East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.180420.200720

Vol. **11**, No. 1, pp. 63-92
February 2021

A Superconvergent Nonconforming Mixed FEM for Multi-Term Time-Fractional Mixed Diffusion and Diffusion-Wave Equations with Variable Coefficients

Huijun Fan[1,2], Yanmin Zhao[2,*], Fenling Wang[2], Yanhua Shi[2] and Yifa Tang[3,4]

[1]*School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, China.*
[2]*School of Science, Xuchang University, Xuchang 461000, China.*
[3]*LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.*
[4]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China.*

Received 18 April 2020; Accepted (in revised version) 20 July 2020.

Abstract. An unconditionally stable fully-discrete scheme on regular and anisotropic meshes for multi-term time-fractional mixed diffusion and diffusion-wave equations (TFMDDWEs) with variable coefficients is developed. The approach is based on a nonconforming mixed finite element method (FEM) in space and classical $L1$ time-stepping method combined with the Crank-Nicolson scheme in time. Then, the unconditionally stability analysis of the fully-discrete scheme is presented. The convergence for the original variable u and the flux $\vec{p} = \mu(\mathbf{x})\nabla u$, respectively, in H^1- and L^2-norms is derived by using the relationship between the projection operator R_h and the interpolation operator I_h. Interpolation postprocessing technique is used to establish superconvergence results. Finally, numerical tests are provided to demonstrate the theoretical analysis.

AMS subject classifications: 65M10, 78A48

Key words: Nonconforming mixed FEM, multi-term time-fractional mixed diffusion and diffusion-wave equations, $L1$ time-stepping method, Crank-Nicolson scheme, convergence and superconvergence.

---

1. Introduction

In recent years, fractional partial differential equations (FPDEs) have attracted a considerable attention. In contrast to integer-order partial differential equations, FPDEs have numerous advantages in applications, including hereditary processes in biology, physics, medicine, finance, fluid mechanics, environmental science, and water waves — cf. Refs. [7,

*Corresponding author. *Email address:* `zhaoym@lsec.cc.ac.cn` (Y. Zhao)

9, 10, 12, 14, 18, 22, 26, 27, 54]. For example, in physics and materials one often uses time-fractional partial differential equations (TFPDEs) such as time-fractional diffusion equations (TFDEs) and time-fractional wave equations (TFWEs). Nevertheless, the use of only TFDEs or only TFWEs does not allow to describe certain processes accurately. In this paper, we focus on a nonconforming mixed FEM for a kind of TFPDEs with variable coefficients, called multi-term TFMDDWEs — viz.

$$\begin{aligned} & u_t + P(D_t^\alpha)u(\mathbf{x},t) + P\left(D_t^\beta\right)u(\mathbf{x},t) \\ & \quad - \nabla\cdot(\mu(\mathbf{x})\nabla u(\mathbf{x},t)) = f(\mathbf{x},t), \quad (\mathbf{x},t)\in\Omega\times(0,T], \\ & u(\mathbf{x},t) = 0, \qquad\qquad\qquad\qquad\qquad\quad (\mathbf{x},t)\in\partial\Omega\times(0,T], \\ & u(\mathbf{x},0) = u_0(\mathbf{x}), u_t(\mathbf{x},0) = \tilde{u}_0(\mathbf{x}), \quad \mathbf{x}\in\Omega, \end{aligned} \tag{1.1}$$

where $\Omega\subset R^2$ is a bounded convex polygonal region with the boundary $\partial\Omega$, $\mathbf{x}=(x,y)$, $u_0(\mathbf{x})$, $\tilde{u}_0(\mathbf{x})$, $\mu(\mathbf{x})$, $f(\mathbf{x},t)$ are sufficiently smooth functions and there are constants μ_1,μ_2 such $0<\mu_1\leq\mu(\mathbf{x})\leq\mu_2$. The operators $P(D_t^\alpha), P(D_t^\beta)$ are defined by

$$P\left(D_t^\alpha\right) = D_t^\alpha + \sum_{i=1}^{r} l_i D_t^{\alpha_i}, \quad l_i>0, \quad r\in\mathbb{N}^+, \quad 0<\alpha_1<\alpha_2<\cdots<\alpha_r<\alpha<1,$$

$$P\left(D_t^\beta\right) = D_t^\beta + \sum_{j=1}^{m} \omega_j D_t^{\beta_j}, \quad \omega_j>0, \quad m\in\mathbb{N}^+, \quad 1<\beta_1<\beta_2<\cdots<\beta_m<\beta<2,$$

where D_t^γ is the left-sided Caputo fractional derivative of order γ with respect to t, i.e.

$$D_t^\gamma u(\mathbf{x},t) := \begin{cases} \dfrac{1}{\Gamma(n-\gamma)}\displaystyle\int_0^t (t-s)^{n-1-\gamma}\frac{\partial^n u(\mathbf{x},s)}{\partial s^n}ds, & \gamma\notin\mathbb{N}^+, \\ \dfrac{\partial^\gamma u(\mathbf{x},t)}{\partial t^\gamma}, & \gamma\in\mathbb{N}^+ \end{cases}$$

with Γ denoting the Gamma function — cf. [22].

Analytical solutions of the most TFDEs and TFWEs are not known. Thus numerical methods for solving TFDEs and TFWEs have attracted increasing attention. For example, Ammi *et al.* [2] studied a fully discrete scheme for TFDEs based on finite difference method in time and FEM in space. In [42], a finite difference method was used to solve a modified TFDE, and the unconditional stability and an optimal-order error estimate of the method were discussed. An efficient Petrov-Galerkin spectral method for a TFDE with its error analysis was presented in [31]. Galerkin methods are a popular approach for solving TF-PDEs — e.g. Sun *et al.* [39] proposed a semi-discrete scheme with a local discontinuous Galerkin discretisation in the spatial variable. Wang and Ren [43] considered a high-order compact finite difference method and numerical analysis for a single TFDE. Meshless methods are also widely used. In particular, Kumar *et al.* [15] studied a radial basis meshless local collocation method. In order to obtain a discrete scheme for time fractional diffusion-wave equation, Liu *et al.* [25] developed a novel finite difference method. Soori and Aminataei [37] presented the stability of a sixth-order compact alternating direction implicit

(ADI) scheme. Zhang and Jiang studied a Crank-Nicolson Legendre spectral method [51] for nonlinear time fractional diffusion-wave equations. Since multi-term TFPDEs describe certain diffusion processes more accurately, various numerical methods, including Galerkin FEMs [3, 17, 50, 56], orthogonal spline collocation methods [45], finite difference methods [5, 16], compact difference methods [28], spectral methods [47, 55], finite volume methods [46] have been developed for such equations. In addition, FEMs [35,36], compact finite difference methods [6, 11], finite difference methods [38], Galerkin spectral element methods [6, 30], singular boundary methods jointed with dual reciprocity methods [29] are also employed to multi-term TFWEs.

Although there is a vast literature on approximation approaches to TFDEs and TFWEs, numerical methods for TFMDDWEs are less developed. Nevertheless, we can note spectral methods [4,24], FEMs [8,44,49], implicit difference methods [32]. In particular, Liu *et al.* [24] constructed an ADI spectral scheme for two-dimensional multi-term TFMDDWEs using Legendre spectral approximation in space and finite difference discretisation in time. Wei *et al.* [44] considered an approximation method based on FEM in space and $L1$-CN scheme in time for two-term time-fractional mixed diffusion-wave equation. Shen *et al.* [32] proposed an implicit difference approximation and used the multivariate Mittag-Leffler function to derive an analytical solution by the method of separation of variables. However, to the best of our knowledge, for multi-term TFMDDWEs there are no publications concerning nonconforming mixed element methods. In this work, we apply the EQ_1^{rot} nonconforming element method to multi-term TFMDDWEs with variable coefficients.

Let us provide a brief information about EQ_1^{rot} elements and mixed FEMs. The degree of freedom of an EQ_1^{rot} element is defined on its edges and the element itself. Since there are only two influencing elements for the internal parameters of each node, EQ_1^{rot} elements have a number of advantages and they has been recently employed to solve partial differential equations on regular or anisotropic meshes [21,33,34,48,52,53]. For example, error estimates for EQ_1^{rot} elements under anisotropic meshes are obtained in [33]. In [53], nonconforming FEMs are applied to time-fractional nonlinear parabolic equation and a novel result of the consistency error estimate for EQ_1^{rot} elements on anisotropic meshes is proved. It is worth noting that mixed FEMs require a lower smoothness with respect to space variable, so that the errors for the original variable and flow can be derived simultaneously. In classical mixed finite element schemes, the mixed finite element spaces have to satisfy the Brezzi-Babuśka condition. This generates numerous difficulties. Recently, mixed FEMs are developed for FPDEs [1, 19, 23, 48]. Thus a new numerical method using mixed finite elements on complex domains for time fractional reaction-diffusion equation was considered in [1], a high accuracy interpolation approximation for distributed-order fractional sub-diffusion equation employing two H^1-Galerkin mixed FEMs was introduced in [19] and a combination of a Galerkin mixed FEM and a time second-order discrete scheme for nonlinear TFDE with fourth-order derivative term was studied in [23].

Another feature of this paper is the simultaneous use of regular and anisotropic meshes. The regular and quasi-uniform assumptions are fundamental conditions for conventional finite elements. However, there are solutions with an essential changes in a certain direction only and in such situations, anisotropic finite elements provide better options [13, 33, 41,49,53]. This paper presents an unconditionally stable fully-discrete scheme for multi-term TFMDDWEs with variable coefficients. We carry out theoretical analysis based on mixed nonconforming FEMs on regular and anisotropic meshes. First, using the anisotropic EQ_1^{rot} finite elements and the $L1$ time-stepping method combined with the Crank-Nicolson scheme, we construct a fully-discrete scheme and prove its unconditional stability. After that, connections between interpolation operator and projection operator are exploited in order to establish convergence and superclose results. Finally, applying interpolation post-processing technique, we prove the superconvergence of the method.

The reminder of the paper is outlined as follows. In Section 2, a fully-discrete scheme using a nonconforming mixed FEM and the classical $L1$ time-stepping method combined with Crank-Nicolson scheme is introduced. Section 3 contains auxiliary results. The unconditional stability, superclose and superconvergence analysis of the fully-discrete scheme are presented in Section 4. Numerical results demonstrating the accuracy and efficiency of the method are the last section.

2. A Mixed Finite Element and a Fully-Discrete Scheme

Let Ω be a rectangular domain and Γ_h a family of rectangular meshes of Ω such that $\bar{\Omega} = \cup_{K\in\Gamma_h} K$. For each $K \in \Gamma_h$, we denote by (x_K, y_K) the center of the element K and $2h_{x,K}$ and $2h_{y,K}$ the length of the edges parallel to the x- and y-axes, respectively. An element K has four vertices A_i and four edges $l_i = \overline{A_iA_{i+1}}$, $i = 1,2,3,4$, where $A_5 := A_1$. Let $h_K := \max\{h_{x,K}, h_{y,K}\}$, $h := \max_{K\in\Gamma_h} h_K$ and $\hat{K} := [-1,1]\times[-1,1]$ be the reference element in $\xi-\eta$ plane with the vertices $\hat{A}_1 := (-1,-1)$, $\hat{A}_2 := (1,-1)$, $\hat{A}_3 := (1,1)$ and $\hat{A}_4 = (-1,1)$. Let $\hat{l}_1 = \overline{\hat{A}_1\hat{A}_2}$, $\hat{l}_2 = \overline{\hat{A}_2\hat{A}_3}$, $\hat{l}_3 = \overline{\hat{A}_3\hat{A}_4}$, $\hat{l}_4 = \overline{\hat{A}_4\hat{A}_1}$. The affine mapping $\mathscr{F}_K : \hat{K} \to K$ is defined by

$$\begin{aligned} x &= x_K + h_{x,K}\xi, \\ y &= y_K + h_{y,K}\eta. \end{aligned}$$

On $\hat{K}$ we consider the finite element $(\hat{K}, \hat{P}, \hat{\Sigma})$, where

$$\hat{P} = \text{span}\big\{1, \xi, \eta, \varphi(\xi), \varphi(\eta)\big\}, \quad \hat{\Sigma} = \big\{\hat{v}_i, i = 1,2,3,4,5\big\},$$

and

$$\begin{aligned} \hat{v}_i &= \frac{1}{|\hat{l}_i|}\int_{\hat{l}_i} \hat{v}\, d\hat{s}, \quad i = 1,2,3,4, \\ \hat{v}_5 &= \frac{1}{|\hat{K}|}\int_{\hat{K}} \hat{v}\, d\xi d\eta, \varphi(t) = \frac{1}{2}\big(3t^2 - 1\big). \end{aligned}$$

Then for any $\hat{v} \in H^1(\hat{K})$, we can write

$$\hat{I}\hat{v} = \hat{v}_5 + \frac{1}{2}\big(\hat{v}_2 - \hat{v}_4\big)\xi + \frac{1}{2}(\hat{v}_3 - \hat{v}_1)\eta + \frac{1}{2}\big(\hat{v}_2 + \hat{v}_4 - 2\hat{v}_5\big)\varphi(\xi) + \frac{1}{2}(\hat{v}_3 + \hat{v}_1 - 2\hat{v}_5)\varphi(\eta).$$

Here and in what follows, $H^r(\Theta)$ refers to the usual Sobolev space on the domain Θ, equipped with the norm $\|\cdot\|_r = \|\cdot\|_{r,\Theta}$ and seminorm $|\cdot|_r = |\cdot|_{r,\Theta}$.

The associated finite element spaces are defined by

$$\begin{aligned}
\vec{W}_h &:= \left\{\vec{q} = (q_1,q_2) \in \left(L^2(\Omega)\right)^2; \vec{q}|_K \in Q_{10}(K)\times Q_{01}(K)\right\},\\
V_h &:= \left\{v_h; v_h|_K = \hat{v}\circ F_K^{-1},\ \hat{v}\in\hat{P}, \forall K\in\Gamma_h, \int_F [v_h]ds = 0,\ F\subset\partial K\right\},
\end{aligned}$$

where $Q_{ij} := \mathrm{span}\{x^r y^s, 0\le r\le i, 0\le s\le j\}$, $[v_h]$ denotes the jump of v_h across the edge F which is an internal edge, and $[v_h] = v_h$ for $F\subset\partial\Omega$. Besides,

$$\|\cdot\|_h = \left(\sum_K |\cdot|_{1,K}^2\right)^{1/2}$$

is a norm on V_h.

