fractal and fractional

Article

A Mixed Finite Element Method for the Multi-Term Time-Fractional Reaction–Diffusion Equations

Jie Zhao [1], Shubin Dong [2] and Zhichao Fang [2,*]

[1] School of Statistics and Mathematics, Inner Mongolia University of Finance and Economics, Hohhot 010070, China; zhaoj@imufe.edu.cn
[2] School of Mathematical Sciences, Inner Mongolia University, Hohhot 010021, China; madongshubin@126.com
* Correspondence: zcfang@imu.edu.cn

Abstract: In this work, a fully discrete mixed finite element (MFE) scheme is designed to solve the multi-term time-fractional reaction–diffusion equations with variable coefficients by using the well-known $L1$ formula and the Raviart–Thomas MFE space. The existence and uniqueness of the discrete solution is proved by using the matrix theory, and the unconditional stability is also discussed in detail. By introducing the mixed elliptic projection, the error estimates for the unknown variable u in the discrete $L^\infty(L^2(\Omega))$ norm and for the auxiliary variable $\boldsymbol{\lambda}$ in the discrete $L^\infty\left((L^2(\Omega))^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norms are obtained. Finally, three numerical examples are given to demonstrate the theoretical results.

Keywords: multi-term time-fractional reaction–diffusion equations; mixed finite element method; $L1$ formula; unconditional stability; error estimate

Citation: Zhao, J.; Dong, S.; Fang, Z. A Mixed Finite Element Method for the Multi-Term Time-Fractional Reaction–Diffusion Equations. *Fractal Fract.* **2024**, *8*, 51. https://doi.org/10.3390/fractalfract8010051

Academic Editor: Jordan Hristov

Received: 5 December 2023
Revised: 4 January 2024
Accepted: 8 January 2024
Published: 12 January 2024

1. Introduction

Fractional calculus and fractional partial differential equations (FPDEs) have been confirmed to be very important tools in describing some anomalous phenomena and processes with memory and nonlocal properties [1–6]. Moreover, some underlying and complex processes can be described more appropriately by multi-term FPDEs [7–9], as they contains multiple fractional derivative or calculus terms. In recent years, many numerical methods have been increasingly used by scholars to solve multi-term FPDEs. Liu et al. [10] constructed some finite difference (FD) schemes to solve the multi-term time-fractional wave-diffusion equations by using two fractional predictor–corrector methods. Dehghan et al. [11] devised two high-order numerical schemes to solve the multi-term time-fractional diffusion-wave equations by using the compact FD method and Galerkin spectral technique. Ren and Sun [12] established an efficient compact FD scheme for the multi-term time-fractional diffusion-wave equation by using the $L1$ formula. Zheng et al. [13] proposed a high-order space–time spectral method for the multi-term time-fractional diffusion equations by using the Legendre polynomials in the temporal direction and the Fourier-like basis functions in the spatial direction. Du and Sun [14] constructed an FD scheme for multi-term time-fractional mixed diffusion and wave equations by using the $L2-1_\sigma$ formula. Hendy and Zaky [15] proposed a spectral method for a coupled system of nonlinear multi-term time–space fractional diffusion equations by using the $L1$ formula on a time-graded mesh. Liu et al. [16] developed an ADI Legendre spectral method for solving a multi-term time-fractional Oldroyd-B fluid-type diffusion equation. Wei and Wang [17] constructed a higher-order numerical scheme for the multi-term variable-order time-fractional diffusion equation by using the local discontinuous Galerkin method. She et al. [18] considered a spectral method for solving the multi-term time-fractional diffusion problem by using a modified $L1$ formula.

Meanwhile, many scholars selected the finite element (FE) method for solving the multi-term FPDEs and have achieved excellent results. Jin et al. [19] developed an FE method for a multi-term time-fractional diffusion equation and considered the case of smooth and nonsmooth initial data. Li et al. [20] proposed an FE method to solve a higher-dimensional multi-term fractional diffusion equation on nonuniform time meshes. Zhou et al. [21] developed a weak Galerkin FE method for solving multi-term time-fractional diffusion equations by using a convolution quadrature formula. Bu et al. [22] proposed a space–time FE method for solving the multi-term time–space fractional diffusion equation based on the suitable graded time mesh. Feng et al. [23] proposed an FE method for a multi-term time-fractional mixed subdiffusion and diffusion-wave equation on the convex domain by using mixed L-type schemes. Meng and Stynes [24] considered an $L1$ FE method for a multi-term time-fractional initial-boundary value problem on the temporal graded mesh. Yin et al. [25] constructed a class of efficient time-stepping FE schemes for multi-term time-fractional reaction–diffusion-wave equations by using the shifted convolution quadrature method. Huang et al. [26] proposed an α-robust FE method for a multi-term time-fractional diffusion problem on a graded mesh by using the $L1$ formula. Liu et al. [27] proposed an FE method for solving a multi-term variable-order time-fractional diffusion equation and developed an efficient parallel-in-time algorithm to reduce the computational costs.

In this work, we will construct a fully discrete mixed finite element (MFE) scheme for the following multi-term time-fractional reaction–diffusion (TFRD) equations with variable coefficients:

$$\begin{cases} P(D_t)u(\boldsymbol{x},t) - \mathrm{div}(\mathcal{A}(\boldsymbol{x})\nabla u(\boldsymbol{x},t)) + p(\boldsymbol{x})u(\boldsymbol{x},t) = f(\boldsymbol{x},t), & (\boldsymbol{x},t) \in \Omega \times J, \\ u(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t) \in \partial\Omega \times \overline{J}, \\ u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), & \boldsymbol{x} \in \overline{\Omega}, \end{cases} \tag{1}$$

where $\Omega \subset R^2$ is a convex and bounded polygon region with boundary $\partial\Omega$, $J = (0,T]$ with $0 < T < \infty$. We assume that the source function $f(\boldsymbol{x},t)$, initial data $u_0(\boldsymbol{x})$, and non-negative coefficient $p(\boldsymbol{x})$ are smooth enough. Specifically, for the symmetric diffusion coefficient matrix $\mathcal{A}(\boldsymbol{x})$, we should assume that there exist two constants A_0, $A_1 > 0$ such that

$$A_0\boldsymbol{z}^T\boldsymbol{z} \leq \boldsymbol{z}^T\mathcal{A}(\boldsymbol{x})\boldsymbol{z} \leq A_1\boldsymbol{z}^T\boldsymbol{z}, \forall \boldsymbol{z} \in R^2, \forall \boldsymbol{x} \in \overline{\Omega}.$$

Moreover, the multi-term time-fractional derivative $P(D_t)u(\boldsymbol{x},t)$ is defined by

$$P(D_t)u(\boldsymbol{x},t) = \sum_{i=1}^{m} b_i D_t^{\alpha_i} u(\boldsymbol{x},t), 0 < \alpha_m < \alpha_{m-1} < \cdots < \alpha_1 < 1,$$

where $b_i \quad (i = 1, 2, \cdots, m)$ are the positive real numbers and $D_t^{\alpha_i}u$ is the Caputo time-fractional derivative as follows:

$$D_t^{\alpha_i}u(\boldsymbol{x},t) = \frac{1}{\Gamma(1-\alpha_i)} \int_0^t \frac{\partial u(x,s)}{\partial s} \frac{1}{(t-s)^{\alpha_i}} \mathrm{d}s,$$

where $\Gamma(\cdot)$ denotes the Γ-function.

It should be noted that the MFE method, as an important numerical calculation method, has been widely used to solve FPDEs [28–32], and some scholars have also used this method to solve the multi-term FPDEs [33–35]. In [33], Shi et al. proposed an H^1 -Galerkin mixed finite element (MFE) method for the multi-term time-fractional diffusion equations and gave a superconvergence result. In [34], Li et al. proposed an MFE method for the multi-term time-fractional diffusion and diffusion-wave equations by using an MFE space contained in $\left(L^2(\Omega)\right)^d \times H_0^1(\Omega)$, where $d = 2, 3$. In [35], Cao et al. constructed a nonconforming MFE scheme for the multi-term time-fractional mixed diffusion and diffusion-wave equations. Motivated by the above excellent works, we will construct a fully discrete MFE scheme for the multi-term TFRD equation (1) by using the Raviart–Thomas MFE space and the $L1$ formula, analyze the existence, uniqueness, and unconditional stability in detail, and give error estimates for u (in discrete $L^\infty\left(L^2(\Omega)\right)$ norm) and auxiliary variable $\boldsymbol{\lambda}$ (in discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ and discrete $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norms). Finally, we give numerical experiments to demonstrate the efficiency of the proposed method.

The remainder of this paper is arranged as follows. In Section 2, we construct a fully discrete MFE scheme for the multi-term TFRD equations by using the Raviart–Thomas MFE space and the $L1$ -formula. In Section 3, we give a fractional Grönwall inequality and analyze the existence and uniqueness of the discrete solution. We derive the unconditional stability results and a priori error estimates in detail in Sections 4 and 5, respectively. Finally, three numerical examples are given to verify the theoretical results.

2. Mixed Finite Element Method

We introduce the flux $\boldsymbol{\lambda}(\boldsymbol{x},t) = -\mathcal{A}(\boldsymbol{x})\nabla u(\boldsymbol{x},t)$ as the auxiliary variable. Then, the original problem (1) can be rewritten as follows:

$$\begin{cases} (a)P(D_t)u(\boldsymbol{x},t) + \mathrm{div}\boldsymbol{\lambda}(\boldsymbol{x},t) + p(\boldsymbol{x})u(\boldsymbol{x},t) = f(\boldsymbol{x},t), & (\boldsymbol{x},t) \in \Omega \times J, \\ (b)\mathcal{A}^{-1}(x)\boldsymbol{\lambda}(\boldsymbol{x},t) + \nabla u(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t) \in \Omega \times J, \\ (c)u(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t) \in \partial\Omega \times \overline{J}, \\ (d)u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), & \boldsymbol{x} \in \overline{\Omega}. \end{cases} \tag{2}$$

Let $V = L^2(\Omega)$ and $\boldsymbol{W} = \boldsymbol{H}(\mathrm{div},\Omega) = \left\{\boldsymbol{w} \in \left(L^2(\Omega)\right)^2 : \mathrm{div}\boldsymbol{w} \in L^2(\Omega)\right\}$. Then, we obtain the mixed variational formulation of (2): find $(u,\boldsymbol{\lambda}) \in V \times \boldsymbol{W}$ such that

$$\begin{cases} (a)(P(D_t)u,v) + (\mathrm{div}\boldsymbol{\lambda},v) + (pu,v) = (f,v), & \forall v \in V, \\ (b)\left(\mathcal{A}^{-1}\boldsymbol{\lambda},\boldsymbol{w}\right) - (u,\mathrm{div}\boldsymbol{w}) = 0, & \forall \boldsymbol{w} \in \boldsymbol{W}, \\ (c)u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), & \forall \boldsymbol{x} \in \overline{\Omega}. \end{cases} \tag{3}$$

Let K_h be a quasi-uniform triangulation of the domain Ω, h_T be the diameter of the triangle $T \in K_h$ and denote $h = \max\{h_T\}$. We select the Raviart–Thomas MFE space $V_h \times \boldsymbol{W}_h \subset V \times \boldsymbol{W}$, that is,

$$\begin{aligned} V_h(K) &= \{v_h \in V : v_h|_T \in P_r(T), \forall T \in K_h\}, \\ \boldsymbol{W}_h(K) &= \left\{\boldsymbol{w}_h \in \boldsymbol{W} : \boldsymbol{w}_h\Big|_T \in (P_r(T))^2 \oplus (\boldsymbol{x}P_r(T)), \forall T \in K_h\right\}, \end{aligned}$$

where the notation $\oplus$ indicates a direct sum, $\boldsymbol{x}P_r(T) = (x_1P_r(T), x_2P_r(T)), \boldsymbol{x} = (x_1,x_2)$ and $r \geq 0$ is a given integer.

