\ &= \left(\tilde{\mathbb{Q}}^{n-\sigma}, v_h\right) - \left({}^C D^\alpha_{t_{n-\sigma}}\rho, v_h\right) + \left(M(x, \|\nabla\bar{u}_h^{n-1,\sigma}\|^2)(\nabla\hat{W}^{n,\sigma} - \nabla W^{n-\sigma}), \nabla v_h\right) \\ &+ \left((M(x, \|\nabla\bar{u}_h^{n-1,\sigma}\|^2) - M(x, \|\nabla u^{n-\sigma}\|^2))\nabla W^{n-\sigma}, \nabla v_h\right) \\ &+ \left(\nabla\tilde{q}^{n-\sigma}(W), \nabla v_h\right) + \sum_{j=1}^{n-1}\tilde{w}_{nj}B(t_{n-\sigma}, t_j, \theta^j, v_h). \end{aligned} \tag{7.19}$$

Put $v_h = \theta^n$ in (7.19) it follows

$$\begin{aligned}\tilde{\mathbb{D}}^{\alpha}_{t_{n-\frac{\alpha}{2}}}\|\theta^n\|^2 + \|\nabla\theta^n\|^2 \lesssim \|\tilde{\mathbb{Q}}^{n-\sigma}\|^2 + \|\ {}^{C}D^{\alpha}_{t_{n-\sigma}}\rho\|^2 + \|\nabla\hat{\mathbb{L}}^{n,\sigma}\|^2 + \|\nabla\bar{\rho}^{n-1,\sigma}\|^2 \\ + \|\nabla\bar{\mathbb{L}}^{n-1,\sigma}\|^2 + \|\nabla\tilde{q}^{n-\sigma}(W)\|^2 \\ + \|\nabla\theta^{n-1}\|^2 + \|\nabla\theta^{n-2}\|^2 + \sum_{j=1}^{n-1}\tilde{w}_{nj}\|\nabla\theta^j\|^2.\end{aligned} \tag{7.20}$$

Further, apply the approximation properties (7.12), (5.5), (5.9), (7.13), (7.14), and (7.15) to arrive at

$$\begin{aligned}\tilde{\mathbb{D}}^{\alpha}_{t_{n-\frac{\alpha}{2}}}\|\theta^n\|^2 + \|\nabla\theta^n\|^2 \lesssim \left(\|\theta^n\|^2 + \|\nabla\theta^{n-1}\|^2 + \|\nabla\theta^{n-2}\|^2 + \sum_{j=1}^{n-1}k_1\|\nabla\theta^j\|^2\right) \\ + \left(k^{3-\alpha} + h^2 + k^2 + h + k^2 + k^2 + h^2 + h^2\right)^2.\end{aligned} \tag{7.21}$$

Now, follow the similar arguments as in Theorem 3.4 and Theorem 7.1 to obtain

$$\|\theta^m\|^2 + k^{\alpha}\sum_{n=1}^{m}\tilde{p}^{(m)}_{m-n}\|\nabla\theta^n\|^2 \lesssim \left(h + k^2\right)^2.$$

Finally, triangle inequality and estimate (5.5) complete the proof. □

8 Numerical results

In this section, we implement the theoretical results obtained from fully discrete formulations $(\mathcal{E}^{\alpha})$ and $(\mathcal{F}^{\alpha})$ for the problem $(\mathcal{P}^{\alpha})$. For the space discretization linear hat basis functions say $\{\psi_1, \psi_2, \ldots, \psi_J\}$ for J dimensional subspace X_h of $H^1_0(\Omega)$ are used, then numerical solution u^n_h $(n \geq 1)$ for the considered $(\mathcal{P}^{\alpha})$ at any time t_n in $[0, T]$ is written as