Consider the interpolation operators I_h, Π_h and the projection operator R_h defined by

$$\begin{aligned}
&I_h : v\in H^1(\Omega)\longrightarrow I_h v\in V_h, && I_h|_K = I_K,\quad I_K v = \left(\hat{I}\hat{v}\right)\circ F_K^{-1},\\
&\Pi_h : \vec{q}\in \left(H^1(\Omega)\right)^2 \longrightarrow \Pi_h\vec{q}\in\vec{W}_h, && \Pi_h|_K = \Pi_K,\\
&\int_{l_i} (\vec{q}-\Pi_K\vec{q})\cdot\vec{n}ds = 0, && i = 1,2,3,4,\\
&R_h : v\in H_0^1(\Omega)\longrightarrow R_h v\in V_h, && \text{for}\quad u\in H_0^1(\Omega),\\
&\left(\mu(\mathbf{x})\nabla(u-R_h u),\nabla v_h\right) = 0, && \forall v_h\in V_h.
\end{aligned}$$

Here and in what follows, $\vec{n}$ denotes a unit normal vector to ∂K.

Let $0 = t_0 < t_1 < \cdots < t_N = T$ be the uniform partition of the time interval with the time step $\tau = T/N$ and $t_n = n\tau$, $n = 0,1,\ldots,N$. If $\varphi(t) := \varphi(\mathbf{x},t) = u_t(\mathbf{x},t)$ is a smooth function on $[0,T]$, then we write:

$$\begin{aligned}
&\varphi^n = \varphi(t_n), && n = 0,1,\ldots,N,\\
&\varphi^{n-1/2} = \frac{\varphi^n+\varphi^{n-1}}{2}, && n = 1,\ldots,N,\\
&\partial_t\varphi^0 = 0,\quad \partial_t\varphi^{k-1/2} = \frac{\varphi^k-\varphi^{k-1}}{\tau}, && k = 1,\ldots,N,\\
&\tilde{D}_t^\alpha u^{n-1/2} = \frac{\tau^{1-\alpha}}{2\Gamma(2-\alpha)}\left(\sum_{k=1}^n \tilde{b}_{\alpha,n-k}\partial_t u^{k-1/2} + \sum_{k=1}^{n-1}\tilde{b}_{\alpha,n-k-1}\partial_t u^{k-1/2}\right),\\
&\hat{D}_t^\beta u^{n-1/2} = \frac{\tau^{1-\beta}}{\Gamma(3-\beta)}\left[\hat{b}_{\beta,0}\partial_t u^{n-1/2} + \sum_{k=1}^{n-1}\left(\hat{b}_{\beta,n-k}-\hat{b}_{\beta,n-k-1}\right)\partial_t u^{k-1/2} - \hat{b}_{\beta,n-1}\tilde{u}_0\right],\\
&\tilde{P}(\tilde{D}_t^\alpha)u^{n-1/2} = \left(\tilde{D}_t^\alpha + \sum_{i=1}^r l_i\tilde{D}_t^{\alpha_i}\right)u^{n-1/2},
\end{aligned}$$

$$\hat{P}\left(\hat{D}_t^{\beta}\right)u^{n-1/2}=\left(\hat{D}_t^{\beta}+\sum_{j=1}^{m}\omega_j\hat{D}_t^{\beta_j}\right)u^{n-1/2},$$

where $\tilde{b}_{\alpha,k}=(k+1)^{1-\alpha}-k^{1-\alpha}$, $\hat{b}_{\beta,k}=(k+1)^{2-\beta}-k^{2-\beta}$, $k=0,1,\dots,N-1$.

For convenience, we denote

$$\tilde{B}_{\alpha,k}=\tilde{b}_{\alpha,k}+\sum_{i=1}^{r}\frac{\Gamma(2-\alpha)}{\Gamma(2-\alpha_i)}l_i\tilde{b}_{\alpha_i,k}\tau^{\alpha-\alpha_i},$$
$$\hat{B}_{\beta,k}=\hat{b}_{\beta,k}+\sum_{j=1}^{m}\frac{\Gamma(3-\beta)}{\Gamma(3-\beta_j)}\omega_j\hat{b}_{\beta_j,k}\tau^{\beta-\beta_j}.$$

It is easily seen that

$$\tilde{P}\left(\tilde{D}_t^{\alpha}\right)u^{n-1/2}=\frac{\tau^{1-\alpha}}{2\Gamma(2-\alpha)}\left(\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\partial_t u^{k-1/2}+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\partial_t u^{k-1/2}\right),$$
$$\hat{P}\left(\hat{D}_t^{\beta}\right)u^{n-1/2}=\frac{\tau^{1-\beta}}{\Gamma(3-\beta)}\left[\hat{B}_{\beta,0}\partial_t u^{n-1/2}+\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k}-\hat{B}_{\beta,n-k-1}\right)\partial_t u^{k-1/2}-\hat{B}_{\beta,n-1}\tilde{u}_0\right].$$

Moreover, direct calculations yield

$$\hat{B}_{\beta,0}=1+\sum_{j=1}^{m}\frac{\Gamma(3-\beta)}{\Gamma(3-\beta_j)}\omega_j\tau^{\beta-\beta_j},\quad \hat{B}_{\beta,0}>\hat{B}_{\beta,1}>\cdots>\hat{B}_{\beta,N-1}>0.$$

Setting $\vec{p}(\mathbf{x},t)=\mu(\mathbf{x})\nabla u(\mathbf{x},t)$, we write the problem (1.1) as

$$\begin{aligned}
&\vec{p}(\mathbf{x},t)=\mu(\mathbf{x})\nabla u(\mathbf{x},t), && (\mathbf{x},t)\in\Omega\times(0,T],\\
&u_t+P\left(D_t^{\alpha}\right)u(\mathbf{x},t)+P\left(D_t^{\beta}\right)u(\mathbf{x},t)-\nabla\cdot\vec{p}(\mathbf{x},t)=f(\mathbf{x},t), && (\mathbf{x},t)\in\Omega\times(0,T],\\
&u(\mathbf{x},t)=0, && (\mathbf{x},t)\in\partial\Omega\times(0,T],\\
&u(\mathbf{x},0)=u_0(\mathbf{x}),\quad u_t(\mathbf{x},0)=\tilde{u}_0(\mathbf{x}), && \mathbf{x}\in\Omega.
\end{aligned}\tag{2.1}$$

The variational formulation of the problem (2.1) is: Find $\{u(\mathbf{x},t),\vec{p}(\mathbf{x},t)\}:(0,T]\longrightarrow H_0^1(\Omega)\times(L^2(\Omega))^2$ such that

$$\begin{aligned}
&(u_t,v)+\left(P\left(D_t^{\alpha}\right)u(\mathbf{x},t),v\right)+\left(P\left(D_t^{\beta}\right)u(\mathbf{x},t),v\right)\\
&\qquad+\left(\vec{p}(\mathbf{x},t),\nabla v\right)=\left(f(\mathbf{x},t),v\right), && \forall v\in H_0^1(\Omega),\\
&\left(\mu(\mathbf{x})\nabla u(\mathbf{x},t),\vec{q}\right)-\left(\vec{p}(\mathbf{x},t),\vec{q}\right)=0, && \forall\vec{q}\in\left(L^2(\Omega)\right)^2,\\
&u(\mathbf{x},0)=u_0(\mathbf{x}),\quad u_t(\mathbf{x},0)=\tilde{u}_0(\mathbf{x}), && \mathbf{x}\in\Omega.
\end{aligned}\tag{2.2}$$

Let $(f,g)=\int_{\Omega}f\,g\,d\Omega$ and $\|f\|_0=(f,f)^{1/2}$ be, respectively, the inner product and norm on the space $L^2(\Omega)$. We now consider the following nonconforming fully-discrete scheme:

Find $\{U^n,\vec{P}^n\}\in V_h\times\vec{W}_h$, such that

$$\begin{aligned}
&\left(\partial_t U^{n-1/2},v_h\right)_h+\left(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)U^{n-1/2},v_h\right)_h+\left(\hat{P}\left(\hat{D}_t^{\beta}\right)U^{n-1/2},v_h\right)_h\\
&\qquad\qquad+\left(\vec{P}^{n-1/2},\nabla_h v_h\right)_h=\left(f^{n-1/2},v_h\right)_h, \quad \forall v_h\in V_h,\\
&\left(\mu(\mathbf{x})\nabla_h U^{n-1/2},\vec{\omega}_h\right)_h-\left(\vec{P}^{n-1/2},\vec{\omega}_h\right)_h=0, \qquad \forall\vec{\omega}_h\in\vec{W}_h,\\
&U^0=R_hu_0(\mathbf{x}), \quad \tilde{U}_0=R_h\tilde{u}_0(\mathbf{x}), \qquad \mathbf{x}\in\Omega,
\end{aligned} \tag{2.3}$$

where ∇_h is the gradient operator defined by

$$\left(\nabla_h\varphi(\mathbf{x}),\nabla_h\phi(\mathbf{x})\right)_h:=\sum_K\int_K\nabla\varphi(\mathbf{x})\cdot\nabla\phi(\mathbf{x})d\mathbf{x}.$$

3. Auxiliary Results

Let us recall some properties of EQ_1^{rot} elements.

Lemma 3.1 (cf. Shi *et al.* [33]). *If $u\in H^2(\Omega)$ and the corresponding mesh is anisotropic, then*

$$\|u-I_hu\|_0+h\|u-I_hu\|_h\le Ch^2\|u\|_2, \tag{3.1}$$

$$\left(\nabla_h(u-I_hu),\nabla_h v_h\right)_h=0, \quad \forall v_h\in V_h. \tag{3.2}$$

Lemma 3.2. *If $u\in H_0^1(\Omega)\cap H^2(\Omega)$, then*

$$\|R_hu-u\|_0+h\|\nabla_h(R_hu-u)\|_0\le Ch^2\|u\|_2.$$

Proof. Noting that

$$\left(\mu(\mathbf{x})\nabla_h(u-R_hu),\nabla_h(I_hu-R_hu)\right)_h=0,$$

we get

$$\begin{aligned}
\mu_1\|\nabla_h(u-R_hu)\|_0^2&\le\left\|\mu^{1/2}(\mathbf{x})\nabla_h(u-R_hu)\right\|_0^2\\
&=\left(\mu(\mathbf{x})\nabla_h(u-R_hu),\nabla_h(u-I_hu)\right)_h\\
&\le\mu_2\|\nabla_h(u-R_hu)\|_0\|\nabla_h(u-I_hu)\|_0.
\end{aligned}$$

It follows from (3.1) that

$$\|\nabla_h(u-R_hu)\|_0\le C\|\nabla_h(u-I_hu)\|_0\le Ch\|u\|_2. \tag{3.3}$$

In order to establish an L^2-norm estimate, we consider the following auxiliary problem:

$$\begin{aligned}
-\nabla_h\cdot\left(\mu(\mathbf{x})\nabla_h\psi\right)&=u-R_hu, \quad \text{in } \Omega,\\
\psi&=0, \qquad\qquad \text{on } \partial\Omega.
\end{aligned} \tag{3.4}$$

The regularity of the solution of the Eq. (3.4) implies

$$\|\psi\|_2 \le C\|u-R_h u\|_0. \tag{3.5}$$

The definition of R_h, (3.3) and (3.5) give

$$\begin{aligned}\|u-R_h u\|_0^2 = &\big(\mu(\mathbf{x})\nabla_h(\psi - I_h\psi), \nabla_h(u-R_h u)\big)_h + \big(\mu(\mathbf{x})\nabla_h(u-R_h u), \nabla_h I_h\psi\big)_h \\ &-\sum_K \int_{\partial K} \mu(\mathbf{x})\frac{\partial\psi}{\partial\vec{n}}(u-R_h u)ds \\ &\le Ch\|\psi\|_2\|\nabla_h(u-R_h u)\|_0 + Ch\|\psi\|_2\|u-R_h u\|_h \\ &\le Ch^2\|u\|_2\|u-R_h u\|_0,\end{aligned}$$

and, hence, $\|u-R_h u\|_0 \le Ch^2\|u\|_2$. □

Lemma 3.3. *If $u\in H_0^1(\Omega)\cap H^2(\Omega)$, then*

$$\|R_h u - I_h u\|_h \le Ch^2\|u\|_2.$$

Proof. Let $|K|$ denote the area of K and

$$\overline{\mu(\mathbf{x})}|_K := \frac{1}{|K|}\int_K \mu(\mathbf{x})dxdy.$$

Then

$$|\mu(\mathbf{x}) - \overline{\mu(\mathbf{x})}| \le Ch\|\mu(\mathbf{x})\|_{1,\infty}. \tag{3.6}$$

Recalling the definition of R_h and using (3.1), (3.2), (3.6) and semi-norm properties, we obtain

$$\begin{aligned}&\left\|\mu^{1/2}(\mathbf{x})(R_h u - I_h u)\right\|_h^2 \\ =&\sum_K\int_K (\mu(\mathbf{x}) - \overline{\mu(\mathbf{x})})\nabla(u-I_h u)\nabla(R_h u - I_h u)dxdy \\ &+\sum_K\int_K \overline{\mu(\mathbf{x})}\nabla(u-I_h u)\nabla(R_h u - I_h u)dxdy \\ \le& Ch^2\|u\|_2\|R_h u - I_h u\|_h.\end{aligned} \tag{3.7}$$

Since $\mu(\mathbf{x})\ge\mu_1$, the application of (3.7) finishes the proof. □

Lemma 3.4 (cf. Refs. [20,34])**.** *If for any fixed $t\in(0,T]$ one has $\vec{p}(t)\in(H^2(\Omega))^2$ and $u\in H^1(\Omega)$, then for every $\vec{\omega}_h\in\vec{W}_h$ the following relations hold:*

$$\begin{aligned}&\big(\vec{p}-\Pi_h\vec{p}, \vec{\omega}_h\big)_h = \mathcal{O}\big(h^2\big)|\vec{p}|_2\|\vec{\omega}_h\|_0, \\ &\big(\nabla_h(u-I_h u), \vec{\omega}_h\big)_h = 0.\end{aligned}$$

Lemma 3.5. *If $u \in H^4(\Omega) \cap H_0^1(\Omega)$ and $v_h \in V_h$, then*

$$\left|\sum_K \int_{\partial K} \mu(\mathbf{x}) \frac{\partial u}{\partial \vec{n}} v_h ds\right| \le Ch^2 \|u\|_4 \|v_h\|_0,$$

where C is a constant independent of h and τ.