Let $\tau = T/N$ and $t_n = n\tau$ for $n = 0,1,2,\cdots,N$, where N is a positive integer. For the parameters α_i and $i = 1,2,\cdots,m$, the Caputo time-fraction derivative $D_t^{\alpha_i}u(\boldsymbol{x},t)$ at $t = t_n$ is approximated by using the well-known $L1$ formula [36,37] as follows:

$$\begin{aligned} D_t^{\alpha_i}u(\boldsymbol{x},t_n) =& \frac{1}{\Gamma(1-\alpha_i)}\int_0^{t_n}\frac{\partial u(x,s)}{\partial s}\frac{1}{(t_n-s)^{\alpha_i}}ds \\ =& \frac{1}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\frac{u^{k+1}-u^k}{\tau}\left[(t_n-t_k)^{1-\alpha_i} - (t_n-t_{k+1})^{1-\alpha_i}\right] + Q_{\alpha_i}^n(x) \\ =& \frac{1}{\Gamma(2-\alpha_i)}\left[d_{\alpha_i,1}^n u^n + \sum_{k=1}^{n-1}\left(d_{\alpha_i,k+1}^n - d_{\alpha_i,k}^n\right)u^{n-k} - d_{\alpha_i,n}^n u^0\right] + Q_{\alpha_i}^n(x) \\ =& \frac{1}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n}\widetilde{d}_{\alpha_i,k}^n u^k + Q_{\alpha_i}^n(x), \end{aligned} \tag{4}$$

where $d^n_{\alpha_i,k} = \frac{(t_n-t_{n-k})^{1-\alpha_i}-(t_n-t_{n-k+1})^{1-\alpha_i}}{\tau}$, $\widetilde{d}^n_{\alpha_i,0} = -d^n_{\alpha_i,n}$, $\widetilde{d}^n_{\alpha_i,n} = -d^n_{\alpha_i,1}$, and $\widetilde{d}^n_{\alpha_i,k} = d^n_{\alpha_i,n-k+1} - d^n_{\alpha_i,n-k}(0 < k \leq n-1)$. Setting $D^{\alpha_i}_N u^n = \frac{1}{\Gamma(2-\alpha_i)} \sum\limits_{k=0}^{n} \widetilde{d}^n_{\alpha_i,k} u^k$, we have $D^{\alpha_i}_t u(\boldsymbol{x}, t_n) = D^{\alpha_i}_N u^n + Q^n_{\alpha_i}(x)$, where $Q^n_{\alpha_i}(x)$ is the truncation error.

Based on the above definitions, and setting u^n_h and $\boldsymbol{\lambda}^n_h$ to be the discrete solutions of u and $\boldsymbol{\lambda}$ at $t = t_n$, respectively, then we can design a fully discrete MFE scheme for the original problem (1): find $\left(u^n_h, \lambda^n_h\right) \in V_h \times \boldsymbol{W}_h$ such that

$$\begin{cases} (a)\left(\sum\limits_{i=1}^{m} b_i D^{\alpha_i}_N u^n_h, v_h\right) + \left(\mathrm{div}\boldsymbol{\lambda}^n_h, v_h\right) + \left(p u^n_h, v_h\right) = (f^n, v_h), & \forall v_h \in V_h, \\ (b)\left(\mathcal{A}^{-1}\boldsymbol{\lambda}^n_h, \boldsymbol{w}_h\right) - \left(u^n_h, \mathrm{div}\boldsymbol{w}_h\right) = 0, & \forall \boldsymbol{w}_h \in \boldsymbol{W}_h, \end{cases} \tag{5}$$

where $\left(u^0_h, \boldsymbol{\lambda}^0_h\right) \in V_h \times \boldsymbol{W}_h$ satisfies

$$\begin{cases} (a)\left(\mathcal{A}^{-1}\boldsymbol{\lambda}^0_h, \boldsymbol{w}_h\right) - \left(u^0_h, \mathrm{div}\boldsymbol{w}_h\right) = 0, & \forall \boldsymbol{w}_h \in \boldsymbol{W}_h, \\ (b)\left(\mathrm{div}\boldsymbol{\lambda}^0_h, v_h\right) = \left(\mathrm{div}\boldsymbol{\lambda}_0, v_h\right), & \forall v_h \in V_h, \end{cases} \tag{6}$$

where $\boldsymbol{\lambda}_0(\boldsymbol{x}) = -\mathcal{A}(\boldsymbol{x})\nabla u_0(\boldsymbol{x})$.

Remark 1. *(I) In the MFE scheme (5)–(6), we particularly emphasize the calculation of initial values $\left(u^0_h, \boldsymbol{\lambda}^0_h\right)$, as this calculation will be used in stability and convergence analyses. Moreover, from the mixed elliptic projection R_h defined in Section 5, we can see that $\left(u^0_h, \boldsymbol{\lambda}^0_h\right) = (R_h u_0, R_h \boldsymbol{\lambda}_0)$.*

(II) Compared with the standard FE methods, it is well known that the MFE method can not only reduce the smoothness requirement of the finite element space, but also simultaneously calculate multiple physical quantities. These advantages are very important and popular in practical applications.

3. Existence and Uniqueness

In this section, we shall prove the existence and uniqueness for the MFE scheme (5)–(6). We first give some lemmas, which are important in subsequent theoretical analysis.

Lemma 1 ([38])**.** *There exist two positive constants μ_0 and μ_1 such that*

$$\mu_0 \| \boldsymbol{w} \|^2 \leq \| \boldsymbol{w} \|^2_{\mathcal{A}^{-1}} \leq \mu_1 \| \boldsymbol{w} \|^2, \text{where} \| \boldsymbol{w} \|^2_{\mathcal{A}^{-1}} = \left(\mathcal{A}^{-1}\boldsymbol{w}, \boldsymbol{w}\right), \forall \boldsymbol{w} \in \boldsymbol{W}.$$

Lemma 2 ([28])**.** *Let $\{\boldsymbol{z}^n\}^{\infty}_{n=0}$ be a sequence on $\boldsymbol{W}_h$. Then, the following identity holds:*

$$\begin{aligned} \sum_{k=0}^{n} \widetilde{d}^n_{\alpha_i,k}\left(\mathcal{A}^{-1}\boldsymbol{z}^k, \boldsymbol{z}^n\right) =& \frac{1}{2}\Big[\widetilde{d}^n_{\alpha_i,n}\left(\mathcal{A}^{-1}\boldsymbol{z}^n, \boldsymbol{z}^n\right) + \sum_{k=0}^{n-1} \widetilde{d}^n_{\alpha_i,k}\left(\mathcal{A}^{-1}\boldsymbol{z}^k, \boldsymbol{z}^k\right) \\ &+ \sum_{k=0}^{n-1} \widetilde{d}^n_{\alpha_i,k}\left(\mathcal{A}^{-1}\left(\boldsymbol{z}^n - \boldsymbol{z}^k\right), \boldsymbol{z}^n - \boldsymbol{z}^k\right)\Big]. \end{aligned}$$

Lemma 3. *Let $\left\{\varphi^k : 0 \leq k \leq N\right\}$ be a non-negative sequence, $\left\{\boldsymbol{\xi}^k : 0 \leq k \leq N\right\}$ be a nondecreasing positive sequence, and $C_0 \geq 1$ be a constant, which satisfy*

$$\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}^n_{\alpha_i,n} \varphi^n \leq -C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}^n_{\alpha_i,k} \varphi^k + \boldsymbol{\xi}^n, 1 \leq n \leq N. \tag{7}$$

Then, we have

$$\varphi^n \leq C^n_0\Big(\varphi^0 + \frac{1}{\sum\limits_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d^n_{\alpha_i,n}} \boldsymbol{\xi}^n\Big), 1 \leq n \leq N. \tag{8}$$

Further, we can further write (8) as

$$\varphi^n \leq C_0^n(\varphi^0 + \sum_{i=1}^{m} \frac{\Gamma(1-\alpha_i)t_n^{\alpha_i}}{b_i}\xi^n), 1 \leq n \leq N. \tag{9}$$

Proof. When $n = 1$ in (7), we have

$$\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,1}^{1} \varphi^1 \leq -C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,0}^{1} \varphi^0 + \xi^1. \tag{10}$$

Noting that $\widetilde{d}_{\alpha_i,0}^{n} = -d_{\alpha_i,n}^{n}, \widetilde{d}_{\alpha_i,n}^{n} = d_{\alpha_i,1}^{n}$, we have

$$\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,1}^{1} \varphi^1 \leq C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,1}^{1} \varphi^0 + \xi^1. \tag{11}$$

Then, we can obtain

$$\varphi^1 \leq C_0(\varphi^0 + \frac{1}{\sum\limits_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,1}^{1}} \xi^1). \tag{12}$$

It means that the conclusion (8) is valid for the case of $n = 1$. Assume that (8) is valid for $n = 1, 2, \cdots, r$. We now need to prove that it also holds for $n = r + 1$. Selecting $n = r + 1$ in (7), we obtain

$$\begin{aligned}
&\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,j+1}^{j+1} \varphi^{j+1} \\
\leq & -C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{j} \widetilde{d}_{\alpha_i,k}^{j+1} \varphi^k + \xi^{j+1} \\
= & C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=1}^{j} \left(d_{\alpha_i,j-k+1}^{j+1} - d_{\alpha_i,j-k+2}^{j+1} \right) \varphi^k + C_0 \sum_{i=1}^{m} b_i d_{\alpha_i,j+1}^{j+1} \varphi^0 + \xi^{j+1} \\
= & C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{j-1} \left(d_{\alpha_i,k+1}^{j+1} - d_{\alpha_i,k+2}^{j+1} \right) \varphi^{j-k} + C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,j+1}^{j+1} \varphi^0 + \xi^{j+1}.
\end{aligned} \tag{13}$$

Noting that $0 < d_{\alpha_i,k+1}^{k+1} < d_{\alpha_i,k}^{k}$ and $0 < d_{\alpha_i,k+1}^{n} < d_{\alpha_i,k}^{n}, (k = 0, 1 \cdots j)$, we have

$$\begin{aligned}
&\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,j+1}^{j+1} \varphi^{j+1} \\
\leq & C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{j-1} \left(d_{\alpha_i,k+1}^{j+1} - d_{\alpha_i,k+2}^{j+1} \right) [C_0^{j-k}(\varphi^0 + \frac{1}{\sum\limits_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,j-k}^{j-k}} \xi^{j-k})] \\
& + C_0 \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \mathrm{d}_{\alpha_i,j+1}^{j+1} (\varphi^0 + \frac{1}{\sum\limits_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,j+1}^{j+1}} \xi^{j+1}) \\
\leq & C_0^{j+1} \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} [\sum_{k=0}^{j-1} \left(d_{\alpha_i,k+1}^{j+1} - d_{\alpha_i,k+2}^{j+1} \right) + d_{\alpha_i,j+1}^{j+1}] (\varphi^0 + \frac{1}{\sum\limits_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,j+1}^{j+1}} \xi^{j+1}) \\
\leq & C_0^{j+1} \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,1}^{j+1} (\varphi^0 + \frac{1}{\sum\limits_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} d_{\alpha_i,j+1}^{j+1}} \xi^{j+1}).
\end{aligned} \tag{14}$$

Therefore, using the mathematical induction method, we can complete the proof of (8). From [37], we know $n^{\alpha_i}\left(n^{1-\alpha_i}-(n-1)^{1-\alpha_i}\right)\geq 1-\alpha_i$, and then

$$d^n_{\alpha_i,n}=\frac{(n\tau)^{1-\alpha_i}-(n\tau-\tau)^{1-\alpha_i}}{\tau}=\frac{\left(n^{1-\alpha_i}-(n-1)^{1-\alpha_i}\right)}{\tau^{\alpha_i}}\geq\frac{(1-\alpha_i)}{\tau^{\alpha_i}n^{\alpha_i}}. \tag{15}$$

Thus, making use of (8) and (15), we can complete the proof of (9). □

Next, we give the existence and uniqueness results for the MFE scheme (5).

