Finite element analysis of time-fractional integro-differential equation of Kirchhoff type for non-homogeneous materials

Lalit Kumar | Sivaji Ganesh Sista | Konijeti Sreenadh

1Department of Mathematics, Indian Institute of Technology Bombay, Mumbai, India
2Department of Mathematics, Indian Institute of Technology Delhi, New Delhi, India

In this paper, we study a time-fractional initial-boundary value problem of Kirchhoff type involving memory term for non-homogeneous materials. As a consequence of energy argument, we derive $L^\infty(0, T; H^1_0(\Omega))$ bound as well as $L^2(0, T; H^2(\Omega))$ bound on the solution of the considered problem by defining two new discrete Laplacian operators. Using these a priori bounds, existence and uniqueness of the weak solution to the considered problem are established. Further, we study semi discrete formulation of the problem by discretizing the space domain using a conforming finite element method (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 Kirchhoff type nonlinearity. Finally, we develop a new linearized L1 Galerkin FEM to obtain numerical solution of the problem under consideration. This method has a convergence rate of $O(h + k^{2-\alpha})$, where $\alpha(0 < \alpha < 1)$ is the fractional derivative exponent and $h$ and $k$ are the discretization parameters in the space and time directions, respectively. This convergence rate is further improved to second order in the time direction by proposing a novel linearized L2-1,\(\sigma\) Galerkin FEM. We conduct a numerical experiment to validate our theoretical claims.

KEYWORDS
fractional derivative, finite element method (FEM), fractional Crank-Nicolson scheme, integro-differential equation, nonlocal

MSC CLASSIFICATION
34K30, 26A33, 65R10, 60K50

1 | INTRODUCTION

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

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

with initial and boundary conditions

\[ u(x, 0) = u_0(x) \text{ in } \Omega, \]
\[ u(x, t) = 0 \text{ on } \partial \Omega \times [0, T], \]

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 [1, 2] as

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

with

\[ \mu(t) = \frac{t^{1-\alpha}} {\Gamma(1-\alpha)}. \tag{1.2} \]

Here, \( \Gamma(\cdot) \) denotes the usual gamma function. Nonlocal diffusion coefficient \( M \), initial data \( u_0 \), source term \( f \), and memory operator \( b(x, t, s) \) are known functions that will be prescribed in Section 2.

There are many physical and biological processes in which the mean-square displacement of the particle motion grows only sublinearly with time \( t \), instead of linear growth, for instance, acoustic wave propagation in viscoelastic materials [3], cancer invasion system [4], and anomalous diffusion transport [5], which cannot be described accurately by classical models having integer-order derivatives. In modelling such types of phenomena, fractional calculus is a key tool due to the nonlocal and hereditary properties of fractional derivatives [6, 7]. Therefore, the study of fractional differential equations has evolved immensely in recent years. We refer the readers to [8, 9] for more physical and biological background with applications of fractional derivatives.

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 [10] studied the following problem:

\[ \frac{\partial^2 u}{\partial t^2} - M \left( x, \int_\Omega |\nabla u(x, t)|^2 dx \right) \Delta u = f(x, t) \text{ in } \Omega \times [0, T], \tag{1.3} \]

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, [11–13]. Also, the stationary case is investigated by many authors in [14–16], for example, nonlocal perturbation of stationary Kirchhoff problem in [14] and Kirchhoff equations with magnetic field in [16].

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, for example, theory of viscoelasticity, nuclear reactor dynamics, heat conduction in materials with memory, and epidemics modelling with delays. These types of phenomena are modeled by integro-differential equations [17–21]. The above mentioned phenomena motivate us to study time-fractional integro-differentialevolution for non-homogeneous materials.

Mathematical problems involving time-fractional derivatives have been studied by many researchers, for instance, see [12, 22–24]. 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 partial differential equations. In the literature, these numerical algorithms mainly consist of finite difference in time and different approximations in space, for example, finite difference methods [25–28], spectral methods [29, 30], finite element methods [31–35], virtual element methods [36, 37], discontinuous Gelerkin methods [38, 39], and mixed finite element methods [40, 41].

Authors in [30] studied the following linear time-fractional PDE in one space direction:

\[ C D_t^\alpha u - \frac{\partial^2 u}{\partial x^2} = f(x, t) \ x \in \Lambda = (0, 1), \ t \in (0, T), \]
\[ u(x, 0) = u_0(x) \ x \in \Lambda = (0, 1), \]
\[ u(0, t) = u(1, t) = 0 \ 0 \leq t \leq T, \tag{1.4} \]
where \( C \Delta^\alpha u \) is the Caputo fractional derivative that is defined by

\[
C \Delta^\alpha u = \int_0^t \mu(t-s) \frac{\partial u}{\partial s}(s) \, ds.
\]

They developed a numerical algorithm based on the L1 scheme in time and Galerkin spectral method in space. This numerical scheme has a convergence rate of \( O(l^{1-m}+k^{2-\alpha}) \) in the energy norm for solutions in \( C^2([0, T]; H^m(\Lambda) \cap H^1_0(\Lambda)) \), where \( l \) is the degree of the polynomial space considered in the Galerkin spectral method. Author in [25] proposed a modification of the L1 scheme in the time direction and difference scheme in the space direction for some linear extension to the problem (1.4). In his work, author proved that the convergence rate is of \( O(h^2 + k^2) \) for solutions belonging to \( C^3([0, T]; C^4(\Lambda)) \). Inspired by the work of [25, 30, 33], we construct two fully discrete numerical schemes for the problem \( (P^n) \) that are based on finite difference in time (L1 scheme and Alikhanov’s L2-1\( _s \) scheme) and conforming finite element method in space. We remark 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 \( (P^n) \).

In this current research, we prove the well-posedness of the weak formulation corresponding to the problem \( (P^n) \). Due to the appearance of Kirchhoff term, we cannot apply Laplace/Fourier transformation in the problem \( (P^n) \); therefore, explicit representation of its solution in terms of Fourier expansion is not possible. To resemble this issue, we use the Galerkin method [12, 13] to show the well-posedness of the weak formulation to the problem \( (P^n) \).

For the semi discrete formulation corresponding to the problem \( (P^n) \), we discretize the space direction by using a conforming FEM and keep the time direction continuous. Well-posedness of the semi discrete formulation is established along the similar lines of the proof of well-posedness of weak formulation to the problem \( (P^n) \). We derive semi discrete error estimates by modifying the standard Ritz-Volterra projection operator [42, 43].

To obtain the numerical solution of the problem \( (P^n) \), we construct two fully discrete formulations by discretizing the semi discrete formulation in the time direction by uniform mesh. First, we develop a new linearized L1 Galerkin FEM. This method comprises of L1 type approximation [30] of the Caputo fractional derivative, linearization technique for the Kirchhoff type nonlinearity, and modified Simpson’s rule [43] 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 has convergence rate of \( O(h + k^{2-\alpha}) \) in the energy norm.

Further, we increase the accuracy of this scheme in the time direction by replacing the L1 scheme with the Alikhanov’s L2-1\( _s \) scheme [25] for the approximation of the Caputo fractional derivative. As a consequence, we propose a new linearized L2-1\( _s \) Galerkin FEM, which has a convergence rate of \( O(h + k^2) \) in the energy norm. These theoretical results are supported by conducting a numerical experiment in MATLAB software.

Key difficulties and new tools for our work are described below:

1. **New discrete Laplacian operators**: The problem \( (P^n) \) has a gradient type nonlinearity in the Kirchhoff term. Thus, we need to derive a priori bound on gradient of the weak solution to the problem \( (P^n) \). Due to the presence of Kirchhoff term and memory operator with variable coefficients, we cannot use standard discrete Laplacian operator [44]. To overcome this issue, we define two new discrete Laplacian operators (4.13) and (4.14) to obtain gradient bound on the weak solution to the problem \( (P^n) \).

2. **Modified Ritz-Volterra projection operator**: The Kirchhoff term causes some difficulties in using standard Ritz-Volterra projection operator [42] to derive semi discrete error estimates. We overcome these difficulties by introducing a modified Ritz-Volterra projection operator (5.3). This projection operator follows the best approximation properties same as that of standard Ritz-Volterra projection operator.

3. **Linearization techniques**: The fully discrete formulation of the considered problem produces a coupled system of nonlinear algebraic equations. In general, numerical schemes based on the Newton method are adopted to solve this system [45]. The Kirchhoff term leads to the highly non-sparse Jacobian of this system [45]. For solving this system, we require high computational cost and huge computer storage. We reduce these costs by developing linearization techniques (3.8) and (3.21) for the nonlinearity occurred in Kirchhoff term.

4. **Modified Simpson’s rule**: 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 (3.9) and (3.23) [43].
1.1 | Turning to the layout of this paper

In Section 2, we provide some notations, assumptions, and preliminary results that will be used throughout this work. In Section 3, we state the main results of this article. Section 4 contains the proof of well-posedness of the weak formulation to the problem \((P^w)\). 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 convergence 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,\(\sigma\) 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 \(\Omega\) with the norm:

\[
\|g\|_{L^1(\Omega)} = \int_{\Omega} |g(x)| \, dx \text{ for } g \in L^1(\Omega).
\] (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 \(L^1(\Omega)\); that is,

\[
W^{1,1}(\Omega) = \{ g \in L^1(\Omega); Dg \in L^1(\Omega) \}.
\] (2.2)

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

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

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

\[
\|g\|^2 = \int_{\Omega} |g(x)|^2 \, dx \text{ for } g \in L^2(\Omega).
\] (2.4)

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

\[
(g, h) = \int_{\Omega} g(x)h(x) \, dx \text{ for } g, h \in L^2(\Omega).
\] (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 up to order \(m\) are also in \(L^2(\Omega)\), that is,

\[
H^m(\Omega) = \{ g \in L^2(\Omega); D^\beta g \in L^2(\Omega), |\beta| \leq m \}.
\] (2.6)

where \(\beta\) is multi-index. The norm on the space \(H^m(\Omega)\) is induced by the following inner product \((\cdot, \cdot)_m\) as follows:

\[
(g, h)_m = \sum_{|\beta| \leq m} (D^\beta g, D^\beta h) \text{ for } g, h \in H^m(\Omega).
\] (2.7)

We denote \(H^m_0(\Omega), (m \in \{1, 2\})\) be the closure of \(C^\infty_c(\Omega)\) in \(H^m(\Omega)\). The space \(H^m_0(\Omega)\) can be characterized by the functions in \(H^m(\Omega)\) having zero trace on the boundary \(\partial\Omega\) [46, Section 2.7]. The dual space of \(H^m_0(\Omega)\) 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

\[
\int_0^T \|g(s)\|^2_X \, ds < \infty.
\] (2.8)
The norm on the space \( L^2(0, T; X) \) is given by
\[
\|g\|^2_{L^2(0, T; X)} = \int_0^T \|g(t)\|^2_X \, ds \quad \text{for } g \in L^2(0, T; X).
\] (2.9)

We also define a \( L^2_0(0, T; X) \) space consisting of all measurable functions \( g : [0, T] \to X \) such that
\[
\sup_{t \in (0, T)} \left( \frac{1}{\Gamma(a)} \int_0^t (t - s)^{a-1} \|g(s)\|^2_X \, ds \right) < \infty.
\] (2.10)

The norm on the space \( L^2_0(0, T; X) \) is given in [47, (4.5)] as
\[
\|g\|^2_{L^2_0(0, T; X)} = \sup_{t \in (0, T)} \left( \frac{1}{\Gamma(a)} \int_0^t (t - s)^{a-1} \|g(s)\|^2_X \, ds \right) \quad \text{for } g \in L^2_0(0, T; X).
\] (2.11)

One can observe that \( L^2_0(0, T; X) \subset L^2(0, T; X) \). The set of all measurable functions \( g : [0, T] \to X \) such that
\[
\text{ess sup}_{t \in (0, T)} \|g(t)\|_X < \infty
\] (2.12)
is denoted by \( L^\infty(0, T; X) \). The norm on this space is given by
\[
\|g\|_{L^\infty(0, T; X)} = \text{ess sup}_{t \in (0, T)} \|g(t)\|_X \quad \text{for } g \in L^\infty(0, T; X).
\] (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 \( a \leq C b \), 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 \( M : \tilde{\Omega} \times (0, \infty) \to (0, \infty) \) is a Lipschitz continuous function such that there exists a positive constant \( m_0 \), which satisfies
\[
M(x, s) \geq m_0 > 0 \quad \text{for all } (x, s) \in \tilde{\Omega} \times (0, \infty) \quad \text{and} \quad \left( m_0 - 4L_M K^2 \right) > 0,
\]
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 \( b_2 : \tilde{\Omega} \times [0, T] \times [0, T] \to \mathbb{R}^{d \times d} \) is a symmetric and positive definite matrix with entries \( \left[ b^{ij}_2(x, t, s) \right] \), \( b_1 : \tilde{\Omega} \times [0, T] \times [0, T] \to \mathbb{R}^d \) is a vector with entries \( \left[ b^i_1(x, t, s) \right] \), and \( b_0 : \tilde{\Omega} \times [0, T] \times [0, T] \to \mathbb{R} \) is a scalar function. We assume that \( b^{ij}_2, b^i_1, b_0 \) are smooth functions in all variables \( (x, t, s) \in \tilde{\Omega} \times [0, T] \times [0, T] \) 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
\[
B(t, s, u(s), v) := (b_2(x, t, s)\nabla u(s), \nabla v) + (\nabla \cdot (b_1(x, t, s)u(s)), v) + (b_0(x, t, s)u(s), v).
\] (2.14)

Using (H3) and Poincaré inequality, one can prove that there exists a positive constant \( B_0 \) such that
\[
|B(t, s, u(s), v)| \leq B_0 \|\nabla u(s)\| \|\nabla v\| \quad \forall t, s \in [0, T], \quad \text{and} \quad \forall u(s), v \in H_0^1(\Omega).
\] (2.15)
We indicate \( \ast \) as the convolution of two integrable functions \( g \) and \( h \) on \([0, T]\) such that
\[
(g \ast h)(t) = \int_0^t g(t-s)h(s)\,ds \quad \forall t \in [0, T].
\] (2.16)

**Remark 1.** Note that \( \lambda(t) \) defined by \( \lambda(t) := \frac{t^{\alpha-1}}{\Gamma(\alpha)} \) satisfies \( \mu \ast \lambda = 1 \).