Proof. The proof is similar to the corresponding proof in [53]. □

Lemma 3.6. *If $\phi \in (H^3(\Omega))^2$ and $v_h \in V_h$, then*

$$\left|\sum_K \int_{\partial K} \phi \cdot \vec{n} v_h ds\right| \le Ch^2 |\phi|_3 \|v_h\|_0.$$

Proof. It uses Lemma 3.5 and the relation $\nabla_h V_h \subset \vec{W}_h$. □

Lemma 3.7 (cf. Sun [40]). *(1) Assume that $u_{tt}(\mathbf{x}, t) \in L^2(\Omega)$ for any $t \in (0, T]$. Then the term*

$$R_1^{n-1/2} = P\left(D_t^\alpha\right) u^{n-1/2} - \tilde{P}\left(\tilde{D}_t^\alpha\right) u^{n-1/2}$$

admits the estimate

$$\left\|R_1^{n-1/2}\right\|_0 \le C \max_{0\le t\le T} \|u_{tt}(\mathbf{x}, t)\|_0 \tau^{2-\alpha}.$$

(2) Assume that $u_{ttt}(\mathbf{x}, t) \in L^2(\Omega)$ for any $t \in (0, T]$. Then the term

$$R_2^{n-1/2} = P\left(D_t^\beta\right) u^{n-1/2} - \hat{P}\left(\hat{D}_t^\beta\right) u^{n-1/2}$$

admits the estimate

$$\left\|R_2^{n-1/2}\right\|_0 \le C \max_{0\le t\le T} \|u_{ttt}(\mathbf{x}, t)\|_0 \tau^{3-\beta}.$$

The following lemma is proven by direct calculations.

Lemma 3.8. *For any $n = 1, 2, \ldots, N$ one has*

$$\sum_{k=1}^n \frac{\tau}{\lambda \hat{B}_{\beta,k-1}} \le \frac{\Gamma(3-\beta) T^\beta}{C_0}, \quad \lambda \tau \sum_{k=1}^n \hat{B}_{\beta,k-1} \le \frac{T^{2-\beta} C_1}{\Gamma(3-\beta)},$$

where

$$C_0 = (2-\beta) + \sum_{j=1}^m \frac{\Gamma(3-\beta)(2-\beta_j)}{\Gamma(3-\beta_j)} \omega_j T^{\beta-\beta_j},$$

$$C_1 = 1 + \sum_{j=1}^m \frac{\Gamma(3-\beta)}{\Gamma(3-\beta_j)} \omega_j T^{\beta-\beta_j}, \quad \lambda = \frac{\tau^{1-\beta}}{\Gamma(3-\beta)}.$$

Lemma 3.9 (cf. Zhao *et al.* [49]). *Let N be a positive integer and $0<\alpha<1$. For any vector $\boldsymbol{Q}:=[v^1,v^2,v^3,\cdots,v^{N-1},v^N]\in R^N$, the inequality*

$$\sum_{n=1}^{N}\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}v^kv^n+\sum_{n=1}^{N}\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}v^kv^n\geq 0$$

holds.

4. Stability, Convergence and Superconvergence of Fully-Discrete Scheme

We show that the fully discrete scheme above is unconditionally stable. First, we introduce the notation

$$\lambda:=\frac{\tau^{1-\beta}}{\Gamma(3-\beta)},\quad \lambda_1:=\frac{\tau^{1-\alpha}}{\Gamma(2-\alpha)}.$$

Theorem 4.1. *Any solution $\{U^n\}$ in the method* (2.3) *satisfies the estimates*

$$\|U^n\|_0+\|U^n\|_h\leq C\left(\frac{\mu_2}{\mu_1}\|U^0\|_h^2+\frac{T^{2-\beta}C_1}{\mu_1\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2+\frac{\Gamma(3-\beta)T^\beta}{C_0\mu_1}\max_{0\leq t\leq T}\|f(t)\|_0^2\right)^{1/2},$$

$$\|\vec{P}^n\|_0\leq C\left(\frac{\mu_2}{\mu_1}\|U^0\|_h^2+\frac{T^{2-\beta}C_1}{\mu_1\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2+\frac{\Gamma(3-\beta)T^\beta}{C_0\mu_1}\max_{0\leq t\leq T}\|f(t)\|_0^2\right)^{1/2}.$$

Proof. Choosing $v_h=\partial_tU^{n-1/2}$ and $\vec{\omega}_h=\nabla_h\partial_tU^{n-1/2}$ in (2.3) yields

$$\begin{aligned}&\left(\partial_tU^{n-1/2},\partial_tU^{n-1/2}\right)_h+\left(\tilde{P}\left(\tilde{D}_t^\alpha\right)U^{n-1/2},\partial_tU^{n-1/2}\right)_h\\&+\left(\hat{P}\left(\hat{D}_t^\beta\right)U^{n-1/2},\partial_tU^{n-1/2}\right)_h+\left(\mu(\mathbf{x})\nabla_hU^{n-1/2},\nabla_h\partial_tU^{n-1/2}\right)_h\\=&\left(f^{n-1/2},\partial_tU^{n-1/2}\right)_h.\end{aligned}\tag{4.1}$$

Recalling the definitions of $\tilde{P}(\tilde{D}_t^\alpha)U^{n-1/2}$ and $\hat{P}(\hat{D}_t^\beta)U^{n-1/2}$, we write

$$\begin{aligned}&\left(\tilde{P}\left(\tilde{D}_t^\alpha\right)U^{n-1/2},\partial_tU^{n-1/2}\right)_h\\=&\frac{\tau^{1-\alpha}}{2\Gamma(2-\alpha)}\left[\left(\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\partial_tU^{k-1/2},\partial_tU^{n-1/2}\right)_h+\left(\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\partial_tU^{k-1/2},\partial_tU^{n-1/2}\right)_h\right],\\&\left(\hat{P}\left(\hat{D}_t^\beta\right)U^{n-1/2},\partial_tU^{n-1/2}\right)_h\\=&\frac{\tau^{1-\beta}}{\Gamma(3-\beta)}\left(\hat{B}_{\beta,0}\partial_tU^{n-1/2}+\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k}-\hat{B}_{\beta,n-k-1}\right)\partial_tU^{k-1/2}-\hat{B}_{\beta,n-1}\tilde{U}_0,\partial_tU^{n-1/2}\right)_h\\=&\frac{\tau^{1-\beta}}{\Gamma(3-\beta)}\Bigg[\hat{B}_{\beta,0}\left\|\partial_tU^{n-1/2}\right\|_0^2+\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k}-\hat{B}_{\beta,n-k-1}\right)\left(\partial_tU^{k-1/2},\partial_tU^{n-1/2}\right)_h\end{aligned}\tag{4.2}$$

$$-\left(\hat{B}_{\beta,n-1}\tilde{U}_0,\partial_t U^{n-1/2}\right)_h\Bigg], \tag{4.3}$$

$$\left(\mu(\mathbf{x})\nabla_h U^{n-1/2},\nabla_h\partial_t U^{n-1/2}\right)_h=\frac{1}{2\tau}\left(\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^n\right\|_0^2-\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^{n-1}\right\|_0^2\right). \tag{4.4}$$

Substituting (4.2)-(4.4) into (4.1), and multiplying (4.1) by 2τ gives

$$\begin{aligned}
&2\lambda\tau\hat{B}_{\beta,0}\left\|\partial_t U^{n-1/2}\right\|_0^2+2\tau\left\|\partial_t U^{n-1/2}\right\|_0^2+\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^n\right\|_0^2-\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^{n-1}\right\|_0^2\\
=&2\lambda\tau\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k-1}-\hat{B}_{\beta,n-k}\right)\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h\\
&+2\lambda\tau\hat{B}_{\beta,n-1}\left(\tilde{U}_0,\partial_t U^{n-1/2}\right)_h+2\tau\left(f^{n-1/2},\partial_t U^{n-1/2}\right)_h\\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h\right]\\
\le&\lambda\tau\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k-1}-\hat{B}_{\beta,n-k}\right)\left(\|\partial_t U^{k-1/2}\|_0^2+\|\partial_t U^{n-1/2}\|_0^2\right)\\
&+\lambda\tau\hat{B}_{\beta,n-1}\left(\|\tilde{U}_0\|_0^2+\|\partial_t U^{n-1/2}\|_0^2\right)+2\tau\left|\left(f^{n-1/2},\partial_t U^{n-1/2}\right)_h\right|\\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h\right].
\end{aligned}$$

Denote

$$\rho^0:=\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^0\right\|_0^2,\quad \rho^n:=\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^n\right\|_0^2+\lambda\tau\sum_{k=1}^{n}\hat{B}_{\beta,n-k}\left\|\partial_t U^{k-1/2}\right\|_0^2,$$

we obtain

$$\begin{aligned}
\rho^n\le&\rho^{n-1}+\lambda\tau\hat{B}_{\beta,n-1}\|\tilde{U}_0\|_0^2+2\tau\left|\left(f^{n-1/2},\partial_t U^{n-1/2}\right)_h\right|\\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h\right]\\
\le&\rho^{n-2}+\lambda\tau\hat{B}_{\beta,n-1}\|\tilde{U}_0\|_0^2+2\tau\left|\left(f^{n-1/2},\partial_t U^{n-1/2}\right)_h\right|\\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{n-1/2}\right)_h\right]\\
&+\lambda\tau\hat{B}_{\beta,n-2}\|\tilde{U}_0\|_0^2+2\tau\left|\left(f^{n-1-1/2},\partial_t U^{n-1-1/2}\right)_h\right|\\
&-\lambda_1\tau\left[\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{n-1-1/2}\right)_h+\sum_{k=1}^{n-2}\tilde{B}_{\alpha,n-k-2}\left(\partial_t U^{k-1/2},\partial_t U^{n-1-1/2}\right)_h\right]\\
\le&\cdots\le\rho^0+\lambda\tau\sum_{l=1}^{n}\hat{B}_{\beta,l-1}\|\tilde{U}_0\|_0^2+2\tau\sum_{l=1}^{n}\left|\left(f^{l-1/2},\partial_t U^{l-1/2}\right)_h\right|
\end{aligned}$$

$$-\lambda_1\tau\sum_{l=1}^{n}\left[\sum_{k=1}^{l}\tilde{B}_{\alpha,l-k}\left(\partial_t U^{k-1/2},\partial_t U^{l-1/2}\right)_h+\sum_{k=1}^{l-1}\tilde{B}_{\alpha,l-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{l-1/2}\right)_h\right].$$

The Cauchy-Schwarz inequality implies

$$\begin{aligned}\rho^n\le\rho^0&+\lambda\tau\sum_{l=1}^{n}\hat{B}_{\beta,l-1}\|\tilde{U}_0\|_0^2+\sum_{l=1}^{n}\frac{\tau\left\|f^{l-1/2}\right\|_0^2}{\lambda\hat{B}_{\beta,n-l}}+\sum_{l=1}^{n}\lambda\tau\hat{B}_{\beta,n-l}\left\|\partial_t U^{l-1/2}\right\|_0^2\\&-\lambda_1\tau\sum_{l=1}^{n}\left[\sum_{k=1}^{l}\tilde{B}_{\alpha,l-k}\left(\partial_t U^{k-1/2},\partial_t U^{l-1/2}\right)_h+\sum_{k=1}^{l-1}\tilde{B}_{\alpha,l-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{l-1/2}\right)_h\right].\end{aligned}\tag{4.5}$$

Lemmas 3.8 and 3.9 show that

$$\lambda\tau\sum_{l=1}^{n}\hat{B}_{\beta,l-1}\|\tilde{U}_0\|_0^2\le\frac{T^{2-\beta}C_1}{\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2,\tag{4.6}$$

$$\sum_{l=1}^{n}\frac{\tau\left\|f^{l-1/2}\right\|_0^2}{\lambda\hat{B}_{\beta,n-l}}\le\frac{\Gamma(3-\beta)T^{\beta}}{C_0}\max_{0\le t\le T}\|f(t)\|_0^2,\tag{4.7}$$

$$-\lambda_1\tau\sum_{l=1}^{n}\left[\sum_{k=1}^{l}\tilde{B}_{\alpha,l-k}\left(\partial_t U^{k-1/2},\partial_t U^{l-1/2}\right)_h+\sum_{k=1}^{l-1}\tilde{B}_{\alpha,l-k-1}\left(\partial_t U^{k-1/2},\partial_t U^{l-1/2}\right)_h\right]\le 0.\tag{4.8}$$

Substituting (4.6)-(4.8) into (4.5) leads to the inequality

$$\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^n\right\|_0^2\le\left\|\mu^{1/2}(\mathbf{x})\nabla_h U^0\right\|_0^2+\frac{T^{2-\beta}C_1}{\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2+\frac{\Gamma(3-\beta)T^{\beta}}{C_0}\max_{0\le t\le T}\|f(t)\|_0^2,$$

and since $0<\mu_1\le\mu(\mathbf{x})\le\mu_2$, it yields

$$\|U^n\|_h\le\left(\frac{\mu_2}{\mu_1}\|U^0\|_h^2+\frac{T^{2-\beta}C_1}{\mu_1\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2+\frac{\Gamma(3-\beta)T^{\beta}}{C_0\mu_1}\max_{0\le t\le T}\|f(t)\|_0^2\right)^{1/2}.$$

Recalling that there is a constant C such that $\|U^n\|_0\le C\|U^n\|_h$, we obtain

$$\|U^n\|_0\le C\left(\frac{\mu_2}{\mu_1}\|U^0\|_h^2+\frac{T^{2-\beta}C_1}{\mu_1\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2+\frac{\Gamma(3-\beta)T^{\beta}}{C_0\mu_1}\max_{0\le t\le T}\|f(t)\|_0^2\right)^{1/2}.$$