Theorem 1. *The MFE scheme (5) has a unique solution.*

Proof. Let $V_h=\text{span}\{\phi_1,\phi_2,\cdots,\phi_{M_1}\}$ and $\boldsymbol{W}_h=\text{span}\{\boldsymbol{\psi}_1,\boldsymbol{\psi}_2,\cdots,\boldsymbol{\psi}_{M_2}\}$. Then, u_h^n and $\boldsymbol{\lambda}_h^n$ can be written as

$$u_h^n=\sum_{i=1}^{M_1}\widetilde{u}_i^n\phi_i,\boldsymbol{\lambda}_h^n=\sum_{j=1}^{M_2}\widetilde{s}_j^n\boldsymbol{\psi}_j. \tag{16}$$

Substituting 16 into (5) and selecting $v_h=\phi_i\ \ (i=1,2,\cdots,M_1)$ and $\boldsymbol{w}_h=\boldsymbol{\psi}_j\ \ (j=1,2,\cdots,M_2)$, we have

$$\begin{bmatrix}\sum\limits_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^n_{\alpha_i,n}B_1+B_3 & D^T\\ -D & B_2\end{bmatrix}\begin{bmatrix}U^n\\ L^n\end{bmatrix}=\begin{bmatrix}F^n-\sum\limits_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum\limits_{k=0}^{n-1}\widetilde{d}^n_{\alpha_i,k}B_1U^k\\ 0\end{bmatrix}, \tag{17}$$

where

$$\begin{aligned}&U^n=\left(\widetilde{u}_1^n,\widetilde{u}_2^n,\cdots,\widetilde{u}_{M_1}^n\right)^T, && L^n=\left(\widetilde{s}_1^n,\widetilde{s}_2^n,\cdots,\widetilde{s}_{M_2}^n\right)^T,\\ &B_1=\left((\phi_i,\phi_j)\right)_{M_1\times M_1}, && B_2=\left((\mathcal{A}^{-1}\boldsymbol{\psi}_i,\boldsymbol{\psi}_j)\right)_{M_2\times M_2},\\ &B_3=\left((p\phi_i,\phi_j)\right)_{M_1\times M_1}, && D=\left((\text{div}\boldsymbol{\psi}_i,\phi_j)\right)_{M_2\times M_1},\\ &F^n=((f^n,\phi_i))_{M_1\times 1},\end{aligned}$$

Noting that B_1 and B_2 are symmetric positive definite matrices and B_3 is a symmetric semi-positive matrix, we have

$$\begin{bmatrix}E & -D^TB_2^{-1}\\ 0 & E\end{bmatrix}\begin{bmatrix}\sum\limits_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^n_{\alpha_i,n}B_1+B_3 & D^T\\ -D & B_2\end{bmatrix}=\begin{bmatrix}G & 0\\ -D & B_2\end{bmatrix}. \tag{18}$$

where $G=\sum\limits_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^n_{\alpha_i,n}B_1+B_3+D^TB_2^{-1}D$. It is easy to see that G is invertible, so the coefficient matrix of linear Equation (17) is invertible. This means that the MFE scheme (5) has a unique solution. □

Remark 2. *For Lemma 3, when $C_0=1$, a similar conclusion can be seen from the proof of Theorem 3.1 in [20]. When $C_0>1$, some special applications can be seen from [39]. It should be noted that this lemma can be considered a fractional Grönwall inequality without any other conditions for its existence, which will play a crucial role in the subsequent proof process of stability and convergence analyses.*

4. Stability Analysis

In this section, we will discuss the unconditional stability for the MFE scheme (5)–(6).

Theorem 2. *Let $\left(u_h^n, \boldsymbol{\lambda}_h^n\right)_{n=1}^N$ be the solutions of the MFE scheme (5). Then, there exists a constant $C > 0$ independent of h and N such that*

$$\| u_h^n \| \leq \| u_h^0 \| + \sum_{i=1}^{m} \frac{\Gamma(1-\alpha_i)t_n^{\alpha_i}}{b_i} \sup_{t\in[0,T]} \| f(t) \| \triangleq U_h^{\diamond},$$

$$\| \boldsymbol{\lambda}_h^n \| \leq C\left(\| \boldsymbol{\lambda}_h^0 \| + \left(\sum_{i=1}^{m} \frac{\Gamma(1-\alpha_i)t_n^{\alpha_i}}{b_i}\right)^{1/2} \left(\sup_{t\in[0,T]} \| f(t) \| + \| p \|_\infty U_h^{\diamond}\right)\right).$$

Proof. Taking $v_h = u_h^n$ and $w_h = \boldsymbol{\lambda}_h^n$ in (5), we have

$$\left(\sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n, u_h^n\right) + \left(\mathcal{A}^{-1}\boldsymbol{\lambda}_h^n, \boldsymbol{\lambda}_h^n\right) + (pu_h^n, u_h^n) = (f^n, u_h^n). \tag{19}$$

Using Lemma 1 and the definition of $D_N^{\alpha_i} u_h^n$, we have

$$\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,n}^n (u_h^n, u_h^n) + \mu_0 \| \boldsymbol{\lambda}_h^n \|^2 \leq \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}_{\alpha_i,k}^n \left(u_h^k, u_h^n\right) + (f^n, u_h^n). \tag{20}$$

Applying the Cauchy–Schwarz inequality yields

$$\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,n}^n \| u_h^n \| \leq \sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}_{\alpha_i,k}^n \| u_h^k \| + \| f^n \|. \tag{21}$$

Using Lemma 3, we obtain

$$\| u_h^n \| \leq \| u_h^0 \| + \sum_{i=1}^{m} \frac{\Gamma(1-\alpha_i)t_n^{\alpha_i}}{b_i} \sup_{t\in[0,T]} \| f(t) \| \triangleq U_h^{\diamond}. \tag{22}$$

Next, using (5) (b) and (6) (a), we have

$$\left(A^{-1} \sum_{i=1}^{m} b_i D_N^{\alpha_i} \boldsymbol{\lambda}_h^n, \boldsymbol{w}_h\right) - \left(\sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n, \mathrm{div}\boldsymbol{w}_h\right) = 0, \forall \boldsymbol{w}_h \in \boldsymbol{W}_h. \tag{23}$$

Choosing $v_h = \sum\limits_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n$ and $\boldsymbol{w}_h = \boldsymbol{\lambda}_h^n$ in (5) (a) and (23), respectively, we obtain

$$\| \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n \|^2 + \left(A^{-1} \sum_{i=1}^{m} b_i D_N^{\alpha_i} \boldsymbol{\lambda}_h^n, \boldsymbol{\lambda}_h^n\right) + \left(pu_h^n, \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n\right) = \left(f^n, \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n\right). \tag{24}$$

Using Lemma 2 in (24) yields

$$\begin{aligned} &\| \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n \|^2 + \frac{1}{2} \sum_{i=0}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \left[\widetilde{d}_{\alpha_i,n}^n \left(A^{-1}\boldsymbol{\lambda}_h^n, \boldsymbol{\lambda}_h^n\right) + \sum_{k=0}^{n-1} \widetilde{d}_{\alpha_i,k}^n \left(A^{-1}\boldsymbol{\lambda}_h^k, \boldsymbol{\lambda}_h^k\right) \right. \\ &\left. - \sum_{k=0}^{n-1} \widetilde{d}_{\alpha_i,k}^n \left(A^{-1}\left(\boldsymbol{\lambda}_h^n - \boldsymbol{\lambda}_h^k\right), \boldsymbol{\lambda}_h^n - \boldsymbol{\lambda}_h^k\right)\right] = \left(f^n, \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n\right) - \left(pu_h^n, \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n\right). \end{aligned} \tag{25}$$

Because of $\widetilde{d}^n_{\alpha_i,k} < 0, 0 < k \le n-1$, we have

$$\begin{aligned}
&\| \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n \|^2 + \frac{1}{2}\sum_{i=0}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}^n_{\alpha_i,n}\left(A^{-1}\boldsymbol{\lambda}_h^n, \boldsymbol{\lambda}_h^n\right) \\
\le& -\frac{1}{2}\sum_{i=0}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}^n_{\alpha_i,k}\left(A^{-1}\boldsymbol{\lambda}_h^k, \boldsymbol{\lambda}_h^k\right) \\
&+\left(f^n, \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n\right) - \left(p u_h^n, \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n\right).
\end{aligned} \tag{26}$$

Apply the Cauchy–Schwarz inequality and the Young inequality in (26) to obtain

$$\begin{aligned}
&\| \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n \|^2 + \frac{1}{2}\sum_{i=0}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}^n_{\alpha_i,n}\left(A^{-1}\boldsymbol{\lambda}_h^n, \boldsymbol{\lambda}_h^n\right) \\
\le& -\frac{1}{2}\sum_{i=0}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}^n_{\alpha_i,k}\left(A^{-1}\boldsymbol{\lambda}_h^k, \boldsymbol{\lambda}_h^k\right) \\
&+\frac{1}{2}\| \sum_{i=1}^{m} b_i D_N^{\alpha_i} u_h^n \|^2 + \| f^n \|^2 + \| p \|_\infty^2 \| u_h^n \|^2.
\end{aligned} \tag{27}$$

Using Lemma 3 in (27), we obtain

$$\| \boldsymbol{\lambda}_h^n \| \le C\left(\| \boldsymbol{\lambda}_h^0 \| + \left(\sum_{i=1}^{m} \frac{\Gamma(1-\alpha_i) t_n^{\alpha_i}}{b_i}\right)^{1/2} \left(\sup_{t\in[0,T]} \| f(t) \| + \| p \|_\infty U_h^\diamond \right)\right). \tag{28}$$

Thus, we complete the proof. □

5. Convergence Analysis

In this section, we will present the convergence results. For this purpose, we first introduce the mixed elliptic projection $(R_h u, R_h \boldsymbol{\lambda}) \in V_h \times \boldsymbol{W}_h$ defined by

$$\begin{cases}
(a)\left(\mathcal{A}^{-1}(\boldsymbol{\lambda} - R_h\boldsymbol{\lambda}), \boldsymbol{w}_h\right) - (u - R_h u, \mathrm{div}\boldsymbol{w}_h) = 0, & \forall \boldsymbol{w}_h \in \boldsymbol{W}_h, \\
(b)(\mathrm{div}(\boldsymbol{\lambda} - R_h\boldsymbol{\lambda}), v_h) = 0, & \forall v_h \in V_h.
\end{cases} \tag{29}$$

Then, the above projection satisfies the classical estimates as follows.

Lemma 4 ([40,41])**.** *There exists a constant $C > 0$ independent of h and N such that*

$$\begin{aligned}
&\| \boldsymbol{\lambda} - R_h\boldsymbol{\lambda} \| \le Ch^{r+1} \| \boldsymbol{\lambda} \|_{r+1}, for \boldsymbol{\lambda} \in \left(H^{r+1}(\Omega)\right)^2, \\
&\| \mathrm{div}(\boldsymbol{\lambda} - R_h\boldsymbol{\lambda}) \| \le Ch^{r+1} \| \mathrm{div}\boldsymbol{\lambda} \|_{r+1}, for \mathrm{div}\boldsymbol{\lambda} \in H^{r+1}(\Omega), \\
&\| u - R_h u \| \le Ch^{r+1}\left(\| u \|_{r+1} + \| \boldsymbol{\lambda} \|_{r+1}\right), for u \in H^{r+1}(\Omega), \boldsymbol{\lambda} \in \left(H^{r+1}(\Omega)\right)^2.
\end{aligned}$$

For the truncation error $Q^n_{\alpha_i}\ \ (i = 1, 2, \cdots, m)$ of the $L1$ formula, from [36,37], we give the following estimates.

Lemma 5. *Let $u \in \mathcal{C}^2\left(\overline{J}, L^2(\Omega)\right)$. Then, we have*

$$\begin{aligned}
&\| Q^n_{\alpha_i} \| \le CN^{-(2-\alpha_i)}, i = 1, 2, \cdots, m, \\
&\| \sum_{i=1}^{m} b_i Q^n_{\alpha_i} \| \le CN^{-(2-\alpha_1)},
\end{aligned}$$

where $C > 0$ is a constant independent of h and N.

Now, we write the errors $u(t_n) - u_h^n = u(t_n) - R_h u(t_n) + R_h u(t_n) - u_h^n = \rho^n + \theta^n$ and $\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_h^n = \boldsymbol{\lambda}(t_n) - R_h\boldsymbol{\lambda}(t_n) + R_h\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_h^n = \boldsymbol{\xi}^n + \boldsymbol{\eta}^n$. From (3) and (5), making use of the mixed elliptic projection R_h, we have the following error equations:

$$\begin{cases} (a)\left(\sum\limits_{i=1}^{m} b_i D_N^{\alpha_i}(\theta^n+\rho^n), v_h\right) + (\mathrm{div}\boldsymbol{\eta}^n, v_h) + (p(\theta^n+\rho^n), v_h) \\ = -\left(\sum\limits_{i=1}^{m} b_i Q_{\alpha_i}^n, v_h\right), & \forall v_h \in V_h, \\ (b)\left(\mathcal{A}^{-1}\boldsymbol{\eta}^n, \boldsymbol{w}_h\right) - (\theta^n, \mathrm{div}\boldsymbol{w}_h) = 0, & \forall \boldsymbol{w}_h \in \boldsymbol{W}_h. \end{cases} \tag{30}$$

Noting that $\left(u_h^0, \boldsymbol{\lambda}_h^0\right) = (R_h u_0, R_h\boldsymbol{\lambda}_0)$, we have $\theta^0 = 0$ and $\boldsymbol{\eta}^0 = \mathbf{0}$. We next give the convergence results for the MFE scheme (5)–(6).

