East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.2023-085.070723

Vol. **xx**, No. x, pp. 1-38
xxx 2024

Error Analysis of BDF-Galerkin FEMs for Thermally Coupled Incompressible MHD with Temperature Dependent Parameters

Shuaijun Liu, Pengzhan Huang* and Yinnian He

College of Mathematics and System Sciences, Xinjiang University, Urumqi 830017, P.R. China.

Received 17 March 2023; Accepted (in revised version) 7 July 2023.

Abstract. In this paper, we consider the electromagnetically and thermally driven flow which is modeled by evolutionary magnetohydrodynamic equations and heat equation coupled through generalized Boussinesq approximation with temperature-dependent coefficients. Based on a third-order backward differential formula for temporal discretization, mixed finite element approximation for spatial discretization and extrapolated treatments in linearization for nonlinear terms, a linearized backward differentiation formula type scheme for the considered equations is proposed and analysed. Optimal L^2-error estimates for the proposed fully discretized scheme are obtained by the temporal-spatial error splitting technique. Numerical examples are presented to check the accuracy and efficiency of the scheme.

AMS subject classifications: 65M12, 65M60

Key words: Thermally coupled magnetohydrodynamic, Boussinesq approximation, temperature dependent coefficient, linearized BDF scheme, convergence.

1. Introduction

The hydrodynamical behavior of conducting fluids subject to external magnetic fields can be well described by magnetohydrodynamic (MHD) equations which are governed by a combination of Navier-Stokes equations and Maxwell's equations. Due to wide applications of MHD systems in astronomy, geophysics and engineering, such as metallurgical engineering, contactless electromagnetic stirring, the design of cooling systems with liquid metals for a nuclear reactor, damping convective flow in metal-like melt and so on [6, 37], it is important to find accurate effective numerical methods for their solution.

Besides, the buoyancy effect can not be ignored in the momentum equation since the fluid will produce temperature difference during conduction. Therefore, MHD systems are

*Corresponding author. *Email addresses:* SJ_Liu123@163.com (S. Liu), hpzh@xju.edu.cn (P. Huang), heyn@mail.xjtu.edu.cn (Y. He)

usually coupled with the heat equation. Furthermore, in many applications, the properties of fluids and materials in MHD system may not be constant. For example, viscosity, thermal diffusivity and magnetic diffusivity can be strongly dependent on temperature, which will lead to the original system becoming a stronger nonlinearity and coupling system [5,14,48].

Let Ω be a connected bounded open set in $R^d, d=2$ or 3, either convex or having a $C^{1,1}$ boundary $\partial\Omega$ and $[0,T]$ an interval of R. In this paper, we deal with a time-dependent thermally coupled incompressible MHD flow. More exactly, we consider the following incompressible Navier-Stokes and Maxwell equations coupled with the heat equation by the well-known Boussinesq approximation

$$\begin{aligned} &\partial_t \mathbf{u} - \nabla\cdot\big(\nu(\theta)\nabla\mathbf{u}\big) + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p + s\mathbf{b}\times\mathrm{curl}\mathbf{b} - \beta(\theta)\theta\mathbf{j} = \mathbf{f} && \text{in } \Omega\times(0,T],\\ &\partial_t \mathbf{b} + \mathrm{curl}\big(\mu(\theta)\mathrm{curl}\mathbf{b}\big) - \mathrm{curl}(\mathbf{u}\times\mathbf{b}) = \mathbf{0} && \text{in } \Omega\times(0,T],\\ &\partial_t\theta - \nabla\cdot\big(\kappa(\theta)\nabla\theta\big) + (\mathbf{u}\cdot\nabla)\theta = \psi && \text{in } \Omega\times(0,T],\\ &\mathrm{div}\,\mathbf{u} = 0, \quad \mathrm{div}\,\mathbf{b} = 0 && \text{in } \Omega\times(0,T], \end{aligned} \tag{1.1}$$

where the fluid viscous diffusivity ν, the magnetic diffusivity μ, the thermal conductivity κ and the thermal expansion coefficient β depend on the temperature. The unknowns are velocity field $\mathbf{u}$, temperature θ, pressure p, and magnetic field $\mathbf{b}$. Functions $\mathbf{f}$ and ψ are respectively a known body force and a heat source, $\nu(\theta)$ denotes the fluid viscous diffusivity, $\kappa(\theta)$ the thermal conductivity, $\beta(\theta)$ the thermal expansion coefficient, and $\mu(\theta) := 1/\eta_0\delta(\theta)$ the magnetic diffusivity, where η_0 and δ denote the magnetic permeability and the electrical conductivity. Besides, $\mathbf{j}$ is the unit vector in the direction opposite to the gravitation, and $s := 1/\eta_0\rho_0$ the coupling coefficient, where ρ_0 is the reference density.

The system (1.1) is considered in conjunction with the following initial values and boundary conditions:

$$\begin{aligned} &\mathbf{u}(\mathbf{x},0) = \mathbf{u}_0, \mathbf{b}(\mathbf{x},0) = \mathbf{b}_0, \theta(\mathbf{x},0) = \theta_0 && \text{for all} \quad \mathbf{x}\in\Omega,\\ &\mathbf{u}|_{\partial\Omega_T} = 0, && \text{(no-slip condition)},\\ &\mathbf{b}\cdot\mathbf{n}|_{\partial\Omega_T} = 0, \quad \mathrm{curl}\mathbf{b}\times\mathbf{n}|_{\partial\Omega_T} = \mathbf{0}, && \text{(perfectly conducting wall)},\\ &\theta|_{\partial\Omega_T} = 0, && \end{aligned} \tag{1.2}$$

where $\mathbf{n}$ is the outer unit normal of $\partial\Omega$, $\partial\Omega_T = \partial\Omega\times(0,T]$ and the initial magnetic induction $\mathbf{b}_0$ satisfies $\mathrm{div}\,\mathbf{b}_0 = 0$.

It is worth noting that in recent years, various analytical investigations and have been carried out and numerous efficient numerical methods for solving MHD systems have been developed. For stationary MHD problems, error estimates of finite element methods are given in [9, 12, 15, 16, 43, 46, 52]. On the other hand, for time-dependent MHD systems, the convergence analysis and error estimates of first- and second-order-in-time for fully discrete finite element methods are respectively established in [11, 17, 28, 30, 38, 54, 56] and [8,19,41,47,55,57]. For example, a fully discrete linearized H^1-conforming Lagrange finite element method for a 2D vector potential MHD model is proposed in [25]. Wang *et al*. [47] introduced a numerical scheme based on the modified Crank-Nicolson finite element projection method and obtained optimal error estimates for MHD equations. Kanbar *et al.* [20] developed a two-dimensional second order unstaggered finite volume central scheme for the system of MHD equations with gravitational source term. Wu *et al*. [53] considered a discrete scheme for solving ideal MHD equations with random inputs based on the finite volume WENO method and the third-order total variation decreasing Runge Kutta method.

In their pioneering works, Meir *et al*. [31–33] studied stationary incompressible MHD equations coupled to thermally equation. Then the Arrow-Hurwicz iterative finite element method was designed for this stationary problem [21]. Ding *et al*. [7] proposed the Crank-Nicolson extrapolation scheme in temporal direction and a mixed finite element in spatial direction for the nonstationary MHD coupled heat equation. Due to the properties of fluids and materials are temperature dependent, the thermally coupled MHD equations with temperature-dependent coefficients have been also extensively studied in the literature. Qiu [40] has proved the existence and uniqueness of weak solutions and discrete weak solutions for the stationary thermally coupled MHD equations with temperature-dependent parameters. Further, a fully discrete Euler semi-implicit scheme and a partitioned time-stepping scheme based on Crank-Nicolson extrapolation are studied by Qiu [39] and Ravindran [42], respectively. In order to enhance the accuracy of time discretization, the backward differentiation formula (BDF) is a usually chosen. This is a multistep method which can achieve a high order accuracy. As a special case of the BDF method, the third-order backward differentiation formula (BDF3) has been widely exploited in numerical methods of time evolution partial differential equations, including parabolic equations [49], Navier-Stokes equations [3, 51], nonlinear thermistor equations [10], and the nonlinear Sobolev equation [50].

In this paper, we construct and study a linearized BDF type scheme for thermally coupled MHD equations with temperature-dependent coefficients. It comprises the BDF3 for time derivative terms, extrapolated treatment in linearization for the coupling, nonlinear terms preserving antisymmetric structure, and a mixed finite element approximation for spatial discretization. Moreover, using temporal-spatial error splitting arguments [24] and the telescope formula for the BDF temporal discretization operator [26], we obtain optimal L^2 error estimates for the velocity and magnetic fields, and determine the temperature without any restriction on the time step and the mesh size.

This paper is organized as follows. Section 2 introduces notations and assumptions. A linearized fully discrete BDF3 finite element scheme and the main error estimates are establishes in Section 3. Section 4 presents an a priori estimate for temporal error and suitable regularity of the solutions for time semi-discrete system and a primary spatial error estimate for the finite element solutions. Optimal L^2 error estimates for the temperature and velocity and magnetic fields are derived in Section 5. Numerical results presented in Section 6 support the theoretical analysis. A short conclusion is drawn in Section 7.

2. Preliminaries

Let us first recall standard notation — cf. [1]. For integer $k \geq 0$, the space $C^k(\Omega)$ denotes the space of functions with k times continuously differentiable in Ω, and the space $C^{k,1}(\Omega)$

consists of functions in $C^k(\Omega)$ that are Lipschitz continuous. Besides, for $1 \leq p \leq \infty$, we define the usual Sobolev space $W^{k,p}(\Omega)$ norm and Lebegue space $L^p(\Omega)$ norm by $\|\cdot\|_{W^{k,p}}$ and $\|\cdot\|_{L^p}$, respectively. In particular, when $p=2$, we denote $W^{k,2}(\Omega)$ by $H^k(\Omega)$ endowed with the norm $\|\cdot\|_k$. Further, for a Banach space X, we denote by $L^p(0,T;X)$ the temporal-spatial function space

$$\|v\|_{L^p(0,T;X)} := \left(\int_0^T \|v\|_X^p dt \right)^{1/p}, \quad 1 \leq p < \infty$$

and $\operatorname{ess\,sup}_{t\in[0,T]} \|v\|_X$ if $p = \infty$. The symbol $C(0,T;X)$ denotes the set of continuous function $v : [0,T] \to X$.

To set up a variational formulation of problem (1.1) with (1.2), we introduce the following function spaces:

$$\begin{aligned}
&\mathbf{X} := \left\{ \mathbf{v} \in H^1(\Omega)^d : \ \mathbf{v}|_{\partial\Omega} = 0 \right\}, && Q := \left\{ q \in L^2(\Omega) : \int_\Omega q dx = 0 \right\}, \\
&\mathbf{W} := \left\{ \mathbf{w} \in H^1(\Omega)^d : \mathbf{w}\cdot\mathbf{n}|_{\partial\Omega} = 0 \right\}, && M := \{ \phi \in H^1(\Omega) : \phi|_{\partial\Omega} = 0 \}, \\
&\mathbf{X}_0 := \left\{ \mathbf{v} \in \mathbf{X} : \ \operatorname{div}\mathbf{v} = 0 \right\}, && \mathbf{W}_0 := \left\{ \mathbf{b} \in \mathbf{W} : \operatorname{div}\mathbf{b} = 0 \right\}.
\end{aligned}$$

Besides, we define the following bilinear forms and trilinear forms, for $\mathbf{u},\mathbf{v} \in \mathbf{X}$, $\mathbf{b},\mathbf{w} \in \mathbf{W}$, $p \in Q$ and $\theta, \phi \in M$,

$$\begin{aligned}
&B_u\big(\nu(\theta);\mathbf{u},\mathbf{v}\big) = \int_\Omega \nu(\theta)\nabla\mathbf{u} : \nabla\mathbf{v}\, d\mathbf{x}, \quad B_\theta\big(\kappa(\theta);\theta,\phi\big) = \int_\Omega \kappa(\theta)\nabla\theta\cdot\nabla\phi\ d\mathbf{x}, \\
&B_b\big(\mu(\theta);\mathbf{b},\mathbf{w}\big) = \int_\Omega \mu(\theta)\mathrm{curl}\mathbf{b}\cdot\mathrm{curl}\mathbf{w} + \mu(\theta)\mathrm{div}\,\mathbf{b}\cdot\mathrm{div}\,\mathbf{w}\, d\mathbf{x}, \\
&D_f\big(\beta(\theta);\theta,\mathbf{v}\big) = \int_\Omega \beta(\theta)\theta\mathbf{j}\cdot\mathbf{v}\, d\mathbf{x}, \quad B_f(\mathbf{v},p) = \int_\Omega \mathrm{div}\,\mathbf{v}\cdot p\ d\mathbf{x}, \\
&T_u(\mathbf{u},\mathbf{v},\mathbf{w}) = \int_\Omega (\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{w} + \frac{1}{2}(\nabla\cdot\mathbf{u})\mathbf{v}\cdot\mathbf{w}\, d\mathbf{x}, \\
&T_{b1}(\mathbf{b},\mathbf{w},\mathbf{v}) = \int_\Omega \mathbf{b}\times\mathrm{curl}\mathbf{w}\cdot\mathbf{v}\, d\mathbf{x}, \quad T_{b_2}(\mathbf{v},\mathbf{b},\mathbf{w}) = \int_\Omega \mathrm{curl}(\mathbf{v}\times\mathbf{b})\cdot\mathbf{w}\, d\mathbf{x}, \\
&T_\theta(\mathbf{u},\theta,\phi) = \int_\Omega (\mathbf{u}\cdot\nabla)\theta\phi + \frac{1}{2}(\mathrm{div}\,\mathbf{u})\theta\phi\ d\mathbf{x}.
\end{aligned}$$

Assume that the bilinear form $B_f(\cdot,\cdot)$ satisfies the following inf-sup condition — cf. [44]:

$$\sup_{\mathbf{v}\in\mathbf{X},\mathbf{v}\neq\mathbf{0}} \frac{|B_f(\mathbf{v},q)|}{\|\mathbf{v}\|_1} \geq \beta_0\|q\|_{L^2} \quad \text{for all} \quad q \in M,$$

where β_0 is a positive constant which depends on Ω.

The corresponding variational weak formulation for (1.1)-(1.2) can be now formulated as follows: Find $(\mathbf{u},p,\mathbf{b},\theta) \in \mathbf{X}\times Q\times\mathbf{W}\times M$, such that for all $t \in (0,T]$ and $(\mathbf{v},q,\mathbf{w},\phi) \in$

$\mathbf{X}\times Q\times\mathbf{W}\times M$ we have

$$\begin{aligned}&(\partial_t\mathbf{u},\mathbf{v})+B_u(\nu(\theta);\mathbf{u},\mathbf{v})+T_u(\mathbf{u},\mathbf{u},\mathbf{v})+sT_{b1}(\mathbf{b},\mathbf{b},\mathbf{v})\\&\quad-B_f(\mathbf{v},p)-D_f(\beta(\theta);\theta,\mathbf{v})=(\mathbf{f},\mathbf{v}),\\&(\partial_t\mathbf{b},\mathbf{w})+B_b(\mu(\theta);\mathbf{b},\mathbf{w})-T_{b2}(\mathbf{u},\mathbf{b},\mathbf{w})=\mathbf{0},\\&(\partial_t\theta,\phi)+B_\theta(\kappa(\theta);\theta,\phi)+T_\theta(\mathbf{u},\theta,\phi)=(\psi,\phi),\\&B_f(\mathbf{u},q)=0.\end{aligned}\tag{2.1}$$

Let us also recall the following inequalities and relations which will be frequently used in the error estimations — cf. [1,12,13,17,34]:

- For $\mathbf{v}\in\mathbf{X}$ or $\mathbf{W}$, it holds
$$\|\nabla\mathbf{v}\|_{L^2}^2\geq c_1\|\mathbf{v}\|_1^2.$$
- For $\mathbf{w}\in\mathbf{W}$, the following inequality holds:
$$c_2\|\mathbf{w}\|_1^2\leq\|\mathrm{div}\,\mathbf{w}\|_{L^2}^2+\|\mathrm{curl}\mathbf{w}\|_{L^2}^2.$$
- For $\mathbf{v}\in\mathbf{X}$, and $1\leq m\leq 6$, the following inequality holds:
$$\|\mathbf{v}\|_{L^m}\leq c_3\|\nabla\mathbf{v}\|_{L^2}.\tag{2.2}$$
- For $\mathbf{v}\in\mathbf{X}\cap H^2(\Omega)^d$, the Agmon's inequality holds
$$\|\mathbf{v}\|_{L^\infty}\leq c\|\mathbf{v}\|_1^{1/2}\|\mathbf{v}\|_2^{1/2}\leq c\|\mathbf{v}\|_2.\tag{2.3}$$
- For $\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbf{X}$, $\theta,\phi\in M$, $T_u(\mathbf{u},\mathbf{v},\mathbf{w})$ and $T_\theta(\mathbf{u},\theta,\phi)$ which preserve the anti-symmetric property satisfy
$$T_u(\mathbf{u},\mathbf{v},\mathbf{w})=-T_u(\mathbf{u},\mathbf{w},\mathbf{v}),\quad T_\theta(\mathbf{u},\theta,\phi)=-T_\theta(\mathbf{u},\phi,\theta).\tag{2.4}$$
- By using $(\mathbf{w}\times\mathrm{curl}\mathbf{b},\mathbf{v})=(\mathbf{v}\times\mathbf{w},\mathrm{curl}\mathbf{b})$, for $\mathbf{u}\in\mathbf{X},\ \mathbf{b}\in\mathbf{W}$, it follows that
$$T_{b1}(\mathbf{b},\mathbf{b},\mathbf{u})-T_{b2}(\mathbf{u},\mathbf{b},\mathbf{b})=0.\tag{2.5}$$
- For $\mathbf{b},\mathbf{w}\in\mathbf{X}$ and $\mathbf{v}\in H_0^2(\Omega)^d$, the following equalities hold:
$$\begin{aligned}&\mathrm{curl}(\mathbf{b}\times\mathbf{w})=(\mathbf{w}\cdot\nabla)\mathbf{b}-(\mathrm{div}\,\mathbf{w})\mathbf{b}+(\mathrm{div}\,\mathbf{b})\mathbf{w}-(\mathbf{b}\cdot\nabla)\mathbf{w},\\&\mathrm{curlcurl}\mathbf{v}=\nabla(\mathrm{div}\,\mathbf{v})-\Delta\mathbf{v}.\end{aligned}\tag{2.6}$$
- Let a and b be two real numbers. Then for any positive real number ϵ, we have the Young inequality
$$ab\leq\frac{1}{2\epsilon}a^2+\frac{\epsilon}{2}b^2.\tag{2.7}$$

In this work, c and C (with or without a subscript) denote general positive constants independent of the mesh size and time step. They may stand for different values at different occurrences. Further, let $0=t_0<t_1<\cdots<t_{N-1}<t_N=T$ be a uniform partition of the time interval $[0,T]$ with time step $\tau=T/N$ and $t_n=n\tau$, $0\le n\le N$. We denote the three-step backward difference operator $D_\tau H(t_n)$ and the extrapolation operator $\widetilde{H}(t_n)$ for $3\le n\le N$ as follows:

$$
\begin{aligned}
&D_\tau H(t_n)=\frac{1}{6\tau}\big(11H(t_n)-18H(t_{n-1})+9H(t_{n-2})-2H(t_{n-3})\big),\\
&\widetilde{H}(t_n)=3H(t_{n-1})-3H(t_{n-2})+H(t_{n-3}).
\end{aligned}
$$

Using the Taylor expansion with the integral remainder gives

$$
\begin{aligned}
&D_\tau H(t_n)=\partial_t H(t_n)+\frac{1}{6\tau}\int_{t_{n-3}}^{t_n}\left(3(t-t_{n-1})_+^3-\frac{3}{2}(t-t_{n-2})_+^3+\frac{1}{3}(t-t_{n-3})^3\right)\partial_t^4 H dt,\\
&\widetilde{H}(t_n)=H(t_n)+\frac{1}{2}\int_{t_{n-3}}^{t_n}\left(3(t-t_{n-1})_+^2-3(t-t_{n-2})_+^2+(t-t_{n-3})^2\right)\partial_t^3 H dt,
\end{aligned}
$$

where

$$
(t-t_{n-1})_+=\max\{(t-t_{n-1}),0\},\quad (t-t_{n-2})_+=\max\{(t-t_{n-2}),0\}.
$$

The Cauchy-Schwarz inequality implies

$$
\big\|\partial_t H(t_n)-D_\tau H(t_n)\big\|_{L^2}\le c\tau^{5/2}\big\|\partial_t^4 H\big\|_{L^2(t_{n-3},t_n;L^2(\Omega))}, \tag{2.8}
$$

$$
\big\|H(t_n)-\widetilde{H}(t_n)\big\|_{L^2}\le c\tau^{5/2}\big\|\partial_t^3 H\big\|_{L^2(t_{n-3},t_n;L^2(\Omega))}. \tag{2.9}
$$

Besides, set

$$
\overline{H}(t_n)=\frac{H(t_n)+H(t_{n-1})}{2},\quad \underline{H}(t_n)=\frac{3}{2}H(t_{n-1})-\frac{1}{2}H(t_{n-2}).
$$

Next, we introduce the telescope formula, the discrete Grönwall inequality and the Gagliardo-Nirenberg inequality, which will be frequently used in error analysis.