It is easily seen that

$$\left(\mu(\mathbf{x})\nabla_h U^n,\vec{\omega}_h\right)_h-\left(\vec{P}^n,\vec{\omega}_h\right)_h=0.$$

It follows that for $\vec{\omega}_h=\vec{P}^n$, we have

$$\|\vec{P}^n\|_0^2=\left(\mu(\mathbf{x})\nabla_h U^n,\vec{P}^n\right)_h\le\|\mu(\mathbf{x})\nabla_h U^n\|_0\|\vec{P}^n\|_0,$$

which yields the inequality

$$\|\vec{P}^n\|_0\le\|\mu(\mathbf{x})\nabla_h U^n\|_0.$$

Finally, we can write

$$\|\vec{P}^n\|_0 \le C\left(\frac{\mu_2}{\mu_1}\|U^0\|_h^2+\frac{T^{2-\beta}C_1}{\mu_1\Gamma(3-\beta)}\|\tilde{U}_0\|_0^2+\frac{\Gamma(3-\beta)T^\beta}{C_0\mu_1}\max_{0\le t\le T}\|f(t)\|_0^2\right)^{1/2}. \qquad \square$$

In order to establish the superclose and superconvergence results, we consider the residues

$$\begin{aligned} u^n-U^n&=\left(u^n-R_hu^n\right)+\left(R_hu^n-U^n\right)=\eta^n+\xi^n,\\ \vec{p}^n-\vec{P}^n&=\left(\vec{p}^n-\Pi_h\vec{p}^n\right)+\left(\Pi_h\vec{p}^n-\vec{P}^n\right)=\vec{\vartheta}^n+\vec{\theta}^n \end{aligned}$$

and use (2.2) and (2.3) to write the error equation as

$$\begin{aligned} &\left(\partial_t u^{n-1/2}-\partial_t U^{n-1/2},v_h\right)_h+\left(\tilde{P}\left(\tilde{D}_t^\alpha\right)\left(u^{n-1/2}-U^{n-1/2}\right),v_h\right)_h\\ &\quad+\left(\hat{P}\left(\hat{D}_t^\beta\right)\left(u^{n-1/2}-U^{n-1/2}\right),v_h\right)_h+\left(\vec{p}^{n-1/2}-\vec{P}^{n-1/2},\nabla_h v_h\right)_h\\ &=-\left(R_1^{n-1/2},v_h\right)_h-\left(R_2^{n-1/2},v_h\right)_h-\left(R_3^{n-1/2},v_h\right)_h\\ &\quad+\sum_K\int_{\partial K}v_h\vec{p}^{n-1/2}\cdot\vec{n}ds,\quad \forall v_h\in V_h,\\ &\left(\mu(\mathbf{x})\nabla_h\left(u^{n-1/2}-U^{n-1/2}\right),\vec{\omega}_h\right)_h-\left(\vec{p}^{n-1/2}-\vec{P}^{n-1/2},\vec{\omega}_h\right)_h=0,\quad \forall\vec{\omega}_h\in\vec{W}_h, \end{aligned} \tag{4.9}$$

where $R_3^{n-1/2}=u_t^{n-1/2}-\partial_t u^{n-1/2}$.

Theorem 4.2. *Let $\{u(t_n),\vec{p}^n\}$ and $\{U^n,\vec{P}^n\}$ be, respectively, the solutions of the equations (2.2) and (2.3) at the point $t=t_n$. If $u\in H^4(\Omega),u_t\in H^2(\Omega),u_{tt}\in L^2(\Omega),u_{ttt}\in L^2(\Omega),\vec{p}\in(H^3(\Omega))^2$ for any $t\in(0,T]$, then*

$$\begin{aligned} &\|u^n-U^n\|_0=\mathscr{O}\left(h^2+\tau^{\min\{2-\alpha,\,3-\beta\}}\right), &&\|I_hu^n-U^n\|_h=\mathscr{O}\left(h^2+\tau^{\min\{2-\alpha,\,3-\beta\}}\right),\\ &\|\Pi_h\vec{p}^n-\vec{P}^n\|_0=\mathscr{O}\left(h^2+\tau^{\min\{2-\alpha,\,3-\beta\}}\right), &&\|\vec{p}^n-\vec{P}^n\|_0=\mathscr{O}\left(h+\tau^{\min\{2-\alpha,\,3-\beta\}}\right) \end{aligned}$$

for any $n=1,2,\dots,N$.

Proof. Choosing $v_h=\partial_t\xi^{n-1/2}$ and $\vec{\omega}_h=\nabla_h\partial_t\xi^{n-1/2}$ in (4.9) and adding the second equation to the first yields

$$\begin{aligned} &\left\|\partial_t\xi^{n-1/2}\right\|_0^2+\left(\tilde{P}\left(\tilde{D}_t^\alpha\right)\xi^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\ &\quad+\left(\hat{P}\left(\hat{D}_t^\beta\right)\xi^{n-1/2},\partial_t\xi^{n-1/2}\right)_h+\left(\mu(\mathbf{x})\nabla_h\xi^{n-1/2},\nabla_h\partial_t\xi^{n-1/2}\right)_h\\ &=-\left(\partial_t\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-\left(\tilde{P}\left(\tilde{D}_t^\alpha\right)\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\ &\quad-\left(\hat{P}\left(\hat{D}_t^\beta\right)\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-\left(\mu(\mathbf{x})\nabla_h\eta^{n-1/2},\nabla_h\partial_t\xi^{n-1/2}\right)_h\\ &\quad-\left(R_1^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-\left(R_2^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-\left(R_3^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\ &\quad+\sum_K\int_{\partial K}\partial_t\xi^{n-1/2}\vec{p}^{n-1/2}\cdot\vec{n}ds. \end{aligned} \tag{4.10}$$

Moreover, writing

$$\begin{aligned}
&\left(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\xi^{n-1/2},\partial_t\xi^{n-1/2}\right)_h \qquad (4.11)\\
&=\frac{\tau^{1-\alpha}}{2\Gamma(2-\alpha)}\left[\left(\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h+\left(\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h\right],\\
&\left(\hat{P}\left(\hat{D}_t^{\beta}\right)\xi^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&=\frac{\tau^{1-\beta}}{\Gamma(3-\beta)}\left(\hat{B}_{\beta,0}\partial_t\xi^{n-1/2}+\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k}-\hat{B}_{\beta,n-k-1}\right)\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&=\frac{\tau^{1-\beta}}{\Gamma(3-\beta)}\left[\hat{B}_{\beta,0}\left\|\partial_t\xi^{n-1/2}\right\|_0^2+\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k}-\hat{B}_{\beta,n-k-1}\right)\left(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h\right], \qquad (4.12)\\
&\left(\mu(\mathbf{x})\nabla_h\xi^{n-1/2},\nabla_h\partial_t\xi^{n-1/2}\right)_h=\frac{1}{2\tau}\left(\left\|\mu^{1/2}(\mathbf{x})\nabla_h\xi^{n}\right\|_0^2-\left\|\mu^{1/2}(\mathbf{x})\nabla_h\xi^{n-1}\right\|_0^2\right), \qquad (4.13)
\end{aligned}$$

substituting (4.11)-(4.13) into (4.10) and multiplying the result by 2τ gives

$$\begin{aligned}
&2\lambda\tau\hat{B}_{\beta,0}\left\|\partial_t\xi^{n-1/2}\right\|_0^2+\left\|\mu^{1/2}(\mathbf{x})\nabla_h\xi^{n}\right\|_0^2-\left\|\mu^{1/2}(\mathbf{x})\nabla_h\xi^{n-1}\right\|_0^2+2\tau\left\|\partial_t\xi^{n-1/2}\right\|_0^2\\
&=2\lambda\tau\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k-1}-\hat{B}_{\beta,n-k}\right)\left(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h-2\tau\left(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&\quad-2\tau\left(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-2\tau\left(\mu(\mathbf{x})\nabla_h\eta^{n-1/2},\nabla_h\partial_t\xi^{n-1/2}\right)_h\\
&\quad-2\tau\left(\partial_t\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-2\tau\left(R_1^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-2\tau\left(R_2^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&\quad-2\tau\left(R_3^{n-1/2},\partial_t\xi^{n-1/2}\right)_h+2\tau\sum_K\int_{\partial K}\partial_t\xi^{n-1/2}\vec{p}^{n-1/2}\cdot\vec{n}ds\\
&\quad-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\left(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h\right]\\
&\le\lambda\tau\sum_{k=1}^{n-1}\left(\hat{B}_{\beta,n-k-1}-\hat{B}_{\beta,n-k}\right)\left(\left\|\partial_t\xi^{k-1/2}\right\|_0^2+\left\|\partial_t\xi^{n-1/2}\right\|_0^2\right)\\
&\quad-2\tau\left(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-2\tau\left(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&\quad-2\tau\left(\mu(\mathbf{x})\nabla_h\eta^{n-1/2},\nabla_h\partial_t\xi^{n-1/2}\right)_h-2\tau\left(\partial_t\eta^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&\quad-2\tau\left(R_1^{n-1/2},\partial_t\xi^{n-1/2}\right)_h-2\tau\left(R_2^{n-1/2},\partial_t\xi^{n-1/2}\right)_h\\
&\quad-2\tau\left(R_3^{n-1/2},\partial_t\xi^{n-1/2}\right)_h+2\tau\sum_K\int_{\partial K}\partial_t\xi^{n-1/2}\vec{p}^{n-1/2}\cdot\vec{n}ds\\
&\quad-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\left(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h+\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\left(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1/2}\right)_h\right]. \qquad (4.14)
\end{aligned}$$

Setting

$$\rho^0 := \left\|\mu^{1/2}(\mathbf{x})\nabla_h\xi^0\right\|_0^2, \quad \rho^n := \left\|\mu^{1/2}(\mathbf{x})\nabla_h\xi^n\right\|_0^2 + \lambda\tau\sum_{k=1}^{n}\hat{B}_{\beta,n-k}\left\|\partial_t\xi^{k-1/2}\right\|_0^2$$

and recalling the definition of R_h, we write (4.14) as

$$\begin{aligned}
&\rho^n + 2\tau\left\|\partial_t\xi^{n-1/2}\right\|_0^2 \\
\le\ &\rho^{n-1} - 2\tau\Big(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\left(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{n-1/2}, \partial_t\xi^{n-1/2}\right)_h \\
&-2\tau\Big(\partial_t\eta^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\Big(R_1^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\Big(R_2^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h \\
&-2\tau\Big(R_3^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h + 2\tau\sum_K\int_{\partial K}\partial_t\xi^{n-1/2}\vec{p}^{\,n-1/2}\cdot\vec{n}ds \\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\Big(\partial_t\xi^{k-1/2}, \partial_t\xi^{n-1/2}\Big)_h + \sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\Big(\partial_t\xi^{k-1/2}, \partial_t\xi^{n-1/2}\Big)_h\right]. \quad (4.15)
\end{aligned}$$

In order to estimate ρ^n, we have to evaluate the fourth term in the right-hand side of (4.15). It follows from Lemma 3.2 that

$$\left|-2\tau\left(\partial_t\eta^{n-1/2}, \partial_t\xi^{n-1/2}\right)_h\right| \le Ch^4\int_{t_{n-1}}^{t_n}\|u_t\|_2^2dt + 2\tau\left\|\partial_t\xi^{n-1/2}\right\|_0^2, \quad (4.16)$$

and substituting (4.16) into (4.15) gives

$$\begin{aligned}
\rho^n \le\ &\rho^{n-1} - 2\tau\Big(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\left(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{n-1/2}, \partial_t\xi^{n-1/2}\right)_h \quad (4.17)\\
&+Ch^4\int_{t_{n-1}}^{t_n}\|u_t\|_2^2dt - 2\tau\Big(R_1^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\Big(R_2^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h \\
&-2\tau\Big(R_3^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h + 2\tau\sum_K\int_{\partial K}\partial_t\xi^{n-1/2}\vec{p}^{\,n-1/2}\cdot\vec{n}ds \\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\Big(\partial_t\xi^{k-1/2}, \partial_t\xi^{n-1/2}\Big)_h + \sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\Big(\partial_t\xi^{k-1/2}, \partial_t\xi^{n-1/2}\Big)_h\right] \\
\le\ &\rho^{n-2} - 2\tau\Big(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\Big(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h \\
&+Ch^4\int_{t_{n-1}}^{t_n}\|u_t\|_2^2dt - 2\tau\Big(R_1^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h - 2\tau\Big(R_2^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h \\
&-2\tau\Big(R_3^{n-1/2}, \partial_t\xi^{n-1/2}\Big)_h + 2\tau\sum_K\int_{\partial K}\partial_t\xi^{n-1/2}\vec{p}^{\,n-1/2}\cdot\vec{n}ds \\
&-\lambda_1\tau\left[\sum_{k=1}^{n}\tilde{B}_{\alpha,n-k}\Big(\partial_t\xi^{k-1/2}, \partial_t\xi^{n-1/2}\Big)_h + \sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\Big(\partial_t\xi^{k-1/2}, \partial_t\xi^{n-1/2}\Big)_h\right]
\end{aligned}$$

$$
\begin{aligned}
&-2\tau\Big(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{n-1-1/2},\partial_t\xi^{n-1-1/2}\Big)_h-2\tau\Big(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{n-1-1/2},\partial_t\xi^{n-1-1/2}\Big)_h\\
&+Ch^4\int_{t_{n-2}}^{t_{n-1}}\|u_t\|_2^2dt-2\tau\Big(R_1^{n-1-1/2},\partial_t\xi^{n-1-1/2}\Big)_h-2\tau\Big(R_2^{n-1-1/2},\partial_t\xi^{n-1-1/2}\Big)_h\\
&-2\tau\Big(R_3^{n-1-1/2},\partial_t\xi^{n-1-1/2}\Big)_h+2\tau\sum_K\int_{\partial K}\partial_t\xi^{n-1-1/2}\vec{p}^{n-1-1/2}\cdot\vec{n}ds\\
&-\lambda_1\tau\left[\sum_{k=1}^{n-1}\tilde{B}_{\alpha,n-k-1}\Big(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1-1/2}\Big)_h+\sum_{k=1}^{n-2}\tilde{B}_{\alpha,n-k-2}\Big(\partial_t\xi^{k-1/2},\partial_t\xi^{n-1-1/2}\Big)_h\right]\\
\le{}&\rho^0-2\tau\sum_{l=1}^{n}\Big(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h-2\tau\sum_{l=1}^{n}\Big(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h\\
&+Ch^4\sum_{l=1}^{n}\int_{t_{l-1}}^{t_l}\|u_t\|_2^2dt-2\tau\sum_{l=1}^{n}\Big(R_1^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h-2\tau\sum_{l=1}^{n}\Big(R_2^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h\\
&-2\tau\sum_{l=1}^{n}\Big(R_3^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h+2\tau\sum_{l=1}^{n}\sum_K\int_{\partial K}\partial_t\xi^{l-1/2}\vec{p}^{l-1/2}\cdot\vec{n}ds\\
&-\lambda_1\tau\sum_{l=1}^{n}\left[\sum_{k=1}^{l}\tilde{B}_{\alpha,l-k}\Big(\partial_t\xi^{k-1/2},\partial_t\xi^{l-1/2}\Big)_h+\sum_{k=1}^{l-1}\tilde{B}_{\alpha,l-k-1}\Big(\partial_t\xi^{k-1/2},\partial_t\xi^{l-1/2}\Big)_h\right].
\end{aligned}
$$