**Proof.** Using the definition of \( \ast \) (2.16) between \( \mu \) and \( \lambda \), we have
\[
(\mu \ast \lambda)(t) = \int_0^t \mu(t-s)\lambda(s)\,ds.
\] (2.17)
By the definition of \( \mu \) (1.2) and \( \lambda \), we obtain
\[
(\mu \ast \lambda)(t) = \int_0^t \mu(t-s)\lambda(s)\,ds = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha}s^{\alpha-1}\,ds.
\] (2.18)
Put \( s = tz \) in Equation (2.18), we get
\[
(\mu \ast \lambda)(t) = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^1 (1-z)^{-\alpha}z^{\alpha-1}\,dz.
\] By the definition of beta function \( \mathbb{B} \) and its relation with gamma function, we obtain
\[
(\mu \ast \lambda)(t) = \frac{\mathbb{B}(1-\alpha, \alpha)}{\Gamma(\alpha)\Gamma(1-\alpha)} = \frac{\Gamma(\alpha)\Gamma(1-\alpha)}{\Gamma(\alpha)\Gamma(1-\alpha)} = 1.
\] (2.19)

**Lemma 2.1** ([48, 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
\[
\left( \frac{d}{dt} (\tilde{k} \ast v)(t), v(t) \right)_H = \frac{1}{2} \frac{d}{dt} \|v\|^2_H(t) + \frac{1}{2} \tilde{k}(t)\|v\|^2_H(t) + \frac{1}{2} \int_0^t [-\tilde{k}'(s)] \|v(t) - v(t-s)\|^2_H \,ds \text{ a.e. } t \in (0, T).
\] (2.20)

**Lemma 2.2** ([49, Theorem 8]). Let \( u, v \) be two nonnegative integrable functions on \([a, b]\) and \( g \) be a continuous function in \([a, b]\). Assume that \( v \) is nondecreasing in \([a, b]\) and \( g \) is nonnegative and nondecreasing in \([a, b]\). If
\[
u(t) \leq v(t) + g(t) \int_a^t (t-s)^{\alpha-1}u(s)\,ds \text{ for } \alpha \in (0, 1) \text{ and } \forall t \in [a, b],
\]
then
\[
u(t) \leq v(t)E_\alpha \left[g(t)\Gamma(\alpha)(t-a)^\alpha\right] \text{ for } \alpha \in (0, 1) \text{ and } \forall t \in [a, b],
\]
where \( E_\alpha(\cdot) \) is the one parameter Mittag-Leffler function [7, Section 1.2].

**Lemma 2.3** ([50]). Consider the following initial value problem:
\[
\frac{d^\alpha}{dt^\alpha} y(t) = g(t, y(t)), \; t \in (0, T), \; \alpha \in (0, 1),
\]
\[
y(0) = y_0.
\] (2.21)
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 \( g : D^* \rightarrow \mathbb{R} \) 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 (2.21).
where

\[ T^* = \begin{cases} t^*; & \text{if } t^* > \frac{1}{M^*}; \\
\min \left\{ t^*, \left( \frac{K^T(1+\alpha)}{M^*} \right)^{\frac{1}{\alpha}} \right\}; & \text{else.} \end{cases} \quad (2.22) \]

**Lemma 2.4** ([47, 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 \( W \subset L^1_{\text{loc}}(0,T;X) \) satisfies the following:

1. There exists a constant \( C_1 > 0 \) such that for all \( u \in W \),

\[ \sup_{t \in (0,T)} \left( \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \|u(s)\|_{L^2_X}^2 \, ds \right) \leq C_1. \quad (2.23) \]

2. There exists a constant \( C_2 > 0 \) such that for all \( u \in W \),

\[ \|\partial_t^\alpha u\|_{L^2(0,T;Z)} \leq C_2. \quad (2.24) \]

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

**Lemma 2.5** ([1]). Let \( \mu \) be the kernel defined in (1.2). Then there exists a sequence of kernels \( \mu_n \) in \( W^{1,1}(0,T) \) such that \( \mu_n \) is nonnegative and nonincreasing in \( (0,\infty) \). Also,

\[ \mu_n \to \mu \text{ in } L^1(0,T) \text{ as } n \to \infty, \quad (2.25) \]

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

\[ \frac{d}{dt} (\mu_n * u) \to \frac{d}{dt} (\mu * u) \text{ in } L^2(0,T;L^2(\Omega)) \text{ as } n \to \infty. \quad (2.26) \]

### 3 | MAIN RESULTS

The weak formulation corresponding to the problem \((P^u)\) is to find \( u \in Z \) such that the following equations hold for all \( v \) in \( H^1_0(\Omega) \) and a.e. \( t \) in \( (0,T) \):

\[ (\partial_t^\alpha u, v) + \left( M \left( x, \|\nabla u\|^2 \right) \nabla u, \nabla v \right) = (f, v) + \int_0^t B(t,s,u(s),v) \, ds \text{ in } \Omega \times (0,T), \]

\[ u(x,0) = u_0(x) \text{ in } \Omega, \quad (W^u) \]

where the solution space \( Z \) is defined as

\[ Z := \left\{ u; u \in L^\infty \left( 0,T;L^2(\Omega) \right) \cap L^2 \left( 0,T;H^1_0(\Omega) \right) \text{ and } \partial_t^\alpha u \in L^2 \left( 0,T;L^2(\Omega) \right) \right\}. \quad (3.1) \]

**Theorem 3.1** (Well-posedness of the weak formulation \((W^u)\)). Under Hypotheses \((H1)\), \((H2)\), and \((H3)\), the problem \((W^u)\) admits a unique solution that satisfies the following a priori bounds:

\[ \|u\|_{L^\infty(0,T;H^1_0(\Omega))} + \|u\|_{L^2(0,T;H^1_0(\Omega))} \leq \left( \|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))} \right) \quad (3.2) \]

\[ \|u\|_{L^\infty(0,T;H^1_0(\Omega))} + \|u\|_{L^2(0,T;H^1_0(\Omega))} \leq \left( \|\nabla u_0\| + \|f\|_{L^\infty(0,T;L^2(\Omega))} \right) \quad (3.3) \]

For the semi discrete formulation of the problem \((P^u)\), we discretize the domain in the space variable by a conforming FEM [44] and keep the time variable continuous. Let \( T_h \) be a shape regular (non overlapping), quasi-uniform triangulation...
of the domain \( \Omega \) and \( h \) be the discretization parameter in the space direction. We define a finite dimensional subspace \( X_h \) of \( H^1_0(\Omega) \) as

\[
X_h := \{ v_h \in C(\bar{\Omega}) : v_h|_r \text{ is a linear polynomial for all } r \in \mathcal{T}_h \text{ and } v_h = 0 \text{ on } \partial \Omega \}.
\]

The semi discrete formulation for the problem \( (P^s) \) 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 \((0, T)\):

\[
\frac{\partial^\alpha}{\partial t^\alpha} u_h + (Mu_h, v_h) + (M(x, \|\nabla u_h\|_2^2) \nabla u_h, \nabla v_h) = (f, v_h) + \int_0^t B(t, s, u_h(s), v_h)ds \text{ in } \mathcal{T}_h \times (0, T),
\]

\[
u_h(x, 0) = u_h^0 \text{ in } \mathcal{T}_h,
\]

where the initial condition \( u_h^0 \) is a suitable projection of \( u_0 \) 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 \((S^s)\)).** 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 \((S^s)\):

\[
\|u - u_h\|_{L^\infty(0,T; L^2(\Omega))} + \|u - u_h\|_{L^2(0,T; H^1(\Omega))} \lesssim h.
\]

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 \((P^s)\) for that we divide the time interval \([0, T]\) into sub intervals of uniform step size \( k \) and

\[
t_n = nk \text{ for } n = 0, 1, 2, 3, \ldots , N \text{ with } t_N = T.
\]

On this uniform mesh, we approximate the Caputo fractional derivative by the L1 scheme [30], Kirchhoff type nonlinearity by linearization, and memory term by modified Simpson’s rule as follows:

**L1 approximation scheme** [30]: In this scheme, Caputo fractional derivative is approximated at the point \( t_n \) using linear interpolation formula as follows:

\[
\frac{C^\alpha}{\Gamma(1 - \alpha)} \int_0^{t_n} \frac{1}{(t_n - s)^\alpha} \frac{du}{ds} ds = \frac{1}{\Gamma(1 - \alpha)} \sum_{j=1}^n u^j - u^{j-1} \int_{t_{j-1}}^{t_j} \frac{1}{(t_n - s)^\alpha} ds + \mathbb{Q}^n
\]

\[
= \frac{k^{-\alpha}}{\Gamma(2 - \alpha)} \sum_{j=1}^n a_{n-j} (u^j - u^{j-1}) + \mathbb{Q}^n
\]

\[
= \mathscr{D}^\alpha_t u^n + \mathbb{Q}^n,
\]

where

\[
a_i = (i + 1)^{1-a} - i^{1-a}, \quad i \geq 0,
\]

and \( u^j = u(x, t_j) \), \( \mathbb{Q}^n \) is the truncation error.

**Linearization:** For nonlinearity in the diffusion coefficient, we use the following linearized approximation of \( u \) at \( t_n \), which is given by

\[
u^n \approx 2u^{n-1} - u^{n-2} \quad \text{for } n \geq 2
\]

\[
:= \bar{u}^{n-1}.
\]

**Modified Simpson’s rule** [43]: Let \( m_1 = \lfloor k^{-1/2} \rfloor \), where \( \lfloor \cdot \rfloor \) denotes the greatest integer function. Set \( k_1 = m_1k \) and \( \bar{t}_j = jk_1 \) for \( 1 \leq j \leq n \). Let \( j_n \) be the largest even integer such that \( \bar{t}_{j_n} < t_n \) and introduce the quadrature points

\[
\bar{t}_j^n = \begin{cases} jk_1, & 0 \leq j \leq j_n, \\ \bar{t}_j^n + (j - j_n)k, & j_n \leq j \leq J_n, \end{cases}
\]
where \( \tilde{t}_a^n = t_{n-1} \). Then the quadrature error for the memory term at \( t_n \) is given by

\[
(q^n, v) = \int_0^{t_n} B(t_n, s, u(s), v) \, ds - \sum_{j=0}^{n-1} w_{nj} B(t_n, t_j, u_j, v) \quad \forall v \in H_0^1(\Omega) \quad \forall n \geq 1.
\]  

(3.9)

Here, \( w_{nj} \) are called quadrature weights such that for the function \( g(t_j) := B(t_n, t_j, u(t_j), v) \), we have

\[
\sum_{j=0}^{n-1} w_{nj} g(t_j) = \frac{k_1}{3} \sum_{j=1}^{j_n/2} \left[ g\left( \tilde{t}_{j,1}^n \right) + 4g\left( \tilde{t}_{j,2}^n \right) + g\left( \tilde{t}_{j,3}^n \right) \right] + k \sum_{j=2}^{j_n+1} \left[ g\left( \tilde{t}_{j}^n \right) + g\left( \tilde{t}_{j-1}^n \right) \right] + k g\left( t_n \right).
\]

(3.10)

On the basis of approximations (3.6), (3.8), and (3.9), we develop the following linearized L1 Galerkin FEM. **Linearized L1 Galerkin FEM:** Find \( u_h^n(n = 1, 2, 3, \ldots, N) \) in \( X_h \) with

\[
\tilde{t}_a^n = 2u_h^{n-1} - u_h^{n-2} \quad \text{for} \quad n \geq 2.
\]

(3.11)

such that the following equations hold for all \( v_h \) in \( X_h \):

For \( n \geq 2,

\[
\left( \mathcal{D}_i u_h^n, v_h \right) + \left( M \left( x, \| \nabla B_h^{a-1} \| \right)^2 \right) \nabla u_h^n, \nabla v_h \right) = \left( f^n, v_h \right) + \sum_{j=1}^{n-1} w_{nj} B\left( t_n, t_j, u_j, v_h \right).
\]

\( (\mathcal{E}_a) \)

For \( n = 1,

\[
\left( \mathcal{D}_i u_h^1, v_h \right) + \left( M \left( x, \| \nabla u_h^1 \| \right)^2 \right) \nabla u_h^1, \nabla v_h \right) = \left( f^1, v_h \right) + k B\left( t_1, t_0, u_h^0, v_h \right).
\]

with the 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 \( (\mathcal{E}_a) \), we need the following discrete kernel corresponding to the kernel \( a_j, j \geq 0 \) (3.7).