Theorem 3. *Let $(u^n, \boldsymbol{\lambda}^n) \in V \times \boldsymbol{W}$ and $\left(u_h^n, \boldsymbol{\lambda}_h^n\right) \in V_h \times \boldsymbol{W}_h$ be the solutions of (3) and (5), respectively. Assume that $u, div\boldsymbol{\lambda} \in C^2\left(\bar{J}, H^{r+1}(\Omega)\right)$, $\boldsymbol{\lambda} \in C^2\left(\bar{J}, \left(H^{r+1}(\Omega)\right)^2\right)$. Then, we have*

$$\max_{1\le n\le N} \| u(t_n) - u_h^n \| + \max_{1\le n\le N} \| \boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_h^n \| \le C\left(h^{r+1} + N^{-(2-\alpha_1)}\right),$$
$$\max_{1\le n\le N} \| \boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_h^n \|_{\boldsymbol{H}(\mathrm{div},\Omega)} \le C\left(1 + N^{\frac{\alpha_m}{2}}\right)\left(h^{r+1} + N^{-(2-\alpha_1)}\right),$$

where $C > 0$ is a constant independent of h and N.

Proof. Taking $v_h = \theta^n$ and $\boldsymbol{w}_h = \boldsymbol{\eta}^n$ in (30), we can obtain

$$\begin{aligned} &\left(\sum_{i=1}^{m} b_i D_N^{\alpha_i}\theta^n, \theta^n\right) + \left(\mathcal{A}^{-1}\boldsymbol{\eta}^n, \boldsymbol{\eta}^n\right) + (p\theta^n, \theta^n) \\ =&-\left(\sum_{i=1}^{m} b_i Q_{\alpha_i}^n, \theta^n\right) - (p\rho^n, \theta^n) - \left(\sum_{i=1}^{m} b_i D_N^{\alpha_i}\rho^n, \theta^n\right). \end{aligned} \tag{31}$$

Noting that $p(\boldsymbol{x}) \ge 0$, using the Lemma 1 and the definition of $D_N^{\alpha_i} u_h^n$, we have

$$\begin{aligned} &\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,n}^n (\theta^n, \theta^n) + \mu_0 \| \boldsymbol{\eta}^n \|^2 \\ \le&-\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}_{\alpha_i,k}^n \left(\theta^k, \theta^n\right) - \left(\sum_{i=1}^{m} b_i Q_{\alpha_i}^n, \theta^n\right) - (p\rho^n, \theta^n) - \left(\sum_{i=1}^{m} b_i D_N^{\alpha_i}\rho^n, \theta^n\right). \end{aligned} \tag{32}$$

Applying the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} &\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \widetilde{d}_{\alpha_i,n}^n \| \theta^n \|^2 + \mu_0 \| \boldsymbol{\eta}^n \|^2 \\ \le&-\sum_{i=1}^{m} \frac{b_i}{\Gamma(2-\alpha_i)} \sum_{k=0}^{n-1} \widetilde{d}_{\alpha_i,k}^n \| \theta^k \| \| \theta^n \| \\ &+(\| \sum_{i=1}^{m} b_i Q_{\alpha_i}^n \| + \| p \|_\infty \| \rho \| + \| \sum_{i=1}^{m} b_i D_N^{\alpha_i}\rho^n \|) \| \theta^n \|, \end{aligned} \tag{33}$$

and then

$$\begin{aligned}&\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^{n}_{\alpha_i,n}\parallel\theta^n\parallel^2\\ \leq&-\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\widetilde{d}^{n}_{\alpha_i,k}\parallel\theta^k\parallel+(\parallel\sum_{i=1}^{m}b_iQ^n_{\alpha_i}\parallel+\parallel p\parallel_\infty\parallel\rho\parallel+\parallel\sum_{i=1}^{m}b_iD_N^{\alpha_i}\rho^n\parallel).\end{aligned}\tag{34}$$

Using Lemmas 4 and 5, we obtain

$$\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^{n}_{\alpha_i,n}\parallel\theta^n\parallel\leq-\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\widetilde{d}^{n}_{\alpha_i,k}\parallel\theta^k\parallel+C\Big(N^{-(2-\alpha_1)}+h^{r+1}\Big).\tag{35}$$

Noting that $\theta^0=0$ and using Lemma 3, we obtain

$$\parallel\theta^n\parallel\leq C\Big(h^{r+1}+N^{-(2-\alpha_1)}\Big).\tag{36}$$

Now, from (30) (b), we obtain

$$\left(\mathcal{A}^{-1}\sum_{i=1}^{m}b_iD_N^{\alpha_i}\boldsymbol{\eta}^n,\boldsymbol{w}_h\right)-\left(\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n,\mathrm{div}\boldsymbol{w}_h\right)=0,\forall\boldsymbol{w}_h\in\boldsymbol{W}_h.\tag{37}$$

Choosing $v_h=\sum\limits_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n$ and $\boldsymbol{w}_h=\boldsymbol{\eta}^n$ in (30) (a) and (37), respectively, we can obtain

$$\begin{aligned}&\parallel\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\parallel^2+\left(\mathcal{A}^{-1}\sum_{i=1}^{m}b_iD_N^{\alpha_i}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\right)\\ =&-\left(\sum_{i=1}^{m}b_iD_N^{\alpha_i}\rho^n,\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right)-\left(p(\rho^n+\theta^n),\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right)-\left(\sum_{i=1}^{m}b_iQ^n_{\alpha_i},\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right).\end{aligned}\tag{38}$$

Using Lemma 2, we have

$$\begin{aligned}&\parallel\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\parallel^2+\frac{1}{2}\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}[\widetilde{d}^{n}_{\alpha_i,n}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Big)-\sum_{k=0}^{n-1}\widetilde{d}^{n}_{\alpha_i,k}\Big(\mathcal{A}^{-1}\Big(\boldsymbol{\eta}^n-\boldsymbol{\eta}^k,\boldsymbol{\eta}^n-\boldsymbol{\eta}^k\Big)]\\ =&-\frac{1}{2}\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\widetilde{d}^{n}_{\alpha_i,k}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^k,\boldsymbol{\eta}^k\Big)-\left(\sum_{i=1}^{m}b_iD_N^{\alpha_i}\rho^n,\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right)\\ &-\left(\sum_{i=1}^{m}b_iQ^n_{\alpha_i},\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right)-\left(p(\rho^n+\theta^n),\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right).\end{aligned}\tag{39}$$

Noting that $\widetilde{d}^{n}_{\alpha_i,k}<0, 0<k\leq n-1$ and using Lemma 1, we obtain

$$\begin{aligned}&\parallel\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\parallel^2+\frac{1}{2}\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^{n}_{\alpha_i,n}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Big)\\ \leq&-\frac{1}{2}\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\widetilde{d}^{n}_{\alpha_i,k}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^k,\boldsymbol{\eta}^k\Big)-\left(\sum_{i=1}^{m}b_iD_N^{\alpha_i}\rho^n,\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right)\\ &-\left(\sum_{i=1}^{m}b_iQ^n_{\alpha_i},\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right)-\left(p(\rho^n+\theta^n),\sum_{i=1}^{m}b_iD_N^{\alpha_i}\theta^n\right).\end{aligned}\tag{40}$$

Applying the Cauchy–Schwarz and the Young inequality in (40) yields

$$\begin{aligned}
&\| \sum_{i=1}^{m} b_i D_N^{\alpha_i}\theta^n \|^2 + \frac{1}{2}\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^n_{\alpha_i,n}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Big)\\
\le& -\frac{1}{2}\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\widetilde{d}^n_{\alpha_i,k}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^k,\boldsymbol{\eta}^k\Big)+\frac{1}{2}\| \sum_{i=1}^{m} b_i D_N^{\alpha}\theta^n \|^2\\
&+2\Bigg(\| \sum_{i=1}^{m} b_i D_N^{\alpha_i}\rho^n \| + \| \sum_{i=1}^{m} b_i Q^n_{\alpha_i} \|^2 + \| p \|^2_\infty(\| \rho^n \|^2+\| \theta^n \|^2)\Bigg).
\end{aligned} \tag{41}$$

Using Lemmas 4 and 5, we obtain

$$\begin{aligned}
&\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\widetilde{d}^n_{\alpha_i,n}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Big)\\
\le& -\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\widetilde{d}^n_{\alpha_i,k}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^k,\boldsymbol{\eta}^k\Big)+C\Big(N^{-2(2-\alpha_1)}+h^{2r+2}\Big).
\end{aligned} \tag{42}$$

Noting that $\boldsymbol{\eta}^0=\mathbf{0}$ and using Lemma 3, we obtain

$$\| \boldsymbol{\eta}^n \| \le C\Big(h^{r+1}+N^{-(2-\alpha_1)}\Big). \tag{43}$$

We now estimate $\| \boldsymbol{\lambda}^n-\boldsymbol{\lambda}^n_h \|_{\boldsymbol{H}(\mathrm{div},\Omega)}$. Taking $v_h=\sum\limits_{i=1}^{m} b_i D_N^{\alpha_i}\theta^n$ and $\boldsymbol{w}_h=\boldsymbol{\eta}^n$ in (30) (a) and (37), respectively, we have

$$\begin{aligned}
\| \mathrm{div}\boldsymbol{\eta}^n \|^2=&-\Bigg(\mathcal{A}^{-1}\sum_{i=1}^{m} b_i D_N^{\alpha_i}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Bigg)-\Bigg(\sum_{i=1}^{m} b_i D_N^{\alpha_i}\rho^n,\mathrm{div}\boldsymbol{\eta}^n\Bigg)\\
&-(p(\rho^n+\theta^n),\mathrm{div}\boldsymbol{\eta}^n)-\Bigg(\sum_{i=1}^{m} b_i Q^n_{\alpha_i},\mathrm{div}\boldsymbol{\eta}^n\Bigg).
\end{aligned} \tag{44}$$

For the term $-\Bigg(\mathcal{A}^{-1}\sum\limits_{i=1}^{m} b_i D_N^{\alpha_i}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Bigg)$, noting that $\sum\limits_{k=0}^{n-1}\Big(-\widetilde{d}^n_{\alpha_i,k}\Big)=T^{-\alpha_i}N^{\alpha_i}$, we obtain

$$\begin{aligned}
-\Bigg(\mathcal{A}^{-1}\sum_{i=1}^{m} b_i D_N^{\alpha_i}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Bigg)=&-\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\Bigg(\sum_{k=0}^{n-1}\widetilde{d}^n_{\alpha_i,k}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^k,\boldsymbol{\eta}^n\Big)+\widetilde{d}^n_{\alpha_i,n}\Big(\mathcal{A}^{-1}\boldsymbol{\eta}^n,\boldsymbol{\eta}^n\Big)\Bigg)\\
\le&\mu_1\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}\sum_{k=0}^{n-1}\Big(-\widetilde{d}^n_{\alpha_i,k}\Big)\| \boldsymbol{\eta}^k \|\| \boldsymbol{\eta}^n \|\\
\le&C\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}T^{-\alpha_i}N^{\alpha_i}\Big(N^{-(2-\alpha_1)}+h^{r+1}\Big)^2.
\end{aligned} \tag{45}$$

Then, it holds from (45) that

$$\begin{aligned}
\| \mathrm{div}\boldsymbol{\eta}^n \|^2=&2\Bigg(\| \sum_{i=1}^{m} b_i D_N^{\alpha_i}\rho^n \| + \| \sum_{i=1}^{m} b_i Q^n_{\alpha_i} \|^2 + \| p \|^2_\infty(\| \rho^n \|^2+\| \theta^n \|^2)\Bigg)\\
&+C\sum_{i=1}^{m}\frac{b_i}{\Gamma(2-\alpha_i)}T^{-\alpha_i}N^{\alpha_i}\Big(N^{-(2-\alpha_1)}+h^{r+1}\Big)^2+\frac{1}{2}\| \mathrm{div}\boldsymbol{\eta}^n \|^2.
\end{aligned} \tag{46}$$

