Superconvergence analysis of FEM for a multidimensional time fractional integro-differential problem with weakly singular kernels

[†]Dilip Sarkar[1], Pratibhamoy Das[1], Ricardo Ruiz-Baier[2,3]

[1]Department of Mathematics, Indian Institute of Technology Patna, Patna, 801106, Bihar, India.

[2]School of Mathematics, Monash University, 9 Rainforest Walk, Melbourne, 3800, Victoria, Australia.

[3]Universidad Adventista de Chile, Casilla 7-D, Chillán, Chile.

Contributing authors: dilip_2321ma07@iitp.ac.in; pratibhamoy@iitp.ac.in; ricardo.ruizbaier@monash.edu;

[†]All authors contributed equally to this work.

Abstract

We analyze the superconvergence approach of a multi-term time-fractional weakly singular integro diffusion problem and the spatial domain belongs to $\mathbb{R}^{\boldsymbol{d}}$ $(\boldsymbol{d = 1, 2, 3})$. A finite element approach based on mapped piecewise bilinear elements is employed to approximate the spatial derivatives. The $\boldsymbol{L1}$ scheme is used for the fractional-derivative terms, and an interpolation-based approximation is employed for the singular kernels on a temporally graded mesh. A new discrete fractional Gronwall inequality is proved to establish optimal error bounds in $\boldsymbol{L^{\infty}(H^1)}$ norm. For axis-aligned rectangular $(\boldsymbol{d = 2})$ or cubical $(\boldsymbol{d = 3})$ meshes as well as for general simplicial 2D or 3D meshes, a simple postprocessing step is proposed to enhance spatial accuracy with less regularity assumptions on the solution. Numerical experiments confirm the theoretical convergence rates of high accuracy for multidimensional problems, including 3D problems.

Keywords: Finite element methods, Discrete Gronwall inequality, Product interpolation, $\boldsymbol{L1}$ schemes, $\boldsymbol{H^1}$ stability analysis, Superconvergence, Weakly singular kernels, Higher dimensional problems (2D & 3D)

MSC Classification: 65M12 , 65M15 , 65M60 , 35R09 , 35R11 , 45K05

1 Introduction

Fractional integro-differential equations are essential for describing memory and hereditary effects, such as in certain regimes of viscoelasticity [1], bioengineering [2], heat transfer, and fluid dynamics [3]. A vast literature exists on numerical techniques for fractional differential equations; however, for the specific case of multi-term fractional derivatives combined with multi-term fractional-order integral operators, contributions are comparatively scarce. Further, the authors discussed different discretizations for the fractional operator. Chen *et al.* [4] employed the hp-DG method for both the fractional integral and the derivative to establish a priori error bounds. In [5], the authors investigate various fractional models, including time-fractional and space-time fractional models, and present their numerical treatments using

the spectral method. For more details on the fractional operator, see, for example, [6–10] and the references therein. In [11, 12], the authors employed a finite difference scheme for spatial discretization and the so-called $L1$ scheme [13] for the Caputo fractional derivative. For the fractional integral terms, the trapezoidal rule and the midpoint Euler method were utilized. The proposed scheme achieves second-order spatial accuracy in the maximum norm and a temporal convergence rate of order $\mathcal{O}(M^{-(2-\alpha)})$ with Caputo fractional dominant order α. Liu *et al.* [14] employed a polynomial interpolation technique to discretize the fractional integral term with weakly singular kernels on a graded mesh, and established a second-order rate of convergence in time. Superconvergence analysis for fractional diffusion problems remains scarce in the existing literature. Huang *et al.* [15] investigate the superconvergence analysis of the time-fractional variable-order subdiffusion problem with a weakly singular solution and established optimal error estimates in the $L^\infty(L^2)$ and $L^\infty(H^1)$ norms. In [16], the authors employed the discontinuous Galerkin method to establish superconvergence results for a one-dimensional time-fractional reaction-diffusion problem with periodic boundary conditions and established optimal error estimates in the L^2 norm and H^1 semi-norm. After that, the authors employed the β-robust superconvergence analysis of the finite element method (FEM) for the distributed-order time-fractional diffusion problem in [17]. Based on the existing literature, to the best of our knowledge, no study has considered the combination of multi-term fractional Caputo derivatives with multi-term fractional-order integral operators for the superconvergence analysis.

In the present work, we consider a multidimensional diffusion problem with multi-term time-fractional Caputo derivatives and multi-term fractional order integral terms with weakly singular kernels. We discretize the Caputo fractional derivative using the $L1$-scheme on a graded mesh in time. Based on the graded mesh component (parameter) $\sigma \geq (2-\alpha_1)/\alpha_1$, we show theoretically and confirm numerically the optimal rate of convergence $\mathcal{O}(M^{-(2-\alpha_1)})$, where α_1 denotes the dominant fractional order in the Caputo derivative. To discretize the integral term, we employ the product interpolation of the weakly singular kernel on a temporal graded mesh. This approach effectively captures rapid variations in the solution and facilitates the construction of numerical schemes that achieve stable second-order time accuracy.

In addition, a novel discrete Gronwall inequality is established for the proposed problem. As an application of this result, the stability of the numerical scheme is rigorously analyzed in the $H^1(\Omega)$ norm, based on polynomial interpolations. In the convergence analysis, we introduce a post-processing operator defined on both macroelements and patch-based macroelements. This helps us establish a superconvergence result and obtain a second-order rate of convergence in space with respect to the $H^1(\Omega)$ norm. In addition, we discuss how to derive convergence results with a lower regularity assumption.

The remainder of this work has been structured in the following manner. In Section 2 we provide the precise statement of the governing equations, introduce the required notation for the paper, and state preliminary properties of space discretizations. In Section 3 we define the time discretization of the Caputo fractional derivatives using the $L1$ scheme and product interpolation. The time-advancing scheme is combined with a finite element discretization in Section 4, where we also present the proof of an auxiliary Gronwall inequality. We show stability of the fully discrete scheme and the corresponding convergence rates. The proof of superconvergence is discussed briefly in Section 5 with low regularity assumption in the convergence analysis, and we provide a set of illustrative numerical examples in Section 6.

2 Problem statement and preliminaries

Let $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ be a bounded polygonal domain. Consider a fixed $T > 0$. We define the multi-term time fractional-integral weakly singular diffusion problem as:

$$\begin{aligned} &\sum_{i=1}^{k} {}^c\mathcal{D}_t^{\alpha_i} u(\boldsymbol{x},t) - \Delta u(\boldsymbol{x},t) - \gamma \sum_{r=1}^{p} \mathcal{I}^{\lambda_r} u(\boldsymbol{x},t) = f(\boldsymbol{x},t), \qquad (\boldsymbol{x},t) \in \mathfrak{D} := \Omega \times (0,T], \\ &u(\boldsymbol{x},0) = u_0, \quad \boldsymbol{x} \in \Omega \quad \text{and } u|_{\partial\Omega} = 0, \quad t > 0, \end{aligned} \tag{2.1}$$

where $\gamma > 0$ is a constant, and the weakly singular kernel is defined as $\mathcal{I}^{\lambda_r} u(\boldsymbol{x},t) = \int_0^t \frac{(t-z)^{\lambda_r - 1}}{\Gamma(\lambda_r)} u(\boldsymbol{x},z)\,\mathrm{d}z$, where $0 < \lambda_r < \alpha_1 < 1$ holds for all $r = 1, \ldots, p$. For fixed $\alpha_i \in (0,1)$, the fractional order Caputo derivative is defined for all $u(\boldsymbol{x},t)$ that are absolutely continuous on the time

domain $[0,T]$ as

$$ {}^{c}\mathcal{D}_t^{\alpha_i} u(\boldsymbol{x},t) := \frac{1}{\Gamma(1-\alpha_i)} \int_0^t (t-z)^{-\alpha_i} \frac{\partial u(\boldsymbol{x},z)}{\partial z}\, \mathrm{d}z. $$

The fractional orders $\alpha_i,\ \lambda_r$ are in decreasing order for all i and r, respectively, and the functions $f \in C(\overline{\mathfrak{D}})$ and $u_0 \in C(\overline{\Omega})$. ${}^{c}\mathcal{D}_t^{\alpha_i}$ denote the Caputo fractional derivative of order α_i for $i = 1,\ldots,k$. By the concept of the eigenvalues and eigenfunctions, we will define the fractional Laplacian domain, where eigenvalues (β_q) and eigenfunctions (φ_q) satisfy the problem $\Delta\varphi_q = \beta_q\varphi_q$, on Ω with $\varphi_q = 0$, on $\partial\Omega$ for $q = 1,2,\ldots$. Therefore, from the sectorial theory [18, 19], we get the fractional domain

$$ \mathcal{D}(\Delta^w) = \{\vartheta \in L^2(\Omega) : \sum_{q=1}^{\infty} \beta_q^{2w} |(\vartheta,\varphi_q)|^2 < \infty\}, $$

for each $w \in \mathbb{R}$ and the corresponding norm is $\|\vartheta\|^2_{\Delta^w} := \sum\limits_{q=1}^{\infty} \beta_q^{2w} |(\vartheta,\varphi_q)|^2$, where $(\vartheta,\varphi_q) = \int_\Omega \vartheta\,\varphi_q\, \mathrm{d}\boldsymbol{x}$ for $d = 1,2,3$.