**Lemma 3.3** ([31]). Let \( p_n \) be a sequence defined by

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

(3.12)

Then \( p_n \) satisfies

\[
0 < p_n < 1 \quad \text{for} \quad n \geq 1,
\]

(3.13)

\[
\sum_{j=k}^{n} p_{n-j} a_{j-k} = 1 \quad 1 \leq k \leq n,
\]

(3.14)

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

(3.15)

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

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

(3.16)
At this point, one can see that convergence rate is of $O(k^{2-\alpha})$ in the time direction. To improve this convergence rate, a new linearized L2-1$_\sigma$ Galerkin FEM is proposed. In this method, we replace the L1 approximation of the Caputo fractional derivative with L2-1$_\sigma$ $\left(\sigma = \frac{\alpha}{2}\right)$ scheme [25] at $t_{n-\sigma}$, where $t_{n-\sigma}$ is given by

$$t_{n-\sigma} = (1 - \sigma)t_n + \sigma t_{n-1}, \quad n = 1, 2, \ldots, N. \quad (3.17)$$

For nonlinearity in the diffusion coefficient and memory term, we again use linearization technique and modified Simpson’s rule at $t_{n-\sigma}$, respectively.

**L2-1$_\sigma$ approximation scheme** [25]: In this scheme, Caputo fractional derivative is approximated at the point $t_{n-\sigma}$ as follows:

$$\frac{d^\alpha}{dt^\alpha} u_{n-\sigma} = \tilde{u}_{n-\sigma} + \tilde{Q}^{n-\sigma}, \quad (3.18)$$

where

$$\tilde{u}_{n-\sigma} = \frac{k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{n} \tilde{c}^{(n)}_{n-j} (u^j - u^{j-1}), \quad (3.19)$$

with weights $\tilde{c}^{(n)}_{n-j}$ satisfying $\tilde{c}^{(1)}_{0} = \tilde{a}_0$ for $n = 1$ and for $n \geq 2$,

$$\tilde{c}^{(n)}_{j} = \begin{cases} \tilde{a}_0 + \tilde{b}_1, & j = 0, \\ \tilde{a}_j + \tilde{b}_{j+1} - \tilde{b}_j, & 1 \leq j \leq n - 2, \\ \tilde{a}_j - \tilde{b}_j, & j = n - 1, \end{cases} \quad (3.20)$$

where

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

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

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

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

For Kirchhoff term,

$$u^{n-\sigma} \approx (2 - \sigma)u^{n-1} - (1 - \sigma)u^{n-2}, \quad \text{for } n \geq 2$$

$$:= \tilde{u}^{n-1-\sigma}. \quad (3.21)$$

For diffusion term,

$$u^{n-\sigma} \approx (1 - \sigma)u^{n} + (\sigma)u^{n-1}, \quad \text{for } n \geq 1$$

$$:= \tilde{u}^{n-\sigma}. \quad (3.22)$$

**Modified Simpson’s rule**: With a small modification in (3.9), we obtain the following quadrature error corresponding to the memory term on $[0, t_{n-\sigma}]$:

$$\left( \tilde{q}^{n-\sigma}, v \right) = \int_0^{t_{n-\sigma}} B(t_{n-\sigma}, s, u(s), v) ds - \sum_{j=0}^{n-1} \tilde{w}_{nj} B(t_{n-\sigma}, t_j, u^j, v) \quad \forall v \in H^1(\Omega), \quad \text{for } n \geq 1. \quad (3.23)$$

Here, $\tilde{w}_{nj}$ are quadrature weights such that for the function $g_{\sigma}(t_j) := B(t_{n-\sigma}, t_j, u(t_j), v)$, we have

$$\sum_{j=0}^{n} \tilde{w}_{nj} g_{\sigma}(t_j) = k \sum_{j=1}^{J_n/2} \left[ g_{\sigma}(\tilde{t}_j^0) + 4 g_{\sigma}(\tilde{t}_j^{n/2}) + g_{\sigma}(\tilde{t}_j^{n/2}) \right]$$

$$+ k \sum_{j=J_n/2}^{J_n} \left[ g_{\sigma}(\tilde{t}_j^0) + g_{\sigma}(\tilde{t}_j^{n/2}) \right] + (1 - \sigma) k g_{\sigma}(\tilde{t}_0^n). \quad (3.24)$$

By combining the approximations (3.18), (3.21), (3.22), and (3.23), we construct the following linearized L2-1$_\sigma$ Galerkin FEM.
Linearized L2-1, Galerkin FEM: Find \( u_h^n (n = 1, 2, \ldots, N) \) in \( X_h \) with

\[
\hat{u}_h^{n-1, \sigma} = (2 - \sigma) u_h^{n-1} - (1 - \sigma) u_h^{n-2} \quad \text{for} \quad n \geq 2, \tag{3.25}
\]

and

\[
\hat{u}_h^{n, \sigma} = (1 - \sigma) u_h^n + (\sigma) u_h^{n-1} \quad \text{for} \quad n \geq 1, \tag{3.26}
\]

such that the following equations hold for all \( v_h \) in \( X_h \).

For \( n \geq 2 \),

\[
\begin{align*}
\left( \overline{D}^{\alpha}_{n, \sigma}, u_h^n, v_h \right) + \left( M \left( x, \| \nabla \hat{u}_h^{n-1, \sigma} \| \right)^2 \nabla \hat{u}_h^{n, \sigma}, v_h \right) &= \sum_{j=1}^{n-1} \overline{w}_{nj} B \left( t_{n-\sigma}, t_j, u_h^j, v_h \right) \\
&\quad + \left( f^{n-\sigma}, v_h \right).
\end{align*}
\]

For \( n = 1 \),

\[
\begin{align*}
\left( \overline{D}^{\alpha}_{1, \sigma}, u_h^1, v_h \right) + \left( M \left( x, \| \nabla \hat{u}_h^{1, \sigma} \| \right)^2 \nabla \hat{u}_h^{1, \sigma}, v_h \right) &= (1 - \sigma) kB \left( t_{1-\sigma}, t_0, u_h^0, v_h \right) \\
&\quad + \left( f^{1-\sigma}, v_h \right),
\end{align*}
\]

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

Similar to the Lemma 3.3, we have the following discrete kernel corresponding to the kernel \( \tilde{c}_j^m \) (3.20).

**Lemma 3.5** ([51]). Define

\[
P^{(n)}_0 = \frac{1}{\tilde{c}_0^{(n)}}, \quad P^{(n)}_j = \frac{1}{\tilde{c}_0^{(n-j)}} \sum_{k=0}^{j-1} \left( \tilde{c}_j^{(n-k)} - \tilde{c}_j^{(n-k)} \right) P^{(n)}_k \quad \text{for} \quad 1 \leq j \leq n - 1. \tag{3.27}
\]

Then \( P^{(n)}_j \) satisfies

\[
0 < P^{(n)}_{n-j} < \frac{1}{\tilde{c}_0} \quad \text{for} \quad 1 \leq j \leq n - 1, \tag{3.28}
\]

\[
\sum_{j=k}^{n} P^{(n)}_{n-j} \tilde{c}_j^{(n-k)} = 1, \quad 1 \leq k \leq n \leq N, \tag{3.29}
\]

\[
\Gamma(2 - \alpha) \sum_{j=1}^{n} P^{(n)}_{n-j} \leq \frac{n^\alpha}{\Gamma(1 + \alpha)}, \quad 1 \leq n \leq N. \tag{3.30}
\]

**Theorem 3.6** (Convergence estimate for the numerical scheme \((P^\sigma)\)). Suppose that Hypotheses (H1), (H2), and (H3) hold. Then the fully discrete scheme \((P^\sigma)\) is conditionally convergent with the following convergence rate:

\[
\max_{1 \leq k \leq N} \left\| u(t_n) - u_h^n \right\| + \left( k^\alpha \sum_{n=1}^{N} \left\| \nabla u(t_n) - \nabla u_h^n \right\|^2 \right)^{1/2} \lesssim (h + k^2). \tag{3.31}
\]

4 | WELL-POSEDNESS OF THE WEAK FORMULATION

In this section, we prove the well-posedness of the weak formulation \((W^\sigma)\) using the Galerkin method. In this method, we reduce the weak formulation onto a finite dimensional subspace of \( H_0^1(\Omega) \). The theory of fractional differential equations ([50]) ensures the existence of Galerkin sequences of weak solutions. Using two new discrete Laplacian operators and energy argument, we derive a priori bounds on every Galerkin sequence. We make use of these a priori bounds in Aubin-Lions type compactness lemma ([31]) to prove that the Galerkin sequence converges to the weak solution of the problem \((P^\sigma)\).

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 ([52, Section 6.5]). For each fixed positive integer \( m \), consider a finite dimensional subspace \( V_m \) of
such that the following equations hold for all \( v_m \in \mathbb{V}_m \) and a.e. \( t \in (0, T) \):

\[
\left( \partial_t^\alpha u_m, v_m \right) + \left( M (x, \| \nabla u_m \|^2) \right) \nabla u_m, \nabla v_m = \langle f, v_m \rangle + \int_0^t B(t, s, u_m(s), v_m(s)) ds,
\]

\[
\tag{4.2} u_m(\cdot, 0) = \sum_{j=1}^m (u_0, \phi_j) \phi_j.
\]

By Riesz-Fischer theorem [53, Theorem 2.29],

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

Put the values of \( u_m \) and \( u_m(0) \) in (4.2); we obtain a coupled system of fractional order differential equations. 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 \( \mu * (u_m - u_m(0)) \) at \( t = 0 \) [1, Theorem 3.1]. These local solutions \( u_m(t) \) are extended to the whole interval by using the following a priori bounds.

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

\[
\left( \frac{d}{dt} (\mu * u_m), u_m \right) + \left( M (x, \| \nabla u_m \|^2) \right) \nabla u_m, \nabla u_m
\]

\[
= \left( \frac{d}{dt} (\mu * u_m(0)), u_m \right) + \langle f, u_m \rangle + \int_0^t B(t, s, u_m(s), u_m(t)) ds.
\]

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

\[
\left( \frac{d}{dt} (\mu_n * u_m), u_m \right) + \left( M (x, \| \nabla u_m \|^2) \right) \nabla u_m, \nabla u_m
\]

\[
= \left( h_{mn}, u_m \right) + \left( \frac{d}{dt} (\mu_n * u_m(0)), u_m \right) + \langle f, u_m \rangle + \int_0^t B(t, s, u_m(s), u_m(t)) ds.
\]

with

\[
\tag{4.5} h_{mn} = \frac{d}{dt} (\mu_n * (u_m - u_m(0))) - \frac{d}{dt} (\mu * (u_m - u_m(0))).
\]

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

\[
\frac{d}{dt} (\mu_n * \| u_m \|^2)(t) + \frac{\alpha}{\Gamma(1 - \alpha)} \int_0^t \frac{\| u_m(t) - u_m(t - s) \|^2}{s^{1+\alpha}} ds + \| \nabla u_m \|^2
\]

\[
\leq \| h_{mn} \|^2 + \mu_n(t) \| u_m(0) \|^2 + \| f \|^2 + \| u_m \|^2 + t \int_0^t \| \nabla u_m(s) \|^2 ds.
\]

(4.6)

By convolving Equation (4.7) with the kernel \( \lambda(t) = \frac{t^{\mu-1}}{\Gamma(\alpha)} \) and letting \( n \to \infty \) in (4.7), we have

\[
\| u_m(t) \|^2 + (\lambda * \| \nabla u_m \|^2)(t) \leq \| u_m(t) \|^2 + \langle \lambda * u_m \| f \|^2 \rangle(t)
\]

\[
+ \lambda * \left( \| u_m \|^2 + t \int_0^t \| \nabla u_m(s) \|^2 ds \right).
\]
In (4.8), we have used the fact that

\[ \lambda \frac{d}{dt} (\mu_n \ast \|u_m\|^2) (t) = \frac{d}{dt} (\mu_n \ast \lambda \ast \|u_m\|^2) (t) = \|u_m\|^2, \]

with \( (\lambda \ast \mu_n)(t) \rightarrow (\lambda \ast \mu)(t) = 1 \) and \( (\lambda \ast h_{mn})(t) \rightarrow 0 \) as \( n \rightarrow \infty \) in \( L^1(0, T) \).