Using Lemmas 4 and 5, we have

$$\| \mathrm{div}\boldsymbol{\eta}^n \| \le C\Big(1+N^{\frac{\alpha_m}{2}}\Big)\Big(h^{r+1}+N^{-(2-\alpha_1)}\Big). \tag{47}$$

Then, we finish the proof. □

Remark 3. *(I) For variables u and* $\boldsymbol{\lambda}$*, we define the discrete norms of the errors as follows:*

$$\begin{aligned}
&\| u - u_h \|_{\hat{L}^\infty(L^2(\Omega))} = \max_{1\le n\le N} \| u(t_n) - u_h^n \|,\\
&\| \boldsymbol{\lambda} - \boldsymbol{\lambda}_h \|_{\hat{L}^\infty((L^2(\Omega))^2)} = \max_{1\le n\le N} \| \boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_h^n \|,\\
&\| \boldsymbol{\lambda} - \boldsymbol{\lambda}_h \|_{\hat{L}^\infty(\boldsymbol{H}(\mathrm{div},\Omega))} = \max_{1\le n\le N} \| \boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_h^n \|_{\boldsymbol{H}(\mathrm{div},\Omega)}.
\end{aligned}$$

From Theorem 3, we obtain the optimal a priori error estimate results for u in the discrete $L^\infty(L^2(\Omega))$ *norm and* $\boldsymbol{\lambda}$ *in the discrete* $L^\infty\left((L^2(\Omega))^2\right)$ *norm and obtain the suboptimal error estimate for* $\boldsymbol{\lambda}$ *in the discrete* $L^\infty(\boldsymbol{H}(div,\Omega))$ *norm. In the actual calculation in the next section, we achieve the optimal convergence rates for variables u and* $\boldsymbol{\lambda}$ *based on the above discrete norms.*

(II) It should be pointed out that the solutions of many FPDEs have an initial layer at $t = 0$ *(see [42,43]). To overcome this difficulty, some scholars have adopted nonuniform mesh methods and achieved excellent results [24,26,42,44–46]. Moreover, it is noted that* $\left\{\xi^k : 0 \le k \le N\right\}$ *in Lemma 3 is required to be a nondecreasing positive sequence, so the error estimates for the MFE scheme (5)–(6) with the temporal nonuniform method should adopt some other techniques. It is gratifying that the numerical results in Example 3 show that the MFE scheme (5)–(6) with the temporal graded mesh is feasible and effective.*

6. Numerical Examples

In this section, we given three test examples to verify the effectiveness and convergence accuracy of the proposed MFE scheme (5)–(6) and adopt the lowest-order Raviart–Thomas MFE space for variables u and $\boldsymbol{\lambda}$ in the numerical experiments.

Example 1. *Consider the following two-term TFRD equation:*

$$\begin{cases}
D_t^{\alpha_1} u(\boldsymbol{x},t) + D_t^{\alpha_2} u(\boldsymbol{x},t) - \Delta u(\boldsymbol{x},t) + p(\boldsymbol{x})u(\boldsymbol{x},t) = f(\boldsymbol{x},t), & (\boldsymbol{x},t) \in \Omega \times J,\\
u(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t) \in \partial\Omega \times \overline{J},\\
u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), & \boldsymbol{x} \in \overline{\Omega},
\end{cases} \tag{48}$$

where $J = (0,1]$*,* $\Omega = (0,1)^2$*,* $p(\boldsymbol{x}) = 1 + x_1^2 + x_2^2$*,* $\boldsymbol{x} = (x_1,x_2) \in \Omega$*,* $u(\boldsymbol{x},0) = 0$*, and the source function f is taken by*

$$\begin{aligned}
f(\boldsymbol{x},t) =& \left(\frac{\Gamma(3+\alpha_1+\alpha_2)}{\Gamma(3+\alpha_2)} t^{2+\alpha_2} + \frac{\Gamma(3+\alpha_1+\alpha_2)}{\Gamma(3+\alpha_1)} t^{2+\alpha_1} + \left(2\pi^2 + p(\boldsymbol{x})\right) t^{2+\alpha_1+\alpha_2} \right)\\
&\times \sin(\pi x_1)\sin(\pi x_2).
\end{aligned}$$

And we can find the analytical solutions for variables u and $\boldsymbol{\lambda}$ *as follows:*

$$\begin{aligned}
u(\boldsymbol{x},t) &= t^{2+\alpha_1+\alpha_2}\sin(\pi x_1)\sin(\pi x_2),\\
\boldsymbol{\lambda}(\boldsymbol{x},t) &= -\pi t^{2+\alpha_1+\alpha_2}(\cos(\pi x_1)\sin(\pi x_2), \sin(\pi x_1)\cos(\pi x_2)).
\end{aligned}$$

In the numerical simulation, we select fractional parameters $\alpha_1 = 0.9, 0.7, 0.5$ and $\alpha_2 = 0.1, 0.4$ in Equation (48) and know that among these different fractional parameters, the convergence rates are only related to the largest fractional parameter α_1 from Theorem 3. By taking $N = 5, 8, 10, 16$ and the corresponding $h = \sqrt{2}/N^{2-\alpha_1}$, we give the error results and convergence rates in Tables 1–3 for the MFE scheme (5)–(6), which show that the convergence rates in the temporal direction for u (in the discrete $L^\infty(L^2(\Omega))$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left((L^2(\Omega))^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norms) are close to $2-\alpha_1$. Moreover, in order to test convergence rates in the spatial direction, by fixing $N = 100$ and taking $h = \sqrt{2}/4, \sqrt{2}/8, \sqrt{2}/16, \sqrt{2}/32$, we give the error results and convergence rates in Tables 4–6, which show that the convergence rates in the spatial direction for u (in the discrete $L^\infty\left(L^2(\Omega)\right)$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norms) are close to 1.

Table 1. Numerical results with $h \approx \sqrt{2}/N^{2-\alpha_1}$ and $\alpha_1 = 0.9$ in Example 1.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	5	8.6930×10^{-2}	–	3.3630×10^{-1}	–	$1.7545 \times 10^{+0}$	–
	8	5.2424×10^{-2}	1.0760	2.0229×10^{-1}	1.0814	$1.0569 \times 10^{+0}$	1.0784
	10	4.0389×10^{-2}	1.1687	1.5578×10^{-1}	1.1709	8.1390×10^{-1}	1.1708
	16	2.3908×10^{-2}	1.1157	9.2174×10^{-2}	1.1165	4.8141×10^{-1}	1.1172
0.4	5	8.7218×10^{-2}	–	3.3776×10^{-1}	–	$1.7631 \times 10^{+0}$	–
	8	5.2624×10^{-2}	1.0750	2.0332×10^{-1}	1.0799	$1.0619 \times 10^{+0}$	1.0786
	10	4.0555×10^{-2}	1.1674	1.5663×10^{-1}	1.1692	8.1786×10^{-1}	1.1704
	16	2.4011×10^{-2}	1.1153	9.2701×10^{-2}	1.1160	4.8371×10^{-1}	1.1175

Table 2. Numerical results with $h \approx \sqrt{2}/N^{2-\alpha_1}$ and $\alpha_1 = 0.7$ in Example 1.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	5	6.5262×10^{-2}	–	2.5184×10^{-1}	–	$1.3146 \times 10^{+0}$	–
	8	3.4916×10^{-2}	1.3308	1.3451×10^{-1}	1.3344	7.0291×10^{-1}	1.3321
	10	2.6204×10^{-2}	1.2863	1.0092×10^{-1}	1.2875	5.2740×10^{-1}	1.2873
	16	1.4175×10^{-2}	1.3073	5.4581×10^{-2}	1.3077	2.8520×10^{-1}	1.3080
0.4	5	6.5379×10^{-2}	–	2.5244×10^{-1}	–	$1.3184 \times 10^{+0}$	–
	8	3.4998×10^{-2}	1.3296	1.3493×10^{-1}	1.3328	7.0499×10^{-1}	1.3318
	10	2.6269×10^{-2}	1.2858	1.0125×10^{-1}	1.2869	5.2895×10^{-1}	1.2875
	16	1.4212×10^{-2}	1.3070	5.4771×10^{-2}	1.3073	2.8602×10^{-1}	1.3081

Table 3. Numerical results with $h \approx \sqrt{2}/N^{2-\alpha_1}$ and $\alpha_1 = 0.5$ in Example 1.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	5	4.7518×10^{-2}	–	1.8311×10^{-1}	–	9.5625×10^{-1}	–
	8	2.2765×10^{-2}	1.5657	8.7633×10^{-2}	1.5679	4.5800×10^{-1}	1.5663
	10	1.6367×10^{-2}	1.4786	6.2996×10^{-2}	1.4792	3.2925×10^{-1}	1.4791
	16	8.1859×10^{-3}	1.4742	3.1504×10^{-2}	1.4743	1.6465×10^{-1}	1.4744
0.4	5	4.7568×10^{-2}	–	1.8337×10^{-1}	–	9.5799×10^{-1}	–
	8	2.2802×10^{-2}	1.5645	8.7824×10^{-2}	1.5662	4.5893×10^{-1}	1.5658
	10	1.6396×10^{-2}	1.4781	6.3145×10^{-2}	1.4785	3.2993×10^{-1}	1.4790
	16	8.2016×10^{-3}	1.4738	3.1585×10^{-2}	1.4739	1.6499×10^{-1}	1.4744

Table 4. Numerical results with $\tau = T/N = 1/100$ and $\alpha_1 = 0.9$ in Example 1.

α_2	h	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	$\sqrt{2}/4$	1.2927×10^{-1}	–	5.0254×10^{-1}	–	$2.5970 \times 10^{+0}$	–
	$\sqrt{2}/8$	6.5234×10^{-2}	0.9867	2.5171×10^{-1}	0.9975	$1.3118 \times 10^{+0}$	0.9853
	$\sqrt{2}/16$	3.2696×10^{-2}	0.9965	1.2589×10^{-1}	0.9996	6.5762×10^{-1}	0.9963
	$\sqrt{2}/32$	1.6361×10^{-2}	0.9989	6.2964×10^{-2}	0.9996	3.2908×10^{-1}	0.9988
0.4	$\sqrt{2}/4$	1.2926×10^{-1}	–	5.0244×10^{-1}	–	$2.5971 \times 10^{+0}$	–
	$\sqrt{2}/8$	6.5231×10^{-2}	0.9866	2.5169×10^{-1}	0.9973	$1.3119 \times 10^{+0}$	0.9853
	$\sqrt{2}/16$	3.2696×10^{-2}	0.9964	1.2589×10^{-1}	0.9995	6.5766×10^{-1}	0.9962
	$\sqrt{2}/32$	1.6363×10^{-2}	0.9987	6.2973×10^{-2}	0.9994	3.2913×10^{-1}	0.9987

Table 5. Numerical results with $\tau = T/N = 1/100$ and $\alpha_1 = 0.7$ in Example 1.

α_2	h	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	$\sqrt{2}/4$	1.2929×10^{-1}	–	5.0266×10^{-1}	–	$2.5969 \times 10^{+0}$	–
	$\sqrt{2}/8$	6.5241×10^{-2}	0.9867	2.5174×10^{-1}	0.9976	$1.3118 \times 10^{+0}$	0.9853
	$\sqrt{2}/16$	3.2698×10^{-2}	0.9966	1.2590×10^{-1}	0.9997	6.5757×10^{-1}	0.9963
	$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2954×10^{-2}	0.9999	3.2900×10^{-1}	0.9991
0.4	$\sqrt{2}/4$	1.2928×10^{-1}	–	5.0258×10^{-1}	–	$2.5970 \times 10^{+0}$	–
	$\sqrt{2}/8$	6.5239×10^{-2}	0.9867	2.5173×10^{-1}	0.9975	$1.3118 \times 10^{+0}$	0.9853
	$\sqrt{2}/16$	3.2697×10^{-2}	0.9966	1.2590×10^{-1}	0.9996	6.5758×10^{-1}	0.9963
	$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2954×10^{-2}	0.9999	3.2900×10^{-1}	0.9990

Table 6. Numerical results with $\tau = T/N = 1/100$ and $\alpha_1 = 0.5$ in Example 1.