Theorem 2.1. *Let $u_0 \in \mathcal{D}(\Delta^{s+2})$, $\partial_t^v f(\bullet,t) \in \mathcal{D}(\Delta^s)$ and $\|u_0\|_{\Delta^{s+2}} + \|\partial_t^v f(\bullet,t)\|_{\Delta^{s+1}} \le K$, for $v = 1,2$, where the integer $s > 0$ and K is independent of v and t. Then the solution of (2.1) is unique and it satisfies the bounds $\|\partial_t^v u(\bullet,t)\|_s \le C(1 + t^{\alpha_1 - v})$ for $v = 0,1,2$, and $\|{}^{c}\mathcal{D}_t^{\alpha_i} u(\bullet,t)\|_s \le C,\ \forall i$ and $t \in (0,T]$.*

Proof By employing separation of variables and eigenvalue expansions, the above bounds can be derived for domains $\Omega \subset \mathbb{R}$ (see, e.g., [20]). On the other hand, a similar technique as in [13] can be used to extend the argument to domains $\Omega \subset \mathbb{R}^d$, for $d = 2,3$. □

Note 2.1. *Theorem 2.1 shows that the fractional integral terms $\mathcal{I}^{\lambda_r} u$ satisfy the bound $\|\mathcal{I}^{\lambda_r} u(\bullet,t)\|_s \le C,\quad \forall\, r$ and $\forall\, t \in (0,T]$, where C is a constant independent of r and t.*

Notation: We write $\|\bullet\|$ as the $L^2(\Omega)$ norm and $\|\bullet\|_q$ as the $H^q(\Omega)$ norm for any integer $q > 0$. Let $H_0^1(\Omega)$ be the usual Sobolev space whose traces vanish on the boundary $\partial\Omega$. C is considered to be a generic constant.

We define $\mathcal{Q}_1$ as the space of polynomials on $\Omega \subset \mathbb{R}^d$ with d variables that are linear $(d = 1)$, bilinear $(d = 2)$ and trilinear $(d = 3)$. Let $E = [0,1]^d$ denote the reference element. Consider a quasiuniform partition $\mathcal{T}_h$ of Ω into elements E_e, $e = 1,\ldots,N$, where each element is obtained from E by a mapping $\mathfrak{f} \in \mathcal{Q}_1$, *i.e.*, $E_e = \mathfrak{f}(E)$. Now, we construct the discretize space by the mapping of $\mathcal{Q}_1$ functions on E_e (see [21]). Thus

$$ \mathcal{V}_h := \left\{ v_h \in H^1(\Omega);\ v_h|_{E_e} = \mathfrak{T} \circ \mathfrak{f}^{-1},\ \text{where } \mathfrak{f} : \hat{E} \to E_e \text{ and } \mathfrak{T} \in \mathcal{Q}_1(\hat{E}) \right\}. $$

Based on the space $\mathcal{V}_h$, we can define the space $\mathcal{V}_{0h} = \{v_h \in \mathcal{V}_h \text{ with } v_h|_{\partial\Omega} = 0\}$ where $h := \max\limits_{e=1}^{N}(\mathrm{diam}(E_e))$.

Now, we define some basic operators: Let P_h be the L^2 projection from $L^2(\Omega)$ to $\mathcal{V}_{0h}$. Then it satisfies the inequality

$$ \|\nabla P_h v\| \le \mathbf{C}\|v\|, \tag{2.2} $$

for all $v \in H_0^1(\Omega)$. Let R_h be the Ritz projection from $H_0^1(\Omega)$ to $\mathcal{V}_{0h}$. In addition

$$ \|u - R_h u\| + h\|u - R_h u\| \le Ch^2 |u|_2,\ \forall u \in H^2(\Omega) \cap H_0^1(\Omega). \tag{2.3} $$

We define the discrete Laplacian $\Delta_h : \mathcal{V}_{0h} \to \mathcal{V}_{0h}$, by

$$ (\Delta_h u, v) = -(\nabla u, \nabla v),\ \forall u, v \in \mathcal{V}_{0h}. \tag{2.4} $$

From above, we can write $\Delta_h R_h u = P_h u$ for all $u \in H^1(\Omega)$ (see, e.g., [22]). Let us consider the interpolation operator $\mathscr{I}_h : H^2(\Omega) \to \mathcal{V}_{0h}$ defined as $\mathscr{I}_h v(b_i) = v(b_i)$ where b_i are the vertices of the element $E_e \in \mathcal{T}_h$ for $i = 1,2,\ldots,2^d$. The following convergence holds

$$ \|R_h u - \mathscr{I}_h u\|_1 \le Ch^2 \|u\|_3,\ \forall\ u \in H_0^1(\Omega) \cap H^3(\Omega). $$

3 Continuous-in-space discrete problem

Now, we discretize the time domain $[0,T]$ using the graded mesh technique. Let $M>0$ be an integer. To tackle the singularity near $t=0$, we use the temporal mesh points as $t_m = T(m/M)^\sigma$, for $m=0,,M$, where $\sigma \geq 1$ is a user-chosen graded mesh component. Let us consider the step size $\hat{\tau}_m = t_m - t_{m-1}$ for $m=1,\ldots,M$. It is easy to see that $\hat{\tau}_m \leq Cm^{-1}t_m$, for $m=1,\ldots,M$. We employ the $L1$ scheme for the fractional derivative and the product interpolation (PI) technique for integral terms. Therefore, the discretization form of the Caputo fractional derivative is

$$\begin{aligned}
{}^c\mathcal{D}_t^{\alpha_i} u(\boldsymbol{x},t) &= \frac{1}{\Gamma(1-\alpha_1)} \sum_{j=0}^{m-1} \frac{u^{j+1}-u^j}{\tau_{j+1}} \int_{t_j^{j+1}} (t_m - z)^{-\alpha_i}\, \mathrm{d}z \\
&= {}^c\mathcal{D}_M^{\alpha_i} u^m + \mathcal{R}_1^m \\
&= \frac{1}{\Gamma(2-\alpha_1)} \sum_{j=0}^{m-1} \frac{u^{j+1}-u^j}{\tau_{j+1}} [(t_n - t_j)^{1-\alpha_i} - (t_n - t_{j+1})^{1-\alpha_i}] \\
&= \xi_{m,1}^i u^m - \xi_{m,m}^i u^0 + \sum_{j=1}^{m-1} (\xi_{m,j+1}^i - \xi_{m,j}^i) u^{m-j} + \mathcal{R}_1^m, \qquad (3.1)
\end{aligned}$$

where $\xi_{m,j}^i = \frac{d_{m,j}^i}{\Gamma(1-\alpha_i)}$ and $d_{m,j}^i = [(t_m - t_{m-j})^{1-\alpha_i} - (t_m - t_{m-j+1})^{1-\alpha_i}]/\tau_{m-j+1}$ for $j=1,\ldots,m$. We define the second-order backward difference formula (BDF2) on the graded mesh to discretize the integral terms $\mathcal{I}^{\lambda_r}$. By [8], the linear multi-step (K-step) method can be written as $\rho(\phi) = \sum_{i=0}^{K} a_i \phi^{K-i}$. Therefore, we discretize the integral terms on the graded mesh (see [23, 24]) as follows:

$$\begin{aligned}
\mathcal{I}^\lambda \phi(t_m) &= \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} \frac{(t_m - z)^{\lambda-1}}{\Gamma(\lambda)} [\frac{z - t_{m-1}}{\hat{\tau}_m} \phi(t_m) + \frac{t_m - z}{\hat{\tau}_m} \phi(t_{m-1})]\, \mathrm{d}z + \mathcal{R}_2^m, \\
&= \frac{\tau_1^\lambda}{\Gamma(2+\lambda)} \left(w_m^\lambda \phi(t_0) + \sum_{j=1}^{m-1} w_{m,j}^\lambda \phi(t_j) + w_{m,m}^\lambda \phi(t_m) \right) + \mathcal{R}_2^m, \ \lambda \in (0,1), \ 1 \leq m \leq M, \qquad (3.2)
\end{aligned}$$

where the corresponding weights are

$$w_m^\lambda = (m^\sigma - 1)^{\lambda+1} - m^{\sigma\lambda}(m^\sigma - \lambda - 1), \quad w_{m,j}^\lambda = \psi_{m,j}(\lambda,\sigma) - \psi_{m,j+1}(\lambda,\sigma)$$

and

$$w_{m,m}^\lambda = (m^\sigma - (m-1)^\sigma)^\lambda,$$

with

$$\psi_{m,j}(\lambda,\sigma) = \frac{(m^\sigma - (j-1)^\sigma)^{\lambda+1} - (m^\sigma - j^\sigma)^{\lambda+1}}{j^\sigma - (j-1)^\sigma}.$$

Therefore, $\mathcal{I}^\lambda \phi(t_m) = \mathcal{I}_M^\lambda \phi(t_m) + \mathcal{R}_2^m$.

The following lemma provides bounds on the truncation error terms $\mathcal{R}_1^m$ and $\mathcal{R}_2^m$ associated with the $L1$ and PI schemes (3.1)-(3.2). This result will be instrumental in the subsequent post-processing analysis.

Lemma 3.1 (Truncation error estimates [14, 24] (see also, e.g., [11, 12])). *Assume the solution $u(\bullet,t)$ of (2.1) satisfies the hypotheses of Theorem 2.1. Then*

$$|\mathcal{R}_2^m| + |\mathcal{R}_1^m| \leq C \left(M^{-\min(\sigma\lambda,\ 2)} + M^{-\min(\sigma\alpha_1,\ 2-\alpha_1)} \right), \quad 1 \leq m \leq M.$$

Remark 3.1. *On a uniform temporal mesh, the $L1$ scheme satisfies the estimate (see [25, 26]) $|\mathcal{R}_1^m| \leq C\,\tau\, t_m^{-1}$. Moreover, when a polynomial interpolation (PI) scheme is applied to approximate the integral term on a uniform mesh, the resulting formulation reduces to the classical midpoint (Euler) scheme, as discussed in [12]. Consequently, the temporal convergence rate on a uniform mesh is of order $\mathcal{O}(M^{-1})$. Hence, both the $L1$ scheme and the PI scheme achieve first-order accuracy in time on a uniform mesh. This is confirmed numerically in the computational examples reported later in Table 6.4.*

Lemma 3.2. *Let us consider $u^m \in L^2(\Omega)$ and $1 \le m \le M$. Then, the following estimate holds true*

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} u^m - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} u^m, u^m\right) \ge \left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \|u^m\| - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \|u^m\|\right) \|u^m\|.$$

Proof Let $m \in \{1, 2, \ldots, M\}$. The definition of ${}^c\mathcal{D}_M^{\alpha_i} u^m$ is as follows:

$$\left({}^c\mathcal{D}_M^{\alpha_i} u^m, u^m\right) = \frac{d^i_{m,1}}{\Gamma(2-\alpha)}(u^m, u^m) - \frac{d^i_{m,m}}{\Gamma(2-\alpha)}(u^0, u^m) - \frac{1}{\Gamma(2-\alpha)} \sum_{j=1}^{m-1} (d^i_{m,j} - d^i_{m,j+1})(u^{m-j}, u^m).$$