Denote \( \bar{u}_m(t) := \|u_m(t)\|^2 + (\lambda \ast \|\nabla u_m\|^2) (t) \) and \( \bar{v}_m(t) := \|u_m(0)\|^2 + (\lambda \ast f\|^2) (t) \) which convert Equation (4.8) into

\[
\bar{u}_m(t) \leq \bar{v}_m(t) + \lambda \ast \left( \|u_m\|^2 + t \int_0^t \|\nabla u_m(s)\|^2 ds \right) \\
\leq \bar{v}_m(t) + \lambda \ast \left( \|u_m\|^2 + t \int_0^t (t - s)^{2-\alpha} \|\nabla u_m(s)\|^2 ds \right) \\
\leq \bar{v}_m(t) + t^{2-\alpha} \int_0^t (t - s)^{2-\alpha} \|u_m(s)\|^2 ds.
\]

(4.11)

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

\[
\|u_m\|_{L^2(0,T;L^2(\Omega))} + \|u_m\|_{L^2(0,T;\mathcal{V}_m(\Omega))} \leq \|u_m(0)\| + \|f\|_{L^2(0,T;L^2(\Omega))} \\
\leq \|u_0\| + \|f\|_{L^2(0,T;L^2(\Omega))} \leq \|\nabla u_0\| + \|f\|_{L^2(0,T;L^2(\Omega))}.
\]

(4.12)

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

\[
(-\Delta^M_m u_m,v_m) := (M(\cdot,\|\nabla u_m\|^2) \nabla u_m, \nabla v_m) \quad \forall u_m,v_m \in \mathcal{V}_m, \ t \in (0,T),
\]

(4.13)

and

\[
(-\Delta^b_2 u_m,v_m) := (b_2(\cdot, t,s) \nabla u_m, \nabla v_m) \quad \forall u_m,v_m \in \mathcal{V}_m, \ t,s \in (0,T).
\]

(4.14)

Since the diffusion coefficient is positive and \(b_2\) is a symmetric positive definite matrix, therefore, \(\Delta^M_m\) and \(\Delta^b_2\) are well defined. We make use of these definitions (4.13) and (4.14) to convert the Equation (4.2) into

\[
\left( \partial_t^H u_m, v_m \right) + \left( -\Delta^M_m u_m, v_m \right) \\
= (f,v_m) + \int_0^t \left( -\Delta^b_2 u_m(s), v_m \right) ds + \int_0^t (\nabla \cdot (b_1(\cdot, t,s) u_m(s)), v_m) ds \\
+ \int_0^t (b_0(\cdot, t,s) u_m(s), v_m) ds.
\]

(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

\[
\left( \partial_t^H \nabla u_m, M(\cdot,\|\nabla u_m\|^2) \nabla u_m \right) + \|\Delta^M_m u_m\|^2 \\
\leq \int_0^t \|\Delta^b_2 u_m(s)\|^2 ds + \int_0^t \|\nabla u_m(s)\|^2 ds \\
+ \int_0^t \|u_m(s)\|^2 ds + \|f\|^2.
\]

(4.16)

Estimate \(|(b_2(\cdot, t,s) \nabla u_m(s), \nabla v_m)| \leq \| \nabla u_m(s) \| \| \nabla v_m \| \) implies...
\[ \left\| \Delta_{m}^{b_{m}} u_{m}(s) \right\| = \sup_{v_{m} \in V_{m} \subset H_{0}^{1}(\Omega)} \frac{\langle \partial_{2}(x, t, s) \nabla u_{m}(s), \nabla v_{m} \rangle}{\left\| \nabla v_{m} \right\|} \leq \left\| \nabla u_{m}(s) \right\|. \]  

(4.17)

Hypothesis (H2) and estimate (4.17) yield

\[ \left( \frac{d}{dt} [\mu \star (u_{m} - u_{m}(0))] (t), \nabla u_{m} \right) + \left\| \Delta_{m}^{M} u_{m} \right\|^{2} \leq \int_{0}^{t} \left( \left\| \nabla u_{m}(s) \right\|^{2} + \left\| u_{m}(s) \right\|^{2} \right) ds + \left\| f \right\|^{2}. \]  

(4.18)

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

\[ \left\| u \right\|_{L^{2}(0, T; H_{o}^{1}(\Omega))} + \left( \lambda \ast \left\| \Delta_{m}^{M} u_{m} \right\|^{2} \right) (t) \leq \left\| \nabla u_{m}(0) \right\|^{2} + \left\| f \right\|_{L_{2}(0, T; L^{2}(\Omega))}^{2} \]  

\[ \leq \left\| \nabla u_{0} \right\|^{2} + \left\| f \right\|_{L_{2}(0, T; L^{2}(\Omega))}^{2}. \]  

(4.19)

Finally, substitute \( v_{m} = \partial_{t}^{2} u_{m} \) in (4.15) to have

\[ \left\| \partial_{t}^{2} u_{m} \right\|^{2} + (-\Delta_{m}^{M} u_{m}, \partial_{t}^{2} u_{m}) \leq \left\| f \right\|^{2} + \int_{0}^{t} \left( \left\| \Delta_{m}^{b_{m}} u_{m}(s) \right\|^{2} + \left\| \nabla u_{m}(s) \right\|^{2} \right) ds \]  

\[ + \int_{0}^{t} \left\| u_{m}(s) \right\|^{2} ds. \]  

(4.20)

Proceeding further as estimate (4.19) is proved, the following is concluded

\[ \left\| \partial_{t}^{2} u_{m} \right\|_{L_{2}(0, T; L^{2}(\Omega))} + \left\| u_{m} \right\|_{L^{2}(0, T; H_{0}^{1}(\Omega))} \leq \left\| \nabla u_{m}(0) \right\|^{2} + \left\| f \right\|_{L_{2}(0, T; L^{2}(\Omega))} \]  

\[ \leq \left\| \nabla u_{0} \right\|^{2} + \left\| f \right\|_{L_{2}(0, T; L^{2}(\Omega))}. \]  

(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 \( L^{2} (0, T; H_{0}^{1}(\Omega)) \) and \( \partial_{t}^{2} u_{m} \rightharpoonup \partial_{t}^{2} u \) in \( L^{2} (0, T; L^{2}(\Omega)) \). In the light of estimates (4.19) and (4.21), we apply the compactness Lemma 2.4 to conclude \( u_{m} \rightharpoonup u \) in \( L^{2} (0, T; H_{0}^{1}(\Omega)) \). Using the fact that \( M(x, \left\| \nabla u_{m} \right\|^{2}) \) 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 \((P^n)\).

(Initial Condition) The weak solution \( u \) satisfies the following equation for all \( v \in H_{0}^{1}(\Omega) \):

\[ \left( \frac{d}{dt} [\mu \star (u - u_{0})](t), v \right) = (M(x, \left\| \nabla u \right\|^{2}) \nabla u, \nabla v) + \int_{0}^{t} B(t, s, u(s), v) ds. \]  

(4.22)

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

\[ - \int_{0}^{T} ((\mu \ast (u - u_{0}))(t), v) \phi'(t) dt + \int_{0}^{T} (M(x, \left\| \nabla u \right\|^{2}) \nabla u, \nabla v) \phi(t) dt \]  

\[ = \int_{0}^{T} (f, v) \phi(t) dt + \int_{0}^{T} \int_{0}^{T} B(t, s, u(s), v) \phi(t) ds dt \]  

\[ + ((\mu \ast (u - u_{0}))(0), \phi(0)). \]  

(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 \( (\mu \ast (u_{m} - u_{m}(0))) \) has vanishing trace at \( t = 0 \), we have
In this section, we discuss the well-posedness of the semi discrete formulation and its error estimate. The projection operator satisfies the following stability and best approximation properties.

(6.7) For the case \( \alpha \) inequality to obtain

\[ \sup_{t} \int_{0}^{t} B(t, s, u_{m}(s), v) \phi(t) ds dt = 0. \]

Since \( \phi(0) \) is arbitrary, we have \( (\mu \ast (u - u_{0}))(0), \phi(0)) = 0 \), which implies \( u = u_{0} \) at \( t = 0 \) for \( a \in \left( \frac{1}{2}, 1 \right) \) [2, Proposition 6.7]. For the case \( \alpha \in \left( 0, \frac{1}{2} \right) \), we need to impose more compatibility conditions on data [2, Theorem 1.3].

(Uniqueness) Suppose that \( u_{1} \) and \( u_{2} \) are solutions of the weak formulation \( (W \alpha) \); then \( z = u_{1} - u_{2} \) satisfies the following equation for all \( v \in H_{0}^{1}(\Omega) \) and a.e. \( t \in (0, T) \):

\[
\left( \frac{d}{dt} (\mu \ast z)(t), v \right) + \left( M (x, \| \nabla u_{1} \|^2) \nabla z, \nabla v \right) = \left( M (x, \| \nabla u_{2} \|^2) - M (x, \| \nabla u_{1} \|^2) \right) \nabla u_{2}, \nabla v) + \int_{0}^{t} B(t, s, z(s), v) ds.
\]

Put \( v = z(t) \) in (4.26) and use (H2), (H3), and a priori bound (3.3) on \( u_{1}, u_{2} \) along with Cauchy-Schwarz and Young's inequality to obtain

\[
\left( \frac{d}{dt} (\mu \ast z)(t), z(t) \right) + (m_{0} - 4L_{M}K^{2}) \| \nabla z \|^2 \leq \int_{0}^{t} \| \nabla z(s) \|^2 ds.
\]

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

5 | SEMI DISCRETE FORMULATION AND ITS ERROR ESTIMATE

In this section, we discuss the well-posedness of the semi discrete formulation \( (S \alpha) \) and derive its error estimate 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 \( (S \alpha) \), which satisfies the following a priori bounds:

\[
\| u_{h} \|_{L^{2}(0, T; L^{2}(\Omega))} + \| u_{h} \|_{L^{2}(0, T; H_{0}^{1}(\Omega))} \leq \left( \| \nabla u_{0} \| + \| f \|_{L^{2}(0, T; L^{2}(\Omega))} \right),
\]

\[
\| \partial_{t}^{2} u_{h} \|_{L^{2}(0, T; L^{2}(\Omega))} + \| u_{h} \|_{L^{2}(0, T; H_{0}^{1}(\Omega))} \leq \left( \| \nabla u_{0} \| + \| f \|_{L^{2}(0, T; L^{2}(\Omega))} \right).
\]

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

For the semi discrete error estimate, we define a new Ritz-Volterra type projection operator \( W : [0, T] \to X_{h} \) by

\[
(M (x, \| \nabla u \|^2) \nabla (u - W), \nabla v_{h}) = \int_{0}^{t} B(t, s, u(s) - W(s), v_{h}) ds \quad \forall v_{h} \in X_{h}.
\]

This modified Ritz-Volterra projection operator \( W \) is well defined by the positivity of the Kirchhoff term \( M \) [54]. This projection operator satisfies the following stability and best approximation properties.
Lemma 5.2 ([55]). Let \( W \) be the modified Ritz-Volterra projection operator defined in (5.3). Then \( \| \nabla W \| \) is bounded for every \( t \) in \([0, T]\), that is,

\[
\| \nabla W \| \leq \| \nabla u \|. \tag{5.4}
\]

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 \( (P^s) \) such that [25, 30]

\[
\| u(t) \|_2 \leq C \text{ and } \| u_t(t) \|_2 \leq C \forall t \in [0, T]. \tag{5.5}
\]

Theorem 5.3. Suppose that the solution \( u \) of the problem \( (P^s) \) satisfies (5.5). Then the modified Ritz-Volterra projection operator has the following best approximation properties:

\[
\| \rho(t) \| + h^2 \| \nabla \rho(t) \| \leq h^2 \forall t \in [0, T],
\]

\[
\| \rho_t(t) \| + h^2 \| \nabla \rho_t(t) \| \leq h^2 \forall t \in [0, T], \tag{5.6}
\]

where \( \rho := u - W \).

**Proof.** For the proof of this theorem, we refer the readers to [42, 54]. \( \square \)

Using these best approximation properties, we prove error estimate for the semi discrete formulation \( (S^s) \) as stated in Theorem 3.2.

### 5.1 Proof of Theorem 3.2

**Proof.** Denote \( \theta := W - u_h \) such that \( u - u_h = \rho + \theta \). Put \( u_h = W - \theta \) in the formulation \( (S^s) \) to have

\[
\left( \partial_t^s (W - \theta), v_h \right) + \left( M(x, \| \nabla u_h \|^2) \nabla (W - \theta), \nabla v_h \right) = \left( f, v_h \right) + \int_0^t B(t, s, W(s) - \theta(s), v_h) \, ds \forall v_h \in X_h.
\]

Weak formulation \( (W^s) \) and the definition (5.3) of the modified Ritz-Volterra projection operator \( W \) yield

\[
\left( \partial_t^s \theta, v_h \right) + \left( M(x, \| \nabla u_h \|^2) \nabla \theta, \nabla v_h \right) = - \left( \partial_t^s \rho, v_h \right) + \int_0^t B(t, s, \theta(s), v_h) \, ds + \left( (M(x, \| \nabla u_h \|^2 - M(x, \| \nabla u \|^2)) \nabla W, \nabla v_h \right). \tag{5.7}
\]

Set \( v_h = \theta(t) \) in (5.7) and employ (H2) and (H3) to obtain

\[
\left( \partial_t^s \theta, \theta \right) + m_0 \| \nabla \theta \|^2 = \| \partial_t^s \rho \| \| \theta(t) \| + \| \nabla \theta \| \int_0^t \| \nabla \theta(s) \| \, ds
\]

\[
+ L_M(\| \nabla u_h \| + \| \nabla u \|)(\| \nabla \rho \| + \| \nabla \theta \|)\| \nabla W \| \| \nabla \theta \|. \tag{5.8}
\]

By utilizing the proved a priori bounds on \( \| \nabla u \| \) (3.3), \( \| \nabla u_h \| \) (5.2), and \( \| \nabla W \| \) (5.4) together with Cauchy-Schwarz and Young’s inequality, we obtain

\[
\left( \partial_t^s \theta, \theta \right) + (m_0 - 4L_MK^2) \| \nabla \theta \|^2 \leq \| \partial_t^s \rho \|^2 + \| \theta \|^2 + \int_0^t \| \nabla \theta(s) \|^2 \, ds + \| \nabla \rho \|^2. \tag{5.9}
\]

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

\[
\| \theta \|^2_{L^2(0,T;L^2(\Omega))} + (\lambda \ast \| \nabla \theta \|^2)(t) \leq \left( \lambda \ast \left( \| \nabla \rho \|^2 + \| \partial_t^s \rho \|^2 \right) \right)(t) + \| \nabla \theta(0) \|^2. \tag{5.10}
\]

For absolutely continuous function \( \rho \), we have \( \partial_t^s \rho = \mathcal{C}D_t^s \rho \) [50]. Therefore,
\[ \| \partial_t^\alpha \rho \| = \left\| \mathcal{D}_t^\alpha \rho \right\| = \left\| \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \frac{\partial \rho}{\partial s}(s) \, ds \right\| \lesssim \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \left\| \frac{\partial \rho}{\partial s}(s) \right\| \, ds. \]  

(5.11)

Using approximation property (5.6), we get

\[ \| \partial_t^\alpha \rho \| = \left\| \mathcal{D}_t^\alpha \rho \right\| \lesssim \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} h^2 \, ds \leq h^2. \]  

(5.12)

We choose \( u_h^0 = W(0) \) such that \( \theta(0) = 0 \) and apply approximation properties of modified Ritz-Volterra projection operator (5.6) and (5.11) in (5.10) to conclude

\[ \| \theta \|^2_{L^2(0,T;L^2(\Omega))} + (\lambda \| \nabla \theta \|^2)(t) \lesssim h^2 + h^4 \lesssim h^2. \]

Finally, triangle inequality and estimate (5.6) finish the proof. \( \square \)

**Remark 2.** In this work, we are focusing on the time discretization techniques. Therefore, we have used only low-order finite element method in space. We can extend this analysis for higher order FEM also by doing slight modifications in the finite dimensional subspace \( X_h \) of \( H^1_0(\Omega) \) as in [56] and in the best approximation properties of modified Ritz-Volterra projection operator.