Lemma 2.1 (Telescope Formula for D_τ, cf. Liu [26]). *With the definition of the BDF3 temporal discrete operator D_τ, there exist some constants a_i, $i=1,\ldots,10$ such that*

$$
\begin{aligned}
&\tau\left(D_\tau u^n,2u^n-u^{n-1}\right)\\
=&(a_1u^n)^2-\left(a_1u^{n-1}\right)^2+\left(a_2u^n+a_3u^{n-1}\right)^2-\left(a_2u^{n-1}+a_3u^{n-2}\right)^2\\
&+\left(a_4u^n+a_5u^{n-1}+a_6u^{n-2}\right)^2-\left(a_4u^{n-1}+a_5u^{n-2}+a_6u^{n-3}\right)\\
&+\left(a_7u^n+a_8u^{n-1}+a_9u^{n-2}+a_{10}u^{n-3}\right)^2,
\end{aligned} \tag{2.10}
$$

where $a_1\neq 0$.

Lemma 2.2 (Discrete Grönwall Inequality, cf. Heywood & Rannacher [18]). *Let τ, B and a_k, b_k, c_k, γ_k, $k > 0$, be nonnegative numbers such that*

$$a_n + \tau \sum_{k=0}^{n} b_k \leq \tau \sum_{k=0}^{n} \gamma_k a_k + \tau \sum_{k=0}^{n} c_k + B.$$

Suppose that $\tau\gamma_k < 1$ for all k and set $\delta_k = (1-\tau\gamma_k)^{-1}$. Then

$$a_n + \tau \sum_{k=0}^{n} b_k \leq \exp\left(\tau \sum_{k=0}^{n} \gamma_k \delta_k\right)\left(\tau \sum_{k=0}^{n} c_k + B\right).$$

Lemma 2.3 (Gagliardo-Nirenberg Inequality, cf. Nirenberg [36]). *Let u be a function defined on Ω and $\partial^s u$ be any partial derivative of u of order s. Then*

$$\|\partial^j u\|_{L^p} \leq c\|\partial^m u\|_{L^r}^a \|u\|_{L^q}^{1-a} + c\|u\|_{L^q},$$

for $0 \leq j < m$ and $j/m \leq a \leq 1$ with

$$\frac{1}{p} = \frac{j}{d} + a\left(\frac{1}{r} - \frac{m}{d}\right) + (1-a)\frac{1}{q},$$

except $1 < r < \infty$ and $m - j - d/r$ is a non-negative integer, in which case the above estimate holds only for $j/m \leq a < 1$.

It is known, the stability and convergence of the numerical solutions to different equations are always studied based on different regularity assumptions on the solution and initial data. For examples, for the 3D Navier-Stokes equations with a given smooth initial value, Li [22] proved that the solution to the Navier-Stokes equations with this given initial value must be smooth and unique if the finite element solution can be bounded. The first-order convergence is shown in [23] for the 2D Navier-Stokes equations with L^2 initial data in convex polygonal domains. Throughout this paper, we make the following assumptions with respect to the domain Ω, the given functions $\mathbf{f}, \psi$, the initial data $(\mathbf{u}_0, \mathbf{b}_0, \theta_0)$ and the regularity of the solution $(\mathbf{u}, \mathbf{b}, p, \theta)$.

Assumption 2.1. Assume that the given functions $\mathbf{f}, \psi$ and the initial data $\mathbf{u}_0, \mathbf{b}_0, \theta_0$ satisfy

$$\begin{aligned}
&\mathbf{f} \in L^2\left(0, T; H^{-1}(\Omega)^d\right), \quad \psi \in L^\infty\left(0, T; L^\infty(\Omega)\right), \\
&\mathbf{u}_0 \in L^2(\Omega)^d, \quad \mathbf{b}_0 \in L^2(\Omega)^d, \quad \theta_0 \in L^\infty(\Omega).
\end{aligned}$$

Assumption 2.2. Assume that the given functions $\nu, \mu, \kappa \in C^{2,1}(\overline{\Omega} \times (0, T] \times R; R^+)$, $\beta \in C^{1,1}(\overline{\Omega} \times [0, T] \times R; R^+)$ and

$$\begin{aligned}
&\beta_1 \leq \beta(\mathbf{x}, t, \theta) \leq \beta_2, \quad \beta_1' \leq |\partial_\theta \beta(\mathbf{x}, t, \theta)| \leq \beta_2', \\
&\nu_1 \leq \nu(\mathbf{x}, t, \theta) \leq \nu_2, \quad \mu_1 \leq \mu(\mathbf{x}, t, \theta) \leq \mu_2, \\
&\kappa_1 \leq \kappa(\mathbf{x}, t, \theta) \leq \kappa_2,
\end{aligned}$$

and

$$\nu_1' \le |\partial_\theta \nu(\mathbf{x},t,\theta)| \le \nu_2', \quad \mu_1' \le |\partial_\theta \mu(\mathbf{x},t,\theta)| \le \mu_2', \quad \kappa_1' \le |\partial_\theta \kappa(\mathbf{x},t,\theta)| \le \kappa_2',$$
$$\nu_1'' \le |\partial_\theta^2 \nu(\mathbf{x},t,\theta)| \le \nu_2'', \quad \mu_1'' \le |\partial_\theta^2 \mu(\mathbf{x},t,\theta)| \le \mu_2'', \quad \kappa_1'' \le |\partial_\theta^2 \kappa(\mathbf{x},t,\theta)| \le \kappa_2''$$

with positive constants $\beta_i, \beta_i', \nu_i, \mu_i, \kappa_i, \nu_i', \mu_i', \kappa_i', \nu_i'', \mu_i'', \kappa_i''$, $i=1,2$.

Moreover, the following regularity results for Stokes problem, Maxwell problem and Poisson problem are also needed [4, 13, 44].

Assumption 2.3. The domain Ω has the regularity (Ω is a convex polyhedron or has a C^2 boundary $\partial\Omega$) such that the Stokes problem

$$-\Delta \mathbf{v} + \nabla q = \mathbf{f}, \quad \operatorname{div} \mathbf{v} = 0 \quad \text{in } \Omega, \quad \mathbf{v}|_{\partial\Omega} = \mathbf{0}$$

for $\mathbf{f} \in L^{k+1}(\Omega)^d$ admits a unique solution $(\mathbf{v}, q)$ which satisfies

$$\|\mathbf{v}\|_{W^{2,k+1}} + \|q\|_{W^{1,k+1}} \le c_f \|\mathbf{f}\|_{W^{0,k+1}}, \quad 1 \le k \le \infty. \tag{2.11}$$

The Maxwell equation

$$\operatorname{curlcurl} \mathbf{w} = \mathbf{g}, \quad \operatorname{div} \mathbf{w} = 0 \quad \text{in } \Omega, \quad \mathbf{n} \times \operatorname{curl} \mathbf{w} = 0, \quad \mathbf{w} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega$$

for $\mathbf{g} \in L^{k+1}(\Omega)^d$ and $\operatorname{div} \mathbf{g} = 0$ own a unique solution $\mathbf{w} \in \mathbf{W}$ satisfying

$$\|\mathbf{w}\|_{W^{2,k+1}} \le c_g \|\mathbf{g}\|_{W^{0,k+1}}, \quad 1 \le k \le \infty.$$

Let $\psi \in H^2(\Omega)$ be a unique solution of

$$-\Delta \psi = \phi, \quad \psi|_{\partial\Omega} = 0,$$

where $\phi \in L^{k+1}(\Omega)$. Then

$$\|\psi\|_{W^{2,k+1}} \le c_\phi \|\phi\|_{W^{0,k+1}} \quad \text{for} \quad 1 \le k \le \infty. \tag{2.12}$$

Here, c_f, c_g and c_ϕ only depend on Ω.

Assumption 2.4. Assume that the solution $(\mathbf{u}, p, \mathbf{b}, \theta)$ to (1.1)-(1.2) satisfies

$$\begin{aligned}
&\|\mathbf{u}\|_{L^\infty(0,T;H^{r+1}(\Omega)^d)} + \|\partial_t \mathbf{u}\|_{L^\infty(0,T;H^{r+1}(\Omega)^d)} + \|\partial_t^2 \mathbf{u}\|_{L^\infty(0,T;H^2(\Omega)^d)} \\
&\quad + \left\|\partial_t^3 \mathbf{u}\right\|_{L^\infty(0,T;H^1(\Omega)^d)} + \left\|\partial_t^4 \mathbf{u}\right\|_{L^2(0,T;L^2(\Omega)^d)} \le C, \\
&\|p\|_{L^\infty(0.T;H^{r+1}(\Omega))} + \|\partial_t p\|_{L^\infty(0.T;H^{r+1}(\Omega))} + \left\|\partial_t^2 p\right\|_{L^2(0.T;H^1(\Omega))} \\
&\quad + \left\|\partial_t^3 p\right\|_{L^2(0.T;H^1(\Omega))} \le C, \\
&\|\mathbf{b}\|_{L^\infty(0,T;H^{r+1}(\Omega)^d)} + \|\partial_t \mathbf{b}\|_{L^\infty(0,T;H^{r+1}(\Omega)^d)} + \left\|\partial_t^2 \mathbf{b}\right\|_{L^\infty(0,T;H^2(\Omega)^d)} \\
&\quad + \left\|\partial_t^3 \mathbf{b}\right\|_{L^\infty(0,T;H^1(\Omega)^d)} + \left\|\partial_t^4 \mathbf{b}\right\|_{L^2(0,T;L^2(\Omega)^d)} \le C, \\
&\|\theta\|_{L^\infty(0,T;H^{r+1}(\Omega))} + \|\partial_t \theta\|_{L^\infty(0,T;H^{r+1}(\Omega))} + \left\|\partial_t^2 \theta\right\|_{L^\infty(0,T;H^2(\Omega))} \\
&\quad + \left\|\partial_t^3 \theta\right\|_{L^\infty(0,T;H^1(\Omega))} + \left\|\partial_t^4 \theta\right\|_{L^2(0,T;L^2(\Omega))} \le C,
\end{aligned}$$

where $r \ge 2$ is the order of the corresponding finite element space.

The emphasis of this paper is on optimal error estimates of the BDF-Galerkin finite element method based on the above regularity assumptions.

3. A Fully Discrete BDF3 Finite Element Scheme.

Let $\mathbb{K}_h$ be a quasi-uniform triangulation of the polygonal or polyhedral bounded domain $\Omega \in R^d$ into triangles when $d = 2$ and tetrahedra when $d = 3$, respectively. Additionally, we define some finite element spaces to approximate the velocity field, the magnetic field, the temperature, and the pressure, respectively

$$
\begin{aligned}
&\mathbf{X}_h^r = \left\{\mathbf{v} \in \mathbf{C}^0(\overline{\Omega})^d \cap \mathbf{X} : \mathbf{v}|_{\mathbb{K}} \in P_r(\mathbb{K})^d \quad \text{for all} \quad \mathbb{K} \in \mathbb{K}_h\right\},\\
&\mathbf{W}_h^r = \left\{\mathbf{w} \in \mathbf{C}^0(\overline{\Omega})^d \cap \mathbf{W} : \mathbf{w}|_{\mathbb{K}} \in P_r(\mathbb{K})^d \quad \text{for all} \quad \mathbb{K} \in \mathbb{K}_h\right\},\\
&M_h^r = \left\{\phi \in C^0(\overline{\Omega}) \cap M : \phi|_{\mathbb{K}} \in P_r(\mathbb{K}) \quad \text{for all} \quad \mathbb{K} \in \mathbb{K}_h\right\},\\
&Q_h^{r-1} = \left\{q \in C^0(\overline{\Omega}) \cap Q : q|_{\mathbb{K}} \in P_{r-1}(\mathbb{K}) \quad \text{for all} \quad \mathbb{K} \in \mathbb{K}_h\right\},
\end{aligned}
$$

where $P_r(\mathbb{K})$ is the set of polynomials of degree less than or equal to r over element $\mathbb{K}$. It is well-known that the popular generalized Taylor-Hood FE space $(\mathbf{X}_h^r,\ Q_h^{r-1})$ with $r > 1$ is a stable mixed finite element pair, and the inf-sup condition on the space $(\mathbf{X}_h^r,\ Q_h^{r-1})$ holds. For the case $r = 1$, we use the well-known stable MINI-element [2] to approximate velocity and pressure in this paper.

Besides, we recall the following inverse estimate from [44].

Lemma 3.1. *For any integers l and m $(0 \le l \le m \le 1)$ and any real numbers p and q $(0 \le p \le q \le \infty)$, it satisfies that*

$$
\|\mathbf{v}_h\|_{W^{m,q}} \le C h^{l-m+d(1/q-1/p)} \|\mathbf{v}_h\|_{W^{l,p}} \quad \textit{for all} \quad \mathbf{v}_h \in \mathbf{X}_h^r, \mathbf{W}_h^r \quad \textit{or} \quad M_h^{r-1}.
$$

Now, we give the third-order fully discrete BDF-Galerkin finite element method for the problem (1.1)-(1.2). For simplicity, we denote $\mathbf{f}(t_n)$, $\psi(t_n)$ as $\mathbf{f}^n$, ψ^n, and $\zeta^n(\widetilde{\theta}_h^n)$ as $\zeta(\mathbf{x}, t_n, \widetilde{\theta}_h^n)$, $\zeta = \nu, \beta, \mu$ or κ.

Seek $(\mathbf{u}_h^n, \mathbf{b}_h^n, \theta_h^n, p_h^n) \in \mathbf{X}_h^r \times \mathbf{W}_h^r \times M_h^r \times Q_h^{r-1}$, for $n \ge 3$, such that for all $\mathbf{v}_h \in \mathbf{X}_h^r$, $\mathbf{w}_h \in \mathbf{W}_h^r$, $\phi_h \in M_h^r$, $q_h \in Q_h^{r-1}$, satisfying

$$
\begin{aligned}
&\left(D_\tau \mathbf{u}_h^n, \mathbf{v}_h\right) + B_u\left(\nu^n(\widetilde{\theta}_h^n); \mathbf{u}_h^n, \mathbf{v}_h\right) + T_u\left(\widetilde{\mathbf{u}}_h^n, \mathbf{u}_h^n, \mathbf{v}_h\right) - B_f\left(\mathbf{v}_h, p_h^n\right)\\
&\quad + sT_{b1}\left(\widetilde{\mathbf{b}}_h^n, \mathbf{b}_h^n, \mathbf{v}_h\right) - D_f\left(\beta^n(\widetilde{\theta}_h^n); \theta_h^n, \mathbf{v}\right) = (\mathbf{f}^n, \mathbf{v}_h),\\
&B_f\left(\mathbf{u}_h^n, q_h\right) = 0,\\
&\left(D_\tau \mathbf{b}_h^n, \mathbf{w}_h\right) + B_b\left(\mu^n(\widetilde{\theta}_h^n); \mathbf{b}_h^n, \mathbf{w}_h\right) - T_{b2}\left(\mathbf{u}_h^n, \widetilde{\mathbf{b}}_h^n, \mathbf{w}_h\right) = \mathbf{0},\\
&\left(D_\tau \theta_h^n, \phi_h\right) + B_\theta\left(\kappa^n(\widetilde{\theta}_h^n); \theta_h^n, \phi_h\right) + T_\theta\left(\widetilde{\mathbf{u}}_h^n, \theta_h^n, \phi_h\right) = (\psi^n, \phi_h),
\end{aligned}
\tag{3.1}
$$

where the initial approximations $(\mathbf{u}_h^i, \mathbf{b}_h^i, \theta_h^i, p_h^i)$, for $i = 0, 1, 2$, will be chosen below. The scheme (3.1) is a linear system at each time step while we employ the extrapolation in linearizing the nonlinear terms. By using Lax-Milgram theorem, we can easily get the existence and uniqueness of the solution $(\mathbf{u}_h^n, \mathbf{b}_h^n, \theta_h^n, p_h^n)$ of the linear problem (3.1).