Since $d^i_{m,j} - d^i_{m,j+1} > 0$, by using the Cauchy–Schwarz inequality, we obtain

$$\left({}^c\mathcal{D}_M^{\alpha_i} u^m, u^m\right) \ge \frac{d^i_{m,1}}{\Gamma(2-\alpha)}\|u^m\|^2 - \frac{d^i_{m,m}}{\Gamma(2-\alpha)}\|u^0\|\,\|u^m\| - \frac{1}{\Gamma(2-\alpha)} \sum_{j=1}^{m-1} (d^i_{m,j} - d^i_{m,j+1})\|u^{m-j}\|\,\|u^m\|.$$

Thus, we get $\left(\sum\limits_{i=1}^{k} \mathcal{D}_M^{\alpha_i} u^m, u^m\right) \ge \left(\sum\limits_{i=1}^{k} \mathcal{D}_M^{\alpha_i} \|u^m\|\right) \|u^m\|$. By the definition of the PI-scheme (3.2), the integral term contribution can be written as follows:

$$\left(\mathcal{I}_M^{\lambda_r} u^m, u^m\right) = \frac{\tau_1^{\lambda_r}}{\Gamma(2+\lambda_r)} \left(w_m^{\lambda_r} \left(u^0, u^m\right) + \sum_{j=1}^{m-1} w_{m,j}^{\lambda_r} \left(u^j, u^m\right) + w_{m,m}^{\lambda_r} (u^m, u^m) \right).$$

Therefore, $\mathcal{I}_M^{\lambda_r} u^m = \frac{\tau_1^{\lambda_r}}{\Gamma(2+\lambda_r)} \left(w_m^{\lambda_r} u^0 + \sum\limits_{j=1}^{m-1} w_{m,j}^{\lambda_r} u^j + w_{m,m}^{\lambda_r} u^m \right)$. Finally, noting that $w_j^{\lambda_r} \ge 0$ for all j, we can obtain

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} u^m - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} u^m, u^m\right) \ge \left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \|u^m\| - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \|u^m\|\right) \|u^m\|,$$

which completes the proof. □

4 Space-time discretization using $L1$-PI finite element methods

We now consider (2.1) and apply a finite element discretization on a uniform spatial mesh. The continuous weak form of (2.1) is: find $u(\bullet, t) \in H_0^1(\Omega)$ for $t \in (0, T]$ such that

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_t^{\alpha_i} u, v\right) + (\nabla u, \nabla v) - \gamma \left(\sum_{r=1}^{p} \mathcal{I}^{\lambda_r} u, v\right) = (f, v), \quad \forall v \in H_0^1(\Omega), \text{ and } u(\boldsymbol{x}, 0) = u_0, \tag{4.1}$$

where the inner product is defined as $(\nabla u, \nabla v)_{L^2(\Omega)} := \int_\Omega \sum\limits_{i=1}^{d} \dfrac{\partial u}{\partial x_i} \dfrac{\partial v}{\partial x_i} \, \mathrm{d}\boldsymbol{x}$. From (4.1), our semi-discrete problem is: find $u_h(\bullet, t) \in \mathcal{V}_{0h}$ with $t \in (0, T]$ such that

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_t^{\alpha_i} u_h, v_h\right) + (\nabla u_h, \nabla v_h) - \gamma \left(\sum_{r=1}^{p} \mathcal{I}^{\lambda_r} u_h, v_h\right) = (f, v_h), \qquad \forall v_h \in \mathcal{V}_{0h}, \quad \text{with } u_h^0 = R_h u_0.$$

Therefore, combining the equations (3.1)-(3.2), the fully discrete method ($L1$-PI-FEM) corresponds to: find $u_h^m \in \mathcal{V}_{0h}$ with $t \in (0, T]$ and $m = 1, \ldots, M$, such that

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} u_h^m, v_h\right) + (\nabla u_h^m, \nabla v_h) - \gamma \left(\sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} u_h^m, v_h\right) = (f^m, v_h), \quad \forall v_h \in \mathcal{V}_{0h} \quad \text{with } u_h^0 = R_h u_0. \tag{4.2}$$

It is easy to see that $P_h f^m = f^m$, and $\mathcal{D}_N^{\alpha_i} u_h^m$, $\Delta_h u_h^m$, $\mathcal{I}_M^{\lambda_r} u_h^m$ and $P_h f^m$ are in the space $\mathcal{V}_{0h}$. By using the definition of Δ_h and the fully discrete form ($L1$-PI-FEM) (4.2), we can write more tersely as:

$$\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} u_h^m - \Delta_h u_h^m - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} u_h^m = P_h f^m, \quad m = 1, \ldots, M, \text{ with } u_h^0 = R_h u_0. \tag{4.3}$$

Now, we prove the discrete Gronwall inequality for the more general multi-term fractional-integro diffusion problem with weakly singular kernels. The proof follows the approach from [13], and the result will be employed in the forthcoming convergence analysis.

Theorem 4.1 (A generalized Gronwall inequality). *Consider* $\{\chi^m\}_{m=1}^M \geq 0$ *and* $\{\eta^m\}_{m=1}^M \geq 0$. *Let* $\Phi^0 \geq 0$ *and* $\{\Phi^m\}_{m=1}^M$ *satisfy*

$$\Phi^m \left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \right) \Phi^m \leq \chi^m \Phi^m + (\eta^m)^2 \quad \textit{for} \quad m = 1, \ldots, M.$$

Then

$$\Phi^m \leq \Phi^0 + \Gamma(1-\alpha_1) \max_{1\leq l\leq m} \left(t_l^{\alpha_1} \chi^l \right) + \sqrt{\Gamma(1-\alpha_1)} \left\{ \max_{1\leq l\leq m} (t_l^{\alpha_1})^{1/2} (\eta^l) \right\}, \quad \textit{for} \quad m = 1, \ldots, M.$$

Proof Let us consider $U^l = \Phi^l - \Phi^0$ for $l = 0, \ldots, M$. Consider $U^0 = 0$, and the function $\Pi^l = \sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} U^l - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} U^l$. Let us consider that the $\max_{1\leq j\leq m} U^j = U^\vartheta$ for fixed value $\vartheta \in \{1, \ldots, m\}$ and $m \in \{1, \ldots, M\}$. Therefore, we can write $\max_{1\leq j\leq m} U^j = \Phi^\vartheta$ and $U^0 = 0$. Thus

$$\begin{aligned}
\Pi^\vartheta &= \sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} U^\vartheta - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} U^\vartheta \\
&= \sum_{i=1}^{k} \left(\xi_{\vartheta,1}^i U^\vartheta + \sum_{j=1}^{\vartheta-1} (\xi_{\vartheta,j+1}^i - \xi_{\vartheta,j}^i) U^{\vartheta-j} \right) - \gamma \sum_{r=1}^{p} \left(w_{\vartheta,\vartheta}^r U^\vartheta + \sum_{j=1}^{\vartheta-1} w_{\vartheta,j}^r U^j \right) \\
&\geq \sum_{i=1}^{k} \left(\xi_{\vartheta,1}^i U^\vartheta + \sum_{j=1}^{\vartheta-1} (\xi_{\vartheta,j+1}^i - \xi_{\vartheta,j}^i) U^\vartheta \right) - \gamma \sum_{r=1}^{p} \left(w_{\vartheta,\vartheta}^r U^\vartheta + \sum_{j=1}^{\vartheta-1} w_{\vartheta,j}^r U^\vartheta \right) \\
&= \left(\sum_{i=1}^{k} \xi_{\vartheta,\vartheta}^i - \gamma \sum_{r=1}^{p} \psi_{\vartheta,1}^r(\lambda_r, \sigma) \right) U^\vartheta, \qquad (4.4)
\end{aligned}$$

where we have used $\omega_{\vartheta,j}^r > 0$ and $\xi_{\vartheta,j+1}^i - \xi_{\vartheta,j}^r < 0$ for r, j. Therefore, from equation (4.4), we can write it as $U^\vartheta \leq \Pi^\vartheta / \left(\sum_{i=1}^{k} \xi_{\vartheta,\vartheta}^i - \gamma \sum_{r=1}^{p} \psi_{\vartheta,\vartheta}^r(\lambda_r, \sigma) \right)$. Note that $U^\vartheta < 0$ for $\Pi^\vartheta < 0$. This is trivial to prove. Our aim is to show the nontrivial case for $\Pi^\vartheta \geq 0$. From (4.4), we can express

$$U^\vartheta \leq \frac{\Pi^\vartheta}{\xi_{\vartheta,\vartheta}^1 - \gamma\psi_{\vartheta,\vartheta}^1} = \frac{\Gamma(2-\alpha_1)\Pi^\vartheta}{d_{\vartheta,\vartheta}^1 - \Gamma(2-\alpha_1)\gamma\psi_{\vartheta,\vartheta}^1} \leq \frac{\Gamma(2-\alpha_1)\Pi^\vartheta}{(1-\alpha_1)t_\vartheta^{-\alpha_1} - \Gamma(2-\alpha_1)\gamma\psi_{\vartheta,\vartheta}^1} = \frac{\Gamma(1-\alpha_i)\Pi^\vartheta t_\vartheta^{\alpha_1}}{1 - \Gamma(1-\alpha_1)\gamma t_\vartheta^{\alpha_1}\psi_{\vartheta,\vartheta}^1},$$

since $d_{\vartheta,\vartheta}^1 \geq (1-\alpha_1)t_\vartheta^{-\alpha_1}$. Taking the multiple on both sides by Φ^ϑ, we get

$$\Phi^\vartheta U^\vartheta \leq \Gamma(1-\alpha_1)\Pi^\vartheta t_\vartheta^{\alpha_1} \Phi^\vartheta, \text{ when } \gamma < 1/(\Gamma(1-\alpha_1)t_\vartheta^{\alpha_1}\psi_{\vartheta,\vartheta}^1).$$