Applying the Cauchy-Schwartz inequality, the Taylor expansion, and Lemmas 3.6 and 3.7, we get

$$
\begin{aligned}
&2\tau\sum_{l=1}^{n}\Big(\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h\\
\le{}&\sum_{l=1}^{n}\frac{6\tau\left\|\tilde{P}(\tilde{D}_t^{\alpha})\eta^{l-1/2}\right\|_0^2}{\lambda\hat{B}_{\beta,n-l}}+\sum_{l=1}^{n}\frac{\lambda\tau\hat{B}_{\beta,n-l}\left\|\partial_t\xi^{l-1/2}\right\|_0^2}{6}, \qquad (4.18)\\
&2\tau\sum_{l=1}^{n}\Big(\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h\\
\le{}&\sum_{l=1}^{n}\frac{6\tau\left\|\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{l-1/2}\right\|_0^2}{\lambda\hat{B}_{\beta,n-l}}+\sum_{l=1}^{n}\frac{\lambda\tau\hat{B}_{\beta,n-l}\left\|\partial_t\xi^{l-1/2}\right\|_0^2}{6}, \qquad (4.19)\\
&2\tau\sum_{l=1}^{n}\Big(R_1^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h\\
\le{}&\sum_{l=1}^{n}\frac{C\tau\max_{0\le t\le T}\|u_{tt}(\mathbf{x},t)\|_0^2\tau^{4-2\alpha}}{\lambda\hat{B}_{\beta,n-l}}+\sum_{l=1}^{n}\frac{\lambda\tau\hat{B}_{\beta,n-l}\left\|\partial_t\xi^{l-1/2}\right\|_0^2}{6}, \qquad (4.20)\\
&2\tau\sum_{l=1}^{n}\Big(R_2^{l-1/2},\partial_t\xi^{l-1/2}\Big)_h
\end{aligned}
$$

$$\leq \sum_{l=1}^{n} \frac{C\tau \max_{0\leq t\leq T} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{6-2\beta}}{\lambda \hat{B}_{\beta,n-l}} + \sum_{l=1}^{n} \frac{\lambda\tau\hat{B}_{\beta,n-l} \left\|\partial_t \xi^{l-1/2}\right\|_0^2}{6}, \tag{4.21}$$

$$2\tau \sum_{l=1}^{n} \left(R_3^{l-1/2}, \partial_t \xi^{l-1/2}\right)_h$$

$$\leq \sum_{l=1}^{n} \frac{C\tau \max_{0\leq t\leq T} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{4}}{\lambda \hat{B}_{\beta,n-l}} + \sum_{l=1}^{n} \frac{\lambda\tau\hat{B}_{\beta,n-l} \left\|\partial_t \xi^{l-1/2}\right\|_0^2}{6}, \tag{4.22}$$

$$2\tau \sum_{l=1}^{n} \sum_{K} \int_{\partial K} \partial_t \xi^{l-1/2} \vec{p}^{l-1/2} \cdot \vec{n} ds$$

$$\leq \sum_{l=1}^{n} \frac{Ch^4\tau \left|\vec{p}^{l-1/2}\right|_3^2}{\lambda \hat{B}_{\beta,n-l}} + \sum_{l=1}^{n} \frac{\lambda\tau\hat{B}_{\beta,n-l} \left\|\partial_t \xi^{l-1/2}\right\|_0^2}{6}. \tag{4.23}$$

It follows from Lemmas 3.2 and 3.8 that

$$\sum_{l=1}^{n} \frac{\tau}{\lambda \hat{B}_{\beta,n-l}} \leq \frac{T^{\beta}\Gamma(3-\beta)}{C_0}.$$

Therefore,

$$\begin{aligned} &\sum_{l=1}^{n} \frac{6\tau \left\|\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{l-1/2}\right\|_0^2}{\lambda \tilde{B}_{\beta,n-l}} \leq \frac{6T^{\beta}\Gamma(3-\beta)}{C_0} \max_{1\leq l\leq n} \left\|\tilde{P}\left(\tilde{D}_t^{\alpha}\right)\eta^{l-1/2}\right\|_0^2 \\ \leq\, & Ch^4 \max_{1\leq l\leq n} \left\|P\left(D_t^{\alpha}\right)u^{l-1/2}\right\|_2^2 + C \max_{1\leq l\leq n} \|u_{tt}(\mathbf{x},t)\|_0^2 \tau^{4-2\alpha}, \end{aligned} \tag{4.24}$$

$$\begin{aligned} &\sum_{l=1}^{n} \frac{6\tau \left\|\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{l-1/2}\right\|_0^2}{\lambda \hat{B}_{\beta,n-l}} \leq \frac{6T^{\beta}\Gamma(3-\beta)}{C_0} \max_{1\leq l\leq n} \left\|\hat{P}\left(\hat{D}_t^{\beta}\right)\eta^{l-1/2}\right\|_0^2 \\ \leq\, & Ch^4 \max_{1\leq l\leq n} \left\|P\left(D_t^{\beta}\right)u^{l-1/2}\right\|_2^2 + C \max_{1\leq l\leq n} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{6-2\beta}, \end{aligned} \tag{4.25}$$

$$\sum_{l=1}^{n} \frac{C\tau \max_{0\leq t\leq T} \|u_{tt}(\mathbf{x},t)\|_0^2 \tau^{4-2\alpha}}{\lambda \hat{B}_{\beta,n-l}} \leq \frac{CT^{\beta}\Gamma(3-\beta)}{C_0} \max_{0\leq t\leq T} \|u_{tt}(\mathbf{x},t)\|_0^2 \tau^{4-2\alpha}, \tag{4.26}$$

$$\sum_{l=1}^{n} \frac{C\tau \max_{0\leq t\leq T} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{6-2\beta}}{\lambda \hat{B}_{\beta,n-l}} \leq \frac{CT^{\beta}\Gamma(3-\beta)}{C_0} \max_{0\leq t\leq T} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{6-2\beta}, \tag{4.27}$$

$$\sum_{l=1}^{n} \frac{C\tau \max_{0\leq t\leq T} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{4}}{\lambda \hat{B}_{\beta,n-l}} \leq \frac{CT^{\beta}\Gamma(3-\beta)}{C_0} \max_{0\leq t\leq T} \|u_{ttt}(\mathbf{x},t)\|_0^2 \tau^{4}, \tag{4.28}$$

$$\sum_{l=1}^{n} \frac{Ch^4\tau \left|\vec{p}^{l-1/2}\right|_3^2}{\lambda \hat{B}_{\beta,n-l}} \leq \frac{CT^{\beta}\Gamma(3-\beta)}{C_0} \max_{0\leq t\leq T} |\vec{p}(t)|_3^2 h^4. \tag{4.29}$$

According to Lemma 3.9, we have

$$-\lambda_1\tau\sum_{l=1}^{n}\left[\sum_{k=1}^{l}\tilde{B}_{\alpha,l-k}\Big(\partial_t\xi^{k-1/2},\partial_t\xi^{l-1/2}\Big)_h+\sum_{k=1}^{l-1}\tilde{B}_{\alpha,l-k-1}\Big(\partial_t\xi^{k-1/2},\partial_t\xi^{l-1/2}\Big)_h\right]\leq 0, \quad (4.30)$$

and substituting (4.18)-(4.30) into (4.17) gives

$$\|\xi^n\|_h\leq C\left(h^2+\tau^{\min\{2-\alpha,\ 3-\beta\}}\right). \quad (4.31)$$

Using now Lemma 3.3, the inequality (4.31) and some results from interpolation theory, we obtain

$$\|I_hu^n-U^n\|_h\leq\|I_hu^n-R_hu^n\|_h+\|\xi^n\|_h=\mathcal{O}\left(h^2+\tau^{\min\{2-\alpha,\ 3-\beta\}}\right).$$

Hence,

$$\begin{aligned}\|u^n-U^n\|_h&=\mathcal{O}\left(h+\tau^{\min\{2-\alpha,\ 3-\beta\}}\right),\\ \|u^n-U^n\|_0&=\mathcal{O}\left(h^2+\tau^{\min\{2-\alpha,\ 3-\beta\}}\right).\end{aligned}$$

Since

$$\big(\mu(\mathbf{x})\nabla_h\xi^n,\vec{\omega}_h\big)_h-\big(\vec{\theta}^n,\vec{\omega}_h\big)_h=-\big(\mu(\mathbf{x})\nabla_h\eta^n,\vec{\omega}_h\big)_h+\big(\vec{\vartheta}^n,\vec{\omega}_h\big)_h,$$

for $\vec{\omega}_h=\vec{\theta}^n$ we have

$$\|\vec{\theta}^n\|_0^2=\big(\mu(\mathbf{x})\nabla_h\xi^n,\vec{\theta}^n\big)_h+\big(\mu(\mathbf{x})\nabla_h\eta^n,\vec{\theta}^n\big)_h-\big(\vec{\vartheta}^n,\vec{\theta}^n\big)_h.$$

The Cauchy-Schwarz inequality implies that

$$\Big|\big(\mu(\mathbf{x})\nabla_h\xi^n,\vec{\theta}^n\big)_h\Big|\leq\|\mu(\mathbf{x})\nabla_h\xi^n\|_0\|\vec{\theta}^n\|_0.$$

On the other hand, Lemmas 3.2-3.4 and (3.6) give

$$\begin{aligned}&\Big|\big(\mu(\mathbf{x})\nabla\eta^n,\vec{\theta}^n\big)_h\Big|\\ =&\Big|\sum_K\int_K(\mu(\mathbf{x})-\overline{\mu(\mathbf{x})})\nabla\eta^n\vec{\theta}^n dxdy+\sum_K\int_K\overline{\mu(\mathbf{x})}\nabla(u^n-I_hu^n)\vec{\theta}^n dxdy\\ &+\sum_K\int_K\overline{\mu(\mathbf{x})}\nabla(I_hu^n-R_hu^n)\vec{\theta}^n dxdy\Big|\\ \leq& Ch^2\|u\|_2\|\vec{\theta}^n\|_0.\end{aligned}$$

It follows from Lemma 3.4 that

$$\Big|\big(\vec{\vartheta}^n,\vec{\theta}^n\big)_h\Big|\leq Ch^2|\vec{p}^n|_2\|\vec{\theta}^n\|_0.$$

Combining the above inequalities shows that

$$\|\vec{\theta}^n\|_0=\mathcal{O}\left(h^2+\tau^{\min\{2-\alpha,\ 3-\beta\}}\right),$$

and using some results from interpolation theory, we obtain

$$\left\|\vec{p}^n-\vec{P}^n\right\|_0=\mathscr{O}\left(h+\tau^{\min\{2-\alpha,\ 3-\beta\}}\right). \qquad \square$$

Let Γ_{2h} be an rectangular partition of Ω parallel with axis. Dividing each $\tilde{K}$ into four equal rectangle generates a new rectangular partition Γ_h of Ω, i.e. $\tilde{K}=\bigcup_{i=1}^4 K_i, K_i\in\Gamma_h$, $i=1,2,3,4$, cf. Fig. 1.

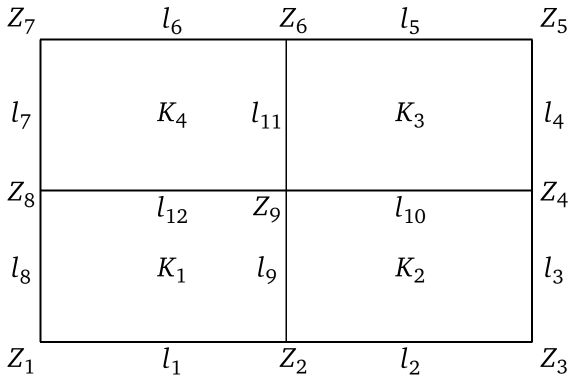

Figure 1: Element $\widetilde{K}$.

Let $L_1=\overline{Z_1Z_3}, L_2=\overline{Z_3Z_5}, L_3=\overline{Z_5Z_7}$ and $L_4=\overline{Z_7Z_1}$ be the four edges of $\tilde{K}$. For global superconvergence, we follow [34] and employ interpolation postprocessing operators I_{2h} and Π_{2h} on $\tilde{K}$ such that

$$\begin{aligned}
&I_{2h}u|_{\tilde{K}}\in P_2(\tilde{K}) && \text{for all}\quad \tilde{K}\in\Gamma_{2h},\\
&\int_{l_i}(I_{2h}u-u)ds=0, && i=1,2,3,4,\\
&\int_{K_1\bigcup K_3}(I_{2h}u-u)dxdy=0,\\
&\int_{K_2\bigcup K_4}(I_{2h}u-u)dxdy=0,
\end{aligned}$$

$$\begin{aligned}
&\Pi_{2h}\vec{w}|_{\tilde{K}}=(\Pi_{2h}w_1|_{\tilde{K}},\Pi_{2h}w_2|_{\tilde{K}})\in Q_1(\tilde{K})\times Q_1(\tilde{K}) && \text{for all}\quad \tilde{K}\in\Gamma_{2h},\\
&\int_{l_i}(\Pi_{2h}w_1-w_1)dy=0, && i=3,4,7,8,\\
&\int_{l_i}(\Pi_{2h}w_2-w_2)dx=0, && i=1,2,5,6,
\end{aligned}$$

where $P_2(\tilde{K})$ and $Q_1(\tilde{K})$ are, respectively, the sets of quadratic and bilinear polynomials on $\tilde{K}$.