Note that the stability and convergence of the second-order BDF is easily proved by classical analysis but it only has the second-order accuracy of time discretization. Hence, in order to achieve high-order temporal accuracy, the BDFk, $k=3,4,5$ is a usually choice. We also notice that the powerful Nevanlinna-Odeh multiplier technique proposed in [35] extends the applicability of the energy method to the non A-stable BDF-k, $k=3,4,5$ methods. Besides, Lubich *et al*. [29] have analyzed the BDF methods up to fifth order for parabolic equations based on G-stability theory and Nevanlinna-Odeh multipliers. Thanks to a new telescope formula for the BDFk, $k=3,4,5$ scheme established by [26] for the parabolic equations, the current results can be extended to higher-order temporal schemes such as forth-order BDF and fifth-order BDF methods.

In order to ensure the third-order accuracy of the full system in time, we need to obtain the initial approximations $(\mathbf{u}_h^n,\mathbf{b}_h^n,\theta_h^n,p_h^n)\in\mathbf{X}_h^r\times\mathbf{W}_h^r\times M_h^r\times Q_h^{r-1}$, $n=0,1,2$, which should be also computed with temporal error of $\mathcal{O}(\tau^3)$. In fact, there are many ways to obtain such high order initial approximations, see [3]. In this paper, we set $\mathbf{u}_h^0=R_h\mathbf{u}_0$, $\mathbf{b}_h^0=F_h\mathbf{b}_0$, $\theta_h^0=T_h\theta_0$, in which R_h,F_h and T_h are defined in Section 4, and the initial values $\mathbf{u}_h^0$, $\mathbf{b}_h^0$ and θ_h^0 meet the following estimates:

$$\begin{aligned}
&\left\|\mathbf{u}_h^0-\mathbf{u}_0\right\|_{L^2}\le ch^{r+1}\|\mathbf{u}_0\|_{r+1},\\
&\left\|\mathbf{b}_h^0-\mathbf{b}_0\right\|_{L^2}\le ch^{r+1}\|\mathbf{b}_0\|_{r+1},\\
&\left\|\theta_h^0-\theta_0\right\|_{L^2}\le ch^{r+1}\|\theta_0\|_{r+1}.
\end{aligned}$$

Next, we compute the other initial approximations by the following four steps:

1. We compute intermediate value $(\widehat{\mathbf{u}}_h^{1/2},\widehat{\mathbf{b}}_h^{1/2},\widehat{\theta}_h^{1/2},\widehat{p}_h^{1/2})\in\mathbf{X}_h^r\times\mathbf{W}_h^r\times M_h^r\times Q_h^{r-1}$ for all $\mathbf{v}_h\in\mathbf{X}_h^r$, $\mathbf{w}_h\in\mathbf{W}_h^r$, $\phi_h\in M_h^r$ and $q_h\in Q_h^{r-1}$ by the application of the classical backward Euler scheme

$$\begin{aligned}
&\left(\frac{\widehat{\mathbf{u}}_h^{1/2}-\mathbf{u}_h^0}{\tau/2},\mathbf{v}_h\right)+T_u\left(\mathbf{u}_h^0,\widehat{\mathbf{u}}_h^{1/2},\mathbf{v}_h\right)+B_u\left(\nu^0(\theta_h^0),\widehat{\mathbf{u}}_h^{1/2},\mathbf{v}_h\right)-B_f\left(\mathbf{v}_h,\widehat{p}_h^{1/2}\right)\\
&\quad+B_f\left(\widehat{\mathbf{u}}_h^{1/2},q_h\right)+sT_{b1}\left(\mathbf{b}_h^0,\widehat{\mathbf{b}}_h^{1/2},\mathbf{v}_h\right)-D_f\left(\beta^0(\theta_h^0),\widehat{\theta}^{1/2},\mathbf{v}_h\right)=\left(\mathbf{f}^{1/2},\mathbf{v}_h\right),\\
&\left(\frac{\widehat{\mathbf{b}}_h^{1/2}-\mathbf{b}_h^0}{\tau/2},\mathbf{w}_h\right)+B_b\left(\mu^0(\theta_h^0),\widehat{\mathbf{b}}_h^{1/2},\mathbf{w}_h\right)-T_{b2}\left(\widehat{\mathbf{u}}_h^{1/2},\mathbf{b}_h^0,\mathbf{w}_h\right)=0,\\
&\left(\frac{\widehat{\theta}_h^{1/2}-\theta_h^0}{\tau/2},\phi_h\right)+B_\theta\left(\kappa^0(\theta_h^0),\widehat{\theta}_h^{1/2},\phi_h\right)+T_\theta\left(\mathbf{u}_h^0,\widehat{\theta}_h^{1/2},\phi_h\right)=\left(\psi^{1/2},\phi_h\right).
\end{aligned}$$

2. For all $\mathbf{v}_h\in\mathbf{X}_h^r,\mathbf{w}_h\in\mathbf{W}_h^r,\phi_h\in M_h^r$ and $q_h\in Q_h^{r-1}$, we obtain intermediate value $(\widehat{\mathbf{u}}_h^1,\widehat{\mathbf{b}}_h^1,\widehat{\theta}_h^1,\widehat{P}_h^1)\in\mathbf{X}_h^r\times\mathbf{W}_h^r\times M_h^r\times Q_h^{r-1}$ by using the Crank-Nicolson scheme as the solution of

$$\begin{aligned}
&\left(\frac{\widehat{\mathbf{u}}_h^1-\mathbf{u}_h^0}{\tau},\mathbf{v}_h\right)+T_u\left(\widehat{\mathbf{u}}_h^{1/2},\frac{\widehat{\mathbf{u}}_h^1+\mathbf{u}_h^0}{2},\mathbf{v}_h\right)+B_u\left(\nu^{1/2}(\widehat{\theta}_h^{1/2}),\frac{\widehat{\mathbf{u}}_h^1+\mathbf{u}_h^0}{2},\mathbf{v}_h\right)\\
&\quad-B_f\left(\mathbf{v}_h,\frac{\widehat{p}_h^1+p_h^0}{2}\right)+B_f\left(\frac{\widehat{\mathbf{u}}_h^1+\mathbf{u}_h^0}{2},q_h\right)+sT_{b1}\left(\widehat{\mathbf{b}}_h^{1/2},\frac{\widehat{\mathbf{b}}_h^1+\mathbf{b}_h^0}{2},\mathbf{v}_h\right)
\end{aligned}$$

$$-D_f\left(\beta^{1/2}(\widehat{\theta}_h^{1/2});\widehat{\theta}_h^1,\mathbf{v}_h\right)=\left(\overline{\mathbf{f}}^1,\mathbf{v}_h\right),$$

$$\left(\frac{\widehat{\mathbf{b}}_h^1-\mathbf{b}_h^0}{\tau},\mathbf{w}_h\right)+B_b\left(\mu^{1/2}(\widehat{\theta}_h^{1/2}),\frac{\widehat{\mathbf{b}}_h^1+\mathbf{b}_h^0}{2},\mathbf{w}_h\right)-T_{b2}\left(\frac{\widehat{\mathbf{u}}_h^1+\mathbf{u}_h^0}{2},\widehat{\mathbf{b}}_h^{1/2},\mathbf{w}_h\right)=0,$$

$$\left(\frac{\widehat{\theta}_h^1-\theta_h^0}{\tau},\phi_h\right)+B_\theta\left(\kappa^{1/2}(\widehat{\theta}_h^{1/2}),\frac{\widehat{\theta}_h^1+\theta_h^0}{2},\phi_h\right)+T_\theta\left(\widehat{\mathbf{u}}_h^{1/2},\frac{\widehat{\theta}_h^1+\theta_h^0}{2},\phi_h\right)=\left(\overline{\psi}^1,\phi_h\right).$$

3. For all $\mathbf{v}_h\in\mathbf{X}_h^r,\mathbf{w}_h\in\mathbf{W}_h^r,\phi_h\in M_h^r,q_h\in Q_h^{r-1}$, we obtain $(\mathbf{u}_h^1,\mathbf{b}_h^1,\theta_h^1,p_h^1)\in\mathbf{X}_h^r\times\mathbf{W}_h^r\times M_h^r\times Q_h^{r-1}$ updated by

$$\begin{aligned}&\left(\frac{\mathbf{u}_h^1-\mathbf{u}_h^0}{\tau},\mathbf{v}_h\right)+T_u\left(\frac{\widehat{\mathbf{u}}_h^1+\mathbf{u}_h^0}{2},\overline{\mathbf{u}}_h^1,\mathbf{v}_h\right)\\&\quad+B_u\left(\frac{\nu^1(\widehat{\theta}_h^1)+\nu^0(\theta_h^0)}{2},\overline{\mathbf{u}}_h^1,\mathbf{v}_h\right)\\&\quad-B_f\left(\mathbf{v}_h,\overline{p}_h^1\right)+B_f\left(\overline{\mathbf{u}}_h^1,q_h\right)+sT_{b1}\left(\frac{\widehat{\mathbf{b}}_h^1+\mathbf{b}_h^0}{2},\overline{\mathbf{b}}_h^1,\mathbf{v}_h\right)\\&\quad-D_f\left(\frac{\beta^1(\widehat{\theta}_h^1)+\beta^0(\theta_h^0)}{2},\overline{\theta}_h^1,\mathbf{v}_h\right)=\left(\overline{\mathbf{f}}^1,\mathbf{v}_h\right),\\&\left(\frac{\mathbf{b}_h^1-\mathbf{b}_h^0}{\tau},\mathbf{w}_h\right)+B_b\left(\frac{\nu^1(\widehat{\theta}_h^1)+\nu^0(\theta_h^0)}{2},\overline{\mathbf{b}}_h^1,\mathbf{w}_h\right)\\&\quad-T_{b2}\left(\overline{\mathbf{u}}_h^1,\frac{\widehat{\mathbf{b}}_h^1+\mathbf{b}_h^0}{2},\mathbf{w}_h\right)=0,\\&\left(\frac{\theta_h^1-\theta_h^0}{\tau},\phi_h\right)+B_\theta\left(\frac{\kappa^1(\widehat{\theta}_0^1)+\kappa^0(\theta_0^0)}{2},\overline{\theta}_h^1,\phi_h\right)\\&\quad+T_\theta\left(\frac{\widehat{\mathbf{u}}_h^1+\mathbf{u}_h^0}{2},\overline{\theta}_h^1,\phi_h\right)=\left(\overline{\psi}^1,\phi_h\right).\end{aligned}\tag{3.2}$$

4. For all $\mathbf{v}_h\in\mathbf{X}_h^r,\mathbf{w}_h\in\mathbf{W}_h^r,\phi_h\in M_h^r,q_h\in Q_h^{r-1}$, the initial value $(\mathbf{u}_h^2,\mathbf{b}_h^2,\theta_h^2,p_h^2)\in\mathbf{X}_h^r\times\mathbf{W}_h^r\times M_h^r\times Q_h^{r-1}$ is determined by

$$\begin{aligned}&\left(\frac{\mathbf{u}_h^2-\mathbf{u}_h^1}{\tau},\mathbf{v}_h\right)+T_u\left(\underline{\mathbf{u}}_h^2,\overline{\mathbf{u}}_h^2,\mathbf{v}_h\right)+B_u\left(\nu^{3/2}(\underline{\theta}_h^2),\overline{\mathbf{u}}_h^2,\mathbf{v}_h\right)-B_f\left(\mathbf{v}_h,\overline{p}_h^2\right)\\&\quad+B_f\left(\overline{\mathbf{u}}_h^2,q_h\right)+sT_{b1}\left(\underline{\mathbf{b}}_h^2,\overline{\mathbf{b}}_h^2,\mathbf{v}_h\right)-D_f\left(\beta^{3/2}(\underline{\theta}_h^2),\overline{\theta}_h^2,\mathbf{v}_h\right)=\left(\overline{\mathbf{f}}^2,\mathbf{v}_h\right),\end{aligned}\tag{3.3a}$$

$$\left(\frac{\mathbf{b}_h^2-\mathbf{b}_h^1}{\tau},\mathbf{w}_h\right)+B_b\left(\mu^{3/2}(\underline{\theta}_h^2),\overline{\mathbf{b}}_h^2,\mathbf{w}_h\right)-T_{b2}\left(\overline{\mathbf{u}}_h^2,\underline{\mathbf{b}}_h^2,\mathbf{w}_h\right)=0,\tag{3.3b}$$

$$\left(\frac{\theta_h^2-\theta_h^1}{\tau},\phi_h\right)+B_\theta\left(\kappa^{3/2}(\underline{\theta}_h^2),\overline{\theta}_h^2,\phi_h\right)+T_\theta\left(\underline{\mathbf{u}}_h^2,\overline{\theta}_h^2,\phi_h\right)=\left(\overline{\psi}^2,\phi_h\right), \tag{3.3c}$$

where $\overline{H}^n=(H^n+H^{n-1})/2$, $n=1,2$, and $\underline{H}^2=3H^1/2-H^0/2$, $H=\theta,\mathbf{u}$ or $\mathbf{b}$.

Now, we present the main result of this work. The proof of it is given in Section 5.

Theorem 3.1. *Suppose that the solution* $(\mathbf{u},\mathbf{b},p,\theta)$ *to the thermally coupled incompressible MHD equations* (1.1)-(1.2) *and the corresponding approximation* $(\mathbf{u}_h^m,\mathbf{b}_h^m,p_h^m,\theta_h^m)$ *of* (3.1)-(3.3) *satisfy Assumption* 2.1-2.4. *Then for* $1\le m\le N$, *there exist two constants* $h_0>0$ *and* $\tau_0>0$ *such that when* $h\le h_0$ *and* $\tau\le\tau_0$, *the following error estimate holds:*

$$\left\|\mathbf{u}(t_m)-\mathbf{u}_h^m\right\|_{L^2}^2+s\left\|\mathbf{b}(t_m)-\mathbf{b}_h^m\right\|_{L^2}^2+\left\|\theta(t_m)-\theta_h^m\right\|_{L^2}^2\le C_0\left(\tau^6+h^{2r+2}\right). \tag{3.4}$$

4. Primary Error Estimates

In this section, we employ the error splitting method to obtain temporal and spatial errors. After that we apply the triangular inequality to show the boundedness of the numerical solutions, which is later used in the proof of Theorem 3.1.

Before giving error analysis, for $\widehat{\nu},\widehat{\mu},\widehat{\kappa}\in C^{0,1}(\overline{\Omega};R^+)$ satisfying Assumption 2.2, we recall the following three projections [39, 42]. Supposing $(\mathbf{u},p)\in\mathbf{X}\times Q$, we define the Stokes projection $(R_h\mathbf{u},J_hp)\in\mathbf{X}_h^r\times Q_h^{r-1}$ by

$$\begin{aligned}&B_u(\widehat{\nu};\mathbf{u}-R_h\mathbf{u},\mathbf{v})+B_f(\mathbf{v}_h,p-J_hp)=0 &&\text{for all}\quad \mathbf{v}\in\mathbf{X}_h^r,\\ &B_f(\mathbf{u}-R_h\mathbf{u},q)=0 &&\text{for all}\quad q\in Q_h^{r-1},\end{aligned}$$

and note that

$$\begin{aligned}&\|(\mathbf{u}-R_h\mathbf{u})\|_{L^2}+h\|\nabla(\mathbf{u}-R_h\mathbf{u})\|_{L^2}+h\|p-J_hp\|_{L^2}\le Ch^{r+1}\big(\|\mathbf{u}\|_{r+1}+\|p\|_r\big),\\ &\|R_h\mathbf{u}\|_{L^\infty}+\|\nabla R_h\mathbf{u}\|_{L^3}\le C\big(\|\mathbf{u}\|_2+\|p\|_1\big).\end{aligned} \tag{4.1}$$

For $\mathbf{b}\in\mathbf{W}$, we define the Maxwell projection $F_h\mathbf{b}\in\mathbf{W}_h^r$ by

$$B_b\big(\widehat{\mu};\mathbf{b}-F_h\mathbf{b},\mathbf{w}_h\big)=0\quad\text{for all}\quad \mathbf{w}_h\in\mathbf{W}_h^r,$$

and note that

$$\begin{aligned}&\|\mathbf{b}-F_h\mathbf{b}\|_{L^2}+h\|\nabla(\mathbf{b}-F_h\mathbf{b})\|_{L^2}\le Ch^{r+1}\|\mathbf{b}\|_{r+1},\\ &\|F_h\mathbf{b}\|_{L^\infty}+\|\nabla F_h\mathbf{b}\|_{L^3}\le C\|\mathbf{b}\|_2.\end{aligned} \tag{4.2}$$

In addition, for $\theta\in M$ we also consider the Ritz projection $T_h\theta\in M_h^r$ defined by

$$B_\theta\big(\widehat{\kappa};\theta-T_h\theta,\phi\big)=0\quad\text{for all}\quad \phi\in M_h^r,$$

and note that

$$\begin{aligned}&\|(\theta-T_h\theta)\|_{L^2}+h\|\nabla(\theta-T_h\theta)\|_{L^2}\le h^{r+1}\|\theta\|_{r+1},\\&\|T_h\theta\|_{L^\infty}+\|\nabla T_h\theta\|_{L^3}\le C\|\theta\|_2.\end{aligned} \tag{4.3}$$

In order to derive error estimates, we will use the following error splits:

$$\begin{aligned}e_{\mathbf{u}}^m&=\mathbf{u}_h^m-\mathbf{u}(t_m)=\big(\mathbf{u}^m-\mathbf{u}(t_m)\big)+\big(\mathbf{u}_h^m-R_h\mathbf{u}^m\big)+\big(R_h\mathbf{u}^m-\mathbf{u}^m\big)=:\delta_{\mathbf{u}}^m+\varepsilon_{\mathbf{u}}^m+\rho_{\mathbf{u}}^m,\\e_{\mathbf{b}}^m&=\mathbf{b}_h^m-\mathbf{b}(t_m)=\big(\mathbf{b}^m-\mathbf{b}(t_m)\big)+\big(\mathbf{b}_h^m-F_h\mathbf{b}^m\big)+\big(F_h\mathbf{b}^m-\mathbf{b}^m\big)=\delta_{\mathbf{b}}^m+\varepsilon_{\mathbf{b}}^m+\rho_{\mathbf{b}}^m,\\e_p^m&:=p_h^m-p(t_m)=\big(p^m-p(t_m)\big)+\big(p_h^m-J_hp^m\big)+\big(J_hp^m-p^m\big)=\delta_p^m+\varepsilon_p^m+\rho_p^m,\\e_\theta^m&=\theta_h^m-\theta(t_m)=\big(\theta^m-\theta(t_m)\big)+\big(\theta_h^m-T_h\theta^m\big)+\big(T_h\theta^m-\theta^m\big)=:\delta_\theta^m+\varepsilon_\theta^m+\rho_\theta^m,\end{aligned}$$

where δ^m,ε^m, and ρ^m, $m\ge 0$ are temporal, spatial, and projection errors, respectively.

Here, we state the convergence results for the initial values.