By using the maximum of Φ^l, U^l, Π^l for $l = 1, \ldots, m$ with a fixed value $\vartheta \in \{1, \ldots, m\}$, we get

$$\begin{aligned}
\max_{1\leq l\leq m} \Phi^l \max_{1\leq l\leq m} (\Phi^l - \Phi^0) &\leq \Gamma(1-\alpha_1) t_\vartheta^{\alpha_1} \Pi^\vartheta \Phi^\vartheta \\
&\leq \Gamma(1-\alpha_1) \max_{1\leq l\leq m} \left\{ t_l^{\alpha_1} \Phi^l \left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \right) \Phi^l \right\} \\
&\leq \Gamma(1-\alpha_1) \max_{1\leq l\leq m} \left\{ t_l^{\alpha_1} \left(\chi^l \Phi^l + (\eta^l)^2 \right) \right\} \\
&\leq \Gamma(1-\alpha_1) \max_{1\leq l\leq m} \left(t_l^{\alpha_1} \chi^l \Phi^l \right) + \Gamma(1-\alpha_1) \left\{ \max_{1\leq l\leq m} (t_l^{\alpha_1})^{1/2} (\eta^l) \right\}^2 \\
&\leq \Gamma(1-\alpha_1) \max_{1\leq l\leq m} \left(t_l^{\alpha_1} \chi^l \right) \left(\max_{1\leq l\leq m} \Phi^l \right) + \Gamma(1-\alpha_1) \left\{ \max_{1\leq l\leq m} (t_l^{\alpha_1})^{1/2} (\eta^l) \right\}^2.
\end{aligned}$$

Therefore, $\Phi^m \le \Phi^0 + \Gamma(1-\alpha_1) \max_{1\le l\le m} \left(t_l^{\alpha_1}\xi^l\right) + \sqrt{\Gamma(1-\alpha_1)}\left\{ \max_{1\le l\le m} (t_l^{\alpha_1})^{1/2}(\eta^l)\right\}$, for $m = 1, \ldots, M$. □

4.1 Stability of the *L*1-PI-FEM

Lemma 4.2 (Local stability in $H^1(\Omega)$)**.** *Let u_h^m be the solution of the problem (4.3). Then*

$$\|u_h^m\|_{H^1(\Omega)} \le \|u_h^0\|_{H^1(\Omega)} + \Gamma(1-\alpha_1)T^{\alpha_1} \max_{1\le l\le m} \|\nabla f^l\|, \text{ for } m = 1, \ldots, M.$$

Proof Let us consider the weak form of (4.3), for a fixed $m = 1, \ldots, M$,

$$-\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} u_h^m, \Delta_h u_h^m\right) + \|\Delta_h u_h^m\|^2 + \gamma\left(\sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} u_h^m, \Delta_h u_h^m\right) = -(P_h f^m, \Delta_h u_h^m).$$

Since $\|\Delta_h u_h^m\|^2$ is non negative, by the definition of the discrete Laplacian Δ_h, we get

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} (\nabla u_h^m), \nabla u_h^m\right) - \gamma\left(\sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} (\nabla u_h^m), \nabla u_h^m\right) \le (\nabla P_h f^m, \nabla u_h^m).$$

By using Lemma 3.2 and $\|\nabla P_h v\| \le C\|v\|$ with Cauchy–Schwarz inequality, we obtain

$$\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \|\nabla u_h^m\| - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \|\nabla u_h^m\|\right) \|\nabla u_h^m\| \le \|\nabla P_h f^m\| \, \|\nabla u_h^m\| \le C\|\nabla f^m\| \, \|\nabla u_h^m\|.$$

Finally, applying Theorem 4.1, we get $\|u_h^m\|_{H^1(\Omega)} \le \|u_h^0\|_{H^1(\Omega)} + \Gamma(1-\alpha_1)T^{\alpha_1} \max_{1\le l\le m} \|\nabla f^l\|$. □

Consider time $t = t_m$ for $m = 0, 1, \ldots, M$. Note u^m and u_h^m denote the solutions to the problems (4.1) and (4.2), respectively. To carry out the error analysis, we split the error at step m as

$$u^m - u_h^m = \Theta^m - \Upsilon^m := R_h u^m - u_h^m + R_h u^m - u^m.$$

Then we can derive the following

$$\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \Theta^m - \Delta_h \Theta^m - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \Theta^m = P_h \left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \Upsilon^m - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \Upsilon^m - \varkappa^m\right), \tag{4.5}$$

where $\varkappa^m = \sum_{i=1}^{k} \left({}^c\mathcal{D}_t^{\alpha_i} - {}^c\mathcal{D}_M^{\alpha_i}\right) u^m - \gamma \sum_{r=1}^{p} \left(\mathcal{I}^{\lambda_r} - \mathcal{I}_M^{\lambda_r}\right) u^m$.

Lemma 4.3. *Consider that u^m and u_h^m are the solutions to the problem (4.1) and (4.2), respectively. Then*

$$\|\nabla R_h u^m - \nabla u_h^m\| \le C\left(h^2 + M^{-\min(2-\alpha_1, \sigma\alpha_1)} + M^{-\min(\sigma\lambda_1, 2)}\right).$$

Proof Multiplying the equation (4.5), by $-\Delta_h \Theta^m$ and after integration over the domain Ω, we can obtain

$$\begin{aligned} &-\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \Theta^m, \Delta_h \Theta^m\right) + (\Delta_h \Theta^m, \Delta_h \Theta^m) + \gamma\left(\sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \Theta^m, \Delta_h \Theta^m\right) \\ &= -\left(P_h\left(\sum_{i=1}^{k} {}^c\mathcal{D}_M^{\alpha_i} \Upsilon^m - \gamma \sum_{r=1}^{p} \mathcal{I}_M^{\lambda_r} \Upsilon^m - \varkappa^m\right), \Delta_h \Theta^m\right), \text{ for } n = 1, \ldots, M. \end{aligned} \tag{4.6}$$

Furthermore, it is clear that

$$\begin{aligned} {}^c\mathcal{D}_M^{\alpha_i} \Upsilon^m &= {}^c\mathcal{D}_M^{\alpha_i} \Upsilon^m - {}^c\mathcal{D}_t^{\alpha_i} \Upsilon^m + {}^c\mathcal{D}_t^{\alpha_i} \Upsilon^m \\ &= \left({}^c\mathcal{D}_t^{\alpha_i} u^m - {}^c\mathcal{D}_M^{\alpha_i} u^m\right) - R_h\left({}^c\mathcal{D}_t^{\alpha_i} u^m - {}^c\mathcal{D}_M^{\alpha_i} u^m\right) + {}^c\mathcal{D}_t^{\alpha_i} \Upsilon^m, \end{aligned} \tag{4.7a}$$

$$\mathcal{I}_M^{\lambda_r} \Upsilon^m = \left(\mathcal{I}^{\lambda_r} u^m - \mathcal{I}_M^{\lambda_r} u^m\right) - R_h\left(\mathcal{I}^{\lambda_r} u^m - \mathcal{I}_M^{\lambda_r} u^m\right) + \mathcal{I}^{\lambda_r} \Upsilon^m. \tag{4.7b}$$

Substituting (4.7a)-(4.7b) back into (4.6), readily yields

$$
\begin{aligned}
&-\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{M}^{\alpha_i}\Theta^m, \Delta_h\Theta^m\right) + (\Delta_h\Theta^m, \Delta_h\Theta^m) + \gamma\left(\sum_{r=1}^{p}\mathcal{I}_{M}^{\lambda_r}\Theta^m, \Delta_h\Theta^m\right)\\
&= -\left(P_h\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{t}^{\alpha_i}\Upsilon^m - \gamma\sum_{r=1}^{p}\mathcal{I}^{\lambda_r}\Upsilon^m - R_h\varkappa^m\right), \Delta_h\Theta^m\right).
\end{aligned}
$$

Next, after recalling the definition of the discrete Laplacian (2.4) and the L^2 projection, we get

$$
\begin{aligned}
&\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{M}^{\alpha_i}(\nabla\Theta^m), \nabla\Theta^m\right) + \|\Delta_h\Theta^m\|^2 - \gamma\left(\sum_{r=1}^{p}\mathcal{I}_{M}^{\lambda_r}(\nabla\Theta^m), \nabla\Theta^m\right)\\
&= -\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{t}^{\alpha_i}\Upsilon^m - \gamma\sum_{r=1}^{p}\mathcal{I}^{\lambda_r}\Upsilon^m, \Delta_h\Theta^m\right) + (R_h\varkappa^m, \Delta_h\Theta^m),\\
&= -\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{t}^{\alpha_i}\Upsilon^m - \gamma\sum_{r=1}^{p}\mathcal{I}^{\lambda_r}\Upsilon^m, \Delta_h\Theta^m\right) + (\nabla P_h(-R_h\varkappa^m), \nabla\Theta^m).
\end{aligned}
$$

Then, owing to Lemma 3.2 and Cauchy–Schwarz inequality, we obtain

$$
\begin{aligned}
&\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{M}^{\alpha_i}\|\nabla\Theta^m\| - \gamma\sum_{r=1}^{p}\mathcal{I}_{M}^{\lambda_r}\|\nabla\Theta^m\|\right)\|\nabla\Theta^m\|\\
&\le \frac{1}{4}\left(\Big\|\sum_{i=1}^{k} {}^{c}\mathcal{D}_{t}^{\alpha_i}\Upsilon^m\Big\|^2 + \gamma\Big\|\sum_{r=1}^{p}\mathcal{I}^{\lambda_r}\Upsilon^m\Big\|^2\right) + \|\nabla P_h(R_h\varkappa^m)\|\|\nabla\Theta^m\|.
\end{aligned}
$$