$$u^n_h = \sum_{i=1}^{J}\alpha^n_i\psi_i, \tag{8.1}$$

where $\alpha^n = (\alpha^n_1, \alpha^n_2, \alpha^n_3, \ldots, \alpha^n_J)$ is to be determined. Further, denote the error estimates that we have proved in Theorems 3.4 and 3.6 by

$$\textbf{Error-1} = \max_{1\leq n\leq N}\|u(t_n) - u^n_h\| + \left(k^{\alpha}\sum_{n=1}^{N}p_{N-n}\|\nabla u(t_n) - \nabla u^n_h\|^2\right)^{1/2},$$

and

$$\textbf{Error-2} = \max_{1\leq n\leq N}\|u(t_n) - u^n_h\| + \left(k^{\alpha}\sum_{n=1}^{N}\tilde{p}^{(N)}_{N-n}\|\nabla u(t_n) - \nabla u^n_h\|^2\right)^{1/2},$$

respectively.

Example 8.1. *We consider the problem* $(\mathcal{P}^\alpha)$ *such that* $(x, y, t) \in \Omega \times [0, T]$ *where* $\Omega = [0, 1] \times [0, 1]$ *and* $T = 1$ *with following data*

1. $M(x, y, \|\nabla u\|^2) = a(x, y) + b(x, y)\|\nabla u\|^2$ *with* $a(x, y) = x^2 + y^2 + 1$ *and* $b(x, y) = xy$. *This type of diffusion coefficient* M *has been studied by medeiros et.al. for the purpose of numerical experiments in [27].*

2. *Moreover, we take the following memory operator as in [3]*
$b_2(t, s, x, y) = -e^{t-s}I;\ \ b_1(t, s, x, y) = b_0(t, s, x, y) = 0.$

3. *Source term* $f(x, t) = f_1(x, t) + f_2(x, t) + f_3(x, t)$ *with*

$$\begin{aligned}
f_1(x, t) &= \frac{2}{\Gamma(3-\alpha)} t^{2-\alpha}(x - x^2)(y - y^2), \\
f_2(x, t) &= 2(x + y - x^2 - y^2)\left(t^2x^2 + t^2y^2 + \frac{xyt^6}{45} - 2t - 2 + 2e^t\right), \\
f_3(x, t) &= t^2(2x - 1)(y - y^2)\left(2x + \frac{yt^4}{45}\right) + t^2(2y - 1)(x - x^2)\left(2y + \frac{xt^4}{45}\right).
\end{aligned}$$

Corresponding to the above data, the exact solution of the problem $(\mathcal{P}^\alpha)$ *is given by* $u = t^2(x - x^2)(y - y^2)$.

This example can be used to model the diffusion of a substance in a domain Ω, see [3].

We obtain errors and convergence rates in the space direction as well as in the time direction for different parameters h, k, and α. The convergence rate is calculated through the following log vs. log formula

$$\text{Convergence rate} = \begin{cases} \frac{\log(E(\tau,h_1)/E(\tau,h_2))}{\log(h_1/h_2)}; & \text{In space direction} \\ \frac{\log(E(\tau_1,h)/E(\tau_2,h))}{\log(\tau_1/\tau_2)}; & \text{In time direction} \end{cases}$$

where $E(\tau, h)$ denotes the error at mesh points τ and h.

Linearized L1 Galerkin FEM: This numerical scheme $(\mathcal{E}^\alpha)$ provides a convergence order of $O(h + k^{2-\alpha})$ (3.11). To observe this order of convergence numerically, we run the MATLAB code at different iterations by setting $h \simeq k^{2-\alpha}$. Here h is taken as the area of the triangle in the triangulation of domain $\Omega = [0, 1] \times [0, 1]$. For next iteration, we join the midpoint of each edge and make another triangulation as presented in Figures 1-3. In this way we collect the numerical results upto five iterations to support our theoretical estimates.