Lemma 4.1. *Let* $(\mathbf{u}_h^j,\mathbf{u}_h^j,\theta_h^j,p_h^j)$*, for* $j=1,2$ *be the solution to* (3.2) *and* (3.3)*, respectively. Then under Assumptions* 2.1-2.4*, the following estimates hold:*

$$\begin{aligned}&\big\|\delta_{\mathbf{u}}^j\big\|_{L^2}^2+s\big\|\delta_{\mathbf{b}}^j\big\|_{L^2}^2+\big\|\delta_\theta^j\big\|_{L^2}^2\\&\quad+\tau\Big(\nu_1\big\|\delta_{\mathbf{u}}^j\big\|_1^2+s\mu_1\big\|\delta_{\mathbf{b}}^j\big\|_1^2+\kappa_1\big\|\delta_\theta^j\big\|_1^2\Big)\\&\quad+\tau^2\Big(\big\|\delta_{\mathbf{u}}^j\big\|_2^2+s\big\|\delta_{\mathbf{b}}^j\big\|_2^2+\big\|\delta_\theta^j\big\|_2^2\Big)\le c\tau^6,\end{aligned}\tag{4.4}$$

$$\max_{0\le i\le 2}\big\{\|\mathbf{u}^i\|_{W^{2,4}}+\|\mathbf{b}^i\|_{W^{2,4}}+\|\theta^i\|_{W^{2,4}}\big\}\le c,$$

$$\big\|\varepsilon_{\mathbf{u}}^j\big\|_{L^2}^2+s\big\|\varepsilon_{\mathbf{b}}^j\big\|_{L^2}^2+\big\|\varepsilon_\theta^j\big\|_{L^2}^2\le ch^{2r+2}.\tag{4.5}$$

Moreover, for $j=0,1,2$*, we have*

$$\big\|e_{\mathbf{u}}^j\big\|_{L^2}^2+s\big\|e_{\mathbf{b}}^j\big\|_{L^2}^2+\big\|e_\theta^j\big\|_{L^2}^2\le C\big(h^{2r+2}+\tau^6\big).$$

The proof of this lemma is similar to proofs in [3, 27], where incompressible MHD equations are solved by a linearized Crank-Nicolson scheme with a linearized backward Euler scheme as the starting approximations (see also [45]).

4.1. Temporal error estimates

In this section, we want to obtain temporal convergence. We first introduce the time semi-discrete scheme. For $3\le n\le N$, we define $(\mathbf{u}^n,\mathbf{b}^n,\theta^n,p^n)\in\mathbf{X}\times\mathbf{W}\times M\times Q$ be the solutions of the following time-discrete system:

$$\begin{aligned}&D_\tau\mathbf{u}^n-\nabla\cdot\Big(\nu^n(\widetilde{\theta}^n)\nabla\mathbf{u}^n\Big)+(\widetilde{\mathbf{u}}^n\cdot\nabla)\mathbf{u}^n+\nabla p^n+s\widetilde{\mathbf{b}}^n\times\operatorname{curl}\mathbf{b}^n-\beta^n(\widetilde{\theta}^n)\theta^n\mathbf{j}=\mathbf{f}^n,\\&\nabla\cdot\mathbf{u}^n=0,\\&D_\tau\mathbf{b}^n+\operatorname{curl}\big(\mu^n(\widetilde{\theta}^n)\operatorname{curl}\mathbf{b}^n\big)-\operatorname{curl}\big(\mathbf{u}^n\times\widetilde{\mathbf{b}}^n\big)=\mathbf{0},\\&D_\tau\theta^n-\nabla\cdot\Big(\kappa^n(\widetilde{\theta}^n)\nabla\theta^n\Big)+\widetilde{\mathbf{u}}^n\cdot\nabla\theta^n=\psi^n.\end{aligned}\tag{4.6}$$

Remark 4.1. Taking divergence on the magnetic equations of time-discrete system, and making use of $\operatorname{div}\operatorname{curl}=0$ and $\operatorname{div}\mathbf{b}_0=0$, we first obtain $\operatorname{div}\mathbf{b}^1=\operatorname{div}\mathbf{b}^2=0$ and then get $\operatorname{div}\mathbf{b}^n=0$ for $3\le n\le N$.

Now, we show the temporal error estimate as follows.

Theorem 4.1. *Let* $(\mathbf{u}^j,\mathbf{b}^j,\theta^j,p^j)$, $3\le j\le N$ *be the solution of* (4.6). *Under Assumptions* 2.1-2.4, *for* $1\le k\le N$ *there exists a constant* $\tau_1>0$ *such that if* $\tau\le\tau_1$, *then*

$$
\begin{aligned}
&\big\|\delta_{\mathbf{u}}^k\big\|_{L^2}^2+s\big\|\delta_{\mathbf{b}}^k\big\|_{L^2}^2+\big\|\delta_{\theta}^k\big\|_{L^2}^2\\
&\quad+\tau\sum_{j=1}^{k}\Big(\nu_1\big\|\delta_{\mathbf{u}}^j\big\|_1^2+s\mu_1\big\|\delta_{\mathbf{b}}^j\big\|_1^2+\kappa_1\big\|\delta_{\theta}^j\big\|_1^2\Big)\\
&\quad+\tau^2\Big(\big\|\delta_{\mathbf{u}}^k\big\|_2^2+s\big\|\delta_{\mathbf{b}}^k\big\|_2^2+\big\|\delta_{\theta}^k\big\|_2^2\Big)\le C_1\tau^6, \qquad (4.7)\\
&\max_{0\le j\le k}\|\mathbf{u}^j\|_{W^{2,4}}+s\max_{0\le j\le k}\|\mathbf{b}^j\|_{W^{2,4}}+\max_{0\le j\le k}\|\theta^j\|_{W^{2,4}}\le C. \qquad (4.8)
\end{aligned}
$$

Proof. We will first prove (4.7) by using the method of mathematical induction. In view of (4.4), the error estimate (4.7) is valid for $k=0,1,2$. Now, we assume that (4.7) holds for $m\le k-1$ for $3\le k\le N$, then

$$
\begin{aligned}
&\big\|\widetilde{\delta}_{\mathbf{u}}^{m+1}\big\|_{L^2}^2+s\big\|\widetilde{\delta}_{\mathbf{b}}^{m+1}\big\|_{L^2}^2+\big\|\widetilde{\delta}_{\theta}^{m+1}\big\|_{L^2}^2\\
&\quad+\tau\sum_{j=1}^{m+1}\Big(\nu_1\big\|\widetilde{\delta}_{\mathbf{u}}^j\big\|_1^2+s\mu_1\big\|\widetilde{\delta}_{\mathbf{b}}^j\big\|_1^2+\kappa_1\big\|\widetilde{\delta}_{\theta}^j\big\|_1^2\Big)\\
&\quad+\tau^2\Big(\big\|\widetilde{\delta}_{\mathbf{u}}^{m+1}\big\|_2^2+s\big\|\widetilde{\delta}_{\mathbf{b}}^{m+1}\big\|_2^2+\big\|\widetilde{\delta}_{\theta}^{m+1}\big\|_2^2\Big)\le CC_1\tau^6,
\end{aligned}
$$

where $C>0$ independent of C_1. In addition, we have

$$
\begin{aligned}
&\|\widetilde{\mathbf{u}}^m\|_2+s\|\widetilde{\mathbf{b}}^m\|_2+\|\widetilde{\theta}^m\|_2\\
\le&\|\widetilde{\mathbf{u}}(t_m)\|_2+s\|\widetilde{\mathbf{b}}(t_m)\|_2+\|\widetilde{\theta}(t_m)\|_2+\|\widetilde{\delta}_{\mathbf{u}}^m\|_2+s\|\widetilde{\delta}_{\mathbf{b}}^m\|_2+\|\widetilde{\delta}_{\theta}^m\|_2\\
\le& C+\sqrt{C}\sqrt{C_1\tau^4}\le C,\quad m\le k,
\end{aligned}
$$

for $C_1\tau^4\le 1$. In order to finish the mathematical induction, we need to find C_1 such that (4.7) also holds for $m\le k$. Subtracting (4.6) from (1.1) leads to

$$
\begin{aligned}
&D_\tau\delta_{\mathbf{u}}^m-\nabla\cdot\big(\nu^m(\widetilde{\theta}^m)\nabla\delta_{\mathbf{u}}^m\big)-\beta^m(\widetilde{\theta}^m)\delta_\theta^m\mathbf{j}+\nabla\delta_p^m=\Gamma_{\mathbf{u}},\\
&\nabla\cdot\delta_{\mathbf{u}}^m=0,\\
&D_\tau\delta_{\mathbf{b}}^m+\operatorname{curl}\big(\mu^m(\widetilde{\theta}^m)\operatorname{curl}\delta_{\mathbf{b}}^m\big)=\Gamma_{\mathbf{b}},\\
&D_\tau\delta_\theta^m-\nabla\cdot\big(\kappa^m(\widetilde{\theta}^m)\nabla\delta_\theta^m\big)=\Gamma_\theta,
\end{aligned}
\qquad (4.9)
$$

where $\Gamma_{\mathbf{u}},\Gamma_{\mathbf{b}},\Gamma_\theta$ are defined as follows. Adding and subtracting suitable terms, we get

$$
\Gamma_{\mathbf{u}}:=\partial_t\mathbf{u}(t_m)-D_\tau\mathbf{u}(t_m)-\nabla\cdot\big(\nu^m(\theta(t_m))\nabla\mathbf{u}(t_m)\big)+\nabla\cdot\big(\nu^m(\widetilde{\theta}^m)\nabla\mathbf{u}(t_m)\big)
$$

$$
\begin{aligned}
&-(\widetilde{\mathbf{u}}^m\cdot\nabla)\mathbf{u}^m+\big(\mathbf{u}(t_m)\cdot\nabla\big)\mathbf{u}(t_m)+s\widetilde{\mathbf{b}}^m\times\operatorname{curl}\mathbf{b}^m\\
&-s\mathbf{b}(t_m)\times\operatorname{curl}\mathbf{b}(t_m)-\beta^m\big(\theta(t_m)\big)\theta(t_m)+\beta^m(\widetilde{\theta}^m)\theta(t_m)\\
=&\big(\partial_t\mathbf{u}(t_m)-D_\tau\mathbf{u}(t_m)\big)-\nabla\cdot\Big(\big(\nu^m(\theta(t_m))-\nu^m(\widetilde{\theta}^m)\big)\nabla\mathbf{u}(t_m)\Big)\\
&-\big(\beta^m(\theta(t_m))-\beta^m(\widetilde{\theta}^m)\big)\theta(t_m)\mathbf{j}-(\widetilde{\mathbf{u}}^m\cdot\nabla)\,\delta_{\mathbf{u}}^m-\big(\widetilde{\delta}_{\mathbf{u}}^m\cdot\nabla\big)\,\mathbf{u}(t_m)\\
&+\big((\mathbf{u}^m-\widetilde{\mathbf{u}}^m)\cdot\nabla\big)\mathbf{u}(t_m)-s\widetilde{\mathbf{b}}^m\times\operatorname{curl}\delta_{\mathbf{b}}^m-s\widetilde{\delta}_{\mathbf{b}}^m\times\operatorname{curl}\mathbf{b}(t_m)\\
&+s(\mathbf{b}^m-\widetilde{\mathbf{b}}^m)\times\operatorname{curl}\mathbf{b}(t_m)=:\sum_{i=1}^{9}\Gamma_{\mathbf{u}}^i,
\end{aligned}
$$

$$
\begin{aligned}
\Gamma_{\mathbf{b}}:=&\,\partial_t\mathbf{b}(t_m)-D_\tau\mathbf{b}(t_m)+\operatorname{curl}\big(\mu^m(\theta(t_m))\operatorname{curl}\mathbf{b}(t_m)\big)\\
&-\operatorname{curl}\big(\mu^m(\widetilde{\theta}^m)\operatorname{curl}\mathbf{b}(t_m)\big)\\
&+\operatorname{curl}(\mathbf{u}^m\times\widetilde{\mathbf{b}}^m)-\operatorname{curl}\big(\mathbf{u}(t_m)\times\mathbf{b}(t_m)\big)\\
=&\big(\partial_t\mathbf{b}(t_m)-D_\tau\mathbf{b}(t_m)\big)+\operatorname{curl}\Big(\big(\mu^m(\theta(t_m))-\mu^m(\widetilde{\theta}^m)\big)\operatorname{curl}\mathbf{b}(t_m)\Big)\\
&+\operatorname{curl}\big(\delta_{\mathbf{u}}^m\times\widetilde{\mathbf{b}}^m\big)+\operatorname{curl}\big(\mathbf{u}(t_m)\times\widetilde{\delta}_{\mathbf{b}}^m\big)\\
&-\operatorname{curl}\big(\mathbf{u}(t_m)\times(\mathbf{b}^m-\widetilde{\mathbf{b}}^m)\big)=:\sum_{i=1}^{5}\Gamma_{\mathbf{b}}^i,
\end{aligned}
$$

$$
\begin{aligned}
\Gamma_\theta:=&\,\partial_t\theta(t_m)-D_\tau\theta(t_m)-\nabla\cdot\big(\kappa^m(\theta(t_m))\nabla\theta(t_m)\big)\\
&+\nabla\cdot\big(\kappa^m(\widetilde{\theta}^m)\nabla\theta(t_m)\big)-\widetilde{\mathbf{u}}^m\cdot\nabla\theta^m+\mathbf{u}(t_m)\cdot\nabla\theta(t_m)\\
=&\big(\partial_t\theta(t_m)-D_\tau\theta(t_m)\big)-\nabla\cdot\Big(\big(\kappa^m(\theta(t_m))-\kappa^m(\widetilde{\theta}^m)\big)\nabla(\theta(t_m))\Big)\\
&-\widetilde{\mathbf{u}}^m\cdot\nabla\delta_\theta^m-\widetilde{\delta}_{\mathbf{u}}^m\cdot\nabla\theta(t_m)+(\mathbf{u}^m-\widetilde{\mathbf{u}}^m)\cdot\nabla\theta(t_m)=:\sum_{i=1}^{5}\Gamma_\theta^i.
\end{aligned}
$$

After that, respectively multiplying (4.9) by $\mathbf{v},q,\mathbf{w}$ and ϕ and using integration by parts gives

$$
\begin{aligned}
&\big(D_\tau\delta_{\mathbf{u}}^m,\mathbf{v}\big)+B_u\big(\nu^m(\widetilde{\theta}^m),\delta_{\mathbf{u}}^m,\mathbf{v}\big)-D_f\big(\beta^m(\widetilde{\theta}^m),\delta_\theta^m,\mathbf{v}\big)-B_f\big(\mathbf{v},\delta_p^m\big)=(\Gamma_{\mathbf{u}},\mathbf{v}),\\
&B_f\big(\delta_{\mathbf{u}}^m,q\big)=0,\\
&\big(D_\tau\delta_{\mathbf{b}}^m,\mathbf{w}\big)+B_b\big(\mu^m(\widetilde{\theta}^m),\delta_{\mathbf{b}}^m,\mathbf{w}\big)=(\Gamma_{\mathbf{b}},\mathbf{w}),\\
&\big(D_\tau\delta_\theta^m,\phi\big)+B_\theta\big(\kappa^m(\widetilde{\theta}^m),\delta_\theta^m,\phi\big)=(\Gamma_\theta,\phi).
\end{aligned}
\tag{4.10}
$$

Choose

$$
(\mathbf{v},q,\mathbf{w},\phi)=\Big(2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1},\delta_p^m,2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1},2\delta_\theta^m-\delta_\theta^{m-1}\Big)
$$

and estimate each term on the right-hand side of (4.10). Taking into account the Hölder and Young inequalities along with (2.8)-(2.9) and (2.4), we obtain

$$
\begin{aligned}
\big(\Gamma_{\mathbf{u}}^1,2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\big)&\le\big\|\partial_t\mathbf{u}(t_m)-D_\tau\mathbf{u}(t_m)\big\|_{L^2}\big\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\big\|_{L^2}\\
&\le c\tau^{5/2}\big\|\partial_t^4\mathbf{u}(t)\big\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\big\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\big\|_{L^2}\\
&\le c\epsilon\tau^5\big\|\partial_t^4\mathbf{u}(t)\big\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\epsilon^{-1}\Big(\big\|\delta_{\mathbf{u}}^m\big\|^2_{L^2}+\big\|\delta_{\mathbf{u}}^{m-1}\big\|^2_{L^2}\Big),
\end{aligned}
$$

$$\begin{aligned}
\left(\Gamma_{\mathbf{u}}^2, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le \left\|\left(\nu^m(\theta(t_m))-\nu^m(\widetilde{\theta}^m)\right)\nabla\mathbf{u}(t_m)\right\|_{L^2}\left\|\nabla\left(2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right)\right\|_{L^2}\\
&\le c|\nu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left(\left\|\theta(t_m)-\widetilde{\theta}(t_m)\right\|_{L^2}+\left\|\widetilde{\theta}(t_m)-\widetilde{\theta}^m\right\|_{L^2}\right)\\
&\quad\times\|\nabla\mathbf{u}(t_m)\|_{L^\infty}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_1\\
&\le c\left(\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}^2+\tau^5\left\|\partial_t^3\theta\right\|_{L^2(t_{m-3},t_m;L^2(\Omega))}^2\right)\\
&\quad+\epsilon\left(\left\|\delta_{\mathbf{u}}^m\right\|_1^2+\left\|\delta_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\left(\Gamma_{\mathbf{u}}^3, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le \left\|\left(\beta^m(\theta(t_m))-\beta^m(\widetilde{\theta}^m)\right)\theta(t_m)\right\|_{L^2}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\le c|\beta|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left(\|\theta(t_m)-\widetilde{\theta}(t_m)\|_{L^2}+\|\widetilde{\theta}(t_m)-\widetilde{\theta}^m\|_{L^2}\right)\\
&\quad\times\|\theta(t_m)\|_{L^\infty}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\le c\epsilon\left(\tau^5\left\|\partial_t^3\theta\right\|_{L^2(t_{m-3},t_m;L^2(\Omega))}^2+\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}^2\right)\\
&\quad+\epsilon^{-1}\left(\left\|\delta_{\mathbf{u}}^m\right\|_{L^2}^2+\left\|\delta_{\mathbf{u}}^{m-1}\right\|_{L^2}^2\right),\\
\left(\Gamma_{\mathbf{u}}^4, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le \|\widetilde{\mathbf{u}}^m\|_{L^\infty}\left\|\nabla\delta_{\mathbf{u}}^m\right\|_{L^2}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\le c\epsilon^{-1}\left(\left\|\delta_{\mathbf{u}}^{m-1}\right\|_{L^2}^2+\left\|\delta_{\mathbf{u}}^m\right\|_{L^2}^2\right)+\epsilon\left\|\delta_{\mathbf{u}}^m\right\|_1^2,\\
\left(\Gamma_{\mathbf{u}}^5, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le \left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}\|\nabla\mathbf{u}(t_m)\|_{L^3}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^6}\\
&\le c\epsilon^{-1}\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}^2+\epsilon\left(\left\|\delta_{\mathbf{u}}^m\right\|_1^2+\left\|\delta_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\left(\Gamma_{\mathbf{u}}^6, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le \|\mathbf{u}^m-\widetilde{\mathbf{u}}^m\|_{L^2}\|\nabla\mathbf{u}(t_m)\|_{L^3}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^6}\\
&\le c\epsilon^{-1}\tau^5\left\|\partial_t^3\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2+\epsilon\left(\left\|\delta_{\mathbf{u}}^m\right\|_1^2+\left\|\delta_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\left(\Gamma_{\mathbf{u}}^8, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le s\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}\|\operatorname{curl}\mathbf{b}(t_m)\|_{L^3}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^6}\\
&\le c\epsilon^{-1}s\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}^2+\epsilon\left(\left\|\delta_{\mathbf{u}}^m\right\|_1^2+\left\|\delta_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\left(\Gamma_{\mathbf{u}}^9, 2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right) &\le s\|\mathbf{b}^m-\widetilde{\mathbf{b}}^m\|_{L^2}\|\operatorname{curl}\mathbf{b}(t_m)\|_{L^3}\left\|2\delta_{\mathbf{u}}^m-\delta_{\mathbf{u}}^{m-1}\right\|_{L^6}\\
&\le c\epsilon^{-1}s\tau^5\left\|\partial_t^3\mathbf{b}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2+\epsilon\left(\left\|\delta_{\mathbf{u}}^m\right\|_1^2+\left\|\delta_{\mathbf{u}}^{m-1}\right\|_1^2\right).
\end{aligned}$$