α_2	h	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	$\sqrt{2}/4$	1.2930×10^{-1}	–	5.0273×10^{-1}	–	$2.5969 \times 10^{+0}$	–
	$\sqrt{2}/8$	6.5243×10^{-2}	0.9868	2.5176×10^{-1}	0.9977	$1.3118 \times 10^{+0}$	0.9852
	$\sqrt{2}/16$	3.2699×10^{-2}	0.9966	1.2591×10^{-1}	0.9997	6.5756×10^{-1}	0.9963
	$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2955×10^{-2}	0.9999	3.2899×10^{-1}	0.9991
0.4	$\sqrt{2}/4$	1.2929×10^{-1}	–	5.0266×10^{-1}	–	$2.5969 \times 10^{+0}$	–
	$\sqrt{2}/8$	6.5242×10^{-2}	0.9867	2.5175×10^{-1}	0.9976	$1.3118 \times 10^{+0}$	0.9853
	$\sqrt{2}/16$	3.2698×10^{-2}	0.9966	1.2590×10^{-1}	0.9997	6.5757×10^{-1}	0.9963
	$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2955×10^{-2}	0.9999	3.2899×10^{-1}	0.9991

Example 2. *Consider the following three-term TFRD equation:*

$$\begin{cases} D_t^{\alpha_1}u(\boldsymbol{x},t) + D_t^{\alpha_2}u(\boldsymbol{x},t) + D_t^{\alpha_3}u(\boldsymbol{x},t) - \Delta u(\boldsymbol{x},t) + p(\boldsymbol{x})u(\boldsymbol{x},t) = f(\boldsymbol{x},t), & (\boldsymbol{x},t) \in \Omega \times J, \\ u\,(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t) \in \partial\Omega \times \bar{J}, \\ u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), & \boldsymbol{x} \in \overline{\Omega}, \end{cases} \tag{49}$$

where the spatial domain Ω, temporal domain J, coefficient $p(\boldsymbol{x})$, and initial data $u(\boldsymbol{x},0)$ are as in Example 1 and the source function f is taken by

$$\begin{aligned} f(\boldsymbol{x},t) = &\bigg(\frac{\Gamma(3+\alpha_1+\alpha_2+\alpha_3)}{\Gamma(3+\alpha_2+\alpha_3)} t^{2+\alpha_2+\alpha_3} + \frac{\Gamma(3+\alpha_1+\alpha_2+\alpha_3)}{\Gamma(3+\alpha_1+\alpha_3)} t^{2+\alpha_1+\alpha_3} \\ &+ \frac{\Gamma(3+\alpha_1+\alpha_2+\alpha_3)}{\Gamma(3+\alpha_1+\alpha_2)} t^{2+\alpha_1+\alpha_2} + \left(2\pi^2 + p(\boldsymbol{x})\right) t^{2+\alpha_1+\alpha_2+\alpha_3} \bigg) \sin(\pi x_1)\sin(\pi x_2). \end{aligned}$$

And we can also find the analytical solutions for variables u and $\boldsymbol{\lambda}$ as follows:

$$\begin{aligned} u(\boldsymbol{x},t) &= t^{2+\alpha_1+\alpha_2+\alpha_3}\sin(\pi x_1)\sin(\pi x_2), \\ \boldsymbol{\lambda}(\boldsymbol{x},t) &= -\pi t^{2+\alpha_1+\alpha_2+\alpha_3}(\cos(\pi x_1)\sin(\pi x_2), \sin(\pi x_1)\cos(\pi x_2)). \end{aligned}$$

In this example, since the Equation (49) contains three Caputo time-fractional derivative terms, we specifically take the fractional parameters $\alpha_1 = 0.9, 0.7, 0.5$ and $(\alpha_2, \alpha_3) = (0.4, 0.2), (0.3, 0.1)$. From Theorem 3, we also point out that the convergence rates are only related to the maximum fractional parameter α_1. In Tables 7–9, for different $N = 5, 8, 10, 16$, we give the error results and convergence rates for the MFE scheme (5)–(6), where the spatial grid sizes are also taken as $h = \sqrt{2}/N^{2-\alpha_1}$. We can also see that the convergence rates in the temporal direction for u (in the discrete $L^\infty(L^2(\Omega))$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left((L^2(\Omega))^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norms) are close to $2-\alpha_1$. Furthermore, in Tables 10–12, we also fix $N = 100$ and take $h = \sqrt{2}/4$, $\sqrt{2}/8$, $\sqrt{2}/16$, $\sqrt{2}/32$, give the error results and convergence rates, and see that the convergence rates in the spatial direction for u and λ in the above corresponding discrete norms are also close to 1.

Table 7. Numerical results with $h \approx \sqrt{2}/N^{2-\alpha_1}$ and $\alpha_1 = 0.9$ in Example 2.

α_2	α_3	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.4	0.2	5	8.7400×10^{-2}	–	3.3868×10^{-1}	–	$1.7681 \times 10^{+0}$	–
		8	5.2741×10^{-2}	1.0747	2.0392×10^{-1}	1.0795	$1.0648 \times 10^{+0}$	1.0790
		10	4.0650×10^{-2}	1.1670	1.5711×10^{-1}	1.1686	8.2004×10^{-1}	1.1704
		16	2.4067×10^{-2}	1.1152	9.2988×10^{-2}	1.1159	4.8494×10^{-1}	1.1177
0.3	0.1	5	8.7138×10^{-2}	–	3.3735×10^{-1}	–	$1.7608 \times 10^{+0}$	–
		8	5.2567×10^{-2}	1.0753	2.0303×10^{-1}	1.0804	$1.0605 \times 10^{+0}$	1.0788
		10	4.0507×10^{-2}	1.1679	1.5638×10^{-1}	1.1698	8.1673×10^{-1}	1.1706
		16	2.3980×10^{-2}	1.1154	9.2546×10^{-2}	1.1162	4.8304×10^{-1}	1.1175

Table 8. Numerical results with $h \approx \sqrt{2}/N^{2-\alpha_1}$ and $\alpha_1 = 0.7$ in Example 2.

α_2	α_3	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.4	0.2	5	6.5465×10^{-2}	–	2.5288×10^{-1}	–	$1.3208 \times 10^{+0}$	–
		8	3.5052×10^{-2}	1.3291	1.3521×10^{-1}	1.3321	7.0629×10^{-1}	1.3319
		10	2.6310×10^{-2}	1.2857	1.0146×10^{-1}	1.2868	5.2989×10^{-1}	1.2877
		16	1.4234×10^{-2}	1.3070	5.4886×10^{-2}	1.3073	2.8651×10^{-1}	1.3083
0.3	0.1	5	6.5343×10^{-2}	–	2.5225×10^{-1}	–	$1.3173 \times 10^{+0}$	–
		8	3.4972×10^{-2}	1.3300	1.3479×10^{-1}	1.3334	7.0434×10^{-1}	1.3321
		10	2.6247×10^{-2}	1.2860	1.0114×10^{-1}	1.2872	5.2845×10^{-1}	1.2876
		16	1.4200×10^{-2}	1.3071	5.4707×10^{-2}	1.3075	2.8575×10^{-1}	1.3082

Table 9. Numerical results with $h \approx \sqrt{2}/N^{2-\alpha_1}$ and $\alpha_1 = 0.5$ in Example 2.

α_2	α_3	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.4	0.2	5	4.7614×10^{-2}	–	1.8360×10^{-1}	–	9.5930×10^{-1}	–
		8	2.2831×10^{-2}	1.5638	8.7972×10^{-2}	1.5654	4.5961×10^{-1}	1.5656
		10	1.6417×10^{-2}	1.4779	6.3255×10^{-2}	1.4782	3.3041×10^{-1}	1.4790
		16	8.2126×10^{-3}	1.4738	3.1641×10^{-2}	1.4738	1.6523×10^{-1}	1.4745
0.3	0.1	5	4.7550×10^{-2}	–	1.8327×10^{-1}	–	9.5743×10^{-1}	–
		8	2.2788×10^{-2}	1.5650	8.7752×10^{-2}	1.5669	4.5859×10^{-1}	1.5661
		10	1.6385×10^{-2}	1.4783	6.3087×10^{-2}	1.4788	3.2967×10^{-1}	1.4791
		16	8.1952×10^{-3}	1.4740	3.1552×10^{-2}	1.4742	1.6486×10^{-1}	1.4745

Table 10. Numerical results with $\tau = T/N = 1/100$ and $\alpha_1 = 0.9$ in Example 2.

α_2	α_3	h	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.4	0.2	$\sqrt{2}/4$	1.2924×10^{-1}	–	5.0230×10^{-1}	–	$2.5973 \times 10^{+0}$	–
		$\sqrt{2}/8$	6.5229×10^{-2}	0.9865	2.5168×10^{-1}	0.9970	$1.3119 \times 10^{+0}$	0.9853
		$\sqrt{2}/16$	3.2696×10^{-2}	0.9964	1.2589×10^{-1}	0.9994	6.5768×10^{-1}	0.9962
		$\sqrt{2}/32$	1.6364×10^{-2}	0.9986	6.2979×10^{-2}	0.9992	3.2916×10^{-1}	0.9986
0.3	0.1	$\sqrt{2}/4$	1.2925×10^{-1}	–	5.0236×10^{-1}	–	$2.5972 \times 10^{+0}$	–
		$\sqrt{2}/8$	6.5231×10^{-2}	0.9865	2.5169×10^{-1}	0.9971	$1.3119 \times 10^{+0}$	0.9853
		$\sqrt{2}/16$	3.2696×10^{-2}	0.9964	1.2589×10^{-1}	0.9995	6.5765×10^{-1}	0.9962
		$\sqrt{2}/32$	1.6362×10^{-2}	0.9988	6.2971×10^{-2}	0.9994	3.2912×10^{-1}	0.9987

Table 11. Numerical results with $\tau = T/N = 1/100$ and $\alpha_1 = 0.7$ in Example 2.

α_2	α_3	h	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.4	0.2	$\sqrt{2}/4$	1.2926×10^{-1}	–	5.0244×10^{-1}	–	$2.5971 \times 10^{+0}$	–
		$\sqrt{2}/8$	6.5237×10^{-2}	0.9866	2.5172×10^{-1}	0.9972	$1.3118 \times 10^{+0}$	0.9853
		$\sqrt{2}/16$	3.2697×10^{-2}	0.9965	1.2590×10^{-1}	0.9996	6.5758×10^{-1}	0.9963
		$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2954×10^{-2}	0.9999	3.2901×10^{-1}	0.9990
0.3	0.1	$\sqrt{2}/4$	1.2927×10^{-1}	–	5.0249×10^{-1}	–	$2.5970 \times 10^{+0}$	–
		$\sqrt{2}/8$	6.5238×10^{-2}	0.9866	2.5172×10^{-1}	0.9973	$1.3118 \times 10^{+0}$	0.9853
		$\sqrt{2}/16$	3.2697×10^{-2}	0.9965	1.2590×10^{-1}	0.9996	6.5758×10^{-1}	0.9963
		$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2954×10^{-2}	0.9999	3.2900×10^{-1}	0.9991

Table 12. Numerical results with $\tau = T/N = 1/100$ and $\alpha_1 = 0.5$ in Example 2.

α_2	α_3	h	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.4	0.2	$\sqrt{2}/4$	1.2927×10^{-1}	–	5.0252×10^{-1}	–	$2.5970 \times 10^{+0}$	–
		$\sqrt{2}/8$	6.5240×10^{-2}	0.9866	2.5173×10^{-1}	0.9973	$1.3118 \times 10^{+0}$	0.9853
		$\sqrt{2}/16$	3.2698×10^{-2}	0.9965	1.2590×10^{-1}	0.9996	6.5757×10^{-1}	0.9963
		$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2955×10^{-2}	0.9999	3.2899×10^{-1}	0.9991
0.3	0.1	$\sqrt{2}/4$	1.2928×10^{-1}	–	5.0257×10^{-1}	–	$2.5970 \times 10^{+0}$	–
		$\sqrt{2}/8$	6.5241×10^{-2}	0.9866	2.5174×10^{-1}	0.9974	$1.3118 \times 10^{+0}$	0.9853
		$\sqrt{2}/16$	3.2698×10^{-2}	0.9966	1.2590×10^{-1}	0.9996	6.5757×10^{-1}	0.9963
		$\sqrt{2}/32$	1.6359×10^{-2}	0.9991	6.2955×10^{-2}	0.9999	3.2899×10^{-1}	0.9991

Based on the numerical results in Tables 1–12 obtained from the above two test examples, we can see that the convergence rates in the spatial and temporal directions for u (in the discrete $L^\infty\left(L^2(\Omega)\right)$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ norm) are in agreement with the theoretical results in Theorem 3, and those for $\boldsymbol{\lambda}$ (in the discrete $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norm) are higher than the theoretical result. These results fully demonstrate that the proposed MFE method for the multi-term TFRD equations is effective.