In turn, utilizing the inequalities (2.2) and (2.3), we readily get

$$
\begin{aligned}
\left(\sum_{i=1}^{k} {}^{c}\mathcal{D}_{M}^{\alpha_i}\|\nabla\Theta^m\| - \gamma\sum_{r=1}^{p}\mathcal{I}_{M}^{\lambda_r}\|\nabla\Theta^m\|\right)\|\nabla\Theta^m\| &\le Ch^4\left(\sum_{i=1}^{k}\|{}^{c}\mathcal{D}_{t}^{\alpha_i}\Upsilon^m\|_2^2 + \gamma\sum_{r=1}^{p}\|\mathcal{I}^{\lambda_r}\Upsilon^m\|_2^2\right)\\
&\quad + \mathbf{C}\|\nabla R_h\varkappa^m\|\|\nabla\Theta^m\|\\
&\le Ch^4 + \mathbf{C}\|\nabla\varkappa^m\|\|\nabla\Theta^m\|,
\end{aligned}
$$

where we used Theorem 2.1 and $\|\nabla R_h F^m\| \le \|\nabla F^m\|$ for all $F \in H_0^1(\Omega)$. An application of Theorem 4.1 and Lemma 3.1 yields the desired estimate. □

5 Superconvergence error analysis of *L*1-PI-FEM

5.1 Rectangular and cubic meshes

For the enhancement of the numerical method, assume that the mesh is rectangle for $d = 2$ and cubical for $d = 3$. First, we present the pre-processing error estimate for the discrete problem (4.2), and then we provide the post-processing error estimate for the problem (4.2). Consider u^m and u_h^m are solutions of the problems (4.1) and (4.2), respectively. Here, we discuss the macroelement E^* which is formed as $d \times d$ elements $E_e \in \mathcal{T}_h$ for $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$. Let us define the biquadratic interpolation on the macroelement E^* for $d = 2$ as follows $\mathscr{I}_{2,h}u(b_{e,i}) = u(b_{e,i})$, where $b_{e,i}$'s are the vertices of the element E_e for $i = 1, 2, \ldots, (d+1)^d$ and $m = 1, \ldots, 2^d$. Similarly, we can define for the quadratic and triquadratic elements for $d = 1$ and $d = 3$, resp. Now we define the interpolation on the macroelement as $\mathscr{I}_{2,h}\mathscr{I}_h u = \mathscr{I}_{2,h}u$, for $u \in H^2(\Omega)$. Note that $\|\mathscr{I}_{2,h}u_h\| \le C\|\nabla u_h\|$ for all $u_h \in \mathcal{V}_{0h}$ (see [27]). Therefore we have the following error estimate

$$
\|\nabla u - \nabla\mathscr{I}_{2,h}u\| \le Ch^2\|u\|_3, \ \forall u \in H^3(\Omega). \tag{5.1}
$$

Lemma 5.1 (Superconvergence of interpolation)**.** *Consider that u^m and u_h^m are the solutions of (4.1) and (4.2) respectively. Then we have*

$$
\|\nabla\mathscr{I}_h u^m - \nabla u_h^m\| \le C\left(h^2 + M^{-\min(2-\alpha_1, \sigma\alpha_1)} + M^{-\min(\sigma\lambda_1, 2)}\right), \ \textit{for } m = 1, \ldots, M.
$$

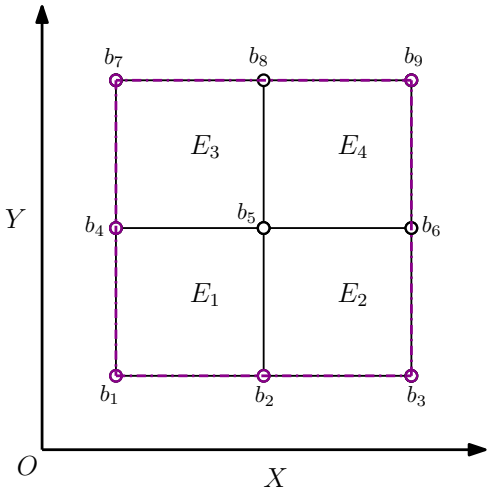

Figure 5.1: For $d = 2$, we consider the macroelement E^* by grouping the four elements E_1, E_2, E_3 and E_4 from the rectangular mesh partition $\mathcal{T}_h$, *i.e.*, $E^* = E_1 \cup E_2 \cup E_3 \cup E_4$. The region enclosed by the magenta dotted boundary represents the macroelement E^*.

Proof Utilize the triangle inequality and combine Theorem 2.1 and Lemma 4.3. □

Theorem 5.2 (Superconvergence of macroelement interpolation)**.** *Consider the solutions of* (4.1) *and* (4.2) *as* u^m *and* u_h^m*, respectively. Then we have*

$$\|\nabla u^m - \nabla \mathscr{I}_{2,h} u_h^m\| \leq C \left(h^2 + M^{-\min(2-\alpha_1, \sigma\alpha_1)} + M^{-\min(\sigma\lambda_1, 2)} \right), \text{ for } m = 1, \ldots, M.$$

Proof Adding and subtracting adequate terms, and using triangle inequality, we obtain the bound

$$\begin{aligned} \|\nabla u^m - \nabla \mathscr{I}_{2,h} u_h^m\| &= \|\nabla u^m - \nabla \mathscr{I}_{2,h} \mathscr{I}_h u_h^m + \nabla \mathscr{I}_{2,h} \mathscr{I}_h u_h^m - \nabla \mathscr{I}_{2,h} u_h^m\| \\ &\leq \|\nabla u^m - \nabla \mathscr{I}_{2,h} \mathscr{I}_h u_h^m\| + \|\nabla \mathscr{I}_{2,h} \mathscr{I}_h u_h^m - \nabla \mathscr{I}_{2,h} u_h^m\| \\ &\leq \|\nabla u^m - \nabla \mathscr{I}_{2,h} \mathscr{I}_h u_h^m\| + \|\nabla \mathscr{I}_{2,h} (\mathscr{I}_h u_h^m - u_h^m)\| \\ &\leq Ch^2 + C(\|\nabla (\mathscr{I}_h u_h^m - u_h^m)\|). \end{aligned}$$

With this, and after combining equation (5.1) with Lemma 5.1, we finally arrive at

$$\|\nabla u^m - \nabla \mathscr{I}_{2,h} u_h^m\| \leq C(h^2 + M^{-\min(2-\alpha_1, \sigma\alpha_1)} + M^{-\min(\sigma\lambda_1, 2)}).$$

□

Remark 5.1. *To establish the error estimate in Lemma 4.3, we take* $s = 2$ *in Theorem 2.1, while for Lemma 5.1 and Theorem 5.2, we take* $s = 3$.

5.2 Triangular and tetrahedral meshes

In the preceding section, the superconvergent postprocessing operator $\mathfrak{I}_{2,h}$ was constructed under the assumption that the mesh $\mathcal{T}_h$ consists of rectangles (in two dimensions) or cuboids (in three dimensions). This restriction arises from the inherent tensor-product structure of the macroelement E^*, which facilitates a natural and straightforward construction of biquadratic for $d = 2$ (respectively triquadratic for $d = 3$) interpolation.

For simplicial partitions, an analogous postprocessing operator $\mathfrak{I}_{2,h}$ can be built using quasi-interpolation on a coarse mesh. The construction presented below follows the well-known framework of Clément-type and Scott–Zhang operators [28, 29], adapted to the specific requirement of quadratic reproduction on a coarsened triangulation. The resulting operator maps functions in $L^2(\Omega)$ onto the space of continuous piecewise quadratics on a coarse simplicial mesh, and it possesses the same stability and approximation properties needed to recover a second-order superconvergence estimate in the H^1 norm.

Let $\mathcal{T}_h$ be a shape-regular simplicial triangulation of the polyhedral domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$. We assume that a conforming, shape-regular coarsening $\mathcal{T}_{2h}$ exists such that every element $K \in \mathcal{T}_{2h}$ is the

union of a uniformly bounded number of elements from $\mathcal{T}_h$. For instance, in two dimensions, one may obtain $\mathcal{T}_{2h}$ by merging four adjacent triangles that share a common vertex (such an approach is illustrated in Figure 5.1 for quadrilateral meshes and in Figure 5.2 for simplicial meshes), while in three dimensions, eight tetrahedra can be grouped into a larger tetrahedron or a macroelement. On the fine mesh, we retain the linear finite element space

$$V_h := \{v_h \in H_0^1(\Omega) \cap C^0(\Omega) : v_h|_K \in \mathbb{P}_1(K),\ \forall K \in \mathcal{T}_h\},$$

and we introduce the quadratic space on the coarse mesh

$$V_{2h} := \{v \in H_0^1(\Omega) \cap C^0(\Omega) : v|_K \in \mathbb{P}_2(K),\ \forall K \in \mathcal{T}_{2h}\}.$$

Vertex patches and weight functions. Denote by V_{2h} the set of vertices of $\mathcal{T}_{2h}$. For each vertex $z \in V_{2h}$, the associated vertex patch is

$$\omega_z := \bigcup \{K \in \mathcal{T}_{2h} : z \in \overline{K}\}.$$

Shape-regularity implies that there exist constants $c, C > 0$, independent of h, such that

$$ch^d \le |\omega_z| \le Ch^d, \qquad ch \le \operatorname{diam}(\omega_z) \le Ch, \quad \text{for } d = 2, 3.$$

For every $z \in V_{2h}$ we select a weight function $\psi_z \in L^\infty(\Omega)$ with the following properties:

$$\operatorname{supp}(\psi_z) \subset \omega_z, \qquad \psi_z \ge 0, \qquad \int_{\omega_z} \psi_z(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1, \qquad \|\psi_z\|_{L^\infty(\omega_z)} \le Ch^{-d}.$$

A canonical choice is the normalized characteristic function $\psi_z = |\omega_z|^{-1}\chi_{\omega_z}$, which satisfies the required sufficient conditions.