Similar considerations for $\Gamma_{\mathbf{b}}^i$ shows

$$\begin{aligned}
\left(\Gamma_{\mathbf{b}}^1, 2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1}\right) &\le \|\partial_t\mathbf{b}(t_m)-D_\tau\mathbf{b}(t_m)\|_{L^2}\left\|2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1}\right\|_{L^2}\\
&\le c\tau^{5/2}\left\|\partial_t^4\mathbf{b}(t)\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\left\|2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1}\right\|_{L^2}\\
&\le c\epsilon\tau^5\left\|\partial_t^4\mathbf{b}(t)\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2+\epsilon^{-1}\left(\left\|\delta_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\delta_{\mathbf{b}}^{m-1}\right\|_{L^2}^2\right),\\
\left(\Gamma_{\mathbf{b}}^2, 2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1}\right) &\le \left\|\left(\mu^m(\theta(t_m))-\mu^m(\widetilde{\theta}^m)\right)\operatorname{curl}\mathbf{b}(t_m)\right\|_{L^2}\left\|\operatorname{curl}\left(2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1}\right)\right\|_{L^2}\\
&\le c|\mu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left(\left\|\theta(t_m)-\widetilde{\theta}(t_m)\right\|_{L^2}+\left\|\widetilde{\theta}(t_m)-\widetilde{\theta}^m\right\|_{L^2}\right)\\
&\quad\times\|\mathbf{b}(t_m)\|_{W^{1,\infty}}\left\|\operatorname{curl}\left(2\delta_{\mathbf{b}}^m-\delta_{\mathbf{b}}^{m-1}\right)\right\|_{L^2}
\end{aligned}$$

$$
\begin{aligned}
&\le c\left(\tau^5\left\|\partial_t^3\theta\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\left\|\widetilde{\delta}^m_\theta\right\|^2_{L^2}\right)+\epsilon\left(\left\|\delta^m_{\mathbf{b}}\right\|^2_1+\left\|\delta^{m-1}_{\mathbf{b}}\right\|^2_1\right),\\
\left(\Gamma^4_{\mathbf{b}},2\delta^m_{\mathbf{b}}-\delta^{m-1}_{\mathbf{b}}\right)&\le\|\mathbf{u}(t_m)\|_{L^\infty}\left\|\widetilde{\delta}^m_{\mathbf{b}}\right\|_{L^2}\left\|\mathrm{curl}\big(2\delta^m_{\mathbf{b}}-\delta^{m-1}_{\mathbf{b}}\big)\right\|_{L^2}\\
&\le c\epsilon^{-1}\left\|\widetilde{\delta}^m_{\mathbf{b}}\right\|^2_{L^2}+\epsilon\left(\left\|\delta^m_{\mathbf{b}}\right\|^2_1+\left\|\delta^{m-1}_{\mathbf{b}}\right\|^2_1\right),\\
\left(\Gamma^5_{\mathbf{b}},2\delta^m_{\mathbf{b}}-\delta^{m-1}_{\mathbf{b}}\right)&\le\|\mathbf{u}(t_m)\|_{L^\infty}\|\mathbf{b}^m-\widetilde{\mathbf{b}}^m\|_{L^2}\left\|\mathrm{curl}\big(2\delta^m_{\mathbf{b}}-\delta^{m-1}_{\mathbf{b}}\big)\right\|_{L^2}\\
&\le c\epsilon^{-1}\tau^5\left\|\partial_t^3\mathbf{b}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\epsilon\left(\left\|\delta^m_{\mathbf{b}}\right\|^2_1+\left\|\delta^{m-1}_{\mathbf{b}}\right\|^2_1\right).
\end{aligned}
$$

Using the Hölder and Young inequalities again and (2.5), we estimate $\Gamma^7_{\mathbf{u}}$ and $\Gamma^3_{\mathbf{b}}$ as

$$
\begin{aligned}
&\left(\Gamma^7_{\mathbf{u}},2\delta^m_{\mathbf{u}}-\delta^{m-1}_{\mathbf{u}}\right)+s\left(\Gamma^3_{\mathbf{b}},2\delta^m_{\mathbf{b}}-\delta^{m-1}_{\mathbf{b}}\right)\\
&=-sT_{b1}\left(\widetilde{\mathbf{b}}^m,\delta^m_{\mathbf{b}},2\delta^m_{\mathbf{u}}-\delta^{m-1}_{\mathbf{u}}\right)+sT_{b2}\left(\delta^m_{\mathbf{u}},\widetilde{\mathbf{b}}^m,2\delta^m_{\mathbf{b}}-\delta^{m-1}_{\mathbf{b}}\right)\\
&=-sT_{b1}\left(\widetilde{\mathbf{b}}^m,\delta^m_{\mathbf{b}},-\delta^{m-1}_{\mathbf{u}}\right)+sT_{b2}\left(\delta^m_{\mathbf{u}},\widetilde{\mathbf{b}}^m,-\delta^{m-1}_{\mathbf{b}}\right)\\
&\le s\|\widetilde{\mathbf{b}}^m\|_{L^\infty}\left\|\mathrm{curl}\delta^m_{\mathbf{b}}\right\|_{L^2}\left\|\delta^{m-1}_{\mathbf{u}}\right\|_{L^2}+s\left(\left\|\delta^m_{\mathbf{u}}\right\|_{L^3}\left\|\nabla\widetilde{\mathbf{b}}^m\right\|_{L^6}+\left\|\nabla\delta^m_{\mathbf{u}}\right\|_{L^2}\left\|\widetilde{\mathbf{b}}^m\right\|_{L^\infty}\right)\left\|\delta^{m-1}_{\mathbf{b}}\right\|_{L^2}\\
&\le c\epsilon^{-1}\left(\left\|\delta^{m-1}_{\mathbf{u}}\right\|^2_{L^2}+s\left\|\delta^{m-1}_{\mathbf{b}}\right\|^2_{L^2}\right)+\epsilon\left(\left\|\delta^m_{\mathbf{u}}\right\|^2_1+s\left\|\delta^m_{\mathbf{b}}\right\|^2_1\right).
\end{aligned}
$$

Bounding Γ^i_θ similarly we have

$$
\begin{aligned}
\left(\Gamma^1_\theta,2\delta^m_\theta-\delta^{m-1}_\theta\right)&\le\|\partial_t\theta(t_m)-D_\tau\theta(t_m)\|_{L^2}\left\|2\delta^m_\theta-\delta^{m-1}_\theta\right\|_{L^2}\\
&\le c\tau^{5/2}\left\|\partial_t^4\theta(t)\right\|_{L^2(t_{m-3},t_m;L^2(\Omega))}\left\|2\delta^m_\theta-\delta^{m-1}_\theta\right\|_{L^2}\\
&\le c\epsilon\tau^5\left\|\partial_t^4\theta(t)\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\epsilon^{-1}\left(\left\|\delta^m_\theta\right\|^2_{L^2}+\left\|\delta^{m-1}_\theta\right\|^2_{L^2}\right),\\
\left(\Gamma^2_\theta,2\delta^m_\theta-\delta^{m-1}_\theta\right)&\le\left\|\big(\kappa^m(\theta(t_m))-\kappa^m(\widetilde{\theta}^m)\big)\nabla\theta(t_m)\right\|_{L^2}\left\|\nabla\big(2\delta^m_\theta-\delta^{m-1}_\theta\big)\right\|_{L^2}\\
&\le c|\kappa|_{C^{0,1}(\overline{\Omega}\times R;R)}\left(\|\theta(t_m)-\widetilde{\theta}(t_m)\|_{L^2}+\|\widetilde{\theta}(t_m)-\widetilde{\theta}^m\|_{L^2}\right)\\
&\quad\times\|\theta(t_m)\|_{W^{1,\infty}}\left\|\big(2\delta^m_\theta-\delta^{m-1}_\theta\big)\right\|_1\\
&\le c\left(\tau^5\left\|\partial_t^3\theta\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\left\|\widetilde{\delta}^m_\theta\right\|^2_{L^2}\right)+\epsilon\left(\left\|\delta^m_\theta\right\|^2_1+\left\|\delta^{m-1}_\theta\right\|^2_1\right),\\
\left(\Gamma^3_\theta,2\delta^m_\theta-\delta^{m-1}_\theta\right)&\le\|\widetilde{\mathbf{u}}^m\|_{L^\infty}\left\|\nabla\delta^m_\theta\right\|_{L^2}\left\|2\delta^m_\theta-\delta^{m-1}_\theta\right\|_{L^2}\\
&\le c\epsilon^{-1}\left(\left\|\delta^{m-1}_\theta\right\|^2_{L^2}+\left\|\delta^m_\theta\right\|^2_{L^2}\right)+\epsilon\left\|\delta^m_\theta\right\|^2_1,\\
\left(\Gamma^4_\theta,2\delta^m_\theta-\delta^{m-1}_\theta\right)&\le\left\|\widetilde{\delta}^m_{\mathbf{u}}\right\|_{L^2}\|\nabla\theta(t_m)\|_{L^3}\left\|2\delta^m_\theta-\delta^{m-1}_\theta\right\|_{L^6}\\
&\le c\epsilon^{-1}\left\|\widetilde{\delta}^m_{\mathbf{u}}\right\|^2_{L^2}+\epsilon\left(\left\|\delta^{m-1}_\theta\right\|^2_1+\left\|\delta^m_\theta\right\|^2_1\right),\\
\left(\Gamma^5_\theta,2\delta^m_\theta-\delta^{m-1}_\theta\right)&\le\|\mathbf{u}^m-\widetilde{\mathbf{u}}^m\|_{L^2}\|\nabla\theta(t_m)\|_{L^3}\left\|2\delta^m_\theta-\delta^{m-1}_\theta\right\|_{L^6}\\
&\le c\epsilon^{-1}\tau^5\left\|\partial_t^3\mathbf{u}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\epsilon\left(\left\|\delta^m_\theta\right\|^2_1+\left\|\delta^{m-1}_\theta\right\|^2_1\right).
\end{aligned}
$$

Meanwhile, we notice that

$$
D_f\left(\beta^m(\widetilde{\theta}^m),\delta^m_\theta,2\delta^m_{\mathbf{u}}-\delta^{m-1}_{\mathbf{u}}\right)
$$

$$\leq \beta_2 \|\delta_\theta^m\|_{L^2} \|2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}\|_{L^2}$$
$$\leq c\beta_2 \left(\|\delta_\theta^m\|_{L^2}^2 + \|\delta_{\mathbf{u}}^m\|_{L^2}^2 + \|\delta_{\mathbf{u}}^{m-1}\|_{L^2}^2 \right),$$

and then, adding up all above estimates leads to

$$\begin{aligned}
&\left(D_\tau \delta_{\mathbf{u}}^m, 2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}\right) + s\left(D_\tau \delta_{\mathbf{b}}^m, 2\delta_{\mathbf{b}}^m - \delta_{\mathbf{b}}^{m-1}\right) + \left(D_\tau \delta_\theta^m, 2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}\right) \\
&\quad + B_u\left(\nu^m(\widetilde{\theta}^m), \delta_{\mathbf{u}}^m, 2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}\right) + sB_b\left(\mu^m(\widetilde{\theta}^m), \delta_{\mathbf{b}}^m, 2\delta_{\mathbf{b}}^m - \delta_{\mathbf{b}}^{m-1}\right) \\
&\quad + B_\theta\left(\kappa^m(\widetilde{\theta}^m), \delta_\theta^m, 2\delta_\theta^m - \delta_\theta^{m-1}\right) - B_f\left(2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}, \delta_p^m\right) + B_f\left(2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}, \delta_p^m\right) \\
&= D_f\left(\beta^m(\widetilde{\theta}^m), \delta_\theta^m, 2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}\right) + \sum_{i=1}^{9}\left(\Gamma_{\mathbf{u}}^i, 2\delta_{\mathbf{u}}^m - \delta_{\mathbf{u}}^{m-1}\right) \\
&\quad + \sum_{i=1}^{5}\left(\Gamma_{\mathbf{b}}^i, 2\delta_{\mathbf{b}}^m - \delta_{\mathbf{b}}^{m-1}\right) + \sum_{i=1}^{5}\left(\Gamma_\theta^i, 2\delta_\theta^m - \delta_\theta^{m-1}\right) \\
&\leq c\tau^5 + c\epsilon^{-1}\Big(\|\widetilde{\delta}_{\mathbf{u}}^m\|_{L^2}^2 + s\|\widetilde{\delta}_{\mathbf{b}}^m\|_{L^2}^2 + \|\widetilde{\delta}_\theta^m\|_{L^2}^2 + \|\delta_{\mathbf{u}}^m\|_{L^2}^2 + s\|\delta_{\mathbf{b}}^m\|_{L^2}^2 \\
&\qquad\qquad + \|\delta_\theta^m\|_{L^2}^2 + \|\delta_{\mathbf{u}}^{m-1}\|_{L^2}^2 + s\|\delta_{\mathbf{b}}^{m-1}\|_{L^2}^2 + \|\delta_\theta^{m-1}\|_{L^2}^2 \Big) \\
&\quad + \epsilon\left(\|\delta_{\mathbf{u}}^m\|_1^2 + \|\delta_{\mathbf{u}}^{m-1}\|_1^2 + s\|\delta_{\mathbf{b}}^m\|_1^2 + s\|\delta_{\mathbf{b}}^{m-1}\|_1^2 + \|\delta_\theta^m\|_1^2 + \|\delta_\theta^{m-1}\|_1^2 \right).
\end{aligned}$$

By choosing a small ϵ, using

$$a(2a-b) = a^2 + \frac{1}{2}\left(a^2 - b^2 + (a-b)^2\right),$$

summing the above inequality from $m = 3$ to k and applying the telescope formula for D_τ in Lemma 2.1, we derive that

$$\begin{aligned}
&\|\delta_{\mathbf{u}}^k\|_{L^2}^2 + s\|\delta_{\mathbf{b}}^k\|_{L^2}^2 + \|\delta_\theta^k\|_{L^2}^2 + \sum_{m=3}^{k} \tau\left(\nu_1 \|\delta_{\mathbf{u}}^m\|_1^2 + \mu_1 s\|\delta_{\mathbf{b}}^m\|_1^2 + \kappa_1 \|\delta_\theta^m\|_1^2 \right) \\
&\leq c\tau^6 + c\tau \sum_{m=0}^{k} \left(\|\delta_{\mathbf{u}}^m\|_{L^2}^2 + s\|\delta_{\mathbf{b}}^m\|_{L^2}^2 + \|\delta_\theta^m\|_{L^2}^2 \right).
\end{aligned} \tag{4.11}$$

Further, rewrite (4.9) as

$$\begin{aligned}
&-\left(\nu^m(\widetilde{\theta}^m)\Delta\delta_{\mathbf{u}}^m\right) + \nabla\delta_p^m = \Gamma_{\mathbf{u}} - D_\tau\delta_{\mathbf{u}}^m + \beta^m(\widetilde{\theta}^m)\delta_\theta^m + \nabla\nu^m(\widetilde{\theta}^m)\cdot\nabla\delta_{\mathbf{u}}^m =: \mathscr{F}^m, \\
&\nabla\cdot\delta_{\mathbf{u}}^m = 0, \\
&\mu^m(\widetilde{\theta}^m)\mathrm{curl}\,\mathrm{curl}\,\delta_{\mathbf{b}}^m = \Gamma_{\mathbf{b}} - D_\tau\delta_{\mathbf{b}}^m - \nabla\mu^m(\widetilde{\theta}^m)\times\mathrm{curl}\,\delta_{\mathbf{b}}^m =: \mathscr{G}^m, \\
&-\kappa^m(\widetilde{\theta}^m)\Delta\delta_\theta^m = \Gamma_\theta - D_\tau\delta_\theta^m + \nabla\kappa^m(\widetilde{\theta}^m)\cdot\nabla\delta_\theta^m =: \mathscr{H}^m.
\end{aligned} \tag{4.12}$$

and

$$\begin{aligned}
\|\Gamma_{\mathbf{u}}\|_{L^2} \leq c\Big(&\left\|\partial_t\mathbf{u}(t_m) - D_\tau\mathbf{u}(t_m)\right\|_{L^2} + \left\|\nabla\cdot\left((\nu^m(\theta(t_m)) - \nu^m(\widetilde{\theta}^m))\nabla\mathbf{u}(t_m)\right)\right\|_{L^2} \\
&+ \left\|\left(\beta^m(\theta(t_m)) - \beta^m(\widetilde{\theta}^m)\right)\theta(t_m)\right\|_{L^2} + \left\|(\widetilde{\mathbf{u}}^m\cdot\nabla)\delta_{\mathbf{u}}^m\right\|_{L^2} + \left\|\left(\widetilde{\delta}_{\mathbf{u}}^m\cdot\nabla\right)\mathbf{u}(t_m)\right\|_{L^2}
\end{aligned}$$

$$
\begin{aligned}
&+\left\|\left((\mathbf{u}^m-\widetilde{\mathbf{u}}^m)\cdot\nabla\right)\mathbf{u}(t_m)\right\|_{L^2}+s\left\|\widetilde{\mathbf{b}}^m\times\operatorname{curl}\delta_{\mathbf{b}}^m\right\|_{L^2}+s\left\|\widetilde{\delta}_{\mathbf{b}}^m\times\operatorname{curl}\mathbf{b}(t_m)\right\|_{L^2}\\
&+s\left\|\left(\mathbf{b}^m-\widetilde{\mathbf{b}}^m\right)\times\operatorname{curl}\mathbf{b}(t_m)\right\|_{L^2}\Big)\\
\le\,& c\Big(\tau^{5/2}\left\|\partial_t^4\mathbf{u}(t)\right\|_{L^2(t_{m-3},t_m,L^2(\Omega)^d)}\\
&+c|\nu|_{C^{1,1}(\overline{\Omega}\times R;R^+)}\left(\|\theta(t_m)-\widetilde{\theta}(t_m)\|_1+\|\widetilde{\delta}_\theta^m\|_1\right)\|\nabla\mathbf{u}(t_m)\|_{L^\infty}\\
&+c|\beta|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left(\|\theta(t_m)-\widetilde{\theta}(t_m)\|_1+\|\widetilde{\delta}_\theta^m\|_1\right)\|\theta(t_m)\|_{L^\infty}+\|\widetilde{\mathbf{u}}^m\|_{L^3}\|\nabla\delta_{\mathbf{u}}^m\|_1\\
&+\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}\|\nabla\mathbf{u}(t_m)\|_{L^\infty}+\tau^{5/2}\left\|\partial_t^3\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\|\nabla\mathbf{u}(t_m)\|_{L^\infty}\\
&+s\|\widetilde{\mathbf{b}}^m\|_{L^6}\left\|\operatorname{curl}\delta_{\mathbf{b}}^m\right\|_{L^2}^{1/2}\left\|\operatorname{curl}\delta_{\mathbf{b}}^m\right\|_1^{1/2}+s\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}\|\operatorname{curl}\mathbf{b}(t_m)\|_{L^\infty}\\
&+s\tau^{5/2}\left\|\partial_t^3\mathbf{b}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2\|\operatorname{curl}\mathbf{b}(t_m)\|_{L^\infty}\Big).
\end{aligned}
$$