Lemma 4.1 (cf. Shi *et al.* [33]). *For all $u \in H^3(\Omega)$, the interpolation operator I_{2h} satisfies the relation $I_{2h}I_h u = I_{2h}u$ and the estimates*

$$\|I_{2h}u - u\|_h \le Ch^2|u|_3,$$
$$\|I_{2h}v_h\|_h \le C\|v_h\|_h \quad \text{for all} \quad v_h \in V_h$$

hold.

Lemma 4.2 (cf. Lin *et al.* [20]). *For all $\vec{p} \in (H^2(\Omega))^2$, the interpolation operator Π_{2h} satisfies the relations $\Pi_{2h}\Pi_h\vec{p} = \Pi_{2h}\vec{p}$ and the estimates*

$$\|\Pi_{2h}\vec{p} - \vec{p}\|_0 \le Ch^2|\vec{p}|_2,$$
$$\|\Pi_{2h}\vec{p}_h\|_0 \le C\|\vec{p}_h\|_0, \quad \text{for all} \quad \vec{p}_h \in \vec{W}_h$$

hold.

Theorem 4.3. *Under the assumption of Theorem* 4.2 *and Lemmas* 4.1,4.2*, the following global superconvergence holds:*

$$\|u^n - I_{2h}U^n\|_h = \mathcal{O}\left(h^2 + \tau^{\min\{2-\alpha,\ 3-\beta\}}\right), \tag{4.32}$$
$$\|\vec{p}^n - \Pi_{2h}\vec{P}^n\|_0 = \mathcal{O}\left(h^2 + \tau^{\min\{2-\alpha,\ 3-\beta\}}\right). \tag{4.33}$$

Proof. It follows from Lemma 4.1 that

$$\|I_{2h}I_h u^n - u^n\|_h = \|I_{2h}u^n - u^n\|_h \le Ch^2\|u^n\|_3, \tag{4.34}$$
$$\begin{aligned}&\|I_{2h}U^n - I_{2h}I_h u^n\|_h \\ =\ &\|I_{2h}(U^n - I_h u^n)\|_h \le C\|U^n - I_h u^n\|_h = \mathcal{O}\left(h^2 + \tau^{\min\{2-\alpha,\ 3-\beta\}}\right).\end{aligned} \tag{4.35}$$

Writing

$$I_{2h}U^n - u^n = I_{2h}U^n - I_{2h}I_h u^n + I_{2h}I_h u^n - u^n$$

and using the triangle inequality and the estimates (4.34) and (4.35) leads to (4.32).

The estimate (4.33) can be derived analogously by exploiting Lemma 4.2. □

5. Numerical Results

We carry out numerical tests to verify the results of theoretical analysis. Let m and n denote the numbers of elements in x- and y-directions, respectively. We will also use the notation

$$h = \sqrt{\frac{1}{m^2} + \frac{1}{n^2}},$$
$$\text{err1} := \|u^n - U^n\|_0, \qquad \text{err2} := \|u^n - U^n\|_h,$$
$$\text{err3} := \|I_h u^n - U^n\|_h, \qquad \text{err4} := \|u^n - I_{2h}U^n\|_h,$$
$$\text{err5} := \|\vec{p}^n - \vec{P}^n\|_0, \qquad \text{err6} := \|\Pi_h\vec{p}^n - \vec{P}^n\|_0,$$
$$\text{err7} := \|\vec{p}^n - \Pi_{2h}\vec{P}^n\|_0.$$

Example 5.1. Let $\Omega = (0,1)\times(0,1), T = 1$, $\mu(\mathbf{x}) = xy + 0.01$ and

$$f(\mathbf{x},t) = \left[3t^2 + \frac{6t^{3-\alpha}}{\Gamma(4-\alpha)} + \frac{6t^{3-\alpha_1}}{\Gamma(4-\alpha_1)} + \frac{6t^{3-\beta}}{\Gamma(4-\beta)} + \frac{6t^{3-\beta_1}}{\Gamma(4-\beta_1)} + 2(xy+0.01)\pi^2 t^3\right] \\ \times \sin\pi x \sin\pi y - x\pi t^3 \sin\pi x \cos\pi y - y\pi t^3 \sin\pi y \cos\pi x.$$

The following problem

$$\begin{aligned} & u_t + D_t^{\alpha} u(\mathbf{x},t) + D_t^{\alpha_1} u(\mathbf{x},t) + D_t^{\beta} u(\mathbf{x},t) + D_t^{\beta_1} u(\mathbf{x},t) \\ & \quad - \nabla\cdot\big(\mu(\mathbf{x})\nabla u(\mathbf{x},t)\big) = f(\mathbf{x},t), \quad && (\mathbf{x},t)\in\Omega\times(0,T], \\ & u(\mathbf{x},t) = 0, && (\mathbf{x},t)\in\partial\Omega\times(0,T], \\ & u(\mathbf{x},0) = 0, \quad u_t(\mathbf{x},0) = 0, && \mathbf{x}\in\Omega \end{aligned}$$

has the solution $u(\mathbf{x},t) = t^3 \sin\pi x \sin\pi y$.

Example 5.2. Let $\Omega = (0,1)\times(0,1), T = 1$, $\mu(\mathbf{x}) = xy + 0.01$, and

$$f(\mathbf{x},t) = \left[2t + \frac{2t^{2-\alpha}}{\Gamma(3-\alpha)} + \frac{2t^{2-\alpha_1}}{\Gamma(3-\alpha_1)} + \frac{2t^{2-\beta}}{\Gamma(3-\beta)} + \frac{2t^{2-\beta_1}}{\Gamma(3-\beta_1)}\right] \\ \times (1-x)(1-y) - t^2y^2(1-2x)(1-y)xy + 2t^2y(xy+0.01)(1-y) \\ - t^2x^2(1-2y)(1-x) + 2t^2x(xy+0.01)(1-x).$$

The following problem

$$\begin{aligned} & u_t + D_t^{\alpha} u(\mathbf{x},t) + D_t^{\alpha_1} u(\mathbf{x},t) + D_t^{\beta} u(\mathbf{x},t) + D_t^{\beta_1} u(\mathbf{x},t) \\ & \quad - \nabla\cdot\big(\mu(\mathbf{x})\nabla u(\mathbf{x},t)\big) = f(\mathbf{x},t), \quad && (\mathbf{x},t)\in\Omega\times(0,T], \\ & u(\mathbf{x},t) = 0, && (\mathbf{x},t)\in\partial\Omega\times(0,T], \\ & u(\mathbf{x},0) = 0, \quad u_t(\mathbf{x},0) = 0, && \mathbf{x}\in\Omega \end{aligned}$$

has the solution $u(\mathbf{x},t) = t^2xy(1-x)(1-y)$.

We divide the domain into regular meshes and anisotropic meshes. Regular meshes are employed in Example 5.1. Tables 1 and 2 show error estimates and the convergence orders in temporal direction for $h^2 \approx \tau^{\min\{2-\alpha,3-\beta\}}$ and $h \approx \tau^{\min\{2-\alpha,3-\beta\}}$. We note that the error estimates in temporal direction are closer to $\min\{2-\alpha, 3-\beta\}$.

In Tables 3-6, we choose $\tau = 10^{-3}$ and present error estimates and the convergence orders for the fully-discrete scheme at different times. We point out that the accuracy of the spatial error estimates err1, err3, err4, err6 and err7 is close to 2, and the accuracy of err2 and err5 is almost one. This agrees with the theoretical analysis.

Anisotropic meshes are used in Example 5.2. Temporal errors and convergence orders are shown in Tables 7 and 8 for $h^2 \approx \tau^{\min\{2-\alpha,3-\beta\}}$ and $h \approx \tau^{\min\{2-\alpha,3-\beta\}}$, $\alpha = 0.6, \alpha_1 = 0.3, \beta = 1.3, \beta_1 = 1.2$. It is easily seen that in temporal direction, the error estimates are close to $\min\{2-\alpha,\ 3-\beta\}$.

Tables 9-12 show spatial errors and convergence orders for anisotropic meshes and $\tau = 10^{-3}$. We note that numerical results agree well with the theoretical analysis. Thus the convergence rates of err1, err3, err4, err6, err7 are close to $\mathscr{O}(h^2)$ and err2, err5 almost reach to $\mathscr{O}(h)$.

Table 1: Example 5.1. Temporal error estimates, $\alpha = 0.5, \alpha_1 = 0.4, \beta = 1.6, \beta_1 = 1.3$.

t_n	τ	$t_n/8$	rate	$t_n/16$	rate	$t_n/32$	rate	$t_n/64$	rate
$t_n = 0.5$	err1	2.143e-3	/	7.869e-4	1.445	2.886e-4	1.447	1.062e-4	1.442
	err3	9.365e-3	/	3.431e-3	1.448	1.255e-3	1.451	4.598e-4	1.448
	err4	1.670e-2	/	6.420e-3	1.379	2.412e-3	1.412	9.191e-4	1.392
	err6	3.470e-3	/	1.297e-3	1.420	4.797e-4	1.434	1.784e-4	1.427
	err7	4.066e-3	/	1.555e-3	1.386	5.835e-4	1.414	2.217e-4	1.396
$t_n = 1$	err1	1.817e-2	/	6.560e-3	1.470	2.464e-3	1.413	9.134e-4	1.431
	err3	6.668e-2	/	2.396e-2	1.477	8.664e-3	1.467	3.124e-3	1.472
	err4	3.657e-1	/	1.326e-1	1.464	5.193e-2	1.352	1.968e-2	1.400
	err6	3.396e-2	/	1.453e-2	1.437	5.615e-3	1.372	2.111e-3	1.411
	err7	8.389e-2	/	3.039e-2	1.465	1.188e-2	1.355	4.499e-3	1.401

Table 2: Example 5.1. Temporal error estimates, $\alpha = 0.5, \alpha_1 = 0.4, \beta = 1.6, \beta_1 = 1.3$.

t_n	τ	err2	rate	err5	rate
$t_n = 0.5$	$t_n/4$	2.172e-2	/	8.029e-3	/
	$t_n/8$	1.001e-2	1.438	2.970e-3	1.435
	$t_n/16$	3.671e-3	1.447	1.091e-3	1.445
	$t_n/32$	1.351e-3	1.442	4.027e-4	1.437
$t_n = 1$	$t_n/8$	9.797e-2	/	3.349e-2	/
	$t_n/16$	3.680e-2	1.412	1.265e-2	1.404
	$t_n/32$	1.361e-2	1.435	4.698e-3	1.429
	$t_n/64$	5.103e-3	1.416	1.775e-3	1.404

Table 3: Example 5.1. Spatial error and convergence order, $\alpha = 0.5, \alpha_1 = 0.3, \beta = 1.5, \beta_1 = 1.1$.

t	$m \times n$	err1	rate	err2	rate	err3	rate	err4	rate
$t_n = 0.2$	4×4	2.050e-4	/	4.015e-3	/	2.472e-4	/	6.428e-3	/
	8×8	5.148e-5	1.994	2.012e-3	0.996	5.967e-5	2.051	1.699e-3	1.919
	16×16	1.288e-5	1.999	1.007e-3	0.999	1.430e-5	2.061	4.321e-4	1.976
	32×32	3.219e-6	2.000	5.036e-4	1.000	2.840e-6	2.330	1.086e-4	1.992
$t_n = 0.4$	4×4	1.631e-3	/	3.213e-2	/	2.145e-3	/	5.128e-2	/
	8×8	4.129e-4	1.982	1.611e-2	0.996	6.338e-4	1.759	1.359e-2	1.916
	16×16	1.035e-4	1.996	8.057e-3	0.999	1.501e-4	2.078	3.454e-3	1.976
	32×32	2.590e-5	1.999	4.029e-3	1.000	3.310e-5	2.181	8.674e-4	1.993
$t_n = 0.6$	4×4	5.530e-3	/	1.085e-1	/	8.579e-3	/	1.727e-1	/
	8×8	1.403e-3	1.978	5.438e-2	0.997	2.556e-3	1.747	4.581e-2	1.914
	16×16	3.524e-4	1.994	2.720e-2	0.999	6.213e-4	2.041	1.164e-2	1.977
	32×32	8.819e-5	1.998	1.360e-2	1.000	1.488e-4	2.062	2.922e-3	1.993
$t_n = 0.8$	4×4	1.320e-2	/	2.575e-1	/	2.300e-2	/	4.084e-1	/
	8×8	3.356e-3	1.975	1.290e-1	0.998	6.876e-3	1.742	1.084e-1	1.914
	16×16	8.433e-4	1.993	6.447e-2	1.000	1.725e-3	1.995	2.753e-2	1.977
	32×32	2.111e-4	1.998	3.223e-2	1.000	4.321e-4	1.997	6.916e-3	1.993

Table 4: Example 5.1. Spatial error and convergence order, $\alpha = 0.5, \alpha_1 = 0.3, \beta = 1.5, \beta_1 = 1.1$.