### 6 | **LINEARIZED L1 GALERKIN FEM**

In this section, first, we derive a priori bounds on the solutions of the proposed fully discrete numerical scheme \( \mathcal{E}^\alpha \). Then we use these a priori bounds to prove the well-posedness of the numerical scheme \( \mathcal{E}^\alpha \) and carry out its convergence analysis.

**Lemma 6.1.** Suppose that (H1), (H2), and (H3) hold. Then the numerical scheme \( \mathcal{E}^\alpha \) is conditionally stable, and the solution \( u_h^1(n \geq 1) \) of this scheme satisfies the following a priori bound:

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

(6.1)

**Proof.** First, we estimate \( u_h^1 \). Put \( v_h = u_h^1 \) for \( n = 1 \) in the formulation \( \mathcal{E}^\alpha \) to get

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

Employing the definition of \( \mathbb{D}_t^\alpha u_h^1 \) (3.6), (H2), and (H3), along with Cauchy-Schwarz inequality and Young’s inequality, we obtain

\[ (1 - \alpha_0) \| u_h^1 \|^2 + m_0 \alpha_0 \| \nabla u_h^1 \|^2 \leq \alpha_0 \| f^1 \|^2 + \frac{\alpha_0 B_0^2 k^2}{m_0} \| \nabla u_h^0 \|^2 + \| u_h^0 \|^2, \]

where \( \alpha_0 = k^n \Gamma(2 - \alpha) \). For sufficiently small \( k \) such that \( k^n < \frac{1}{\Gamma(2 - \alpha)} \), apply Poincaré inequality to conclude

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

(6.2)

Now, we estimate \( u_h^n(n \geq 2) \). Set \( v_h = u_h^n \) for \( n \geq 2 \) in the scheme \( \mathcal{E}^\alpha \) to have
(6.3) \[ (\mathbb{D}_t^n u_h^n, u_h^n) + \left( M \left( x, \sqrt{\nabla u_{n-1}^2} \right) \nabla u_h^n, \nabla u_h^n \right) = (f^n, u_h^n) + \sum_{j=1}^{n-1} w_{nj} B (t_n, t_j, u_h^j, u_h^n). \]

Apply the identity \( (\mathbb{D}_t^n u_h^n, u_h^n) \geq \frac{1}{2} \mathbb{D}_t^n \| u_h^n \|^2 \) \([31, (3.10)]\), \(\text{(H2)}\), and \(\text{(H3)}\) along with Cauchy-Schwarz and Young's inequality to reach at

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

By the definition of \( \mathbb{D}_t^n \| u_h^n \|^2 \) \((6.6)\), Equation \((6.4)\) reduces to

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

Multiply Equation \((6.5)\) by discrete convolution \(P_{m-n}\) defined in Lemma 3.3 and take summation from \(n = 2\) to \(m\) for \(m \in \{2, 3, \ldots, N\}\) to obtain

\[ \sum_{m=2}^{m} P_{m-n} \sum_{j=1}^{n} a_{n-j} \left( \| u_h^j \|^2 - \| u_h^{j-1} \|^2 \right) + k^{\alpha} \Gamma(2-\alpha) \sum_{m=2}^{m} P_{m-n} \| \nabla u_h^n \|^2 \lesssim k^{\alpha} \Gamma(2-\alpha) \left( \sum_{m=2}^{m} P_{m-n} \| f^n \|^2 + \sum_{m=2}^{m} P_{m-n} \| u_h^n \|^2 + \sum_{m=2}^{m} P_{m-n} \sum_{j=1}^{n-1} w_{nj} \| \nabla u_h^n \|^2 \right) \]

\[ + k^{\alpha} \Gamma(2-\alpha) \sum_{m=2}^{m} P_{m-n} k \| \nabla u_h^n \|^2. \]

The estimate \((6.6)\) is rewritten as

\[ \sum_{m=1}^{m} P_{m-n} \sum_{j=1}^{n} a_{n-j} \left( \| u_h^j \|^2 - \| u_h^{j-1} \|^2 \right) + k^{\alpha} \Gamma(2-\alpha) \sum_{m=1}^{m} P_{m-n} \| \nabla u_h^n \|^2 \lesssim P_{m-1} \| u_h^n \|^2 - P_{m-1} \| u_h^0 \|^2 + k^{\alpha} \Gamma(2-\alpha) P_{m-1} \| \nabla u_h^n \|^2 + k^{\alpha} \Gamma(2-\alpha) \sum_{m=2}^{m} P_{m-n} k \| \nabla u_h^n \|^2 \]

\[ + k^{\alpha} \Gamma(2-\alpha) \left( \sum_{m=2}^{m} P_{m-n} \| f^n \|^2 + \sum_{m=2}^{m} P_{m-n} \| u_h^n \|^2 + \sum_{m=2}^{m} \sum_{j=1}^{n} k \sum_{n-j} \| \nabla u_h^n \|^2 \right). \]

Interchanging the summation in the first term of L.H.S of \((6.7)\) and applying the properties of discrete kernel \((3.13)\) and \((3.14)\) with \(a_0 = k^{\alpha} \Gamma(2-\alpha)\), we get

\[ (1 - a_0) \| u_h^n \|^2 + k^{\alpha} \sum_{n=2}^{m} P_{m-n} \| \nabla u_h^n \|^2 \lesssim \| u_h^0 \|^2 + k^{\alpha} \sum_{n=2}^{m} P_{m-n} \| f^n \|^2 + \| u_h^0 \|^2 \]

\[ + \sum_{n=1}^{m-1} \left( k^{\alpha} P_{m-n} \| u_h^n \|^2 + k \sum_{j=1}^{n} k \sum_{n-j} \| \nabla u_h^n \|^2 \right) \]

\[ + k^{\alpha} \Gamma(2-\alpha) \sum_{m=2}^{m} P_{m-n} \| \nabla u_h^n \|^2. \]

For sufficiently small \(k^{\alpha} < \frac{1}{\Gamma(2-\alpha)}\), we apply estimate \((6.2)\) to have
\[ \| u_h^m \|^2 + k^a \sum_{n=1}^{m} p_{m-n} \| \nabla u_h^n \|^2 \lesssim \| f^1 \|^2 + \| \nabla u_0 \|^2 + \max_{1 \leq n \leq N} \| f^n \|^2 k^a \sum_{n=2}^{m} p_{m-n} + \| u_h^0 \|^2 \]
\[ + \sum_{n=1}^{m} \max (k^a p_{m-n}, + k_1) \left( \| u_h^n \|^2 + k^a \sum_{j=1}^{n} p_{n-j} \| \nabla u_h^j \|^2 \right) \]
\[ + k^a \Gamma (2 - \alpha) \sum_{n=2}^{m} p_{m-n} k \| \nabla u_h^n \|^2. \]

Apply Poincaré inequality and property of discrete kernel (3.15) to obtain
\[ \| u_h^m \|^2 + k^a \sum_{n=1}^{m} p_{m-n} \| \nabla u_h^n \|^2 \lesssim \max_{1 \leq n \leq N} \| f^n \|^2 + \| \nabla u_0 \|^2 \]
\[ + \sum_{n=1}^{m-1} \max (k^a p_{m-n}, + k_1) \left( \| u_h^n \|^2 + k^a \sum_{j=1}^{n} p_{n-j} \| \nabla u_h^j \|^2 \right). \]

For sufficiently small k, discrete fractional Grönwall inequality [31] implies
\[ \| u_h^n \|^2 \lesssim \max_{1 \leq n \leq N} \| f^n \|^2 + \| \nabla u_0 \|^2. \]
(6.8)

The estimate (6.8) is true for all \( m \in \{2, \ldots, N\} \); hence, the result (6.1) follows.

Due to the presence of gradient type nonlinearity in our problem \((P^n)\), we need to derive a priori bound on \( \| \nabla u_h^n \| \) \( (n \geq 1) \) also. One can obtain the bound on \( \| \nabla u_h^n \| \) by using the estimate (6.1), but then we have
\[ \| \nabla u_h^n \| \lesssim k^{-a} \forall n \geq 1. \]
(6.9)

In practices, we take very small time step that causes the blow up of \( \| \nabla u_h^n \| \) in (6.9). We overcome this difficulty in the following lemma by making use of two new discrete Laplacian operators defined in (4.13) and (4.14).

**Lemma 6.2.** Suppose that Hypotheses (H1), (H2), and (H3) hold. Then the fully discrete scheme \((E^n)\) is conditionally stable, and its solutions \( u_h^n \) \( (n \geq 1) \) satisfy
\[ \max_{1 \leq n \leq N} \| \nabla u_h^n \|^2 + k^a \sum_{n=1}^{N} p_{N-n} \| \Delta_h^M u_h^n \|^2 \lesssim \max_{1 \leq n \leq N} \| f^n \|^2 + \| \nabla u_0 \|^2. \]
(6.10)

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 operators (4.13) and (4.14), the scheme \((E^n)\) for \( n = 1 \) is rewritten as
\[ (D_t^a u_h^1, v_h) + (-\Delta_h^M u_h^1, v_h) = (f^1, v_h) + k \left\{ \Delta_h^M u_h^0, v_h \right\} + k \left\{ \nabla \cdot (b_1 (x, t, s) u_h^0), v_h \right\} + k \left\{ b_0 (x, t, s) u_h^0, v_h \right\}. \]
(6.11)

Setting \( v_h = -\Delta_h^M u_h^1 \) in (6.11), one obtain
\[ (D_t^a \nabla u_h^1, \nabla u_h^1) + \| \Delta_h^M u_h^1 \|^2 \lesssim \| f^1 \|^2 + k^2 \left( \| \Delta_h^M u_h^0 \|^2 + \| \nabla u_h^0 \|^2 + \| u_h^0 \|^2 \right). \]
(6.12)
Again, identity \( (D_i^n u_h^n, u_h^n) \geq \frac{1}{2} D_i^n \| u_h^n \|^2 \) and estimate (4.17) simplify Equation (6.12) to
\[
\left\| \nabla u_h^1 \right\|^2 + k^\alpha \left\| \nabla^M u_h^1 \right\|^2 \leq k^\alpha \left\| f \right\|^2 + k^\alpha k \left\| \nabla u_h^0 \right\|^2 + \left\| \nabla u_h^2 \right\|^2.
\] (6.13)

For sufficiently small \( k \) as in Lemma 6.1, we deduce
\[
\left\| \nabla u_h^1 \right\|^2 + k^\alpha \left\| \nabla^M u_h^1 \right\|^2 \leq \left\| f \right\|^2 + \left\| \nabla u_h^0 \right\|^2 \leq \left\| f \right\|^2 + \left\| \nabla u_0 \right\|^2.
\] (6.14)

Consider the scheme \( (\mathcal{E}^n) \) for \( n \geq 2 \) with definitions of discrete Laplacian operators (4.13) and (4.14):
\[
(D_i^n u_h^n, v_h) + (\nabla^M u_h^n, \nabla v_h) = (f^n, v_h) + \sum_{j=1}^{n-1} w_n j \left( -\nabla^M u_h^n, v_h \right)
\]
\[
+ \sum_{j=1}^{n-1} w_n j \left( \nabla \cdot (b_l(x, t_n, t_j) u_h^n), v_h \right)
\]
\[
+ \sum_{j=1}^{n-1} w_n j \left( b_0(x, t_n, t_j) u_h^n, v_h \right).
\] (6.15)

Take \( v_h = -\nabla^M u_h^n \) in (6.15) and apply (H2) and (H3) along with Cauchy-Schwarz and Young’s inequality to get
\[
(D_i^n \nabla u_h^n, \nabla u_h^n) + \left\| \nabla^M u_h^n \right\|^2 \leq \left\| f^n \right\|^2 + \sum_{j=0}^{n-1} w_n j \left\| \nabla^M u_h^n \right\|^2
\]
\[
+ \sum_{j=0}^{n-1} w_n j \left\| \nabla u_h^n \right\|^2 + \sum_{j=0}^{n-1} w_n j \left\| u_h^n \right\|^2.
\] (6.16)

By using the estimate (4.17) and the identity \( (D_i^n u_h^n, u_h^n) \geq \frac{1}{2} D_i^n \| u_h^n \|^2 \), Equation (6.16) is converted into
\[
(D_i^n \nabla u_h^n, \nabla u_h^n) + \left|\nabla^M u_h^n \right|^2 \leq \left( \left\| f^n \right\|^2 + \sum_{j=0}^{n-1} w_n j \left\| \nabla u_h^n \right\|^2 \right).
\] (6.17)

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

The priori bound in Lemma (6.2) helps us to prove the well-posedness of the fully discrete numerical scheme \( (\mathcal{E}^n) \). To show the existence of the fully discrete solution \( u_h^n (n \geq 1) \) of the formulation \( (\mathcal{E}^n) \), the following variant of Bröuer fixed point theorem is used.