Thus it follows from (4.11), (4.12) and the above bound that for $m\ge 3$ we have

$$
\begin{aligned}
\tau\|\mathscr{F}^m\|_{L^2}\le\,& c\tau\left(\|\Gamma_{\mathbf{u}}\|_{L^2}+\left\|D_\tau\delta_{\mathbf{u}}^m\right\|_{L^2}+\nu_2'\left\|\nabla\delta_{\mathbf{u}}^m\right\|_{L^2}+\beta_2\left\|\delta_\theta^m\right\|_{L^2}\right)\\
\le\,& c\left(\tau\|\Gamma_{\mathbf{u}}\|_{L^2}+\nu_2'\left\|\nabla\delta_{\mathbf{u}}^m\right\|_{L^2}+\beta_2\left\|\delta_\theta^m\right\|_{L^2}\right)\\
&+c\left(\left\|\delta_{\mathbf{u}}^m\right\|+\left\|\delta_{\mathbf{u}}^{m-1}\right\|+\left\|\delta_{\mathbf{u}}^{m-2}\right\|+\left\|\delta_{\mathbf{u}}^{m-3}\right\|\right)\\
\le\,& c\Big(\tau^3+\tau\left(\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}\right)\\
&\quad+\epsilon^{-1}\tau\left(\left\|\delta_{\mathbf{u}}^m\right\|_1+\left\|\delta_{\mathbf{b}}^m\right\|_1\right)+\tau\left\|\delta_\theta^m\right\|_1\Big)\\
&+\epsilon\tau\left(\left\|\delta_{\mathbf{u}}^m\right\|_2+\left\|\delta_{\mathbf{b}}^m\right\|_2\right)+c\tau\sum_{j=0}^m\left(\left\|\delta_{\mathbf{u}}^j\right\|_{L^2}+s\left\|\delta_{\mathbf{b}}^j\right\|_{L^2}+\left\|\delta_\theta^j\right\|_{L^2}\right).
\end{aligned}
$$

From the regularity of the Stokes problem in Assumption 2.3, we have

$$
\begin{aligned}
&\nu_1\tau\left\|\delta_{\mathbf{u}}^m\right\|_2+\tau\left\|\delta_p^m\right\|_1\\
&\le\tau\|\mathscr{F}^m\|_{L^2}\le c\Big(\tau^3+\tau\left(\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}\right)\\
&\qquad+\epsilon^{-1}\tau\left(\left\|\delta_{\mathbf{u}}^m\right\|_1+\left\|\delta_{\mathbf{b}}^m\right\|_1\right)+\tau\left\|\delta_\theta^m\right\|_1\Big)\\
&\quad+\epsilon\tau\left(\left\|\delta_{\mathbf{u}}^m\right\|_2+\left\|\delta_{\mathbf{b}}^m\right\|_2\right)+c\tau\sum_{j=0}^m\left(\left\|\delta_{\mathbf{u}}^j\right\|_{L^2}+s\left\|\delta_{\mathbf{b}}^j\right\|_{L^2}+\left\|\delta_\theta^j\right\|_{L^2}\right).
\end{aligned}\tag{4.13}
$$

Estimating $\Gamma_{\mathbf{b}}$ similarly we get

$$
\begin{aligned}
\|\Gamma_{\mathbf{b}}\|_{L^2}\le\,& c\Big(\|\partial_t\mathbf{b}(t_m)-D_\tau\mathbf{b}(t_m)\|_{L^2}+\|\operatorname{curl}((\mu^m(\theta(t_m))-\mu^m(\widetilde{\theta}^m))\operatorname{curl}\mathbf{b}(t_m))\|_{L^2}\\
&+\left\|\operatorname{curl}\left(\delta_{\mathbf{u}}^m\times\widetilde{\mathbf{b}}^m\right)\right\|_{L^2}+\left\|\operatorname{curl}\left(\mathbf{u}(t_m)\times\widetilde{\delta}_{\mathbf{b}}^m\right)\right\|_{L^2}+\left\|\operatorname{curl}\left(\mathbf{u}(t_m)\times(\mathbf{b}^m-\widetilde{\mathbf{b}}^m)\right)\right\|_{L^2}\Big)\\
\le\,& c\Big(c|\mu|_{C^{1,1}(\overline{\Omega}\times R;R^+)}\left(\left\|\theta(t_m)-\widetilde{\theta}(t_m)\right\|_1+\left\|\widetilde{\delta}_\theta^m\right\|_1\right)\|\operatorname{curl}\mathbf{b}(t_m)\|_{L^\infty}\\
&+\tau^{5/2}\left\|\partial_t^4\mathbf{b}(t)\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\left\|\delta_{\mathbf{u}}^m\right\|_{L^6}\left\|\nabla\widetilde{\mathbf{b}}^m\right\|_{L^2}^{1/2}\left\|\nabla\widetilde{\mathbf{b}}^m\right\|_1^{1/2}\\
&+\left\|\nabla\delta_{\mathbf{u}}^m\right\|_{L^2}^{1/2}\left\|\nabla\delta_{\mathbf{u}}^m\right\|_1^{1/2}\left\|\widetilde{\mathbf{b}}^m\right\|_{L^6}+\|\mathbf{u}(t_m)\|_{L^\infty}\left\|\operatorname{curl}\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}
\end{aligned}
$$

$$
+\tau^{5/2}\left\|\partial_t^3\mathbf{b}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\|\nabla\mathbf{u}(t_m)\|_{L^\infty}+\|\mathrm{curl}\,\mathbf{u}(t_m)\|_{L^\infty}\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}
+\tau^{5/2}\left\|\partial_t^3\mathbf{b}\right\|_{L^2(t_{m-3},t_m;H^1(\Omega)^d)}\|\mathbf{u}(t_m)\|_{L^\infty}\Big).
$$

Then we have

$$
\begin{aligned}
\tau\|\mathscr{G}^m\|_{L^2}&\le c\tau\left(\|\Gamma_{\mathbf{b}}\|_{L^2}+\left\|D_\tau\delta_{\mathbf{b}}^m\right\|_{L^2}+\mu_2'\left\|\mathrm{curl}\,\delta_{\mathbf{b}}^m\right\|_{L^2}\right)\\
&\le c\left(\tau^3+\epsilon^{-1}\tau\left\|\delta_{\mathbf{u}}^m\right\|_1+\tau\left\|\delta_{\mathbf{b}}^m\right\|_1+\tau\left\|\delta_\theta^m\right\|_1\right)\\
&\quad+\tau\left(\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}\right)\\
&\quad+\epsilon\tau\left\|\delta_{\mathbf{u}}^m\right\|_2+c\tau\sum_{j=0}^m\left(\left\|\delta_{\mathbf{u}}^j\right\|_{L^2}+s\left\|\delta_{\mathbf{b}}^j\right\|_{L^2}+\left\|\delta_\theta^j\right\|_{L^2}\right).
\end{aligned}
$$

From the regularity of the Maxwell problem in Assumption 2.3, we obtain

$$
\begin{aligned}
\mu_1\tau\left\|\delta_{\mathbf{b}}^m\right\|_2\le c\Big(&\tau^3+\tau\left(\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_{\mathbf{b}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}\right)\\
&+\epsilon^{-1}\tau\left\|\delta_{\mathbf{u}}^m\right\|_1+\tau\left\|\delta_{\mathbf{b}}^m\right\|_1+\tau\left\|\delta_\theta^m\right\|_1\Big)\\
&+\epsilon\tau\left(\left\|\delta_{\mathbf{u}}^m\right\|_2+\left\|\delta_{\mathbf{b}}^m\right\|_2\right)+c\tau\sum_{j=0}^m\left(\left\|\delta_{\mathbf{u}}^j\right\|_{L^2}+s\left\|\delta_{\mathbf{b}}^j\right\|_{L^2}+\left\|\delta_\theta^j\right\|_{L^2}\right).
\end{aligned}\tag{4.14}
$$

Estimating Γ_θ in the same way, it follows that

$$
\begin{aligned}
\|\Gamma_\theta\|_{L^2}&\le c\Big(\|\partial_t\theta(t_m)-D_\tau\theta(t_m)\|_{L^2}+\|\nabla\cdot(\kappa^m(\theta(t_m))\nabla\theta(t_m))\|_{L^2}\\
&\quad+\left\|\nabla\cdot\left(\kappa^m(\widetilde{\theta}^m)\nabla\theta(t_m)\right)\right\|_{L^2}+\left\|\widetilde{\mathbf{u}}^m\cdot\nabla\delta_\theta^m\right\|_{L^2}\\
&\quad+\left\|\widetilde{\delta}_{\mathbf{u}}^m\cdot\nabla\theta(t_m)\right\|_{L^2}+\left\|(\mathbf{u}^m-\widetilde{\mathbf{u}}^m)\cdot\nabla\theta(t_m)\right\|_{L^2}\Big)\\
&\le c\Big(\tau^{5/2}\left\|\partial_t^4\theta(t)\right\|_{L^2(t_{m-3},t_m;L^2(\Omega))}\\
&\quad+c|\kappa|_{C^{1,1}(\overline{\Omega}\times R;R^+)}\left(\left\|\theta(t_m)-\widetilde{\theta}(t_m)\right\|_1+\left\|\widetilde{\delta}_\theta^m\right\|_1\right)\|\nabla\theta(t_m)\|_{L^\infty}\\
&\quad+\|\widetilde{\mathbf{u}}^m\|_{L^6}\left\|\nabla\delta_\theta^m\right\|_{L^2}^{1/2}\left\|\nabla\delta_\theta^m\right\|_1^{1/2}+\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}\|\nabla\theta(t_m)\|_{L^\infty}\\
&\quad+\tau^{5/2}\left\|\partial_t^3\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\|\nabla\theta(t_m)\|_{L^\infty}\Big).
\end{aligned}
$$

From the regularity of the Poisson problem in Assumption 2.3, we have

$$
\begin{aligned}
\kappa_1\tau\left\|\delta_\theta^m\right\|_2\le\tau\|\mathscr{H}^m\|_{L^2}&\le c\tau\left(\|\Gamma_\theta\|_{L^2}+\left\|D_\tau\delta_\theta^m\right\|_{L^2}+\kappa_2'\left\|\nabla\delta_\theta^m\right\|_{L^2}\right)\\
&\le c\Big(\tau^3+\tau\left(\left\|\widetilde{\delta}_{\mathbf{u}}^m\right\|_{L^2}+\left\|\widetilde{\delta}_\theta^m\right\|_{L^2}\right)+\epsilon^{-1}\tau\left\|\delta_\theta^m\right\|_1+\tau\left\|\delta_{\mathbf{u}}^m\right\|_1\Big)\\
&\quad+\epsilon\tau\left\|\delta_\theta^m\right\|_2+c\tau\sum_{j=0}^m\left(\left\|\delta_{\mathbf{u}}^j\right\|_{L^2}+s\left\|\delta_{\mathbf{b}}^j\right\|_{L^2}+\left\|\delta_\theta^j\right\|_{L^2}\right).
\end{aligned}\tag{4.15}
$$

Next, adding up (4.13)-(4.15) and (4.11), we obtain

$$
\left\|\delta_{\mathbf{u}}^m\right\|_{L^2}^2+s\left\|\delta_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\delta_\theta^m\right\|_{L^2}^2+\sum_{i=3}^m\tau\left(\nu_1\left\|\delta_{\mathbf{u}}^i\right\|_1^2+\mu_1 s\left\|\delta_{\mathbf{b}}^i\right\|_1^2+\kappa_1\left\|\delta_\theta^i\right\|_1^2\right)
$$

$$
\begin{aligned}
&+\tau^2\Big(\big\|\delta_{\mathbf{u}}^m\big\|_2^2+s\big\|\delta_{\mathbf{b}}^m\big\|_2^2+\big\|\delta_{\theta}^m\big\|_2^2\Big)\\
\le{}& c\tau^6+c\tau\sum_{j=0}^{m}\Big(\big\|\delta_{\mathbf{u}}^j\big\|_{L^2}^2+s\big\|\delta_{\mathbf{b}}^j\big\|_{L^2}^2+\big\|\delta_{\theta}^j\big\|_{L^2}^2\Big).
\end{aligned}
$$

Using the discrete Grönwall inequality, when $\tau\le\tau_1$, for a certain small positive number τ_1, we have

$$
\begin{aligned}
&\big\|\delta_{\mathbf{u}}^m\big\|_{L^2}^2+s\big\|\delta_{\mathbf{b}}^m\big\|_{L^2}^2+\big\|\delta_{\theta}^m\big\|_{L^2}^2+\sum_{i=3}^{m}\tau\Big(\nu_1\big\|\delta_{\mathbf{u}}^i\big\|_1^2+\mu_1 s\big\|\delta_{\mathbf{b}}^i\big\|_1^2+\kappa_1\big\|\delta_{\theta}^i\big\|_1^2\Big)\\
&\quad+\tau^2\Big(\big\|\delta_{\mathbf{u}}^m\big\|_2^2+s\big\|\delta_{\mathbf{b}}^m\big\|_2^2+\big\|\delta_{\theta}^m\big\|_2^2\Big)\\
\le{}& \exp\Big(\frac{T}{1-C\tau_1}\Big)c\tau^6=:C_1\tau^6.
\end{aligned}\tag{4.16}
$$

Thus, we prove that (4.7) is valid for $m=k$ by (4.16) and close the mathematical induction.

On the other hand, it follows from (4.7) that

$$
\begin{aligned}
&\big\|D_\tau\mathbf{u}^m\big\|_2\le\big\|D_\tau\delta_{\mathbf{u}}^m\big\|_2+\|D_\tau\mathbf{u}(t_m)\|_2\le C, &&\|\mathbf{u}^m\|_2\le\big\|\delta_{\mathbf{u}}^m\big\|_2+\|\mathbf{u}(t_m)\|_2\le C,\\
&\big\|D_\tau\mathbf{b}^m\big\|_2\le\big\|D_\tau\delta_{\mathbf{b}}^m\big\|_2+\|D_\tau\mathbf{b}(t_m)\|_2\le C, &&\|\mathbf{b}^m\|_2\le\big\|\delta_{\mathbf{b}}^m\big\|_2+\|\mathbf{b}(t_m)\|_2\le C,\\
&\big\|D_\tau\theta^m\big\|_2\le\big\|D_\tau\delta_{\theta}^m\big\|_2+\|D_\tau\theta(t_m)\|_2\le C, &&\|\theta^m\|_2\le\big\|\delta_{\theta}^m\big\|_2+\|\theta(t_m)\|_2\le C
\end{aligned}
$$

for $1\le m\le N$. To prove (4.8), we rewrite (4.6) as

$$
\begin{aligned}
&-\nu^m(\widetilde{\theta}^m)\Delta\mathbf{u}^m+\nabla p^m\\
={}&\mathbf{f}^m-D_\tau\mathbf{u}^m+\nabla\big(\nu^m(\widetilde{\theta}^m)\big)\cdot\nabla\mathbf{u}^m-(\widetilde{\mathbf{u}}^m\cdot\nabla)\mathbf{u}^m\\
&\quad-s\widetilde{\mathbf{b}}^m\times\operatorname{curl}\mathbf{b}^m+\beta^m(\widetilde{\theta}^n)\theta^m=:\mathscr{F}',\\
&\nabla\cdot\mathbf{u}^m=0,\\
&\mu^n(\widetilde{\theta}^m)\operatorname{curlcurl}\mathbf{b}^m=-D_\tau\mathbf{b}^m-\nabla\mu^m(\widetilde{\theta}^m)\times\operatorname{curl}\mathbf{b}^m+\operatorname{curl}\big(\mathbf{u}^m\times\widetilde{\mathbf{b}}^m\big)=:\mathscr{G}',\\
&-\kappa^m(\widetilde{\theta}^m)\Delta\theta^m=\psi^m-D_\tau\theta^m+\nabla\big(\kappa^m(\widetilde{\theta}^m)\big)\cdot\nabla\theta^n-\widetilde{\mathbf{u}}^m\cdot\nabla\theta^m.
\end{aligned}
$$

With (2.2)-(2.3) and (2.11)-(2.12), one finds that

$$
\begin{aligned}
&\nu_1\|\mathbf{u}^m\|_{W^{2,4}}+\|p^m\|_{W^{1,4}}\\
\le{}& c\|\mathscr{F}'\|_{W^{0,4}}\le c\Big(\|\mathbf{f}^m\|_{L^4}+\big\|D_\tau\mathbf{u}^m\big\|_{L^4}+\big\|\nabla\nu^m(\widetilde{\theta}^m)\big\|_{L^4}\|\nabla\mathbf{u}^m\|_{L^\infty}\\
&\qquad+\|\widetilde{\mathbf{u}}^m\|_{L^\infty}\|\nabla\mathbf{u}^m\|_{L^4}+s\|\widetilde{\mathbf{b}}^m\|_{L^\infty}\|\operatorname{curl}\mathbf{b}^m\|_{L^4}\\
&\qquad+\|\beta^m(\widetilde{\theta}^n)\|_{L^4}\|\theta^m\|_{L^\infty}\Big)\\
\le{}& c+c\epsilon^{-1}\|\nabla\nu^m(\widetilde{\theta}^m)\|_{L^4}^2+\epsilon\|\nabla\mathbf{u}^m\|_{L^\infty}^2,\\
&\mu_1\|\mathbf{b}^m\|_{W^{2,4}}\le c\Big(\big\|D_\tau\mathbf{b}^m\big\|_{L^4}+\big\|\nabla\mu^m(\widetilde{\theta}^m)\big\|_{L^4}\|\operatorname{curl}\mathbf{b}^m\|_{L^\infty}\\
&\qquad+\|\mathbf{u}^m\|_{L^\infty}\|\nabla\widetilde{\mathbf{b}}^m\|_{L^4}+\|\widetilde{\mathbf{b}}^m\|_{L^\infty}\|\nabla\mathbf{u}^m\|_{L^4}\Big)\\
&\qquad\le c+c\epsilon^{-1}\|\nabla\mu^m(\widetilde{\theta}^m)\|_{L^4}^2+\epsilon\|\operatorname{curl}\mathbf{b}^m\|_{L^\infty}^2,
\end{aligned}
$$

and

$$\begin{aligned}\kappa_1\|\theta^m\|_{W^{2,4}} \le\, & c\Big(\|D_\tau\theta^m\|_{L^4}+\|\nabla\kappa^m(\widetilde{\theta}^m)\|_{L^4}\|\nabla\theta^m\|_{L^\infty}\\ &\quad+\|\widetilde{\mathbf{u}}^m\|_{L^\infty}\|\nabla\theta^m\|_{L^4}+\|\psi^m\|_{L^4}\Big)\\ \le\, & c+c\epsilon^{-1}\|\nabla\kappa^m(\widetilde{\theta}^m)\|_{L^4}^2+\epsilon\|\nabla\theta^m\|_{L^\infty}^2.\end{aligned}$$

Finally, by choosing a small ϵ, adding up the above inequalities, using the embedding inequality $\|\nabla\mathbf{v}\|_{L^\infty}\le c\|\mathbf{v}\|_{W^{2,4}}$ in Lemma 2.3, the inequality (4.8) is proved. The proof is complete. □

4.2. Primary spatial error estimates

In this subsection, we mainly present the following theorem which shows the spatial error estimates of the scheme (3.1) with the MINI element [2] for the velocity and pressure, linear element for the magnetic and temperature.