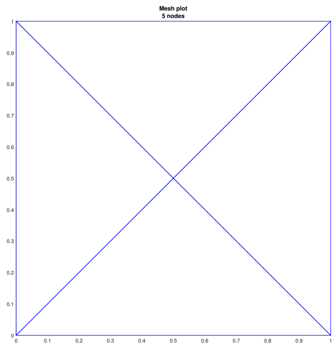

Figure 1: Iteration no. 1

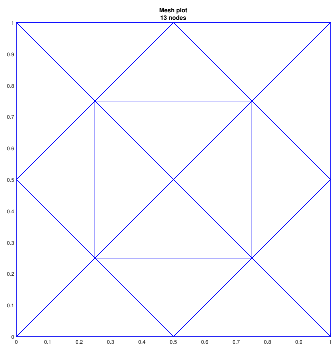

Figure 2: Iteration no. 2

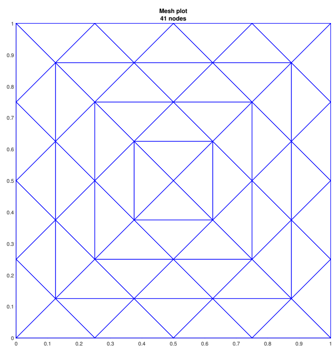

Figure 3: Iteration no. 3

From Tables 1, 2, and 3, we conclude that the convergence rate is linear in space and $(2-\alpha)$ in the time direction as predicted in Theorem 3.4. We also observe that as $\alpha \to 1$ then convergence rates is approaching to 1 in the time direction which coincide with the results established in [17] for classical diffusion case.

Table 1: *Error and Convergence Rates in space-time direction for* $\alpha = 0.25$

Iteration No.	**Error-1**	**Rate in Space**	**Rate in Time**
1	3.04e-02	-	-
2	1.50e-02	1.0175	1.7807
3	7.60e-03	0.9824	1.7192
4	3.82e-03	0.9909	1.7342
5	1.89e-03	1.0166	1.7790

Table 2: *Error and Convergence Rates in space-time direction for* $\alpha = 0.5$

Iteration No.	**Error-1**	**Rate in Space**	**Rate in Time**
1	2.63e-02	-	-
2	1.35e-02	0.9568	1.4352
3	6.53e-03	1.0521	1.5781
4	3.20e-03	1.0285	1.5428
5	1.58e-03	1.0141	1.5212

Table 3: *Error and Convergence Rates in space-time direction for* $\alpha = 0.75$

Iteration No.	**Error-1**	**Rate in Space**	**Rate in Time**
1	2.23e-02	-	-
2	1.09e-02	1.0340	1.2925
3	5.33e-03	1.0345	1.2931
4	2.65e-03	1.0087	1.2960
5	1.29e-03	1.0331	1.2914

Linearized L2-1_σ Galerkin FEM: This numerical scheme has theoretical convergence order of $O(h+k^2)$ (3.21). Here we set $h \simeq k^2$ to conclude the convergence rates in the space-time directions. Here iteration numbers have the same meaning as for L1 Galerkin FEM ($\mathcal{E}^\alpha$). From the Tables 4, 5, and 6, we deduce that the convergence rate is linear in space and quadratic in the time direction which coincide with the estimate proved in Theorem 3.6.

Table 4: *Error and Convergence Rates in space-time direction for* $\alpha = 0.25$

Iteration No.	**Error-2**	**Rate in Space**	**Rate in Time**
1	2.32e-02	-	-
2	1.10e-02	1.0679	2.1359
3	5.38e-03	1.0393	2.0787
4	2.52e-03	1.0904	2.1809
5	1.24e-03	1.0225	2.0451

Table 5: *Error and Convergence Rates in space-time direction for* $\alpha = 0.5$

Iteration No.	**Error-2**	**Rate in Space**	**Rate in Time**
1	2.42e-02	-	-
2	1.16e-02	1.0573	2.1147
3	5.70e-03	1.0329	2.0658
4	2.68e-03	1.0873	2.1746
5	1.32e-03	1.0207	2.0414

Table 6: *Error and Convergence Rates in space-time direction for* $\alpha = 0.75$

Iteration No.	**Error-2**	**Rate in Space**	**Rate in Time**
1	1.97e-02	-	-
2	9.05e-03	1.1284	2.2569
3	4.25e-03	1.0873	2.1746
4	1.94e-03	1.1335	2.2671
5	9.30e-04	1.0614	2.1228

Now, we plot the graph of an approximate solution as well as an exact solution in Figure 4 using linearized L1 Galerkin FEM.