**Theorem 6.3** ([44]). Let \( H \) be a 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 fully discrete formulation \( (\mathcal{E}^n) \).

**Proof.** (Existence) Take \( n = 1 \) in the scheme \( (\mathcal{E}^n) \) and apply the definition of \( D_i^n u_h^1 (3.6) \) with \( \alpha_0 = k^\alpha \Gamma(2 - \alpha) \) to obtain
\[
(u_h^1 - u_h^1, v_h) + \alpha_0 \left( M(x, \left\| \nabla u_h^1 \right\|^2) \nabla u_h^1, \nabla v_h \right) = \alpha_0 (f^1, v_h) + \alpha_0 k B(t_n, t_n, u_h^0, v_h).
\] (6.18)
In the view of (6.18), we define a map $G : X_h \to X_h$ by

$$(G(u^1_h),v_h) = (u^1_h,v_h) - (u^0_h,v_h) + a_0 \left( M \left( x, \|\nabla u^0_h\| \right) \nabla u^1_h, \nabla v_h \right) - a_0 \left( f^1, v_h \right) - a_0 k_B \left( t_0, u^0_h, v_h \right) . \quad (6.19)$$

Using (H2), (H3), and Cauchy-Schwarz inequality together with Poincaré inequality with Poincaré constant $C_p$, we have

$$(G(u^1_h),u^1_h) \geq \|u^1_h\|^2 - \|u^0_h\| \|u^0_h\| - a_0 \|f^1\| \|u^1_h\|^2 + a_0 m_0 \|\nabla u^1_h\|^2 - a_0 k_B \|\nabla u^0_h\| \|u^0_h\| + \|u^1_h\| (\|u^1_h\|-\|u^0_h\|-a_0 \|f^1\|) + a_0 \|\nabla u^0_h\| (m_0 \|\nabla u^1_h\| - k_B \|\nabla u^0_h\|). \quad (6.20)$$

Thus, for $\|u^1_h\| > \|u^1_h\| + a_0 \|f^1\| + \frac{k_B C_p}{m_0} \|\nabla u^0_h\|$, one has $(G(u^1_h),u^1_h) > 0$, and the map $G$ defined by (6.19) is continuous as a consequence of continuity of $M$ and $B$. Hence, existence of $u^1_h$ follows by Theorem 6.3 immediately.

(Uniformity) Suppose that $X^1_h$ and $Y^1_h$ are solutions of the scheme $(E^\alpha)$ for $n = 1$; then $Z^1_h = X^1_h - Y^1_h$ satisfies the following equation for all $v_h$ in $X_h$:

$$(D^\alpha Z^1_h, v_h) + \left( M \left( x, \|\nabla X^1_h\| \right) \nabla Z^1_h, \nabla v_h \right) = \left( \left( M \left( x, \|\nabla Y^1_h\| \right) \right) \left( \nabla X^1_h, \nabla v_h \right) \right) . \quad (6.21)$$

Put $v_h = Z^1_h$ in (6.21) and use (H2) to get

$$\frac{1}{2} D^\alpha \|Z^1_h\|^2 + m_0 \|\nabla Z^1_h\|^2 \leq L_M \|\nabla Z^1_h\| \left( \|\nabla X^1_h\| + \|\nabla Y^1_h\| \right) \left( \nabla Y^1_h, \nabla Z^1_h \right) .$$

Cauchy-Schwarz inequality and a priori bound (6.10) yield

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

Employ (H2) to obtain $\|Z^1_h\| = \|\nabla Z^1_h\| = 0$ that concludes the uniqueness of solution for $n = 1$ in the scheme $(E^\alpha)$.

For $n \geq 2$, the numerical scheme $(E^\alpha)$ is linear with a positive definite coefficient matrix by (H2). As a result, we get the existence and uniqueness of the solution $u^1_h(n \geq 2)$ for the scheme $(E^\alpha)$. \hfill \Box

Now, we derive one of the main results about the convergence rate of the developed numerical scheme $(E^\alpha)$. First, we discuss the approximation properties of the L1 scheme (3.6), linearization technique (3.8), and quadrature error (3.9). 

**Lemma 6.5.** Suppose that $X = H^2(\Omega) \cap H^1_0(\Omega)$ and $u \in C^2([0,T];X)$ be the solution of the problem $(P^\alpha)$. Then the truncation error $Q^\alpha$ defined in (3.6) satisfies

$$\|Q^\alpha\| \lesssim k^{2-\alpha} \text{ for } n \geq 1.$$ \quad (6.22)

**Proof.** For the proof of this lemma, we refer the readers to [30, (3.3)]. \hfill \Box
Lemma 6.6. Suppose that $X = H^2(\Omega) \cap H^1_0(\Omega)$ and $u \in C^2([0, T]; X)$ be the solution of the problem $(P^u)$. Then the linearization error $\mathcal{E}^{n-1} := u^n - \bar{u}^{n-1}$ defined in (3.8) undergoes

$$\|\mathcal{E}^{n-1}\|_1 \lesssim k^2 \text{ for } n \geq 2. \quad (6.23)$$

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

$$u^{n-1} = u^n - ku^n + \frac{k^2}{2} u_t^n(\xi_1) \text{ for some } \xi_1 \in (t_{n-1}, t_n). \quad (6.24)$$

and

$$u^{n-2} = u^n - 2ku^n + \frac{4k^2}{2} u_t^n(\xi_2) \text{ for some } \xi_2 \in (t_{n-2}, t_n). \quad (6.25)$$

Multiplying Equation (6.24) by 2 and subtracting Equation (6.25) from Equation (6.24), we get

$$\|\mathcal{E}^{n-1}\|_1 \lesssim k^2 (||u_t^n(\xi_1)||_1 + ||u_t^n(\xi_2)||_1). \quad (6.26)$$

Since $u \in C^2([0, T]; X)$, it implies that $||u_t^n||_1 \lesssim C \forall t \in [0, T]$. Hence, we conclude the result (6.23). \hfill \Box

Lemma 6.7. Suppose that $X = H^2(\Omega) \cap H^1_0(\Omega)$ and $u \in C^4([0, T]; X)$ be the solution of the problem $(P^u)$. Then the quadrature error defined by (3.9) has the following estimate:

$$\left| (q^n, v) \right| \lesssim k^2 \|\nabla v\| \forall v \in H^1_0(\Omega), \text{ for } n \geq 1. \quad (6.27)$$

Proof. Recalling the approximation (3.9) and denoting $h(s) := B(t_n, s, u(s), v)$, then we have

$$(q^n, v) = \int_0^{t_n} h(s) ds - \sum_{j=0}^{n-1} w_j h(t_j) \forall v \in H^1_0(\Omega), \quad (6.28)$$

where

$$\sum_{j=0}^{n-1} w_n h(t_j) = \frac{k_1}{3} \sum_{j=1}^{j_{n/2}} \left[ h \left( \frac{t^n_j}{2} \right) + 4h \left( \frac{t^n_j}{2} \right) + h \left( \frac{t^n_j}{2} \right) \right]$$

$$+ \frac{k}{2} \sum_{j=0}^{j_n} \left[ h \left( \frac{t^n_j}{2} \right) + h \left( \frac{t^n_j}{2} \right) \right] + kh \left( \frac{t^n_j}{2} \right). \quad (6.29)$$

In this approximation (6.29), we break the interval of integration $[0, t_n]$ into three sub-intervals. On the first interval, we use composite Simpson's 1/3 rule with time step $k_1$ that gives truncation error of order $k_1^4 |h^4(s)|$. In the second interval, a composite trapezoidal rule with step length $k$ is applied that has truncation error of order $k^2 |h^2(s)|$. In the last interval, rectangle rule with time step $k$ is used that gives truncation error of order $k^2 |h^1(s)|$. By combining these truncation errors in $(q^n, v)$, we obtain

$$\left| (q^n, v) \right| \lesssim \left( k_1^4 |h^4(s)| + k^2 |h^2(s)| + k^2 |h^1(s)| \right) \forall s \in [0, T]. \quad (6.30)$$

Using Cauchy-Schwarz inequality, smoothness of the coefficients of memory operator, and solution of the problem $(P^u)$, we have

$$\left| (q^n, v) \right| \lesssim (k_1^4 + k^2 + k^2) \|\nabla v\|. \quad (6.31)$$

Using the fact that $k_1 = O \left( k^{1/3} \right)$ in (6.31), we conclude the result (6.27). \hfill \Box
We prove the convergence estimate of the proposed numerical scheme \((\mathcal{E}^a)\) by assuming that the solution \(u\) of the problem \((P^a)\) satisfies additional regularity used in Lemmas 6.5 to 6.7, that is, \(u \in C^4([0, T]; H^2(\Omega) \cap H_0^1(\Omega))\) [25, 30].

### 6.1 Proof of 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 \((\mathcal{E}^a)\) and use weak formulation \((W^a)\) along with modified Ritz-Volterra projection operator \(W\) at \(t_1\) to get

\[
\left( \mathbb{D}_t^a \theta^1, v_h \right) + \left( M \left( x, \| \nabla u_h^1 \|^2 \right) \nabla \theta^1, \nabla v_h \right) = \left( \mathbb{D}_t^a W^1 - C D_t^a \theta^1 \right) - kB \left( t_1, t_0, W^0, v_h \right) + \int_{t_0}^{t_1} B(t, s, W(s), v_h) ds \tag{6.32}
\]

\[
+ \left( \left[ M \left( x, \| \nabla u_h^1 \|^2 \right) - M \left( x, \| \nabla u^1 \|^2 \right) \right] \nabla W^1, \nabla v_h \right) + kB \left( t_1, t_0, \theta^0, v_h \right).
\]

Set \(v_h = \theta^1\) and \(\theta^0 = 0\) in (6.32), we get

\[
\left( \mathbb{D}_t^a \theta^1, \theta^1 \right) + \left( M \left( x, \| \nabla u_h^1 \|^2 \right) \nabla \theta^1, \nabla \theta^1 \right) = \left( \mathbb{D}_t^a W^1 - C D_t^a \theta^1 \right) + \left( C D_t^a W - C D_t^a \theta^1 \right) + (q^1, \theta^1) \tag{6.33}
\]

\[
+ \left( \left[ M \left( x, \| \nabla u_h^1 \|^2 \right) - M \left( x, \| \nabla u^1 \|^2 \right) \right] \nabla W^1, \nabla \theta^1 \right).
\]

By the definition of \(\mathbb{D}_t^a \theta^1\) (3.6), (H2), and estimate (6.27), we have

\[
(1 - k^a \Gamma(2 - a)) \| \theta^1 \|^2 + k^a (m_0 - 4 L_M K^2) \| \nabla \theta^1 \|^2 \leq k^a \| Q^1 \|^2 + k^a \| C D_t^a \rho \|^2 + k^a k^4 \| \nabla \rho^1 \|^2. \tag{6.34}
\]

For sufficiently small \(k^a < \frac{1}{\Gamma(2 - a)}\), we apply (H2) and the approximation properties (6.22), (5.6), and (5.11) to conclude

\[
\| \theta^1 \|^2 + k^a \| \nabla \theta^1 \|^2 \leq k^a (k^{2-a})^2 + k^a h^4 + k^a k^4 + k^a h^2 \leq (k^{2-a} + h)^2. \tag{6.35}
\]

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

\[
\left( \mathbb{D}_t^a \theta^n, v_h \right) + \left( M \left( x, \| \nabla u_h^{n-1} \|^2 \right) \nabla \theta^n, \nabla v_h \right) = \left( Q^n, v_h \right) - \left( C D_t^a \rho, v_h \right) \tag{6.36}
\]

\[
+ \left( \left[ M \left( x, \| \nabla u_h^{n-1} \|^2 \right) - M \left( x, \| \nabla u^n \|^2 \right) \right] \nabla W^n, \nabla v_h \right) + \left( q^n, v_h \right) + \sum_{j=1}^{n-1} w_{n_j} B \left( t_n, t_j, \theta^j, v_h \right).
\]

Put \(v_h = \theta^n\) in (6.36); it follows that

\[
\mathbb{D}_t^a \| \theta^n \|^2 + \| \nabla \theta^n \|^2 \leq \| Q^n \|^2 + \| C D_t^a \rho \|^2 + \| \theta^n \|^2 + \| \nabla \theta^{n-1} \|^2 + \| \nabla \theta^{n-2} \|^2 \tag{6.37}
\]

\[
+ \| \nabla \theta \|^2 + k^4 + \sum_{j=1}^{n-1} w_{n_j} \| \nabla \theta^j \|^2.
\]
Employ the approximation properties \((6.22), (5.6), (5.11),\) and \((6.23)\) to deduce
\[
\begin{align*}
\|\nabla\theta^n\|^2 + \|\nabla\theta^n\|^2 & \leq (k^{2-\alpha} + h^2 + h + k^2 + k^2)^2 + \|\theta^n\|^2 \\
& + \left(\|\nabla\theta^n\|^2 + \|\nabla\theta^n\|^2 + \sum_{j=1}^{n-1} k_j\|\nabla\theta^j\|^2\right).
\end{align*}
\] (6.38)

Following the similar arguments of the proof of estimate \((6.1)\), we conclude the result \((3.16)\). □

7 | LINEARIZED L2-1σ GALERKIN FEM

In this section, we show that the proposed numerical scheme \((\mathcal{F}^a)\) stated in Theorem 3.6 is conditionally stable and conditionally convergent with the second-order convergence rate in the time direction.