Next, let $\mathcal{N}_{2h}$ be the set of Lagrange nodes of the $\mathbb{P}_2$ finite element on $\mathcal{T}_{2h}$, *i.e.*, the vertices of $\mathcal{T}_{2h}$, the midpoints of edges, and (for $d = 3$) the barycenters of faces. For each node $a \in \mathcal{N}_{2h}$ we fix a vertex $z(a) \in V_{2h}$ such that $a \in \omega_{z(a)}$. A sketch is provided in Figure 5.2. Such an assignment is always feasible and can be chosen a priori. With this, the quasi-interpolation operator $\mathfrak{I}_{2,h} : L^2(\Omega) \to V_{2h}$ is defined by prescribing its nodal values:

$$(\mathfrak{I}_{2,h}u)(a) := \int_{\omega_{z(a)}} u(\boldsymbol{x})\,\psi_{z(a)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad \forall a \in \mathcal{N}_{2h}.$$

These values uniquely determine a function in V_{2h}, *i.e.*, a continuous piecewise quadratic polynomial on $\mathcal{T}_{2h}$. All subsequent estimates will require additional regularity of the argument u, which is guaranteed by the a priori bounds of Theorem 2.1.

Polynomial reproduction and approximation properties. A crucial property of $\mathfrak{I}_{2,h}$ is its exactness on quadratic polynomials:

$$\mathfrak{I}_{2,h}p = p \quad \forall p \in \mathbb{P}_2(\mathcal{T}_{2h}).$$

Indeed, for any node a and any $p \in \mathbb{P}_2(\mathcal{T}_{2h})$, shifting the coordinate system to the center of the patch and using the unit mass of $\psi_{z(a)}$ yields

$$(\mathfrak{I}_{2,h}p)(a) = \int_{\omega_{z(a)}} p(\boldsymbol{x})\psi_{z(a)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = p(a),$$

since local averaging with a symmetric weight reproduces polynomials of degree at most two exactly.

The reproduction property together with the stability on V_h (established below) leads to the following result.

Theorem 5.3 (Optimal error estimate)**.** *Assume that $u \in H^{2+\varepsilon}(\Omega)$, be the solution of (4.1) for some $\varepsilon > 0$. Then there exists a constant C, independent of h and u, such that*

$$\|\nabla(u - \mathfrak{I}_{2,h}u)\| \le Ch^2|u|_{H^{2+\varepsilon}(\Omega)}.$$

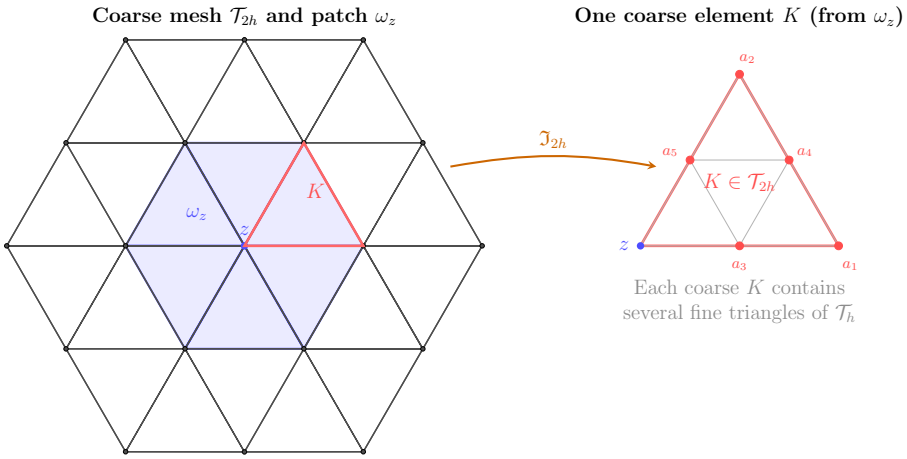

Figure 5.2: Coarse mesh $\mathcal{T}_{2h}$ (left). The blue patch ω_z is the union of all 6 coarse triangles sharing vertex z, and the highlighted red triangle $K = (z, v_1, v_2)$ is one of these 6 triangles. The right plot zooms into this exact triangle K. The fine mesh $\mathcal{T}_h$ (gray lines) subdivides K into 4 smaller triangles. $\mathbb{P}_2$ nodes (red dots) on K include vertices (z, a_1, a_2) and edge midpoints (a_3, a_4, a_5). For every node a on K, we can choose the parent vertex $z(a) = z$, because $K \subset \omega_z$. The operator computes the value at node a by averaging the input function u over the entire blue patch ω_z.

Proof For each $K \in \mathcal{T}_{2h}$, let $p_K \in \mathbb{P}_2(K)$ be the averaged Taylor polynomial of u over K (see, e.g., [30]). Standard polynomial approximation theory gives

$$\|\nabla(u - p_K)\|_{L^2(K)} \leq Ch^2 |u|_{H^{2+\varepsilon}(K)}.$$

Using $\mathfrak{I}_{2,h} p_K = p_K$,

$$u - \mathfrak{I}_{2,h} u = (u - p_K) + \mathfrak{I}_{2,h}(p_K - u).$$

The H^1-stability of $\mathfrak{I}_{2,h}$ (Lemma 5.4) implies

$$\|\nabla \mathfrak{I}_{2,h}(p_K - u)\|_{L^2(K)} \leq C \|\nabla(p_K - u)\|_{L^2(\omega_K)},$$

where ω_K denotes the union of patches intersecting K. Combining these estimates and summing over all $K \in \mathcal{T}_{2h}$ completes the proof. □

Lemma 5.4 (Stability on the fine-mesh space)**.** *For all $u_h \in V_h$, the post-processing operator $\mathfrak{I}_{2,h}$ satisfies the following bound*

$$\|\mathfrak{I}_{2,h} u_h\| \leq C \|\nabla u_h\|.$$

Proof For any node $a \in \mathcal{N}_{2h}$, the Cauchy–Schwarz inequality and the bound $\|\psi_{z(a)}\|_{L^2(\omega_{z(a)})} \leq Ch^{-d/2}$ give

$$|(\mathfrak{I}_{2,h} u_h)(a)| \leq \|\psi_{z(a)}\|_{L^2(\omega_{z(a)})} \|u_h\|_{L^2(\omega_{z(a)})} \leq Ch^{-d/2} \|u_h\|_{L^2(\omega_{z(a)})}.$$

Since u_h is continuous on $\overline{\omega_{z(a)}}$, a local Poincaré inequality (see, e.g., [30, Prop. 4.5.3]) yields

$$\|u_h\|_{L^2(\omega_{z(a)})} \leq Ch \|\nabla u_h\|_{L^2(\omega_{z(a)})}.$$

Consequently, we obtain

$$|(\mathfrak{I}_{2,h} u_h)(a)| \leq Ch^{1-d/2} \|\nabla u_h\|_{L^2(\omega_{z(a)})},$$

and using norm equivalence on finite-dimensional spaces, we arrive at

$$\|\mathfrak{I}_{2,h} u_h\|^2 \leq C \sum_{a \in \mathcal{N}_{2h}} |(\mathfrak{I}_{2,h} u_h)(a)|^2 h^d \leq C \|\nabla u_h\|^2,$$

where the finite overlap of the patches ensures the final bound. □

Application to superconvergence postprocessing with lower regularity. With the operator $\mathfrak{I}_{2,h}$ at hand, the superconvergence analysis for simplicial meshes follows the same lines as in Section 5 for quadrilaterals. One defines the postprocessed solution $\overline{u_h^m} := \mathfrak{I}_{2,h} u_h^m$, where u_h^m is the $L1$-PI-FEM solution on $\mathcal{T}_h$. Then, using the triangle inequality,

$$\|\nabla u^m - \nabla \overline{u_h^m}\| \leq \|\nabla u^m - \nabla \mathfrak{I}_{2,h} u^m\| + \|\nabla \mathfrak{I}_{2,h}(u^m - u_h^m)\|.$$

Table 6.1: Spatial errors $e^{N,M}$, $e^{N,M}_{\mathscr{J}_{2,h}}$ and corresponding order of convergence in H^1 norm for $\alpha_1 = 0.7$ on a time-graded mesh for the example in Section 6.1.

		Before postprocessing		After postprocessing	
N	h	$e^{N,M}$	Spatial rate	$e^{N,M}_{\mathscr{J}_{2,h}}$	Spatial rate
4	5.000e-01	1.3560e+00	–	6.4230e-01	–
8	2.500e-01	6.6758e-01	1.0224	1.6184e-01	1.9886
16	1.250e-01	3.3234e-01	1.0063	4.0540e-02	1.9972
32	6.250e-02	1.6598e-01	1.0016	1.0176e-02	1.9941
64	3.125e-02	8.2968e-02	1.0004	2.5777e-03	1.9811

The first term is bounded by Ch^2 thanks to Theorem 5.3 and the regularity of u^m. The second term is estimated by the H^1-stability of $\mathfrak{I}_{2,h}$ on the space V_h, which can be proved by a scaling argument (see, e.g., [31, Lemma 3.2]) and yields

$$\|\nabla \mathfrak{I}_{2,h} v_h\| \leq C \|\nabla v_h\|, \quad \forall v_h \in V_h.$$

Together with the supercloseness estimate of Lemma 4.3, one obtains the optimal second-order convergence in the H^1 norm, exactly as in Theorem 5.2.

We remark that the construction of $\mathfrak{I}_{2,h}$ presented here is not limited to simplicial meshes. The same idea can be adapted to any shape-regular triangulation for which a coarsening $\mathcal{T}_{2h}$ with a quadratic finite element space is available. The essential ingredients are local averaging over vertex patches and exact reproduction of quadratic polynomials.

6 Numerical experiments

6.1 Convergence against smooth manufactured solutions in 2D

Consider the multi-term time-fractional integro diffusion problem (2.1) on $\Omega = (0,2)^2$, with $k = p = 2$ and with boundary condition $u|_{\partial\Omega} = 0$, $t \in (0,1]$ and initial condition $u(\boldsymbol{x}, 0) = 0$ for $\boldsymbol{x} \in \overline{\Omega}$, where we take different values of α_1, and fix $\alpha_2 = 0.4$, $\lambda_1 = 0.5$, $\lambda_2 = 0.4$, with $\gamma = 10^{-5}$. The source function $f(\boldsymbol{x}, t)$ is constructed in such a way that the exact solution is $u(\boldsymbol{x}, t) = (t^{\alpha_1} + 2t^3)\sin(x)\sin(y)(x-2)(y-2)$.