Example 3. *Consider the two-term TFRD equation in Example 1 with weak regularity solutions near the initial time $t = 0$, where the source function f is taken by*

$$\begin{aligned} f(\boldsymbol{x},t) =& \bigg(\frac{2}{\Gamma(3-\alpha_1)}t^{2-\alpha_1} + \Gamma(1+\alpha_1) + \frac{\Gamma(2+\alpha_2)}{\Gamma(2+\alpha_2-\alpha_1)}t^{1+\alpha_2-\alpha_1} \\ &+ \frac{2}{\Gamma(3-\alpha_2)}t^{2-\alpha_2} + \frac{\Gamma(1+\alpha_1)}{\Gamma(1+\alpha_1-\alpha_2)}t^{\alpha_1-\alpha_2} + \Gamma(2+\alpha_2)t \\ &+ \left(2\pi^2 + p(\boldsymbol{x})\right)\left(t^2 + t^{\alpha_1} + t^{1+\alpha_2}\right)\bigg)\sin(\pi x_1)\sin(\pi x_2). \end{aligned}$$

And we can also find the analytical solutions for variables u and $\boldsymbol{\lambda}$ as follows:

$$\begin{aligned} u(\boldsymbol{x},t) &= \left(t^2 + t^{\alpha_1} + t^{1+\alpha_2}\right)\sin(\pi x_1)\sin(\pi x_2), \\ \boldsymbol{\lambda}(\boldsymbol{x},t) &= -\pi\left(t^2 + t^{\alpha_1} + t^{1+\alpha_2}\right)(\cos(\pi x_1)\sin(\pi x_2), \sin(\pi x_1)\cos(\pi x_2)). \end{aligned}$$

In this example, we will select the graded mesh to discretize the interval $[0,T]$ and set $t_n = T(n/N)^\gamma$, for $n = 0,1,2,\cdots,N$, where constant $\gamma \geq 1$ is the temporal graded mesh parameter. The ideal optimal error estimates for u (in the discrete $L^\infty\left(L^2(\Omega)\right)$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div},\Omega))$ norms) should be $O\left(N^{-\min\{\gamma\alpha_1, 2-\alpha_1\}} + h\right)$. Here, we will mainly test the convergence rates in the temporal direction with the graded mesh parameter $\gamma = 1$ and $(2-\alpha_1)/\alpha_1$. We first conduct numerical experiments with

$\gamma = 1$. Then, the optimal convergence rate in the temporal direction is α_1. For fractional parameters $\alpha_1 = 0.9, 0.7, 0.5$ and $\alpha_2 = 0.1, 0.4$, we take the time mesh parameter $N = 20, 40, 80, 160$ and special spatial grid parameters: (i) when $\alpha_1 = 0.9$, take $h \approx 2\sqrt{2}/N^{\alpha_1}$; (ii) when $\alpha_1 = 0.7$, take $h \approx \sqrt{2}/N^{\alpha_1}$; (iii) when $\alpha_1 = 0.5$, take $h \approx \sqrt{2}/(2N^{\alpha_1})$. Then, we give the numerical results in Tables 13–15, which show that the convergence rates in the temporal direction for u (in the discrete $L^\infty(L^2(\Omega))$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div}, \Omega))$ norms) are close to α_1.

Table 13. Numerical results with $\alpha_1 = 0.9$ and graded mesh parameter $\gamma = 1$ in Example 3.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	20	1.9572×10^{-1}	–	7.5520×10^{-1}	–	$3.9354 \times 10^{+0}$	–
	40	1.1208×10^{-1}	0.8042	4.3165×10^{-1}	0.8070	$2.2539 \times 10^{+0}$	0.8041
	80	6.0396×10^{-2}	0.8920	2.3245×10^{-1}	0.8930	$1.2146 \times 10^{+0}$	0.8920
	160	3.2722×10^{-2}	0.8842	1.2591×10^{-1}	0.8845	6.5806×10^{-1}	0.8842
0.4	20	1.9571×10^{-1}	–	7.5516×10^{-1}	–	$3.9354 \times 10^{+0}$	–
	40	1.1208×10^{-1}	0.8042	4.3164×10^{-1}	0.8069	$2.2540 \times 10^{+0}$	0.8041
	80	6.0396×10^{-2}	0.8920	2.3244×10^{-1}	0.8929	$1.2146 \times 10^{+0}$	0.8920
	160	3.2722×10^{-2}	0.8842	1.2591×10^{-1}	0.8845	6.5806×10^{-1}	0.8842

Table 14. Numerical results with $\alpha_1 = 0.7$ and graded mesh parameter $\gamma = 1$ in Example 3.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	20	1.9573×10^{-1}	–	7.5526×10^{-1}	–	$3.9353 \times 10^{+0}$	–
	40	1.2068×10^{-1}	0.6976	4.6487×10^{-1}	0.7002	$2.4269 \times 10^{+0}$	0.6974
	80	7.4765×10^{-2}	0.6908	2.8779×10^{-1}	0.6918	$1.5035 \times 10^{+0}$	0.6907
	160	4.4872×10^{-2}	0.7365	1.7268×10^{-1}	0.7369	9.0241×10^{-1}	0.7365
0.4	20	1.9572×10^{-1}	–	7.5523×10^{-1}	–	$3.9354 \times 10^{+0}$	–
	40	1.2068×10^{-1}	0.6976	4.6486×10^{-1}	0.7001	$2.4269 \times 10^{+0}$	0.6974
	80	7.4765×10^{-2}	0.6908	2.8779×10^{-1}	0.6918	$1.5035 \times 10^{+0}$	0.6907
	160	4.4872×10^{-2}	0.7365	1.7268×10^{-1}	0.7369	9.0241×10^{-1}	0.7365

Table 15. Numerical results with $\alpha_1 = 0.5$ and graded mesh parameter $\gamma = 1$ in Example 3.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\mathrm{div}))$	Rates
0.1	20	1.7410×10^{-1}	–	6.7141×10^{-1}	–	$3.5006 \times 10^{+0}$	–
	40	1.2069×10^{-1}	0.5286	4.6487×10^{-1}	0.5303	$2.4269 \times 10^{+0}$	0.5285
	80	8.7212×10^{-2}	0.4687	3.3576×10^{-1}	0.4694	$1.7538 \times 10^{+0}$	0.4686
	160	6.2812×10^{-2}	0.4735	2.4175×10^{-1}	0.4739	$1.2632 \times 10^{+0}$	0.4735
0.4	20	1.7410×10^{-1}	–	6.7139×10^{-1}	–	$3.5006 \times 10^{+0}$	–
	40	1.2069×10^{-1}	0.5286	4.6487×10^{-1}	0.5303	$2.4269 \times 10^{+0}$	0.5285
	80	8.7212×10^{-2}	0.4687	3.3576×10^{-1}	0.4694	$1.7538 \times 10^{+0}$	0.4686
	160	6.2812×10^{-2}	0.4735	2.4175×10^{-1}	0.4739	$1.2632 \times 10^{+0}$	0.4735

Next, we conduct numerical experiments with $\gamma = (2 - \alpha_1)/\alpha_1$. Then, the optimal convergence rate is $2 - \alpha_1$. We take the time mesh parameter $N = 5, 8, 10, 16$ and the spatial grid parameter $h = \sqrt{2}/N^{2-\alpha_1}$. Then, we give the numerical results in Tables 16–18 and find that the convergence rates in the temporal direction for u (in the discrete $L^\infty(L^2(\Omega))$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ and $L^\infty(\boldsymbol{H}(\mathrm{div}, \Omega))$ norms) are close to $2 - \alpha_1$. Based on the above discussion, we know that the MFE scheme (5)–(6) with the temporal graded mesh for solving the multi-term TFRD equations with the initial layer is also feasible and effective.

Table 16. Numerical results with $\alpha_1 = 0.9$ and graded mesh parameter $\gamma = \frac{2-\alpha_1}{\alpha_1}$ in Example 3.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\text{div}))$	Rates
0.1	5	2.6024×10^{-1}	–	$1.0062 \times 10^{+0}$	–	$5.2342 \times 10^{+0}$	–
	8	1.5673×10^{-1}	1.0789	6.0412×10^{-1}	1.0855	$3.1526 \times 10^{+0}$	1.0787
	10	1.2067×10^{-1}	1.1718	4.6479×10^{-1}	1.1750	$2.4273 \times 10^{+0}$	1.1718
	16	7.1368×10^{-2}	1.1175	2.7470×10^{-1}	1.1189	$1.4355 \times 10^{+0}$	1.1176
0.4	5	2.6021×10^{-1}	–	$1.0060 \times 10^{+0}$	–	$5.2351 \times 10^{+0}$	–
	8	1.5673×10^{-1}	1.0787	6.0410×10^{-1}	1.0852	$3.1531 \times 10^{+0}$	1.0787
	10	1.2067×10^{-1}	1.1716	4.6479×10^{-1}	1.1748	$2.4276 \times 10^{+0}$	1.1718
	16	7.1372×10^{-2}	1.1174	2.7472×10^{-1}	1.1188	$1.4357 \times 10^{+0}$	1.1176

Table 17. Numerical results with $\alpha_1 = 0.7$ and graded mesh parameter $\gamma = \frac{2-\alpha_1}{\alpha_1}$ in Example 3.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\text{div}))$	Rates
0.1	5	1.9568×10^{-1}	–	7.5499×10^{-1}	–	$3.9360 \times 10^{+0}$	–
	8	1.0462×10^{-1}	1.3323	4.0284×10^{-1}	1.3365	$2.1043 \times 10^{+0}$	1.3323
	10	7.8497×10^{-2}	1.2872	3.0217×10^{-1}	1.2887	$1.5789 \times 10^{+0}$	1.2873
	16	4.2450×10^{-2}	1.3079	1.6336×10^{-1}	1.3086	8.5380×10^{-1}	1.3081
0.4	5	1.9566×10^{-1}	–	7.5492×10^{-1}	–	$3.9370 \times 10^{+0}$	–
	8	1.0462×10^{-1}	1.3320	4.0289×10^{-1}	1.3361	$2.1049 \times 10^{+0}$	1.3323
	10	7.8506×10^{-2}	1.2870	3.0221×10^{-1}	1.2885	$1.5793 \times 10^{+0}$	1.2874
	16	4.2457×10^{-2}	1.3078	1.6339×10^{-1}	1.3084	8.5400×10^{-1}	1.3081

Table 18. Numerical results with $\alpha_1 = 0.5$ and graded mesh parameter $\gamma = \frac{2-\alpha_1}{\alpha_1}$ in Example 3.

α_2	N	u–$\hat{L}^\infty(L^2)$	Rates	λ–$\hat{L}^\infty((L^2)^2)$	Rates	λ–$\hat{L}^\infty(H(\text{div}))$	Rates
0.1	5	1.4253×10^{-1}	–	5.4923×10^{-1}	–	$2.8673 \times 10^{+0}$	–
	8	6.8272×10^{-2}	1.5661	2.6278×10^{-1}	1.5685	$1.3733 \times 10^{+0}$	1.5663
	10	4.9083×10^{-2}	1.4788	1.8890×10^{-1}	1.4794	9.8727×10^{-1}	1.4790
	16	2.4548×10^{-2}	1.4742	9.4464×10^{-2}	1.4744	4.9373×10^{-1}	1.4744
0.4	5	1.4255×10^{-1}	–	5.4933×10^{-1}	–	$2.8688 \times 10^{+0}$	–
	8	6.8306×10^{-2}	1.5654	2.6296×10^{-1}	1.5675	$1.3743 \times 10^{+0}$	1.5659
	10	4.9113×10^{-2}	1.4783	1.8905×10^{-1}	1.4788	9.8804×10^{-1}	1.4787
	16	2.4567×10^{-2}	1.4739	9.4561×10^{-2}	1.4740	4.9416×10^{-1}	1.4742

7. Conclusions

This work presents a Raviart–Thomas MFE method for solving the multi-term TFRD equations with variable coefficients by using the well-known $L1$ formula. The existence, uniqueness, and unconditional stability of the discrete solution are discussed, and the optimal a priori error estimates for u (in the discrete $L^\infty\left(L^2(\Omega)\right)$ norm) and $\boldsymbol{\lambda}$ (in the discrete $L^\infty\left(\left(L^2(\Omega)\right)^2\right)$ norm) and the suboptimal a priori error estimate for $\boldsymbol{\lambda}$ (in the discrete $L^\infty(\boldsymbol{H}(\text{div},\Omega))$ norm) are obtained in this work. In addition, some numerical results are given to demonstrate the effectiveness of the proposed MFE method. In future research, we will try to give theoretical analysis for the MFE method with the temporal graded mesh to solve some FPDEs with the initial layer at $t = 0$ and apply the MFE method to solve more FPDEs in scientific and engineering fields.