**Lemma 7.1.** Under Hypotheses \((H1), (H2),\) and \((H3),\) the fully discrete numerical scheme \((\mathcal{F}^a)\) is conditionally stable, and its solutions \(u^n_h (n \geq 1)\) satisfy the following a priori bound:

\[
\max_{1 \leq n \leq N} \left\|u^n_h\right\|^2 + k^n \sum_{n=1}^{N} p^{(N)}_{N-n} \left\|\nabla u^n_h\right\|^2 \leq \max_{1 \leq n \leq N} \left\|f^{n-\sigma}\right\|^2 + \|\nabla u_0\|^2.
\] (7.1)

**Proof.** For \(n = 1\), the scheme \((\mathcal{F}^a)\) is given by

\[
\left(\bar{D}^{\sigma}_{i-\sigma} u^1_h, v_h\right) + \left(M \left(x, \|\nabla u_h^{1,\sigma}\|\right) \nabla u_h^{1,\sigma}, \nabla v_h\right) = (1 - \sigma) k B \left(t_{i-\sigma}, t_0, u^0_h, v_h\right)
\] (7.2)

Substitute \(v_h = u^1_h\) in \((7.2)\) and apply the definition of \(\bar{D}^{\sigma}_{i-\sigma} u^1_h\) \((3.18)\) to get

\[
\begin{align*}
\frac{k^{-\alpha}(1 - \sigma)^{-\alpha}}{\Gamma(2 - \alpha)} \left(\left|u^1_h - u^0_h, u^1_h\right| + (1 - \sigma) \left(M \left(x, \|\nabla u_h^{1,\sigma}\|\right) \nabla u_h^{1,\sigma}, \nabla u_h^{1,\sigma}\right)
\end{align*}
\] (7.3)

Using \((H2)\) and \((H3)\) along with Cauchy-Schwarz inequality and Young's inequality, we obtain

\[
\begin{align*}
\left(1 - \frac{a_0}{\bar{a}_0}\right) \left\|u^1_h\right\|^2 + k^n \left\|\nabla u_h\right\|^2 & \leq k^n \left\|\nabla u^0_h\right\|^2 + \left\|u^0_h\right\|^2 + k^{2-\alpha} \left\|\nabla u_h\right\|^2 + k^n \|f^{1-\sigma}\|^2.
\end{align*}
\] (7.4)

where \(a_0 = k^n \Gamma(2 - \alpha)\) and \(\bar{a}_0 = (1 - \sigma)^{-1}.\) Take sufficiently small \(k\) such that \(\frac{a_0}{\bar{a}_0} < 1,\) that is, \(k^n < \frac{1}{\Gamma(2 - \alpha)(1 - \sigma)^{-1}}\), to conclude

\[
\begin{align*}
\left\|u^1_h\right\|^2 + k^n \left\|\nabla u^1_h\right\|^2 & \leq \left\|\nabla u^0_h\right\|^2 + \|f^{1-\sigma}\|^2 \leq \left\|\nabla u_h\right\|^2 + \|f^{1-\sigma}\|^2.
\end{align*}
\] (7.5)

For \(u^n_h (n \geq 2)\) in the scheme \((\mathcal{F}^a)\), we have

\[
\left(\bar{D}^{\sigma}_{i-\sigma} u^n_h, v_h\right) + \left(M \left(x, \|\nabla u_h^{n-1,\sigma}\|\right) \nabla u_h^{n-1,\sigma}, \nabla v_h\right) = \sum_{j=1}^{n-1} \tilde{w}_{nj} B \left(t_{n-\sigma}, t_j, u^j_h, v_h\right)
\] (7.6)

Put \(v_h = u^n_h\) in \((7.6)\) to obtain
\begin{equation}
(\tilde{D}^a_{t_{n-1}} u^a_h, u^a_h) + (1 - \sigma) \left( M \left( x, \| \nabla \hat{u}^{n-1,\sigma}_h \| \right) \nabla u_h, \nabla u_h \right)
= \sum_{j=1}^{n-1} \tilde{w}_j B \left( t_{n-j}, t_j, u^j_h, u^j_h \right) + \left( f^{n-j}, u^j_h \right) - \sigma \left( M \left( x, \| \nabla \hat{u}^{n-1,\sigma}_h \| \right) \nabla u_h^{n-1}, \nabla u_h^{n-1} \right). \tag{7.7}
\end{equation}

We use the identity \((\tilde{D}^a_{t_{n-1}} u^a_h, u^a_h) \geq \frac{1}{2} \tilde{D}^a_{t_{n-\sigma}} \left\| u^a_h \right\|^2\) [25, Lemma 1] and apply (H2) and (H3) along with Cauchy-Schwarz and Young’s inequality to reach at
\begin{equation}
\tilde{D}^a_{t_{n-\sigma}} \left\| u^a_h \right\|^2 + \left\| \nabla u^a_h \right\|^2 \leq \| f^{n-\sigma} \|^2 + \sum_{j=0}^{n-1} \tilde{w}_j \left\| \nabla u^j_h \right\|^2 + \left\| \nabla u^{n-1}_h \right\|^2 + \left\| u^a_h \right\|^2. \tag{7.8}
\end{equation}

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 numerical scheme \((F^a)\) is conditionally stable, and the solutions \(u^a_h (n \geq 1)\) of this scheme satisfy the following a priori bound:
\begin{equation}
\max_{1 \leq n \leq N} \left\| \nabla u^a_h \right\|^2 + k^a \sum_{n=1}^{N} \| \Delta^a_h u^a_h \|^2 \leq \max_{1 \leq n \leq N} \| f^{n-\sigma} \|^2 + \left\| \nabla u_0 \right\|^2. \tag{7.9}
\end{equation}

**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 fully discrete formulation \((F^a)\).

**Proof.** For the case \(n \geq 2\), the scheme \((F^a)\) is linear having positive definite coefficient matrix. Thus, existence and uniqueness of solution follow immediately for \(n \geq 2\). For the case \(n = 1\) in the scheme \((F^a)\), the equation is nonlinear so we again use Bröuwer fixed point Theorem 6.3. Consider the case for \(n = 1\) in the problem \((F^a)\) with \(a_0 = k^a \Gamma (2 - \alpha)\) and \(a_0 = (1 - \sigma)^{1-\sigma}\); we have
\begin{equation}
(u_1^h - u_0^h, v_h) + \frac{a_0}{a_0} \left( M \left( x, \| \nabla \hat{u}^{1,\sigma}_h \| \right) \nabla \hat{u}^{1,\sigma}_h, \nabla v_h \right)
= \frac{a_0}{a_0} \left( f^{1-\sigma}, v_h \right) + \frac{a_0}{a_0} k B \left( t_{1-\sigma}, t_0, u_0^h, v_h \right). \tag{7.10}
\end{equation}

Multiply Equation (7.10) by \((1 - \sigma)\) to obtain
\begin{equation}
\left( \hat{u}^{1,\sigma}_h, v_h \right) - (u_0^h, v_h) + (1 - \sigma) \frac{a_0}{a_0} \left( M \left( x, \| \nabla \hat{u}^{1,\sigma}_h \| \right) \nabla \hat{u}^{1,\sigma}_h, \nabla v_h \right)
= (1 - \sigma) \frac{a_0}{a_0} \left( f^{1-\sigma}, v_h \right) + (1 - \sigma) \frac{a_0}{a_0} k B \left( t_{1-\sigma}, t_0, u_0^h, v_h \right). \tag{7.11}
\end{equation}

Further, proceeding analogously to the proof of Theorem 6.4, we conclude the existence of \(\hat{u}^{1,\sigma}_h\). Hence, existence of \(u_1^h\) follows.

Now, we derive the convergence estimate for the proposed numerical scheme \((F^a)\). This estimate is proved with the help of the following lemmas.

**Lemma 7.4.** Suppose that \(X = H^2(\Omega) \cap H^1_0(\Omega)\) and \(u \in C^1([0, T]; X)\) be the solution of the problem \((P^a)\). Then the truncation error \(\tilde{Q}^{n-\sigma}_h\) defined in (3.18) satisfies
\begin{equation}
\| \tilde{Q}^{n-\sigma}_h \| \lesssim k^{3-\alpha} \text{ for } n \geq 1. \tag{7.12}
\end{equation}

**Proof.** One can obtain this estimate as in [25, Lemma 2].
Lemma 7.5. Suppose that $X = H^2(\Omega) \cap H_0^1(\Omega)$ and $u \in C^2([0, T]; X)$ be the solution of the problem $(P^a)$. Then the linearization error $\|L^{n-1, a}\|_1 \leq k^2$ for $n \geq 2$. (7.13)

and

$$\|\tilde{L}^{n, a}\|_1 \leq k^2 \text{ for } n \geq 1. \quad (7.14)$$

Proof. Similar to the proof of Lemma (6.6), we write the Taylor's series expansion of $u^{n-1}$ and $u^{n-2}$ around $u^{n-a}$ to obtain

$$u^{n-1} = u^{n-a} - (1 - \sigma)ku^n + (1 - \sigma)^2 \frac{k^2}{2} u_n(\xi_1) \text{ for some } \xi_1 \in (t_{n-1}, t_{n-a}). \quad (7.15)$$

and

$$u^{n-2} = u^{n-a} - (2 - \sigma)ku^n + (2 - \sigma)^2 \frac{k^2}{2} u_n(\xi_2) \text{ for some } \xi_2 \in (t_{n-2}, t_{n-a}). \quad (7.16)$$

Multiply Equations (7.15) and (7.16) by $(2 - \sigma)$ and $(1 - \sigma)$, respectively. Subtracting Equation (7.16) from Equation (7.15), we get

$$\|\tilde{L}^{n-1, a}\|_1 \leq k^2 (||u_n(\xi_1)||_1 + ||u_n(\xi_2)||_1). \quad (7.17)$$

Since $u \in C^2([0, T]; X)$, it implies that $\|u_n\|_1 \leq C \forall t \in [0, T]$. Hence, we conclude the result (7.13). Similarly, one can prove the estimate (7.14).

Lemma 7.6. Suppose that $X = H^2(\Omega) \cap H_0^1(\Omega)$ and $u \in C^4([0, T]; X)$ be the solution of the problem $(P^a)$. Then the quadrature error defined by (3.23) has the following error estimate:

$$\left|\left(\tilde{q}^{n-a}, v\right)\right| \leq k^2 \|\nabla v\| \forall v \in H_0^1(\Omega) \text{ for } n \geq 1. \quad (7.18)$$

Proof. The quadrature error (3.23) is just a slight modification of the quadrature error (3.9). Thus, the proof of this lemma follows from the similar lines of the proof of Lemma (6.7).