t	$m \times n$	err5	rate	err6	rate	err7	rate
$t_n = 0.2$	4×4	1.544e-2	/	4.553e-4	/	1.437e-3	/
	8×8	7.831e-4	0.979	1.198e-4	1.926	3.937e-4	1.868
	16×16	3.930e-4	0.995	3.017e-5	1.989	9.974e-5	1.981
	32×32	1.967e-3	0.999	7.535e-6	2.001	2.501e-5	1.996
$t_n = 0.4$	4×4	1.236e-2	/	3.698e-3	/	1.148e-2	/
	8×8	6.268e-3	0.979	1.007e-3	1.877	3.131e-3	1.875
	16×16	3.144e-3	0.995	2.568e-4	1.972	7.905e-4	1.986
	32×32	1.573e-3	0.999	6.447e-5	1.994	1.980e-4	1.998
$t_n = 0.6$	4×4	4.177e-2	/	1.298e-2	/	3.869e-2	/
	8×8	2.116e-2	0.981	3.569e-3	1.862	1.050e-2	1.882
	16×16	1.061e-2	0.996	9.128e-4	1.967	2.646e-3	1.988
	32×32	5.310e-3	0.999	2.295e-4	1.992	6.623e-4	1.998
$t_n = 0.8$	4×4	9.915e-2	/	3.180e-2	/	9.161e-2	/
	8×8	5.019e-2	0.982	8.789e-3	1.856	2.476e-2	1.888
	16×16	2.516e-2	0.996	2.250e-3	1.972	6.233e-3	1.990
	32×32	1.259e-2	0.999	5.661e-4	1.991	1.559e-3	1.999

Table 5: Example 5.1. Spatial error and convergence order, $\alpha = 0.6, \alpha_1 = 0.3, \beta = 1.6, \beta_1 = 1.3$.

t	$m \times n$	err1	rate	err2	rate	err3	rate	err4	rate
$t_n = 0.3$	4×4	6.901e-4	/	1.355e-2	/	8.312e-4	/	2.167e-2	/
	8×8	1.739e-4	1.999	6.794e-3	0.996	2.202e-4	1.916	5.734e-3	1.918
	16×16	4.351e-5	1.998	3.399e-3	0.999	5.215e-5	2.078	1.458e-3	1.976
	32×32	1.088e-5	2.000	1.700e-3	1.000	1.044e-5	2.320	3.663e-4	1.993
$t_n = 0.5$	4×4	3.189e-3	/	6.277e-2	/	4.328e-3	/	1.001e-1	/
	8×8	8.073e-4	1.982	3.146e-2	0.996	1.273e-3	1.766	2.654e-2	1.916
	16×16	2.024e-4	1.996	1.574e-2	1.000	3.002e-4	2.084	6.743e-3	1.976
	32×32	5.063e-5	1.999	7.869e-3	1.000	6.650e-5	2.174	1.684e-3	1.993
$t_n = 0.7$	4×4	8.790e-3	/	1.724e-1	/	1.378e-2	/	2.742e-1	/
	8×8	2.230e-3	1.979	8.636e-2	0.997	4.086e-3	1.754	7.273e-2	1.915
	16×16	5.597e-4	1.994	4.319e-2	1.000	9.891e-4	2.047	1.847e-2	1.977
	32×32	1.401e-4	1.999	2.159e-2	1.000	2.364e-4	2.065	4.640e-3	1.993
$t_n = 0.9$	4×4	1.880e-2	/	3.667e-1	/	3.286e-2	/	5.816e-1	/
	8×8	4.779e-3	1.976	1.836e-1	0.998	9.782e-3	1.748	1.544e-1	1.914
	16×16	1.200e-3	1.993	9.179e-2	1.000	2.442e-3	2.002	3.920e-2	1.977
	32×32	3.005e-4	1.998	4.590e-2	1.000	6.105e-4	2.000	9.848e-3	1.993

Table 6: Example 5.1. Spatial error and convergence order, $\alpha = 0.6, \alpha_1 = 0.3, \beta = 1.6, \beta_1 = 1.3$.

t	$m \times n$	err5	rate	err6	rate	err7	rate
$t_n = 0.3$	4×4	5.210e-3	/	1.536e-3	/	4.850e-3	/
	8×8	2.643e-3	0.979	4.091e-4	1.908	1.327e-3	1.870
	16×16	1.326e-3	0.995	1.034e-4	1.984	3.358e-4	1.982
	32×32	6.637e-4	0.999	2.585e-5	2.000	8.419e-5	1.996
$t_n = 0.5$	4×4	2.414e-2	/	7.269e-3	/	2.242e-2	/
	8×8	1.224e-2	0.980	1.980e-3	1.876	6.109e-3	1.876
	16×16	6.141e-3	0.995	5.049e-4	1.972	1.542e-3	1.986
	32×32	3.073e-3	0.999	1.267e-4	1.994	3.863e-4	1.998
$t_n = 0.7$	4×4	6.633e-2	/	2.065e-2	/	6.144e-2	/
	8×8	3.361e-2	0.981	5.678e-3	1.863	1.667e-2	1.882
	16×16	1.685e-2	0.996	1.452e-3	1.968	4.201e-3	1.988
	32×32	8.432e-3	0.999	3.648e-4	1.992	1.052e-3	1.998
$t_n = 0.9$	4×4	1.412e-1	/	4.531e-2	/	1.304e-1	/
	8×8	7.146e-2	0.982	1.251e-2	1.856	3.525e-2	1.888
	16×16	3.582e-2	0.996	3.204e-3	1.966	8.875e-3	1.990
	32×32	1.792e-2	0.999	8.056e-4	1.992	2.220e-3	1.999

Table 7: Example 5.2. Temporal error estimates, $\alpha = 0.6, \alpha_1 = 0.3, \beta = 1.3, \beta_1 = 1.2$.

t_n	τ	$t_n/13$	rate	$t_n/26$	rate	$t_n/52$	rate	$t_n/104$	rate
$t_n = 0.5$	err1	2.476e-5	/	9.523e-6	1.379	3.604e-6	1.402	1.365e-6	1.400
	err3	1.432e-4	/	5.180e-5	1.467	1.836e-5	1.496	6.732e-6	1.448
	err4	1.930e-3	/	7.572e-4	1.350	2.877e-4	1.396	1.104e-4	1.381
	err6	9.879e-5	/	3.876e-5	1.350	1.471e-5	1.398	5.641e-6	1.383
	err7	4.059e-4	/	1.591e-4	1.351	6.034e-5	1.399	2.313e-5	1.383
$t_n = 1$	err1	2.116e-4	/	7.689e-5	1.460	2.972e-5	1.371	1.124e-5	1.403
	err3	1.661e-3	/	5.669e-4	1.550	2.025e-4	1.485	7.246e-5	1.482
	err4	2.132e-2	/	7.738e-3	1.462	3.041e-3	1.348	1.156e-3	1.395
	err6	1.100e-3	/	4.031e-4	1.448	1.585e-4	1.346	6.023e-5	1.396
	err7	4.435e-3	/	1.616e-3	1.457	6.334e-4	1.351	2.402e-4	1.399

Table 8: Example 5.2. Temporal error estimates, $\alpha = 0.6, \alpha_1 = 0.3, \beta = 1.3, \beta_1 = 1.2$.

t_n	τ	err2	rate	err5	rate
$t_n = 1/3$	t_n	2.726e-3	/	1.071e-3	/
	$t_n/2$	9.642e-4	1.499	3.789e-4	1.500
	$t_n/4$	3.710e-4	1.378	1.440e-4	1.396
	$t_n/8$	1.403e-4	1.403	5.395e-5	1.416
$t_n = 1$	$t_n/3$	2.410e-2	/	9.501e-3	/
	$t_n/6$	8.159e-3	1.562	3.286e-3	1.532
	$t_n/12$	3.067e-3	1.412	1.236e-3	1.411
	$t_n/24$	1.142e-3	1.425	4.599e-4	1.425

Table 9: Example 5.2. Spatial error and convergence order, $\alpha = 0.4, \alpha_1 = 0.3, \beta = 1.6, \beta_1 = 1.4$.

t	$m \times n$	err1	rate	err2	rate	err3	rate	err4	rate
$t_n = 0.2$	4×32	8.865e-6	/	9.467e-4	/	4.602e-5	/	1.964e-3	/
	8×64	2.242e-6	1.983	4.821e-4	0.974	1.478e-5	1.639	4.911e-4	2.000
	16×128	5.624e-7	1.995	2.421e-4	0.994	3.946e-6	1.905	1.230e-4	1.998
	32×256	1.415e-7	1.991	1.212e-4	0.998	7.820e-7	2.335	3.083e-5	1.996
$t_n = 0.5$	4×32	6.646e-5	/	5.933e-3	/	5.204e-4	/	1.214e-2	/
	8×64	1.678e-5	1.986	3.016e-3	0.976	1.620e-4	1.684	3.041e-3	1.997
	16×128	4.212e-6	1.994	1.513e-3	0.995	3.611e-5	2.166	7.660e-4	1.989
	32×256	1.061e-6	1.989	7.574e-4	0.999	7.612e-6	2.245	1.924e-4	1.993
$t_n = 0.7$	4×32	1.479e-4	/	1.165e-2	/	1.266e-3	/	2.362e-2	/
	8×64	3.720e-5	1.991	5.915e-3	0.978	3.710e-4	1.770	5.942e-3	1.991
	16×128	9.362e-6	1.990	2.966e-3	0.996	8.125e-5	2.191	1.500e-3	1.986
	32×256	2.361e-6	1.987	1.484e-3	0.999	1.812e-5	2.165	3.768e-4	1.993
$t_n = 0.9$	4×32	2.721e-4	/	1.930e-2	/	2.428e-3	/	3.881e-2	/
	8×64	6.811e-5	1.998	9.781e-3	0.981	6.763e-4	1.844	9.807e-3	1.985
	16×128	1.718e-5	1.987	4.904e-3	0.996	1.487e-4	2.185	2.477e-3	1.985
	32×256	4.333e-6	1.987	2.454e-3	0.999	3.438e-5	2.113	6.225e-4	1.992

Table 10: Example 5.2. Spatial error and convergence order, $\alpha = 0.4, \alpha_1 = 0.3, \beta = 1.6, \beta_1 = 1.4$.

t	$m \times n$	err5	rate	err6	rate	err7	rate
$t_n = 0.2$	4×32	3.750e-4	/	9.224e-5	/	3.827e-4	/
	8×64	1.946e-4	0.946	2.343e-5	1.977	9.878e-5	1.955
	16×128	9.821e-5	0.987	5.890e-6	1.992	2.485e-5	1.990
	32×256	4.922e-5	0.997	1.474e-6	1.998	6.225e-6	1.997
$t_n = 0.5$	4×4	2.349e-3	/	5.887e-4	/	2.404e-3	/
	8×8	1.217e-3	0.948	1.514e-4	1.960	6.199e-4	1.955
	16×16	6.139e-4	0.988	3.819e-5	1.986	1.562e-4	1.989
	32×32	3.076e-4	0.997	9.574e-6	1.996	3.912e-5	1.997
$t_n = 0.7$	4×32	4.611e-3	/	1.167e-3	/	4.723e-3	/
	8×64	2.387e-3	0.950	3.016e-4	1.952	1.219e-3	1.955
	16×128	1.203e-3	0.988	7.621e-5	1.984	3.070e-4	1.989
	32×256	6.030e-4	0.997	1.911e-5	1.995	7.691e-5	1.997
$t_n = 0.9$	4×32	7.630e-3	/	1.946e-3	/	7.824e-3	/
	8×64	3.947e-3	0.951	5.046e-4	1.947	2.019e-3	1.954
	16×128	1.990e-3	0.988	1.277e-4	1.983	5.088e-4	1.989
	32×256	9.967e-4	0.997	3.204e-5	1.995	1.275e-4	1.997

Table 11: Example 5.2. Spatial error and convergence order, $\alpha=0.6, \alpha_1=0.2, \beta=1.4, \beta_1=1.2$.

t	$m \times n$	err1	rate	err2	rate	err3	rate	err4	rate
$t_n=0.3$	4×32	2.220e-5	/	2.133e-3	/	1.569e-4	/	4.387e-3	/
	8×64	5.618e-6	1.982	1.085e-3	0.975	4.993e-5	1.651	1.098e-3	1.999
	16×128	1.420e-6	1.985	5.448e-4	0.994	1.159e-5	2.107	2.760e-4	1.992
	32×256	3.714e-7	1.934	2.726e-4	0.999	2.390e-6	2.278	6.922e-5	1.995
$t_n=0.6$	4×32	1.085e-4	/	8.560e-3	/	9.230e-4	/	1.735e-2	/
	8×64	2.729e-5	1.991	4.345e-3	0.978	2.687e-4	1.780	4.367e-3	1.990
	16×128	6.895e-6	1.985	2.179e-3	0.996	5.923e-5	2.182	1.102e-3	1.987
	32×256	1.766e-6	1.965	1.091e-3	0.999	1.322e-5	2.164	2.767e-4	1.993
$t_n=0.8$	4×32	2.141e-4	/	1.525e-2	/	1.899e-3	/	3.068e-2	/
	8×64	5.356e-5	1.999	7.728e-3	0.980	5.272e-4	1.849	7.750e-3	1.985
	16×128	1.354e-5	1.984	3.875e-3	0.996	1.164e-4	2.179	1.957e-3	1.986
	32×256	3.449e-6	1.973	1.939e-3	0.999	2.685e-5	2.116	4.917e-4	1.993
$t_n=1$	4×32	3.634e-4	/	2.387e-2	/	3.288e-3	/	4.773e-2	/
	8×64	9.041e-5	2.007	1.208e-2	0.982	8.808e-4	1.900	1.209e-2	1.981
	16×128	2.286e-5	1.984	6.055e-3	0.996	1.957e-4	2.170	3.055e-3	1.985
	32×256	5.807e-6	1.977	3.030e-3	0.999	4.608e-5	2.086	7.679e-4	1.992

Table 12: Example 5.2. Spatial error and convergence order, $\alpha=0.6, \alpha_1=0.2, \beta=1.4, \beta_1=1.2$.

t	$m \times n$	err5	rate	err6	rate	err7	rate
$t_n=0.3$	4×32	8.450e-4	/	2.101e-4	/	8.640e-4	/
	8×64	4.381e-4	0.948	5.380e-5	1.966	2.228e-4	1.956
	16×128	2.210e-4	0.987	1.355e-5	1.989	5.612e-5	1.989
	32×256	1.107e-4	0.997	3.391e-6	1.999	1.407e-5	1.995
$t_n=0.6$	4×4	3.387e-3	/	8.560e-4	/	3.470e-3	/
	8×8	1.754e-3	0.950	2.211e-4	1.953	8.952e-4	1.955
	16×16	8.841e-4	0.988	5.587e-5	1.985	2.256e-4	1.989
	32×32	4.430e-4	0.997	1.401e-5	1.996	5.653e-5	1.996
$t_n=0.8$	4×32	6.029e-3	/	1.534e-3	/	6.181e-3	/
	8×64	3.119e-3	0.951	3.978e-4	1.948	1.595e-3	1.954
	16×128	1.572e-3	0.988	1.006e-4	1.983	4.019e-4	1.989
	32×256	7.875e-4	0.997	2.524e-5	1.995	1.007e-4	1.997
$t_n=1$	4×32	9.428e-3	/	2.412e-3	/	9.676e-3	/
	8×64	4.874e-3	0.952	6.271e-4	1.944	2.497e-3	1.954
	16×128	2.456e-3	0.989	1.588e-4	1.982	6.292e-4	1.989
	32×256	1.231e-3	0.997	3.985e-5	1.995	1.577e-4	1.997

6. Summary

We developed an unconditionally stable fully-discrete scheme on regular and anisotropic meshes for multi-term TFMDDWEs with variable coefficients. Our approach is based on a nonconforming mixed FEM in space and classical $L1$ time-stepping method combined with Crank-Nicolson scheme in time. Interpolation and projection operators are used to derive superclose and convergence results. Numerical experiments are carried out for multi-term TFMDDWEs.