Theorem 4.2. *Assume that Assumptions* 2.1-2.4 *hold and* $(\mathbf{u}_h^m,\mathbf{b}_h^m,p_h^m,\theta_h^m)$ *is the solution of* (3.1)-(3.3). *Then for* $1\le k\le N$*, there exist two positive constants* h_1 *and* τ_2 *such that when* $\tau\le\tau_2$ *and* $h\le h_1$ *there holds*

$$\big\|\varepsilon_{\mathbf{u}}^k\big\|_{L^2}^2+s\big\|\varepsilon_{\mathbf{b}}^k\big\|_{L^2}^2+\big\|\varepsilon_\theta^k\big\|_{L^2}^2+\tau\sum_{j=0}^{k}\Big(\nu_1\big\|\varepsilon_{\mathbf{u}}^j\big\|_1^2+s\mu_1\big\|\varepsilon_{\mathbf{b}}^j\big\|_1^2+\kappa_1\big\|\varepsilon_\theta^j\big\|_1^2\Big)\le C_2h^4, \tag{4.17}$$

where $C_2>0$ *is independent of* h *and* τ.

Proof. We use the mathematical induction. It follows from (4.5) that the error estimate (4.17) holds for $k=0,1,2$. Now, we suppose that (4.17) is valid for $m\le k-1$ with $3\le k\le N$. Under this assumption, one has

$$\big\|\widetilde{\varepsilon}_{\mathbf{u}}^{m+1}\big\|_{L^2}^2+s\big\|\widetilde{\varepsilon}_{\mathbf{b}}^{m+1}\big\|_{L^2}^2+\big\|\widetilde{\varepsilon}_\theta^{m+1}\big\|_{L^2}^2+\tau\sum_{j=1}^{m+1}\Big(\nu_1\big\|\widetilde{\varepsilon}_{\mathbf{u}}^j\big\|_1^2+s\mu_1\big\|\widetilde{\varepsilon}_{\mathbf{b}}^j\big\|_1^2+\kappa_1\big\|\widetilde{\varepsilon}_\theta^j\big\|_1^2\Big)\le CC_2h^4,$$

where $C>0$ is independent of h,τ and C_2. Furthermore, from the inverse inequality in Lemma 3.1, for $m\le k-1$, we have

$$\begin{aligned}\big\|\widetilde{\varepsilon}_{\mathbf{u}}^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\varepsilon}_{\mathbf{u}}^{m+1}\big\|_{L^3}&\le C\left(h^{-d/2}+h^{-1-d/6}\right)\big\|\widetilde{\varepsilon}_{\mathbf{u}}^{m+1}\big\|_{L^2}\le CC_2h^{1/2}\le C,\\ \big\|\widetilde{\varepsilon}_{\mathbf{b}}^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\varepsilon}_{\mathbf{b}}^{m+1}\big\|_{L^3}&\le C\left(h^{-d/2}+h^{-1-d/6}\right)\big\|\widetilde{\varepsilon}_{\mathbf{b}}^{m+1}\big\|_{L^2}\le CC_2h^{1/2}\le C,\\ \big\|\widetilde{\varepsilon}_\theta^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\varepsilon}_\theta^{m+1}\big\|_{L^3}&\le C\left(h^{-d/2}+h^{-1-d/6}\right)\big\|\widetilde{\varepsilon}_\theta^{m+1}\big\|_{L^2}\le CC_2h^{1/2}\le C\end{aligned}$$

for $C_2h^{1/2}\le 1$, which result in

$$\begin{aligned}\big\|\widetilde{\mathbf{u}}_h^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\mathbf{u}}_h^{m+1}\big\|_{L^3}\le\, & \big\|R_h\widetilde{\mathbf{u}}^{m+1}\big\|_{L^\infty}+\big\|\nabla R_h\widetilde{\mathbf{u}}^{m+1}\big\|_{L^3}+\big\|\widetilde{\varepsilon}_{\mathbf{u}}^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\varepsilon}_{\mathbf{u}}^{m+1}\big\|_{L^3}\\ \le\, & C|\ln h|\|\mathbf{u}\|_{L^\infty}+C\le C.\end{aligned}$$

Similarly we get

$$\big\|\widetilde{\mathbf{b}}_h^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\mathbf{b}}_h^{m+1}\big\|_{L^3}\leq\big\|F_h\widetilde{\mathbf{b}}^{m+1}\big\|_{L^\infty}+\big\|\nabla F_h\widetilde{\mathbf{b}}^{m+1}\big\|_{L^3}+\big\|\widetilde{\varepsilon}_{\mathbf{b}}^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\varepsilon}_{\mathbf{b}}^{m+1}\big\|_{L^3}\leq C,$$
$$\big\|\widetilde{\theta}_h^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\theta}_h^{m+1}\big\|_{L^3}\leq\big\|T_h\widetilde{\theta}^{m+1}\big\|_{L^\infty}+\big\|\nabla T_h\widetilde{\theta}^{m+1}\big\|_{L^3}+\big\|\widetilde{\varepsilon}_{\theta}^{m+1}\big\|_{L^\infty}+\big\|\nabla\widetilde{\varepsilon}_{\theta}^{m+1}\big\|_{L^3}\leq C.$$

To finish the mathematical induction, we have to show the validity of (4.17) for $m=k$. Test (4.6) by $(\mathbf{v}_h,q_h,\mathbf{w}_h,\phi_h)$ respectively. And then subtract the ensuing equations from (3.1). Adding these equations and applying the definition of Stokes, Maxwell, and Ritz projections, we obtain for $m\leq k$ the following error equation:

$$\begin{aligned}
&\big(D_\tau\varepsilon_{\mathbf{u}}^m,\mathbf{v}_h\big)+s\big(D_\tau\varepsilon_{\mathbf{b}}^m,\mathbf{w}_h\big)+\big(D_\tau\varepsilon_\theta^m,\phi_h\big)\\
&\quad+B_u\big(\nu^m(\widetilde{\theta}_h^m),\varepsilon_{\mathbf{u}}^m,\mathbf{v}_h\big)+sB_b\big(\mu^m(\widetilde{\theta}_h^m),\varepsilon_{\mathbf{b}}^m,\mathbf{w}_h\big)\\
&\quad+B_\theta\big(\kappa^m(\widetilde{\theta}_h^m),\varepsilon_\theta^m,\phi_h\big)+B_f\big(\mathbf{v}_h,\varepsilon_p^m\big)-B_f\big(2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1},q_h\big)\\
=&\,B_u\big(\nu^m(\widetilde{\theta}^m)-\nu^m(\widetilde{\theta}_h^m),\mathbf{u}^m,\mathbf{v}_h\big)+sB_b\big(\mu^m(\widetilde{\theta}^m)-\mu^m(\widetilde{\theta}_h^m),\mathbf{b}^m,\mathbf{w}_h\big)\\
&\quad+B_\theta\big(\kappa^m(\widetilde{\theta}^m)-\kappa^m(\widetilde{\theta}_h^m),\theta^m,\phi_h\big)\\
&\quad-D_f\big(\beta^m(\widetilde{\theta}_h^m),\theta_h^m-\theta^m,\mathbf{v}_h\big)+D_f\big(\beta^m(\widetilde{\theta}_h^m)-\beta^m(\widetilde{\theta}^m),\theta^m,\mathbf{v}_h\big)\\
&\quad+T_u\big(\widetilde{\mathbf{u}}^m-\widetilde{\mathbf{u}}_h^m,\mathbf{u}^m,\mathbf{v}_h\big)+T_u\big(\widetilde{\mathbf{u}}_h^m,\mathbf{u}^m-\mathbf{u}_h^m,\mathbf{v}_h\big)+sT_{b1}\big(\widetilde{\mathbf{b}}^m-\widetilde{\mathbf{b}}_h^m,\mathbf{b}^m,\mathbf{v}_h\big)\\
&\quad+sT_{b1}\big(\widetilde{\mathbf{b}}_h^m,\mathbf{b}^m-\mathbf{b}_h^m,\mathbf{v}_h\big)-sT_{b2}\big(\mathbf{u}^m,\widetilde{\mathbf{b}}^m-\widetilde{\mathbf{b}}_h^m,\mathbf{w}_h\big)-sT_{b2}\big(\mathbf{u}^m-\mathbf{u}_h^m,\widetilde{\mathbf{b}}_h^m,\mathbf{w}_h\big)\\
&\quad+T_\theta\big(\widetilde{\mathbf{u}}^m-\widetilde{\mathbf{u}}_h^m,\theta^m,\phi_h\big)+T_\theta\big(\widetilde{\mathbf{u}}_h^m,\theta^m-\theta_h^m,\phi_h\big)\\
&\quad+\Big[\big(D_\tau(\mathbf{u}^m-R_h\mathbf{u}^m),\mathbf{v}_h\big)+s\big(D_\tau(\mathbf{b}^m-F_h\mathbf{b}^m),\mathbf{w}_h\big)+\big(D_\tau(\theta^m-T_h\theta^m),\phi_h\big)\Big]\\
=:&\sum_{i=1}^{13}\Lambda_i+R'.
\end{aligned}$$

Set

$$(\mathbf{v}_h,q_h,\mathbf{w}_h,\phi_h)=\Big(2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1},2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1},\varepsilon_p^m,2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\Big).$$

By using the Hölder inequality, Young inequality, (4.1)-(4.3), for $m\leq k$, the right-hand side of the above equation can be estimated as follows:

$$\begin{aligned}
\Lambda_1\leq&\big\|\nu^m(\widetilde{\theta}^m)-\nu^m(\widetilde{\theta}_h^m)\big\|_{L^2}\big\|\nabla\mathbf{u}^m\big\|_{L^\infty}\big\|\nabla\big(2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\big)\big\|_{L^2}\\
\leq&\,c|\nu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\Big(\big\|\widetilde{\theta}^m-T_h\widetilde{\theta}^m\big\|_{L^2}+\big\|T_h\widetilde{\theta}^m-\widetilde{\theta}_h^m\big\|_{L^2}\Big)\big\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\big\|_1\\
\leq&\,c\epsilon^{-1}\Big(h^4\big\|\widetilde{\theta}^m\big\|_2^2+\big\|\widetilde{\varepsilon}_\theta^m\big\|_{L^2}^2\Big)+\epsilon\Big(\big\|\varepsilon_{\mathbf{u}}^m\big\|_1^2+\big\|\varepsilon_{\mathbf{u}}^{m-1}\big\|_1^2\Big),\\
\Lambda_2\leq&\,s\big\|\mu^m(\widetilde{\theta}^m)-\mu^m(\widetilde{\theta}_h^m)\big\|_{L^2}\|\mathrm{curl}\,\mathbf{b}^m\|_{L^\infty}\big\|\mathrm{curl}\big(2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\big)\big\|_{L^2}\\
\leq&\,cs|\mu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\Big(\big\|\widetilde{\theta}^m-T_h\widetilde{\theta}^m\big\|_{L^2}+\big\|T_h\widetilde{\theta}^m-\widetilde{\theta}_h^m\big\|_{L^2}\Big)\big\|2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\big\|_1\\
\leq&\,cs\epsilon^{-1}\Big(h^4\big\|\widetilde{\theta}^m\big\|_2^2+\big\|\widetilde{\varepsilon}_\theta^m\big\|_{L^2}^2\Big)+\epsilon s\Big(\big\|\varepsilon_{\mathbf{b}}^m\big\|_1^2+\big\|\varepsilon_{\mathbf{b}}^{m-1}\big\|_1^2\Big),\\
\Lambda_3\leq&\big\|\kappa^m(\widetilde{\theta}^m)-\kappa^m(\widetilde{\theta}_h^m)\big\|_{L^2}\|\nabla\theta^m\|_{L^\infty}\big\|\nabla\big(2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\big)\big\|_{L^2}\\
\leq&\,c|\kappa|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\Big(\big\|\widetilde{\theta}^m-T_h\widetilde{\theta}^m\big\|_{L^2}+\big\|T_h\widetilde{\theta}^m-\widetilde{\theta}_h^m\big\|_{L^2}\Big)\big\|2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\big\|_1\\
\leq&\,c\epsilon^{-1}\Big(h^4\big\|\widetilde{\theta}^m\big\|_2^2+\big\|\widetilde{\varepsilon}_\theta^m\big\|_{L^2}^2\Big)+\epsilon\Big(\big\|\varepsilon_\theta^m\big\|_1^2+\big\|\varepsilon_\theta^{m-1}\big\|_1^2\Big),
\end{aligned}$$

where we have used the embedding inequality $\|\nabla \mathbf{v}\|_{L^\infty} \leq c\|\mathbf{v}\|_{W^{2,4}}$ in Lemma 2.3 and noticed that Theorem 4.1 yields

$$\max_{0\leq j\leq k}\|\mathbf{u}^j\|_{W^{2,4}}\leq C,\quad \max_{0\leq j\leq k}\|\mathbf{b}^j\|_{W^{2,4}}\leq C,\quad \max_{0\leq j\leq k}\|\theta^j\|_{W^{2,4}}\leq C.$$

After that, we have

$$\begin{aligned}
\Lambda_4 &\leq \beta_2\left\|\theta_h^m-\theta^m\right\|_{L^2}\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}\leq c\epsilon^{-1}\left(h^4+\left\|\varepsilon_\theta^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2\right),\\
\Lambda_5 &\leq \left\|\beta^m(\widetilde{\theta}_h^m)-\beta^m(\widetilde{\theta}^m)\right\|_{L^2}\left\|\theta^m\right\|_{L^\infty}\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\leq c|\beta|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left(\left\|\widetilde{\theta}^m-T_h\widetilde{\theta}^m\right\|_{L^2}+\left\|T_h\widetilde{\theta}^m-\widetilde{\theta}_h^m\right\|_{L^2}\right)\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\leq c\epsilon^{-1}\left(h^4\left\|\widetilde{\theta}^m\right\|_2^2+\left\|\widetilde{\varepsilon}_\theta^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2\right).
\end{aligned}$$