Figure 4: Approximate solution(L.H.S) and Exact solution(R.H.S) at $T = 1$ and $\alpha = 0.5$.

9 Conclusions

In this work, we established the well-posedness of the weak formulation corresponding to the time-fractional integro-differential equations of Kirchhoff type for non-homogeneous materials. As a consequence of new Ritz-Volterra type projection operator, semi discrete error estimate in energy norm is derived. Further, to obtain the numerical solution for this class of equations, we have developed and analyzed two different kinds of efficient numerical schemes. First, we constructed a linearized L1 Galerkin FEM and derived the convergence rate of order (2-α) in the time direction. Next, to enhance the convergence order in the time direction we proposed a new linearized L2-1_σ Galerkin FEM which has quadratice rate of convergence in the time direction. Finally, numerical results revealed that theoretical error estimates are sharp.

References

[1] A. A. Alikhanov. A new difference scheme for the time-fractional diffusion equation. *Journal of Computational Physics*, 280:424–438, 2015.

[2] R. Almeida. A Grönwall inequality for a general Caputo fractional operator. *Math. Inequal. Appl.*, 20(4):1089–1105, 2017.

[3] S. Barbeiro, J. A. Ferreira, and L. Pinto. H1-second order convergent estimates for non-Fickian models. *Applied numerical mathematics*, 61(2):201–215, 2011.

[4] J. R. Cannon and Y. Lin. Non-classical H1 projection and Galerkin methods for nonlinear parabolic integro-differential equations. *Calcolo*, 25(3):187–201, 1988.

[5] J. R. Cannon and Y. Lin. A priori L^2 error estimates for finite element methods for nonlinear diffusion equations with memory. *SIAM Journal on Numerical Analysis*, 27(3):595–607, 1990.

[6] K. Diethelm and N. J. Ford. Analysis of fractional differential equations. *Journal of Mathematical Analysis and Applications*, 265(2):229–248, 2002.

[7] L. C. Evans. *Partial differential equations*, volume 19. American Mathematical Society, 2022.

[8] J. A. Ferreira and P. D. Oliveira. Memory effects and random walks in reaction-transport systems. *Applicable Analysis*, 86(1):99–118, 2007.

[9] J. Giacomoni, V. D. Rădulescu, and G. Warnault. Quasilinear parabolic problem with variable exponent: qualitative analysis and stabilization. *Communications in Contemporary Mathematics*, 20(08):1750065, 2018.

[10] Y. Giga and T. Namba. Well-posedness of Hamilton–Jacobi equations with Caputo's time-fractional derivative. *Communications in Partial Differential Equations*, 42(7):1088–1120, 2017.

[11] D. Goel and K. Sreenadh. Kirchhoff equations with Hardy–Littlewood–Sobolev critical nonlinearity. *Nonlinear Analysis*, 186:162–186, 2019.

[12] G. Gripenberg, S. O. Londen, and O. Staffans. *Volterra integral and functional equations*. Number 34. Cambridge University Press, 1990.

[13] T. Gudi. Finite element method for a nonlocal problem of Kirchhoff type. *SIAM Journal on Numerical Analysis*, 50(2):657–668, 2012.

[14] F. Huang and F. Liu. The time-fractional diffusion equation and the advection-dispersion equation. *The ANZIAM Journal*, 46(3):317–330, 2005.