As we know, graded meshes is an effective tool to overcome the initial singularity of time fractional differential equations with non-smooth solutions. In near future, we plan to discuss numerical approximation of a nonconforming mixed FEM for multi-term TFMDDWEs with weak singularity solutions on graded meshes.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (Grant No. 11971416), by the Program for Scientific and Technological Innovation Talents in Universities of Henan Province (Grant No. 19HASTIT025), by the Key Scientific Research Projects in Universities of Henan Province (Grant No. 19B110013) and by the Foundation for University Key Young Teacher of Henan Province (Grant No. 2019GGJS214).

References

[1] M. Abbaszadeh and M. Dehghan, *Analysis of mixed finite element method (MFEM) for solving the generalized fractional reaction-diffusion equation on nonrectangular domains*, Comput. Math. Appl. **78**, 1531–1547 (2019).

[2] M.R.S. Ammi, I. Jamiai and D.F.M. Torres, *A finite element approximation for a class of Caputo time-fractional diffusion equations*, Comput. Math. Appl. **78**, 1334–1344 (2019).

[3] W. Bu, X. Liu, Y. Tang and J. Yang, *Finite element multigrid method for multi-term time fractional advection diffusion equations*, Int. J. Model. Simul. Sci. Comput. **6**, 1540001 (2015).

[4] H. Chen, S. Lü and W. Chen, *A unified numerical scheme for the multi-term time fractional diffusion and diffusion-wave equations with variable coefficients*, J. Comput. Appl. Math. **330**, 380–397 (2018).

[5] M. Cui, *Finite difference schemes for the variable coefficients single and multi-term time-fractional diffusion equations with non-smooth solutions on graded and uniform meshes*, Numer. Math. Theory Methods Appl. **12**, 845–866 (2019).

[6] M. Dehghan, M. Safarpoor and M. Abbaszadeh, *Two high-order numerical algorithms for solving the multi-term time fractional diffusion-wave equations*, J. Comput. Appl. Math. **290**, 174–195 (2015).

[7] S. El-Ganaiai and M.O. Al-Amr, *New abundant wave solutions of the conformable space-time fractional (4+1)-dimensional Fokas equation in water waves*, Comput. Math. Appl. **78**, 2094–2106 (2019).

[8] L. Feng, F. Liu and I. Turner, *Finite difference/finite element method for a novel 2D multi-term time-fractional mixed sub-diffusion and diffusion-wave equation on convex domains*, Commun. Nonlinear Sci. Numer. Simul. **70**, 354–371 (2019).

[9] L. Feng, F. Liu, I. Turner and L. Zheng, *Novel numerical analysis of multi-term time fractional viscoelastic non-Newtonian fluid models for simulating unsteady MHD Couette flow of a generalized Oldroyd-B fluid*, Fract. Calc. Appl. Anal. **21**, 1073–1103 (2018).

[10] L. Frunzo, R. Garra, A. Giusti and V. Luongo, *Modeling biological systems with an improved fractional Gompertz law*, Commun. Nonlinear Sci. Numer. Simul. **74**, 260–267 (2019).

[11] G. Gao and R. Liu, *A compact difference scheme for fourth-order temporal multi-term fractional wave equations and maximum error estimates*, East Asian J. Appl. Math. **9**, 703–722 (2019).

[12] E. Hashemizadeh and A. Ebrahimzadeh, *An efficient numerical scheme to solve fractional diffusion-wave and fractional Klein-Gordon equations in fluid mechanics*, Phys. A **503**, 1189–1203 (2018).

[13] W. Huang, L. Kamenski and J. Lang, *Conditioning of implicit Runge-Kutta integration for finite element approximation of linear diffusion equations on anisotropic meshes*, J. Comput. Appl. Math. (2019), https://doi.org/10.1016/j.cam.2019.112497.

[14] E. Javierre, *Impact of anomalous transport kinetics on the progress of wound healing*, Med. Eng. Phys. **38**, 885–894 (2016).

[15] A. Kumar, A. Bhardwaj and B.V.R. Kumar, *A meshless local collocation method for time fractional diffusion wave equation*, Comput. Math. Appl. **78**, 1851–1861 (2019).

[16] G. Li, C. Sun, X. Jia and D. Du, *Numerical solution to the multi-term time fractional diffusion equation in a finite domain*, Numer. Math. Theory Methods Appl. **9**, 337–357 (2016).

[17] 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**, 1269–1284 (2017).

[18] S. Li and B. Cao, *Anomalous heat diffusion from fractional Fokker-Planck equation*, Appl. Math. Lett. **99**, 105992 (2020).

[19] X. Li and H. Rui, *Two temporal second-order H^1-Galerkin mixed finite element schemes for distributed-order fractional sub-diffusion equations*, Numer. Algorithms **79**, 1107–1130 (2018).

[20] Q. Lin and J. Lin, *Finite Element Methods: Accuracy and Improvement*, Science Press (2006).

[21] Q. Lin, L. Tobiska, and A. Zhou, *Superconvergence and extrapolation of non-conforming low order finite elements applied to the Possion equation*, IMA J. Numer. Anal. **25**, 160–181 (2005).

[22] F. Liu, P. Zhuang and Q. Liu, *Numerical Methods of Fractional Partial Differential Equations and Applications*, Science Press (2015).

[23] N. Liu, Y. Liu, H. Li and J. Wang, *Time second-order finite difference/finite element algorithm for nonlinear time-fractional diffusion problem with fourth-order derivative term*, Comput. Math. Appl. **75**, 3521–3536 (2018).

[24] Z. Liu, F. Liu and F. Zeng, *An alternating direction implicit spectral method for solving two dimensional multi-term time fractional mixed diffusion and diffusion-wave equations*, Appl. Numer. Math. **136**, 139–151 (2019).

[25] Z. Liu, A. Cheng and X. Li, *A novel finite difference discrete scheme for the time fractional diffusion-wave equation*, Appl. Numer. Math. **134**, 17–30 (2018).

[26] B. Mehdinejadiani, *Estimating the solute transport parameters of the spatial fractional advection-dispersion equation using Bees algorithm*, J. Contam. Hydrol. **203**, 51–61 (2017).

[27] J.P. Neto, R.M. Coelho, D. Valério, S. Vinga, D. Sierociuk, W. Malesza, M. Macias and A. Dzieliński, *Simplifying biochemical tumorous bone remodeling models through variable order derivatives*, Comput. Math. Appl. **75**, 3147–3157 (2018).

[28] J. Ren and Z. Sun, *Efficient and stable numerical methods for multi-term time-fractional sub-diffusion equations*, East Asian J. Appl. Math. **4**, 242–266 (2014).

[29] F. Safari and W. Chen, *Coupling of the improved singular boundary method and dual reciprocity method for multi-term time-fractional mixed diffusion-wave equations*, Comput. Math. Appl. **78**, 1594–1607 (2019).

[30] M. Saffarian and A. Mohebbi, *The Galerkin spectral element method for the solution of two-dimensional multiterm time fractional diffusion-wave equation*, Math. Methods Appl. Sci. (2019), https://0-doi.org.carlson.utoledo.edu/10.1002/mma.6049.

[31] J. Shen and C. Sheng, *An efficient space-time method for time fractional diffusion equation*, J. Sci. Comput. **81**, 1088–1110 (2019).

[32] S. Shen, F. Liu and V. Anh, *The analytical solution and numerical solutions for a two-dimensional multi-term time fractional diffusion and diffusion-wave equation*, J. Comput. Appl. Math. **345**, 515–534 (2019).

[33] D. Shi, S. Mao and S. Chen, *An anisotropic nonconforming finite element with some superconvergence results*, J. Comput. Math. **23**, 261–274 (2005).

[34] D. Shi and Y. Zhang, *High accuracy analysis of a new nonconforming mixed finite element scheme for Sobolev equations*, Appl. Math. Comput. **218**, 3176–3186 (2011).

[35] Y. Shi, Y. Zhao, F. Wang and Y. Tang, *Superconvergence analysis of FEM for 2D multi-term time fractional diffusion-wave equations with variable coefficient*, Int. J. Comput. Math. **97**, 1621–1635 (2020).

[36] Z. Shi, Y. Zhao, F. Liu, F. Wang and Y. Tang, *Nonconforming quasi-Wilson finite element method for 2D multi-term time-fractional diffusion-wave equation on regular and anisotropic meshes*, Appl. Math. Comput. **338**, 290–304 (2018).

[37] Z. Soori and A. Aminataei, *Effect of the nodes near boundary points on the stability analysis of sixth-order compact finite difference ADI scheme for the two-dimensional time fractional diffusion-wave equation*, Trans. A. Razmadze Math. Inst. **172**, 582–605 (2018).

[38] H. Sun, X. Zhao and Z. Sun, *The temporal second order difference schemes based on the interpolation approximation for the time multi-term fractional wave equation*, J. Sci. Comput. **73**, 93–121 (2017).

[39] X. Sun and C. Li, *Local discontinuous Galerkin methods for the time tempered fractional diffusion equation*, Appl. Math. Comput. **365**, 124725 (2020).

[40] Z. Sun, *The Method of Order Reduction and its Application to the Numerical Solutions of Partial Differential Equations*, Science Press (2009).

[41] F. Wang, Y. Zhao, C. Chen, Y. Wei and Y. Tang, *A novel high-order approximate scheme for two-dimensional time-fractional diffusion equations with variable coefficient*, Comput. Math. Appl. **78**, 1288–1301 (2019).

[42] H. Wang and X. Zheng, *A modified time-fractional diffusion equation and its finite difference method: regularity and error analysis*, Fract. Calc. Appl. Anal. **22**, 1014–1038 (2019).

[43] Y. Wang and L. Ren, *A high-order L2-compact difference method for Caputo-type time-fractional sub-diffusion equations with variable coefficients*, Appl. Math. Comput. **342**, 71–93 (2019).

[44] Y. Wei, Y. Zhao, Y. Tang, F. Wang, Z. Shi and K. Li, *High accuracy analysis of finite-element method for two-term time-fractional mixed diffusion-wave equations*, Sci. China Inform. **48**, 871–887 (2018).

[45] X. Yang, H. Zhang and D. Xu, *Orthogonal spline collocation scheme for multiterm fractional convection-diffusion equation with variable coefficients*, Numer. Methods Partial Differential Equations **34**, 555–574 (2017).

[46] X. Yang, Q. Zhang, G. Yuan and Z. Sheng, *On positivity preservation in nonlinear finite volume method for multi-term fractional subdiffusion equation on polygonal meshes*, Nonlinear. Dyn. **92**, 595–612 (2018).

[47] M.A. Zaky, *A Legendre spectral quadrature tau method for the multi-term time-fractional diffusion equations*, Comput. Appl. Math. **37**, 3525–3538 (2017).

[48] Y. Zhao, P. Chen, W. Bu, X. Liu and Y. Tang, *Two mixed finite element methods for time-fractional diffusion equations*, J. Sci. Comput. **70**, 407–428 (2017).

[49] Y. Zhao, F. Wang, X. Hu, Z. Shi and Y. Tang, *Anisotropic linear triangle finite element approximation for multi-term time-fractional mixed diffusion and diffusion-wave equations with variable coefficient on 2D bounded domain*, Comput. Math. Appl. **78**, 1705–1719 (2019).

[50] Y. Zhao, Y. Zhang, F. Liu, I. Turner and D. Shi, *Analytical solution and nonconforming finite element approximation for the 2D multi-term fractional subdiffusion equation*, Appl. Math. Model.

40, 8810–8825 (2016).
[51] H. Zhang and X. Jiang, *Unconditionally convergent numerical method for the two-dimensional nonlinear time fractional diffusion-wave equation*, Appl. Numer. Math. **146**, 1–12 (2019).
[52] H. Zhang and J. Wang, *Superconvergence analysis of Crank-Nicolson Galerkin FEMs for a generalized nonlinear Schrödinger equation*, Numer. Methods Partial Differential Equations **34**, 799–820 (2018).
[53] H. Zhang and X. Yang, *Superconvergence analysis of nonconforming finite element method for time-fractional nonlinear parabolic equations on anisotropic meshes*, Comput. Math. Appl. **77**, 2707–2724 (2019).
[54] J. Zhang, X. Fu and H. Morris, *Construction of indicator system of regional economic system impact factors based on fractional differential equations*, Chaos Solitons Fractals **128**, 25–33 (2019).
[55] M. Zheng, F. Liu, V. Anh and I. Turner, *A high-order spectral method for the multi-term time-fractional diffusion equations*, Appl. Math. Model. **40**, 4970–4985 (2015).
[56] J. Zhou, D. Xu and H. Chen, *A weak Galerkin finite element method for multi-term time-fractional diffusion equations*, East Asian J. Appl. Math. **8**, 181–193 (2018).