7.1 | Proof of Theorem 3.6

Proof. Take $u_h^1 = W^1 - \theta^1$ with $\theta^0 = 0$ in the scheme $(P^a)$ for $n = 1$; we have the following error equation for $\theta^1$:

$$(\tilde{L}^{1, a} \theta^1, v_h) + \left( M \left( x, \left\| \nabla \tilde{u}^1 \right\|^2 \right) \nabla \theta^1, \nabla v_h \right) \quad (7.19)$$

Set $v_h = \theta^1$ in (7.19) with $a_0 = k^aT(2 - a)$, $\tilde{a}_0 = (1 - \sigma)^{-1-a}$ and apply (7.18) to have

$$\left(1 - \frac{a_0}{\tilde{a}_0}\right) \|\theta^1\|^2 + k^a \left( m_0 - 4L_{K^2} \right) \|\nabla \theta^1\|^2$$

$$\leq k^a \left\| \tilde{L}^{1-a} \right\|^2 + k^a \left\| C \tilde{D}_{i-x} \rho \right\|^2 + k^a \left\| \nabla \tilde{u}^1 \right\|^2 + k^a \left\| \nabla \tilde{u}^1 \right\|^2 + k^a \left\| \nabla \tilde{u}^1 \right\|^2$$

$$+ k^a \left\| \nabla \tilde{u}^1 \right\|^2. \quad (7.20)$$

For smaller value of $k$ as in Lemma 7.1, apply Hypothesis (H2) and approximation properties (7.12), (5.6), (5.11), and (7.14) to deduce

$$\|\theta^1\|^2 + k^a \|\nabla \theta^1\|^2 \leq (k^{3-a} + h^2 + k^2 + h^2)^2 \leq (k^2 + h)^2. \quad (7.21)$$
For \( n \geq 2 \), substitute \( u^n_h = W^n - \theta^n \) in the scheme \((P^\alpha)\); then \( \theta^n \) satisfies

\[
(\tilde{D}_{t_{n-\sigma}}\theta^n, v_h) + \left( M \left( x, \| \nabla \hat{u}^{n-\sigma}_h \| \right)^2 \right) \nabla \theta^n, \nabla v_h \\
= (\tilde{Q}^{n-\sigma}, v_h) - \left( C D_{t_{n-\sigma}}^\alpha \rho, v_h \right) + \left( M \left( x, \| \nabla \hat{u}^{n-\sigma}_h \| \right)^2 \right) \left( \nabla W^{n,\sigma} - \nabla W^{n-\sigma} \right) \nabla v_h \\
+ \left( M \left( x, \| \nabla \hat{u}^{n-\sigma}_h \| \right)^2 \right) - M \left( x, \| \nabla \theta^n \| \right)^2 \nabla W^{n-\sigma}, \nabla v_h \\
+ (\tilde{q}^{n-\sigma}, v_h) + \sum_{j=1}^{n-1} \bar{w}_n B \left( t_{n-\sigma}, t_j, \theta^j, v_h \right).
\]

Put \( v_h = \theta^n \) in (7.22); it follows that

\[
\tilde{D}_{t_{n-\sigma}}\| \theta^n \|^2 + \| \nabla \theta^n \|^2 \leq \| \tilde{Q}^{n-\sigma} \|^2 + \| CD_{t_{n-\sigma}}^\alpha \rho \|^2 + \| \nabla \hat{u}^{n-\sigma}_h \|^2 + \| \nabla \theta^n \|^2 \left( 1 + \sum_{j=1}^{n-1} \kappa^2 \| \nabla \theta^j \|^2 \right)
\]

Further, apply the approximation properties (7.12), (5.6), (5.11), (7.13), and (7.14) to arrive at

\[
\tilde{D}_{t_{n-\sigma}}\| \theta^n \|^2 + \| \nabla \theta^n \|^2 \leq \left( \| \theta^n \|^2 + \| \nabla \theta^n \|^2 + \sum_{j=1}^{n-1} \kappa^2 \| \nabla \theta^j \|^2 \right) \left( k^3 + h^2 + h + k + h^2 \right)^2.
\]

Follow the similar arguments as in Theorem 3.4 and Lemma 7.1 to obtain

\[
\| \theta^n \|^2 + k^2 \sum_{m=1}^{m} \| \nabla \theta^n \|^2 \leq \left( h + k^3 \right)^2.
\]

Finally, triangle inequality and estimate (5.6) complete the proof. \( \Box \)

8 | NUMERICAL RESULTS

In this section, we implement the theoretical results obtained from fully discrete formulations \((E^\alpha)\) and \((F^\alpha)\) for the problem \((P^\alpha)\). 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. The numerical solution \( u^n_h (n \geq 1) \) for the considered problem \((P^\alpha)\) at any time \( t_n \) in \([0, T]\) is given by

\[
u^n_h = \sum_{i=1}^{J} a^n_i \psi_i, \tag{8.1}
\]

where \( a^n = (a^n_1, a^n_2, a^n_3, \ldots, a^n_J) \) is to be determined. We denote the errors obtained in Theorems 3.4 and 3.6 by

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

and

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

respectively.
Example 8.1. We consider the problem \((P_n)\) in \(\Omega \times [0, T]\), where \(\Omega = [0, 1] \times [0, 1]\) and \(T = 1\). For all \((x, y, t) \in \Omega \times [0, T]\), the following data are given:

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 [57].

2. We take the following memory operator as in [17]:

\[
   b_2(t, s, x, y) = -e^{t-s}I; \quad b_1(t, s, x, y) = b_0(t, s, x, y) = 0.
\]
3. Source term

\[ f(x, t) = f_1(x, t) + f_2(x, t) + f_3(x, t) \]

with

\[ f_1(x, t) = \frac{2}{\Gamma(3-\alpha)} t^{2-\alpha} (x-x^2)(y-y^2), \]
\[ f_2(x, t) = 2(x+y-x^2-y^2) \left( t^2 x^2 + t^2 y^2 + \frac{xy^6}{45} - 2t - 2 + 2e^t \right), \]
\[ f_3(x, t) = t^2(2x-1)(y-y^2) \left( 2x + \frac{yt^4}{45} \right) + t^2(2y-1)(x-x^2) \left( 2y + \frac{xt^4}{45} \right). \]

Corresponding to the above data, the exact solution of the problem \((P^\alpha)\) 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 \(\Omega\) [17].

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

\[ \text{Convergence rate} = \begin{cases} \log(E(\tau, h_1)/E(\tau, h_2)) & \text{in space direction} \\ \log(h_1/h_2) & \text{in time direction} \end{cases} \tag{8.2} \]

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

**Linearized L1 Galerkin FEM:** This numerical scheme \((E^\alpha)\) provides a convergence rate of \(O(h + k^{2-\alpha})\) (3.16) as in Theorem 3.4. To observe this order of convergence numerically, we run the MATLAB code at different iterations by setting \(h \approx k^{2-\alpha}.\) Here, \(h\) is taken as the area of the triangle in the triangulation of domain \(\Omega.\) 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 up to five iterations to support our theoretical estimates.

From Tables 1, 2, and 3, we conclude that the convergence rate is first order in the space direction and \((2 - \alpha)\) in the time direction as predicted in Theorem 3.4. We also observe that as \(\alpha \to 1,\) then convergence rate is approaching to first order in the time direction that coincides with the results established in [55] for classical diffusion case.
TABLE 1  Errors and convergence rates in the space and time directions 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  Errors and convergence rates in the space and time directions 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  Errors and convergence rates in the space and time directions 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 | 1.0368        | 1.2960       |
| 5             | 1.29e-03 | 1.0331        | 1.2914       |

TABLE 4  Errors and convergence rates in the space and time directions 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       |

Remark 3. In the numerical experiments, we have chosen $h \approx k^{2-\alpha}$ and collected the convergence rates in Tables 1, 2, and 3. From these tables, it can be easily observed that the convergence rate in the space direction is 1, which perfectly match with the theoretical result proved in Theorem 3.4. To see the convergence rates in the time direction, we use the following formula:

$$\text{Convergence rate in the time direction} = \frac{\log(E(k_1, h)/E(k_2, h))}{\log(k_1/k_2)}. \quad (8.3)$$

Using the relation $h \approx k^{2-\alpha}$ in (8.3), we get

$$\text{Convergence rate in the time direction} = \frac{\log \left( E \left( h_1^{1-\alpha}, h_1 \right) / E \left( h_2^{1-\alpha}, h_2 \right) \right)}{\log \left( h_1 / h_2 \right)^{(2-\alpha)}}$$

$$= (2 - \alpha) \frac{\log \left( E \left( h_1^{1-\alpha}, h_1 \right) / E \left( h_2^{1-\alpha}, h_2 \right) \right)}{\log(h_1 / h_2)}$$

$$= (2 - \alpha) \text{ (convergence rate in space)} = (2 - \alpha). \quad (8.4)$$

This convergence rate in the time direction is the same as that we have proved in Theorem 3.4.

Now, we see the errors and convergence rates for the linearized L2-1$\alpha$ Galerkin FEM ($F^{\alpha}$).

**Linearized L2-1$\alpha$ Galerkin FEM:** This numerical scheme has theoretical convergence rate of $O(h + k^2)$ (3.31) as proved in Theorem 3.6. Here, we set $h \approx k^2$ to conclude the convergence rates in the space and time directions. Here, iteration numbers have the same meaning as for L1 Galerkin FEM ($F^{\alpha}$). From the Tables 4, 5, and 6, we deduce that the convergence...
TABLE 5 Errors and convergence rates in the space and time directions 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 Errors and convergence rates in the space and time directions 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       |

FIGURE 4 Approximate solution (L.H.S) and exact solution (R.H.S) at $T = 1$ for $\alpha = 0.5$. [Colour figure can be viewed at wileyonlinelibrary.com]

rate is first order in the space direction and second order in the time direction, which coincide with the estimate proved in Theorem 3.6.

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

Remark 4. In the proofs of convergence estimates of the developed numerical schemes, we have assumed that the solution of the problem (P$^*$) is sufficiently smooth. If the solution is not smooth enough, then we may have a loss of accuracy in the time direction. To recover this loss of accuracy, a non-uniform mesh in the time direction is used as in [56]. We can design the numerical schemes on this non-uniform mesh [56], which compensates the loss of accuracy caused by the non-smoothness of the solution.

Remark 5. The convergence estimates for the proposed numerical schemes hold under some restrictions on the time step that makes these schemes conditionally stable and conditionally convergent. In order to get rid of these restrictions, a temporal-spatial error splitting scheme has been adopted in [33]. We are exploring this new technique to prove unconditional convergence estimates for the proposed numerical schemes in this work.
9 | CONCLUSIONS

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

AUTHOR CONTRIBUTIONS

Lalit Kumar: Conceptualization; investigation; methodology; formal analysis; software; validation; writing—review and editing; writing—original draft; resources. Sivaji Ganesh Sista: Supervision; formal analysis; methodology; validation; writing—review and editing; writing—original draft; investigation; conceptualization; resources. Konijeti Sreenadh: Conceptualization; investigation; writing—original draft; methodology; validation; writing—review and editing; formal analysis; supervision; resources.

CONFLICT OF INTEREST STATEMENT

This work does not have any conflicts of interest.

ORCID

Konijeti Sreenadh https://orcid.org/0000-0001-7953-7887

REFERENCES

1. R. Zacher, Weak solutions of abstract evolutionary integro-differential equations in Hilbert spaces, Funkcialaj Ekvacioj 52 (2009), no. 1, 1–18.
2. A. Kubica and M. Yamamoto, Initial-boundary value problem for fractional diffusion equations with time-dependent coefficients, Fractional Calculus Appl. Anal. 21 (2018), no. 2, 112–125.
3. F. Mainardi, Fractional calculus and waves in linear viscoelasticity: an introduction to mathematical models, World Scientific, 2010.
4. J. Manimaran, L. Shangerganesh, A. Debbouche, and V. Antonov, Numerical solutions for time-fractional cancer invasion system with nonlocal diffusion, Front. Phys. 7 (2019), 93.
5. R. Metzler and J. Klafter, The random walk’s guide to anomalous diffusion: a fractional dynamics approach, Phys. Rep. 339 (2000), no. 1, 1–77.
6. A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, Theory and applications of fractional differential equations, Elsevier, 2006, pp. 204.
7. 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.
8. J. P. Bouchaud and A. Georges, Anomalous diffusion in disordered media: statistical mechanics, models and physical applications, Phys. Rep. 195 (1990), no. 4-5, 127–293.
9. B. Ross, Fractional calculus and its applications: proceedings of the international conference held at the University of New Haven, June 1974, Springer, 2006, pp. 457.
10. J. L. Lions, On some questions in boundary value problems of mathematical physics, North-Holland Math. Stud. 30 (1978), 284–346.
11. J. Giacomoni, V. D. Rădulescu, and G. Warnault, Quasilinear parabolic problem with variable exponent: qualitative analysis and stabilization, Commun. Contemp. Math. 20 (2018), no. 8, 1750065.
12. R. Shen, M. Xiang, and V. D. Rădulescu, Time-space fractional diffusion problems: existence, decay estimates and blow-up of solutions, Milan J. Math. 90 (2022), no. 1, 103–129.
13. M. Xiang, V. D. Rădulescu, and B. Zhang, Nonlocal Kirchhoff diffusion problems: local existence and blow-up of solutions, Nonlinearity 31 (2018), no. 7, 3228.
14. D. Goel and K. Sreenadh, Kirchhoff equations with Hardy-Littlewood-Sobolev critical nonlinearity, Nonlinear Anal. 186 (2019), 162–186.
15. V. D. Rădulescu and C. Vetro, Anisotropic Navier Kirchhoff problems with convection and Laplacian dependence, Math. Methods Appl. Sci. 46 (2023), no. 1, 461–478.
16. X. Mingqi, V. D. Rădulescu, and B. Zhang, A critical fractional Choquard-Kirchhoff problem with magnetic field, Commun. Contemp. Math. 21 (2018), no. 4, 1850004.
17. S. Barbeiro, J. A. Ferreira, and L. Pinto, $H^2$ second-order convergent estimates for non-Fickian models, Appl. Numer. Math. 61 (2011), no. 2, 201–215.
55. L. Kumar, S. G. Sista, and K. Sreenadh, *Finite element analysis of parabolic integro-differential equations of Kirchhoff type*, Math. Methods Appl. Sci. **43** (2020), no. 15, 9129–9150.

56. M. Li, C. Huang, and F. Jiang, *Galerkin finite element method for higher dimensional multi-term fractional diffusion equation on non-uniform meshes*, Appl. Anal. **96** (2017), no. 8, 1269–1284.

57. L. A. Medeiros, T. N. Rabello, M. A. Rincón, and M. C. C. Vieira, *On perturbation of the Kirchhoff operator: analysis and numerical simulation*, Commun. Math. Sci. **10** (2012), no. 3, 751–766.

**How to cite this article:** L. Kumar, S. G. Sista, and S. Konijeti, *Finite element analysis of time-fractional integro-differential equation of Kirchhoff type for non-homogeneous materials*, Math. Meth. Appl. Sci. **47** (2024), 2120–2153, DOI 10.1002/mma.9737.