For this example, we consider rectangular partitions in the spatial domain with $N+1$ mesh points in each direction. For the optimal rate of Theorem 5.2, we consider $\sigma = (2-\alpha_1)/\alpha_1$. We define the computational errors as follows: $e^{N,M} := \max\limits_{0\leq m\leq M} \|\nabla u^m - \nabla u_h^m\|$, the reported error is computed without applying any post-processing and $e^{N,M}_{\mathscr{J}_{2,h}} := \max\limits_{0\leq m\leq M} \|\nabla u^m - \nabla \mathscr{J}_{2,h} u_h^m\|$, the error is measured after post-processing for Theorem 5.2. In the computational study of this example, we present the errors in both temporal and spatial directions, along with their corresponding convergence orders by $\frac{\log(E_{\text{prev}}/E)}{\log(h_{\text{prev}}/h)}$. Table 6.1 shows the usual error $e^{N,M}$ and the postprocessing error $e^{N,M}_{\mathscr{J}_{2,h}}$ for $\alpha_1 = 0.7$. We consider $M = 1500$ for spatial dominance. The numerical demonstration shows that the theoretical results are sharp for Theorem 5.2. In Table 6.2 and Table 6.3, we report the temporal errors and their corresponding convergence rates for fractional orders $\alpha_1 = 0.7,\ 0.8$, and 0.9. The numerical results, computed using the error definitions $e^{N,M}$ and $e^{N,M}_{\mathscr{J}_{2,h}}$, closely match the theoretical predictions stated in Lemma 3.1 and Theorem 5.2, respectively. To demonstrate the errors and rates of convergence in Tables 6.2 and 6.3, we choose $N = 2\lceil M^{2-\alpha_1}\rceil$ so that the temporal error dominates. The observed order of convergence is $\mathcal{O}(M^{-(2-\alpha_1)})$ based on the errors $e^{N,M}$ and $e^{N,M}_{\mathscr{J}_{2,h}}$, which agrees with the theoretical predictions established in Lemma 3.1 and Theorem 5.2 on the graded mesh technique. Table 6.4 reports the temporal errors computed using the definition $e^{N,M}$ on a uniform mesh. The observed first-order convergence is in agreement with Remark 3.1 for $M = N$. These computational result (Figure 6.1) confirm the sharpness of the theoretical analysis. For example, with a graded time-mesh we observed a convergence of $O(M^{-1.1})$ for $\alpha_1 = 0.9$, whereas for the case of uniform mesh we only attain a $O(M^{-1})$ (see last row of Tables 6.3 and 6.4, respectively).

Table 6.2: Temporal dominating $e^{N,M}$ errors and corresponding rates of convergence on temporal-graded mesh for the example in Section 6.1.

M	$\alpha_1 = 0.7$		$\alpha_1 = 0.8$		$\alpha_1 = 0.9$	
	$H^1(\Omega)$-norm	Temporal rate	$H^1(\Omega)$-norm	Temporal rate	$H^1(\Omega)$-norm	Temporal rate
10	1.608e-01	–	1.9191e-01	–	2.3098e-01	–
20	6.6067e-02	1.3262	8.4124e-02	1.2127	1.1130e-01	1.0623
30	3.9426e-02	1.2966	5.1957e-02	1.2011	7.0320e-02	1.1377
40	2.7335e-02	1.2894	3.7044e-02	1.1847	5.1948e-02	1.0561
50	2.0444e-02	1.3145	2.8303e-02	1.2129	4.0702e-02	1.0961

Table 6.3: Temporal dominating $e^{N,M}_{\mathcal{I}_{2,h}}$ errors and corresponding rates of convergence on temporal-graded mesh for the example in Section 6.1.

M	$\alpha_1 = 0.7$		$\alpha_1 = 0.8$		$\alpha_1 = 0.9$	
	$H^1(\Omega)$-norm	Temporal rate	$H^1(\Omega)$-norm	Temporal rate	$H^1(\Omega)$-norm	Temporal rate
20	2.0608e-02	–	2.3667e-02	–	2.9269e-02	–
40	8.3125e-03	1.3308	9.9889e-03	1.2561	1.2964e-02	1.1797
80	3.3426e-03	1.3246	4.2558e-03	1.2365	5.8702e-03	1.1454
160	1.3424e-03	1.3213	1.8239e-03	1.2252	2.6892e-03	1.1274
320	5.3993e-04	1.3165	7.8543e-04	1.2169	1.2410e-03	1.1162

Table 6.4: Temporal dominating $e^{N,M}$ errors and corresponding rates of convergence on uniform mesh for the example in Section 6.1.

M	$\alpha_1 = 0.7$		$\alpha_1 = 0.8$		$\alpha_1 = 0.9$	
	$H^1(\Omega)$-norm	Temporal rate	$H^1(\Omega)$-norm	Temporal rate	$H^1(\Omega)$-norm	Temporal rate
10	5.3134e-01	–	5.3186e-01	–	5.3388e-01	–
20	2.6566e-01	1.0001	2.6616e-01	0.9987	2.6766e-01	0.9961
30	1.7712e-01	0.9998	1.7748e-01	0.9995	1.7856e-01	0.9983
40	1.3284e-01	1.0000	1.3311e-01	1.0000	1.3395e-01	0.9994
50	1.0627e-01	1.0001	1.0648e-01	1.0002	1.0716e-01	1.0000

6.2 Convergence against smooth manufactured solutions in 3D

Consider the time-fractional integro diffusion problem (2.1) on $\Omega = (0,1)^3$, with $k = p = 1$ and with initial condition $u(\boldsymbol{x}, 0) = 0$ for $\boldsymbol{x} \in \overline{\Omega}$, where $\alpha_1 = 0.5$, $\lambda_1 = 0.3$, and $\gamma = 0.1$. The source term $f(\boldsymbol{x}, t)$ and the non-homogeneous boundary datum are constructed in such a way that the exact solution is $u(\boldsymbol{x}, t) = t^{2+\alpha_1} \cos(\pi x) \sin(\pi y) \sin(\pi z)$.

For this case, we use two levels of refined uniform tetrahedral meshes. We only address the convergence with respect to spatial mesh refinement. We set $T = 0.5$ and use $M = 50$ graded time steps. The outcome of the error tests is presented in Table 6.5, which confirms (for finer meshes) an asymptotic optimal convergence rate of $\mathcal{O}(h)$ in the $H^1(\Omega)$ norm for the numerical solution and the predicted $\mathcal{O}(h^2)$ rate for the postprocessed solution. A snapshot of the approximate solution at the final time and on the second-last mesh refinement is shown in Figure 6.2, indicating well-resolved profiles.

6.3 Time evolution with no closed-form solution

We next consider a case without an analytical solution to test the model and finite element scheme on a problem that combines non-separable forcing and anisotropic diffusion. The domain is $\Omega = (0,1)^2$, $T = 1$, and $M = 100$ time steps are used. The source term

$$f(\boldsymbol{x}, t) = 100 \sin(\pi x \cos t)\, \sin(\pi y e^t)\, \cos(\pi (x + y) e^{-t})$$

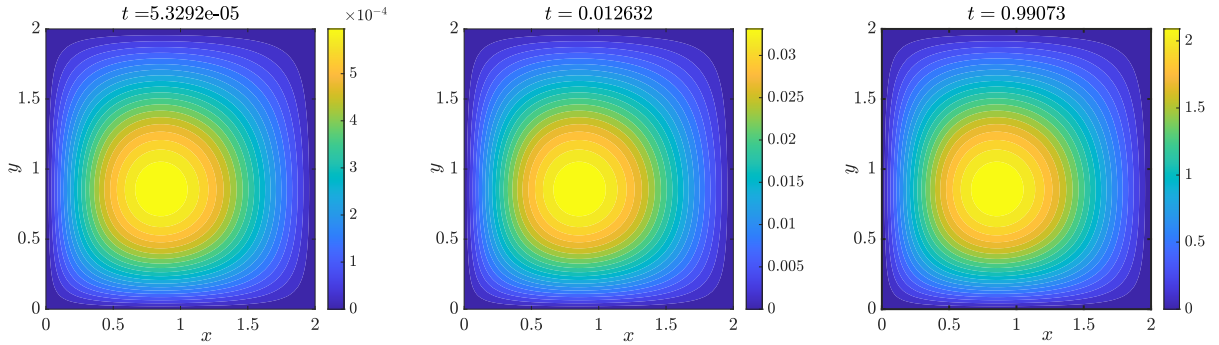

Figure 6.1: Contour plots of the numerical solution $u(\boldsymbol{x}, t)$ at different time levels for the example in Section 6.1.

Table 6.5: Error history for the example in Section 6.2 (in 3D). Number of degrees of freedom, mesh size, individual errors for the approximate solution and the postprocessed solution, and corresponding experimental convergence rates.