Author Contributions: Conceptualization, J.Z. and S.D.; methodology, Z.F.; software, Z.F.; validation, J.Z., S.D. and Z.F.; formal analysis, J.Z. and S.D.; writing—original draft preparation, J.Z. and S.D.; writing—review and editing, J.Z. and Z.F.; funding acquisition, J.Z. and Z.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (11701299), the Scientific Research Projection of Higher Schools of the Inner Mongolia Autonomous Region

(NJZY23055), and the Central Government Guided Local Science and Technology Development Fund Project of China (2022ZY0175).

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors are very grateful to the editors and reviewers for their helpful comments and suggestions on improving this work.

Conflicts of Interest: The authors declares no conflicts of interest.

References

1. Ross, B. *Fractional Calculus and Its Applications*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 1975.
2. Hilfer, R. *Applications of Fractional Calculus in Physics*; World Scientific: Singapore, 2000.
3. Diethelm, K. *The Analysis of Fractional Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2010.
4. Hendy, M.H.; Amin, M.M.; Ezzat, M.A. Two-dimensional problem for thermoviscoelastic materials with fractional order heat transfer. *J. Therm. Stress.* **2019**, *42*, 1298–1315. [CrossRef]
5. Zhou, J.C.; Salahshour, S.; Ahmadian, A.; Senu, N. Modeling the dynamics of COVID-19 using fractal-fractional operator with a case study. *Results Phys.* **2022**, *33*, 105103. [CrossRef] [PubMed]
6. Shymanskyi, V.; Sokolovskyy, I.; Sokolovskyy, Y.; Bubnyak, T. Variational Method for Solving the Time-Fractal Heat Conduction Problem in the Claydite-Block Construction. In *Advances in Computer Science for Engineering and Education. ICCSEEA 2022*; Hu, Z., Dychka, I., Petoukhov, S., He, M., Eds.; Lecture Notes on Data Engineering and Communications Technologies; Springer: Cham, Switzerland, 2022; Volume 134.
7. Bagley, R.L.; Torvik, R.J. On the appearance of the fractional derivative in the behaviour of real materials. *Appl. Mech.* **1984**, *51*, 294–298.
8. Ming, C.Y.; Liu, F.W.; Zheng, L.C.; Turner, I.; Anh, V. Analytical solutions of multi-term time fractional differential equations and application to unsteady flows of generalized viscoelastic fluid. *Comput. Math. Appl.* **2006**, *72*, 2084–2097. [CrossRef]
9. Fetecau, C.; Athar, M.; Fetecau, C. Unsteady flow of a generalized Maxwell fluid with fractional derivative due to a constantly accelerating plate. *Comput. Math. Appl.* **2009**, *57*, 596–603. [CrossRef]
10. Liu, F.W.; Meerschaert, M.M.; McGough, R.J.; Zhuang, P.H.; Liu, Q.X. Numerical methods for solving the multi-term time-fractional wave-diffusion equation. *Fract. Calc. Appl. Anal.* **2013**, *16*, 9–25. [CrossRef]
11. Dehghan, M.; Safarpoor, M.; Abbaszadeh, M. Two high-order numerical algorithms for solving the multi-term time fractional diffusion-wave equations. *J. Comput. Appl. Math.* **2015**, *290*, 174–195. [CrossRef]
12. Ren, J.C.; Sun, Z.Z. Efficient Numerical Solution of the Multi-Term Time Fractional Diffusion-Wave Equation. *East Asian J. Appl. Math.* **2015**, *5*, 1–28. [CrossRef]
13. Zheng, M.; Liu, F.W.; Anh, V.; Turner, I. A high-order spectral method for the multi-term time-fractional diffusion equations. *Appl. Math. Model.* **2016**, *40*, 4970–4985. [CrossRef]
14. Du, R.L.; Sun, Z.Z. Temporal second-order difference methods for solving multi-term time fractional mixed diffusion and wave equations. *Numer. Algorithms* **2021**, *88*, 191–226. [CrossRef]
15. Hendy, A.S.; Zaky, M.A. Graded mesh discretization for coupled system of nonlinear multi-term time-space fractional diffusion equations. *Eng. Comput.* **2022**, *38*, 1351–1363. [CrossRef]
16. Liu, Y.Q.; Yin, X.L.; Liu, F.W.; Xin, X.Y.; Shen, Y.F.; Feng, L.B. An alternating direction implicit Legendre spectral method for simulating a 2D multi-term time-fractional Oldroyd-B fluid type diffusion equation. *Comput. Math. Appl.* **2022**, *113*, 160–173. [CrossRef]
17. Wei, L.L.; Wang, H.H. Local discontinuous Galerkin method for multi-term variable-order time fractional diffusion equation. *Math. Comput. Simulat.* **2023**, *203*, 685–698. [CrossRef]
18. She, M.F.; Li, D.F.; Sun, H.W. A transformed $L1$ method for solving the multi-term time-fractional diffusion problem. *Math. Comput. Simulat.* **2022**, *193*, 584–606. [CrossRef]
19. Jin, B.T.; Lazarov, R.; Liu, Y.K.; Zhou, Z. The Galerkin finite element method for a multi-term time-fractional diffusion equation. *J. Comput. Phys.* **2015**, *281*, 825–843. [CrossRef]
20. Li, M.; Huang, C.M.; Jiang, F.Z. Galerkin finite element method for higher dimensional multi-term fractional diffusion equation on non-uniform meshes. *Appl. Anal.* **2017**, *96*, 1269–1284. [CrossRef]
21. Zhou, J.; Xu, D.; Chen, H.B. A Weak Galerkin Finite Element Method for Multi-Term Time-Fractional Diffusion Equations. *East Asian J. Appl. Math.* **2018**, *8*, 181–193. [CrossRef]
22. Bu, W.P.; Shu, S.; Xue, X.Q.; Xiao, A.G.; Zeng, W. Space-time finite element method for the multi-term time-space fractional diffusion equation on a two-dimensional domain. *Comput. Math. Appl.* **2019**, *78*, 1367–1379. [CrossRef]
23. Feng, L.B.; Liu, F.W.; Turner, I. 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. Simulat.* **2019**, *70*, 354–371. [CrossRef]
24. Meng, X.Y.; Stynes, M. Barrier function local and global analysis of an $L1$ finite element method for a multiterm time-fractional initial-boundary value problem. *J. Sci. Comput.* **2020**, *84*, 5. [CrossRef]

25. Yin, B.L.; Liu, Y.; Li, H.; Zeng, F.H. A class of efficient time-stepping methods for multi-term time-fractional reaction-diffusion-wave equations. *Appl. Numer. Math.* **2021**, *165*, 56–82. [CrossRef]
26. Huang, C.B.; Stynes, M.; Chen, H. An α-robust finite element method for a multi-term time-fractional diffusion problem. *J. Comput. Appl. Math.* **2021**, *389*, 113334. [CrossRef]
27. Liu, H.; Zheng, X.C.; Fu, H.F. Analysis of a multi-term variable-order time-fractional diffusion equation and its Galerkin finite element approximation. *J. Comput. Math.* **2022**, *40*, 814–834. [CrossRef]
28. Zhao, Y.M.; Chen, P.; Bu, W.P.; Liu, X.T.; Tang, Y.F. Two mixed finite element methods for time-fractional diffusion equations. *J. Sci. Comput.* **2017**, *70*, 407–428. [CrossRef]
29. Liu, Y.; Fang, Z.C.; Li, H.; He, S. A mixed finite element method for a time-fractional fourth-order partial differential equation. *Appl. Math. Comput.* **2014**, *243*, 703–717. [CrossRef]
30. Li, X.C.; Yang, X.Y.; Zhang, Y.H. Error estimates of mixed finite element methods for time-fractional Navier-Stokes equations. *J. Sci. Comput.* **2017**, *70*, 500–515. [CrossRef]
31. Liu, Y.; Du, Y.W.; Li, H.; Li, J.C.; He, S. A two-grid mixed finite element method for a nonlinear fourth-order reaction-diffusion problem with time-fractional derivative. *Comput. Math. Appl.* **2015**, *70*, 2474–2492. [CrossRef]
32. Li, X.W.; Tang, Y.L. Interpolated coefficient mixed finite elements for semilinear time fractional diffusion equations. *Fractal Fract.* **2023**, *7*, 482. [CrossRef]
33. Shi, Z.G.; Zhao, Y.M.; Tang, Y.F.; Wang, F.L.; Shi, Y.H. Superconvergence analysis of an H^1-Galerkin mixed finite element method for two-dimensional multi-term time fractional diffusion equations. *Int. J. Comput. Math.* **2018**, *95*, 1845–1857. [CrossRef]
34. Li, M.; Huang, C.M.; Ming, W.Y. Mixed finite-element method for multi-term time-fractional diffusion and diffusion-wave equations. *Comp. Appl. Math.* **2018**, *37*, 2309–2334. [CrossRef]
35. Cao, F.F.; Zhao, Y.M.; Wang, F.L.; Shi, Y.H.; Yao, C.H. Nonconforming mixed FEM analysis for multi-term time-fractional mixed sub-diffusion and diffusion-wave equation with time-space coupled derivative. *Adv. Appl. Math. Mech.* **2023**, *15*, 322–358. [CrossRef]
36. Sun, Z.Z.; Wu, X.N. A fully discrete scheme for a diffusion-wave system. *Appl. Numer. Math.* **2006**, *56*, 193–209. [CrossRef]
37. Lin, Y.M.; Xu, C.J. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.* **2007**, *225*, 1533–1552. [CrossRef]
38. Sinha, R.K.; Ewing, R.E.; Lazarov, R.D. Mixed finite element approximations of parabolic integro-differential equations with nonsmooth initial data. *SIAM J. Numer. Anal.* **2009**, *47*, 3269–3292. [CrossRef]
39. Fang, Z.C.; Zhao, J.; Li, H.; Liu, Y. Finite volume element methods for two-dimensional time fractional reaction-diffusion equations on triangular grids. *Appl. Anal.* **2023**, *102*, 2248–2270. [CrossRef]
40. Douglas, J.; Roberts, J.E. Global Estimates for Mixed Methods for Second Order Elliptic Equations. *Math. Comput.* **1985**, *44*, 39–52. [CrossRef]
41. Durán, R.G. *Mixed Finite Element Methods*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 2008; Volume 1939.
42. Stynes, M.; O'Riordan, E.; Gracia, J.L. Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. *SIAM J. Numer. Anal.* **2017**, *55*, 1057–1079. [CrossRef]
43. Jin, B.T.; Li, B.Y.; Zhou, Z. Numerical analysis of nonlinear subdiffusion equations. *SIAM J. Numer. Anal.* **2018**, *56*, 1–23. [CrossRef]
44. Liao, H.L.; Li, D.F.; Zhang, J.W. Sharp error estimate of the nonuniform $L1$ formula for linear reaction-subdiffusion equations. *SIAM J. Numer. Anal.* **2018**, *56*, 1112–1133. [CrossRef]
45. Ren, J.C.; Liao, H.L.; Zhang, Z.M. Superconvergence error estimate of a finite element method on nonuniform time meshes for reaction-subdiffusion equations. *J. Sci. Comput.* **2020**, *84*, 38. [CrossRef]
46. Huang, C.B.; Stynes, M. Optimal H^1 spatial convergence of a fully discrete finite element method for the time-fractional Allen-Cahn equation. *Adv. Comput. Math.* **2020**, *46*, 63. [CrossRef]