[15] S. Kesavan. Topics in functional analysis and applications. *New Age International, New Delhi*, 1989.

[16] A. Kubica and M. Yamamoto. Initial-boundary value problem for fractional diffusion equations with time-dependent coefficients. *Frac. cal. appl. anal.*, 21(2):112–125, 2018.

[17] L. Kumar, S. G. Sista, and K. Sreenadh. Finite element analysis of parabolic integro-differential equations of Kirchhoff type. *Mathematical Methods in the Applied Sciences*, 43(15):9129–9150, 2020.

[18] D. Li, H. L. Liao, W. Sun, J. Wang, and J. Zhang. Analysis of L1-Galerkin FEMs for time-fractional nonlinear parabolic problems. *Communication in Computational Physics*, 24(1):86–103, 2018.

[19] L. Li and J. G. Liu. Some compactness criteria for weak solutions of time-fractional PDEs. *SIAM Journal on Mathematical Analysis*, 50(4):3963–3995, 2018.

[20] H. L. Liao, W. McLean, and J. Zhang. A discrete Grönwall inequality with applications to numerical schemes for subdiffusion problems. *SIAM Journal on Numerical Analysis*, 57(1):218–237, 2019.

[21] B. V. Limaye. *Linear functional analysis for scientists and engineers*. Springer, 2016.

[22] Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *Journal of Computational Physics*, 225(2):1533–1552, 2007.

[23] J. L. Lions. On some questions in boundary value problems of mathematical physics. In *North-Holland Mathematics Studies*, volume 30, pages 284–346. Elsevier, 1978.

[24] S. Mahata and R. K. Sinha. On the existence, uniqueness and stability results for time-fractional parabolic integro-differential equations. *Journal of Integral Equations and Applications*, 32(4):457–477, 2020.

[25] F. Mainardi. *Fractional calculus and waves in linear viscoelasticity: an introduction to mathematical models.* World Scientific, 2010.

[26] J. Manimaran, L. Shangerganesh, A. Debbouche, and V. Antonov. Numerical solutions for time-fractional cancer invasion system with nonlocal diffusion. *Frontiers in Physics*, 7:93, 2019.

[27] L. A. Medeiros, T. N. Rabello, M. A. Rincon, and M. C. C. Vieira. On perturbation of the Kirchhoff operator: analysis and numerical simulation. *Communications in Mathematical Sciences*, 10(3):751–766, 2012.

[28] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Physics Reports*, 339(1):1–77, 2000.

[29] A. K. Pani, V. Thomée, and L. B. Wahlbin. Numerical methods for hyperbolic and parabolic integro-differential equations. *The Journal of Integral Equations and Applications*, 4(4):533–584, 1992.

[30] I. Podlubny. *Fractional differential equations: An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications.* Elsevier, 1998.

[31] V. D. Rădulescu and C. Vetro. Anisotropic Navier Kirchhoff problems with convection and Laplacian dependence. *Mathematical Methods in the Applied Sciences*, 46(1):461–478, 2023.

[32] R. Shen, M. Xiang, and V. D. Rădulescu. Time-space fractional diffusion problems: Existence, decay estimates and blow-up of solutions. *Milan Journal of Mathematics*, 90(1):103–129, 2022.

[33] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25. Springer Science & Business Media, 2007.

[34] M. Xiang, V. D. Rădulescu, and B. Zhang. A critical fractional Choquard–Kirchhoff problem with magnetic field. *Communications in Contemporary Mathematics*, 21(04):1850004, 2018.

[35] M. Xiang, V. D. Rădulescu, and B. Zhang. Nonlocal Kirchhoff diffusion problems: local existence and blow-up of solutions. *Nonlinearity*, 31(7):3228, 2018.

[36] R. Zacher. Weak solutions of abstract evolutionary integro-differential equations in Hilbert spaces. *Funkcialaj Ekvacioj*, 52(1):1–18, 2009.