		Before postprocessing		After postprocessing	
DoF	h	$e^{N,M}$	Spatial rate	$e^{N,M}_{\mathfrak{I}_{2,h}}$	Spatial rate
27	0.8660	2.88e-01	–	2.95e-01	–
125	0.4330	1.63e-01	0.816	1.07e-01	1.457
729	0.2165	8.50e-02	0.943	3.13e-02	1.777
4913	0.1083	4.30e-02	0.985	8.19e-03	1.935
35937	0.0541	2.15e-02	0.996	2.07e-03	1.984
274625	0.0271	1.09e-02	0.999	5.26e-04	1.996

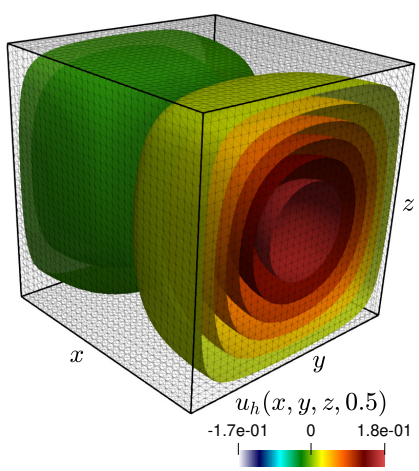

Figure 6.2: Iso-contours of the approximate solution corresponding to Section 6.2, shown at the fifth mesh refinement and at the final time step $t = 0.5$.

couples the spatial and temporal variables in a way that prevents factorization. Instead of a constant and scalar diffusion, the second-order term is now $-\mathrm{div}(K\nabla u)$ with a diffusion tensor

$$K = \begin{pmatrix} 1 & 0.1 \\ 0.2 & 0.05 \end{pmatrix}$$

having principal diffusion direction rotated by about 12° from the x-axis. We discretize the space domain with an unstructured mesh of size $h = 0.01$ and set the fractional model parameters to $k = p = 1$,

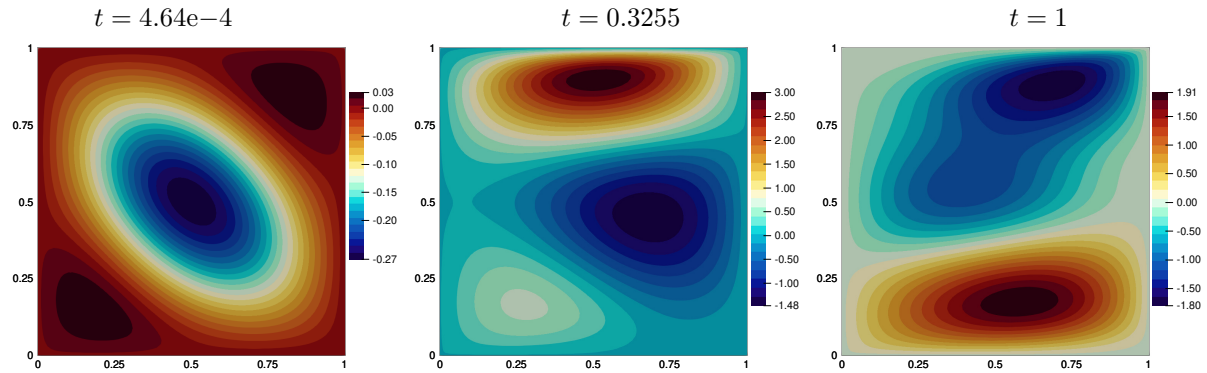

Figure 6.3: Contour plots of the numerical solution $u(\boldsymbol{x}, t)$ at different time levels for the example in Section 6.3.

$\alpha_1 = 0.75$, $\lambda_1 = 0.5$, $\gamma = 0.5$. Homogeneous Dirichlet boundary conditions and zero initial condition are imposed, so the entire evolution is driven by the forcing.

Figure 6.3 shows the numerical solution at the first time step, at the midpoint $n = 50$, and at the final time $t = T$. At earliest stages the solution closely follows the spatial pattern of the forcing, but a slight elongation of the contours along the strong-diffusion direction is already visible. Steeper gradients are preserved across the weak-diffusion axis, while diffusive smoothing acts predominantly along the strong-diffusion axis. The subdiffusive order $\alpha_1 < 1$ introduces memory that slows down the relaxation relative to a standard parabolic problem. The numerical solution at later times shows the time-dependent pattern of the forcing term.

7 Conclusions

In this work, as an application of the novel discrete Gronwall inequality, we establish a superconvergence result for a multi-term fractional diffusion problem having multiterm weakly singular kernels of several orders with low regularity assumptions in the space. The standard finite element method achieves an $\mathcal{O}(h)$ convergence rate in the H^1 norm, while a postprocessing step based on the macroelement interpolation operator $\mathscr{I}_{2,h}$ yields an improved $\mathcal{O}(h^2)$ convergence rate in the same norm in the spatial direction. To resolve the initial layer, we employ a graded temporal mesh, which yields the optimal rate of convergence $\mathcal{O}(M^{\alpha_1 - 2})$ in the temporal direction. The numerical examples confirm the theoretical convergence rates predicted by the analysis in both space and time, especially in higher dimensional models in 2D and 3D. The present methodology will be helpful to analyze several future goals which include the development of a posteriori error estimates and adaptivity [6], the extension to the semi-linear case [7, 10], and the mixed hybridizable discontinuous Galerkin formulations for problems having weakly singular kernel in space directions, also. Also, the application of the presented methodology in the context of poromechanics equations with non-local effects is part of our ongoing work.

Acknowledgments

The first author acknowledges administrative support from the University Grants Commission (UGC), India, and IIT Patna. The second author acknowledges the support of Ministry of Education, Govt. of India for their administrative support. The third author was partially supported by the Australian Research Council through the Future Fellowship grant FT220100496 and Discovery Project grant DP22010316; and by the Center of Advanced Study (CAS) at the Norwegian Academy of Science and Letters under the program Mathematical Challenges in Brain Mechanics.

Data availability

Data sharing is not applicable for this work as no data is generated.

Declarations

The authors declare that they do not have any conflict of interest. We also declare that this work is not submitted elsewhere.

References

[1] Jang, Y., Shaw, S.: A priori error analysis for a finite element approximation of dynamic viscoelasticity problems involving a fractional order integro-differential constitutive law. Adv. Comput. Math. **47**(3), 46 (2021)

[2] Magin, R.L.: Fractional calculus in bioengineering. Begell House, Redding, CT (2006)

[3] Kilbas, A.: Theory and applications of fractional differential equations. North-Holland Mathematics Studies **204** (2006)

[4] Chen, Y., Chen, Z., Huang, Y.: An hp-version of the discontinuous Galerkin method for fractional integro-differential equations with weakly singular kernels. BIT Numer. Math. **64**(3), 1–23 (2024)

[5] Zayernouri, M., Karniadakis, G.E.: Fractional spectral collocation method. SIAM J. Sci. Comput. **36**(1), 40–62 (2014)

[6] Adolfsson, K., Enelund, M., Larsson, S.: Adaptive discretization of an integro-differential equation with a weakly singular convolution kernel. Computer. Methods Appl. Mech. Engrg. **192**(51-52), 5285–5304 (2003)

[7] Chen, H., Qiu, W., Zaky, M.A., Hendy, A.S.: A two-grid temporal second-order scheme for the two-dimensional nonlinear Volterra integro-differential equation with weakly singular kernel. Calcolo **60**(1), 1–13 (2023)

[8] Lubich, C.: Discretized fractional calculus. SIAM J. Math. Anal. **17**(3), 704–719 (1986)

[9] Mahata, S., Sinha, R.K.: Finite element method for fractional parabolic integro-differential equations with smooth and nonsmooth initial data. J. Sci. Comput. **87**(1), 1–7 (2021)

[10] Mustapha, K., Mustapha, H.: A second-order accurate numerical method for a semilinear integro-differential equation with a weakly singular kernel. IMA J. Numer. Anal. **30**(2), 555–578 (2010)

[11] Santra, S.: Analysis of a higher-order scheme for multi-term time-fractional integro-partial differential equations with multi-term weakly singular kernels. Numer. Algorithms **99**(4), 1649–1695 (2025)

[12] Santra, S., Mohapatra, J., Das, P., Choudhuri, D.: Higher order approximations for fractional order integro-parabolic partial differential equations on an adaptive mesh with error analysis. Comput. Math. Appl. **150**, 87–101 (2023)

[13] Kopteva, N.: Error analysis of the L1 method on graded and uniform meshes for a fractional-derivative problem in two and three dimensions. Math. Comput. **88**(319), 2135–2155 (2019)

[14] Liu, T., Zhang, H., Yang, X.: The BDF2-ADI compact difference scheme on graded meshes for the three-dimensional PIDEs with multi-term weakly singular kernels. Numer. Algorithms **00**, 1–34 (2025)

[15] Huang, C., Chen, H.: Superconvergence analysis of finite element methods for the variable-order subdiffusion equation with weakly singular solutions. Appl. Math. Lett. **139**, 108559 (2023)

[16] Huang, C., Stynes, M.: Superconvergence of the direct discontinuous Galerkin method for a time-fractional initial-boundary value problem. Numer. Methods Partial Differ. Equ. **35**(6), 2076–2090 (2019)

[17] Huang, C., Chen, H., An, N.: β-robust superconvergent analysis of a finite element method for the distributed order time-fractional diffusion equation. J. Sci. Comput. **90**(1), 1–20 (2022)

[18] Sakamoto, K., Yamamoto, M.: Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. J. Math. Anal. Appl. **382**(1), 426–447 (2011)

[19] Henry, D.: Geometric Theory of Semilinear Parabolic Equations vol. 840. Springer, Heidelberg (2006)

[20] Huang, C., Liu, X., Meng, X., Stynes, M.: Error analysis of a finite difference method on graded meshes for a multiterm time-fractional initial-boundary value problem. Comput. Methods Appl. Math. **20**(4), 815–825 (2020)

[21] Ganesan, S., Tobiska, L.: Finite Elements: Theory and Algorithms. Cambridge University Press, Delhi (2017)

[22] Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems vol. 25. Springer, Heidelberg (2007)

[23] Atkinson, K., Han, W.: Numerical Solution of Fredholm Integral Equations of the Second Kind. In: Theoretical Numerical Analysis: A Functional Analysis Framework, pp. 473–549. Springer, New York (2009)

[24] Li, C., Chen, A., Ye, J.: Numerical approaches to fractional calculus and fractional ordinary differential equation. J. Comput. Phys. **230**(9), 3352–3368 (2011)

[25] Gracia, J.-L., O'Riordan, E., Stynes, M.: Convergence in positive time for a finite difference method applied to a fractional convection-diffusion problem. Comput. Meth. Appl. Math. **18**(1), 33–42 (2018)

[26] Jin, B., Lazarov, R., Zhou, Z.: An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data. IMA J. Numer. Anal. **36**(1), 197–221 (2016)

[27] Stynes, M., Tobiska, L.: The SDFEM for a convection-diffusion problem with a boundary layer: optimal error analysis and enhancement of accuracy. SIAM J. Numer. Anal. **41**(5), 1620–1642 (2003)

[28] Clément, P.: Approximation by finite element functions using local regularization. Rev. Franç. Inform. Rech. Opér. **9**(R2), 77–84 (1975)

[29] Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. **54**(190), 483–493 (1990)

[30] Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Texts in Applied Mathematics, vol. 15. Springer, New York (2008)

[31] Haijun, W., Zhiming, C.: Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems. Sci. China Math. **49**(10), 1405–1429 (2006)