Using anti-symmetric structure of T_u and T_θ yields

$$\begin{aligned}
\Lambda_6 &\leq \left\|\widetilde{\mathbf{u}}^m-\widetilde{\mathbf{u}}_h^m\right\|_{L^2}\left(\|\nabla\mathbf{u}^m\|_{L^3}+\|\mathbf{u}^m\|_{L^\infty}\right)\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_1\\
&\leq c\left(\|\widetilde{\rho}_{\mathbf{u}}^m\|_{L^2}+\|\widetilde{\varepsilon}_{\mathbf{u}}^m\|_{L^2}\right)\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_1\\
&\leq c\epsilon^{-1}\left(h^4\left\|\widetilde{\mathbf{u}}^m\right\|_2^2+\left\|\widetilde{\varepsilon}_{\mathbf{u}}^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_1^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\Lambda_7 &= T_u\left(\widetilde{\mathbf{u}}_h^m,-\rho_{\mathbf{u}}^m,2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right)+T_u\left(\widetilde{\mathbf{u}}_h^m,-\varepsilon_{\mathbf{u}}^m,-\varepsilon_{\mathbf{u}}^{m-1}\right)\\
&\leq \left\|\widetilde{\mathbf{u}}_h^m\right\|_2\left\|\rho_{\mathbf{u}}^m\right\|_{L^2}\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_1+\left\|\widetilde{\mathbf{u}}_h^m\right\|_2\left\|\varepsilon_{\mathbf{u}}^m\right\|_1\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\leq c\epsilon^{-1}\left(h^4\left\|\mathbf{u}^m\right\|_2^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_1^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\Lambda_8 &\leq s\left\|\widetilde{\mathbf{b}}^m-\widetilde{\mathbf{b}}_h^m\right\|_{L^2}\|\mathrm{curl}\,\mathbf{b}^m\|_{L^3}\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^6}\\
&\leq cs\left(\left\|\widetilde{\rho}_{\mathbf{b}}^m\right\|_{L^2}+\left\|\widetilde{\varepsilon}_{\mathbf{b}}^m\right\|_{L^2}\right)\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_1\\
&\leq cs\epsilon^{-1}\left(h^4\left\|\widetilde{\mathbf{b}}^m\right\|_2^2+\left\|\widetilde{\varepsilon}_{\mathbf{b}}^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_1^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_1^2\right),\\
\Lambda_{10} &\leq s\left\|\widetilde{\mathbf{b}}^m-\widetilde{\mathbf{b}}_h^m\right\|_{L^2}\|\mathbf{u}^m\|_{L^\infty}\left\|\mathrm{curl}\left(2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\right)\right\|_{L^2}\\
&\leq cs\left(\left\|\widetilde{\rho}_{\mathbf{b}}^m\right\|_{L^2}+\left\|\widetilde{\varepsilon}_{\mathbf{b}}^m\right\|_{L^2}\right)\left\|\mathrm{curl}\left(2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\right)\right\|_{L^2}\\
&\leq cs\epsilon^{-1}\left(h^4\left\|\widetilde{\mathbf{b}}^m\right\|_2^2+\left\|\widetilde{\varepsilon}_{\mathbf{b}}^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_{\mathbf{b}}^m\right\|_1^2+\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_1^2\right),\\
\Lambda_{12} &\leq \left\|\widetilde{\mathbf{u}}^m-\widetilde{\mathbf{u}}_h^m\right\|_{L^2}\left(\|\nabla\theta^m\|_{L^3}+\|\theta^m\|_{L^\infty}\right)\left\|2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\right\|_1\\
&\leq c\left(\left\|\widetilde{\rho}_{\mathbf{u}}^m\right\|_{L^2}+\left\|\widetilde{\varepsilon}_{\mathbf{u}}^m\right\|_{L^2}\right)\left\|2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\right\|_1\\
&\leq c\epsilon^{-1}\left(h^4\left\|\widetilde{\mathbf{u}}^m\right\|_2^2+\left\|\widetilde{\varepsilon}_{\mathbf{u}}^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_\theta^m\right\|_1^2+\left\|\varepsilon_\theta^{m-1}\right\|_1^2\right),\\
\Lambda_{13} &= T_\theta\left(\widetilde{\mathbf{u}}_h^m,-\rho_\theta^m-\varepsilon_\theta^m,2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\right)\\
&= T_\theta\left(\widetilde{\mathbf{u}}_h^m,-\rho_\theta^m,2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\right)+T_\theta\left(\widetilde{\mathbf{u}}_h^m,-\varepsilon_\theta^m,-\varepsilon_\theta^{m-1}\right)\\
&\leq \left\|\widetilde{\mathbf{u}}_h^m\right\|_2\left\|\rho_\theta^m\right\|_{L^2}\left\|2\varepsilon_\theta^m-\varepsilon_\theta^{m-1}\right\|_1+\left\|\widetilde{\mathbf{u}}_h^m\right\|_2\left\|\varepsilon_\theta^m\right\|_1\left\|\varepsilon_\theta^{m-1}\right\|_{L^2}\\
&\leq c\epsilon^{-1}\left(h^4\left\|\theta^m\right\|_2^2+\left\|\varepsilon_\theta^{m-1}\right\|_{L^2}^2\right)+\epsilon\left(\left\|\varepsilon_\theta^m\right\|_1^2+\left\|\varepsilon_\theta^{m-1}\right\|_1^2\right).
\end{aligned}$$

In addition, (2.6) and integration by parts give

$$
\begin{aligned}
\Lambda_9+\Lambda_{11} &= -sT_{b1}\left(\widetilde{\mathbf{b}}_h^m,\rho_{\mathbf{b}}^m+\varepsilon_{\mathbf{b}}^m,2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right)\\
&\quad +sT_{b2}\left(\rho_{\mathbf{u}}^m+\varepsilon_{\mathbf{u}}^m,\widetilde{\mathbf{b}}_h^m,2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\right)\\
&= -sT_{b1}\left(\widetilde{\mathbf{b}}_h^m,\rho_{\mathbf{b}}^m,2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right)+sT_{b1}\left(\widetilde{\mathbf{b}}_h^m,\varepsilon_{\mathbf{b}}^m,\varepsilon_{\mathbf{u}}^{m-1}\right)\\
&\quad +sT_{b2}\left(\rho_{\mathbf{u}}^m,\widetilde{\mathbf{b}}_h^m,2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\right)-sT_{b2}\left(\varepsilon_{\mathbf{u}}^m,\widetilde{\mathbf{b}}_h^m,\varepsilon_{\mathbf{b}}^{m-1}\right)\\
&\le cs\left(\left\|\widetilde{\mathbf{b}}_h^m\right\|_{L^\infty}+\left\|\nabla\widetilde{\mathbf{b}}_h^m\right\|_{L^3}\right)\left\|\rho_{\mathbf{b}}^m\right\|_{L^2}\left\|2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right\|_1\\
&\quad +cs\left\|\widetilde{\mathbf{b}}_h^m\right\|_{L^\infty}\left\|\operatorname{curl}\varepsilon_{\mathbf{b}}^m\right\|_{L^2}\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_{L^2}\\
&\quad +cs\left\|\rho_{\mathbf{u}}^m\right\|_{L^2}\left\|\widetilde{\mathbf{b}}_h^m\right\|_{L^\infty}\left\|2\varepsilon_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\right\|_1\\
&\quad +cs\left\|\varepsilon_{\mathbf{u}}^m\right\|_1\left(\left\|\widetilde{\mathbf{b}}_h^m\right\|_{L^\infty}+\left\|\nabla\widetilde{\mathbf{b}}_h^m\right\|_{L^3}\right)\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\le \epsilon\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_1^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_1^2+s\left\|\varepsilon_{\mathbf{b}}^m\right\|_1^2+s\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_1^2\right)\\
&\quad +c\epsilon^{-1}\left(\left\|\rho_{\mathbf{u}}^m\right\|_{L^2}^2+s\left\|\rho_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_{L^2}^2\right).
\end{aligned}
$$

Besides, the term R' is bounded — viz.

$$
\begin{aligned}
R' &\le \epsilon^{-1}\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^m\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_{L^2}^2+\left\|\varepsilon_{\theta}^m\right\|_{L^2}^2+\left\|\varepsilon_{\theta}^{m-1}\right\|_{L^2}^2\right)\\
&\quad +c\epsilon\left(\|D_\tau(\mathbf{u}^m-R_h\mathbf{u}^m)\|_{L^2}^2+\|sD_\tau(\mathbf{b}^m-F_h\mathbf{b}^m)\|_{L^2}^2+\|D_\tau(\theta^m-T_h\theta^m)\|_{L^2}^2\right)\\
&\le c\epsilon^{-1}\left(\left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^m\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_{L^2}^2+\left\|\varepsilon_{\theta}^m\right\|_{L^2}^2+\left\|\varepsilon_{\theta}^{m-1}\right\|_{L^2}^2\right)\\
&\quad +\epsilon h^4\left(\|D_\tau\mathbf{u}^m\|_2^2+s\|D_\tau\mathbf{b}^m\|_2^2+\|D_\tau\theta^m\|_2^2\right).
\end{aligned}
$$

Combining the estimates above yields

$$
\begin{aligned}
&\left(D_\tau\varepsilon_{\mathbf{u}}^m,2\varepsilon_{\mathbf{u}}^m-\varepsilon_{\mathbf{u}}^{m-1}\right)+c_1\nu_1\left\|\varepsilon_{\mathbf{u}}^m\right\|_1^2+\frac{\nu_1}{2}\left(\left\|\nabla\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2-\left\|\nabla\varepsilon_{\mathbf{u}}^{m-1}\right\|_{L^2}^2\right)\\
&\quad +s\left(D_\tau\varepsilon_{\mathbf{b}}^m,2\theta_{\mathbf{b}}^m-\varepsilon_{\mathbf{b}}^{m-1}\right)+c_2\mu_1 s\left\|\varepsilon_{\mathbf{b}}^m\right\|_1^2\\
&\quad +\frac{s\mu_1}{2}\left(\left\|\operatorname{curl}\varepsilon_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\operatorname{div}\varepsilon_{\mathbf{b}}^m\right\|_{L^2}^2-\left\|\operatorname{curl}\varepsilon_{\mathbf{b}}^{m-1}\right\|_{L^2}^2-\left\|\operatorname{div}\varepsilon_{\mathbf{b}}^{m-1}\right\|_{L^2}^2\right)\\
&\quad +\left(D_\tau\varepsilon_{\theta}^m,2\varepsilon_{\theta}^m-\varepsilon_{\theta}^{m-1}\right)+c_1\kappa_1\left\|\varepsilon_{\theta}^m\right\|_1^2+\frac{\kappa_1}{2}\left(\left\|\nabla\varepsilon_{\theta}^m\right\|_{L^2}^2-\left\|\nabla\varepsilon_{\theta}^{m-1}\right\|_{L^2}^2\right)\\
&\le ch^4+c\epsilon^{-1}\left(\left\|\widetilde{\varepsilon}_{\mathbf{u}}^m\right\|_{L^2}^2+s\left\|\widetilde{\varepsilon}_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\widetilde{\varepsilon}_{\theta}^m\right\|_{L^2}^2+\left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\theta}^m\right\|_{L^2}^2\right)\\
&\quad +\epsilon\left(\left\|\varepsilon_{\mathbf{u}}^{m-1}\right\|_1^2+s\left\|\varepsilon_{\mathbf{b}}^{m-1}\right\|_1^2+\left\|\varepsilon_{\theta}^{m-1}\right\|_1^2+\left\|\varepsilon_{\mathbf{u}}^m\right\|_1^2+s\left\|\varepsilon_{\mathbf{b}}^m\right\|_1^2+\left\|\varepsilon_{\theta}^m\right\|_1^2\right).
\end{aligned}
$$

Taking a small ϵ, summing the above inequalities in m from 3 to k, using telescope formula (2.10), and the discrete Grönwall inequality, we get that there is a $\tau_2>0$ such that if $\tau\le\tau_2$, then

$$
\left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^2}^2+s\left\|\varepsilon_{\mathbf{b}}^m\right\|_{L^2}^2+\left\|\varepsilon_{\theta}^m\right\|_{L^2}^2+\tau\sum_{i=3}^m\left(\nu_1\left\|\varepsilon_{\mathbf{u}}^i\right\|_1^2+\nu_2 s\left\|\varepsilon_{\mathbf{b}}^i\right\|_1^2+\kappa_1\left\|\varepsilon_{\theta}^i\right\|_1^2\right)\le\exp\left(\frac{TC}{1-C\tau_1}\right)Ch^4.
$$

Thus the estimate (4.17) holds for $m = k$, if we choose a C_2 such that

$$C_2 \geq \exp\left(\frac{TC}{1-C\tau_1}\right)C.$$

This finishes the proof. $\square$

Combining Theorems 4.1, 4.2, Lemma 4.1 and the projection error estimates (4.1)-(4.3) with the chosen finite element leads to the following corollary.

Corollary 4.1. *Under the assumptions of Theorems* 4.1 *and* 4.2*, for* $0 \leq m \leq N$ *there exist two constants* $\hat{h}_0 > 0$ *and* $\hat{\tau}_0 > 0$ *such that if* $h \leq \hat{h}_0$ *and* $\tau \leq \hat{\tau}_0$*, then the following optimal error estimate holds:*

$$\left\|\mathbf{u}(t_m)-\mathbf{u}_h^m\right\|_{L^2}^2 + s\left\|\mathbf{b}(t_m)-\mathbf{b}_h^m\right\|_{L^2}^2 + \left\|\theta(t_m)-\theta_h^m\right\|_{L^2}^2 \leq C_\dagger\left(\tau^6+h^4\right).$$

Remark 4.2. For sufficiently small τ and h, Theorem 4.1, and the inverse and triangle inequalities show that

$$\begin{aligned}
&\left\|\mathbf{u}_h^m\right\|_{L^\infty} + \left\|\nabla\mathbf{u}_h^m\right\|_{L^3} \leq \left\|\varepsilon_{\mathbf{u}}^m\right\|_{L^\infty} + \left\|\nabla\varepsilon_{\mathbf{u}}^m\right\|_{L^3} + \left\|\rho_{\mathbf{u}}^m\right\|_{L^\infty} + \left\|\nabla\rho_{\mathbf{u}}^m\right\|_{L^3} + c\|\mathbf{u}^m\|_2 \leq C,\\
&\left\|\mathbf{b}_h^m\right\|_{L^\infty} \leq \left\|\varepsilon_{\mathbf{b}}^m\right\|_{L^\infty} + \left\|\rho_{\mathbf{b}}^m\right\|_{L^\infty} + c\|\mathbf{b}^m\|_2 \leq C,\\
&\left\|\theta_h^m\right\|_{L^\infty} + \left\|\nabla\theta_h^m\right\|_{L^3} \leq \left\|\varepsilon_\theta^m\right\|_{L^\infty} + \left\|\nabla\varepsilon_\theta^m\right\|_{L^3} + \left\|\rho_\theta^m\right\|_{L^\infty} + \left\|\nabla\rho_\theta^m\right\|_{L^3} + c\|\theta^m\|_2 \leq C
\end{aligned}$$

for $0 \leq m \leq N$.

Remark 4.3. If some higher-order finite elements ($r \geq 2$) are applied, then the lower-order convergence $\mathcal{O}(\tau^3 + h^2)$ can be obtained in almost the same way as the MINI element. Hence, as the results in Remark 4.2, the boundedness of numerical solutions can still be derived.

5. Optimal Error Estimates

In the previous section, we showed the boundedness of the numerical solutions to (3.1) in a strong norm by a temporal-spatial error splitting arguments. Now we are ready to prove our main result in Theorem 3.1. For the sake of convenience, we define

$$\begin{aligned}
e_{\mathbf{u}}^m &= \left(\mathbf{u}_h^m - R_h\mathbf{u}(t_m)\right) - \left(\mathbf{u}(t_m) - R_h\mathbf{u}(t_m)\right) =: e_{1\mathbf{u}}^m - e_{2\mathbf{u}}^m,\\
e_{\mathbf{b}}^m &= \left(\mathbf{b}_h^m - F_h\mathbf{b}(t_m)\right) - \left(\mathbf{b}(t_m) - F_h\mathbf{b}(t_m)\right) =: e_{1\mathbf{b}}^m - e_{2\mathbf{b}}^m,\\
e_p^m &= \left(p_h^m - J_hp(t_m)\right) - \left(p(t_m) - J_hp(t_m)\right) =: e_{1p}^m - e_{2p}^m,\\
e_\theta^m &= \left(\theta_h^m - T_h\theta(t_m)\right) - \left(\theta(t_m) - T_h\theta(t_m)\right) =: e_{1\theta}^m - e_{2\theta}^m.
\end{aligned}$$

Referring to the projection error estimates, we only need to evaluate the terms $e_{1\mathbf{u}}^m, e_{1\mathbf{b}}^m$, $e_{1\theta}^m, e_{1p}^m$. Thus we first establish the error equations. Subtracting (3.1) from (2.1) and utilizing the Stokes projection, the Maxwell projection and the Ritz projection, we note that the errors $e_{1\mathbf{u}}^m$, $e_{1\mathbf{b}}^m$, $e_{1\theta}^m$ and e_{1p}^m satisfy the following equations:

$$\begin{aligned}
&\left(D_\tau e_{1\mathbf{u}}^m,\mathbf{v}_h\right)+B_u\left(\nu^m(\widetilde{\theta}_h^m),e_{1\mathbf{u}}^m,\mathbf{v}_h\right)-B_f\left(\mathbf{v}_h,e_{1p}^m\right)=\left(\partial_t\mathbf{u}(t_m)-D_\tau R_h\mathbf{u}(t_m),\mathbf{v}_h\right)+(\aleph^m,\mathbf{v}_h),\\
&B_f\left(e_{1\mathbf{u}}^m,q_h\right)=0,\\
&\left(D_\tau\mathbf{e}_{1\mathbf{b}}^m,\mathbf{w}_h\right)+B_b\left(\mu^m(\widetilde{\theta}_h^m),\mathbf{e}_{1\mathbf{b}}^m,\mathbf{w}_h\right)=\left(\partial_t\mathbf{b}(t_m)-D_\tau F_h\mathbf{b}(t_m),\mathbf{w}_h\right)+(\varkappa^m,\mathbf{w}_h),\\
&\left(D_\tau\mathbf{e}_{1\theta}^m,\phi_h\right)+B_\theta\left(\kappa^m(\widetilde{\theta}_h^m),\mathbf{e}_{1\theta}^m,\phi_h\right)=\left(\partial_t\theta(t_m)-D_\tau T_h\theta(t_m),\phi_h\right)+(\wp^m,\phi_h),
\end{aligned}$$

where

$$\begin{aligned}
(\aleph^m,\mathbf{v}_h):=&\,B_u\big(\nu^m(\theta(t_m)),\mathbf{u}(t_m),\mathbf{v}_h\big)-B_u\left(\nu^m(\widetilde{\theta}_h^m),\mathbf{u}(t_m),\mathbf{v}_h\right)\\
&+T_u\big(\mathbf{u}(t_m),\mathbf{u}(t_m),\mathbf{v}_h\big)-T_u\left(\widetilde{\mathbf{u}}_h^m,\mathbf{u}_h^m,\mathbf{v}_h\right)\\
&+sT_{b1}\big(\mathbf{b}(t_m),\mathbf{b}(t_m),\mathbf{v}_h\big)-sT_{b1}\left(\widetilde{\mathbf{b}}_h^m,\mathbf{b}_h^m,\mathbf{v}_h\right)\\
&-D_f\big(\beta^m(\theta(t_m)),\theta(t_m),\mathbf{v}_h\big)+D_f\left(\beta^m(\widetilde{\theta}_h^m),\theta_h^m,\mathbf{v}_h\right),\\
(\varkappa^m,\mathbf{w}_h):=&\,B_b\big(\mu^m(\theta(t_m)),\mathbf{b}(t_m),\mathbf{w}_h\big)-B_b\left(\mu^m(\widetilde{\theta}_h^m),\mathbf{b}(t_m),\mathbf{w}_h\right)\\
&-T_{b_2}\big(\mathbf{u}(t_m),\mathbf{b}(t_m),\mathbf{w}_h\big)+T_{b2}\left(\mathbf{u}_h^m,\widetilde{\mathbf{b}}_h^m,\mathbf{w}_h\right),\\
(\wp^m,\mathbf{v}_h):=&\,B_\theta\big(\kappa^m(\theta(t_m)),\theta(t_m),\phi_h\big)-B_u\left(\kappa^m(\widetilde{\theta}_h^m),\theta(t_m),\phi_h\right)\\
&+T_\theta\big(\mathbf{u}(t_m),\theta(t_m),\phi_h\big)-T_\theta\big(\widetilde{\mathbf{u}}_h^m,\theta_h^m,\phi_h\big).
\end{aligned}$$

Choosing

$$\mathbf{v}_h=2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\in\mathbf{V}_h,\quad \mathbf{w}_h=2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1},\quad \phi_h=2e_{1\theta}^m-e_{1\theta}^{m-1},$$

we write

$$\begin{aligned}
&\left(D_\tau e_{1\mathbf{u}}^m,2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)+B_u\left(\nu^m(\widetilde{\theta}^m),e_{1\mathbf{u}}^m,2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\\
=&\left(\partial_t\mathbf{u}(t_m)-D_\tau R_h\mathbf{u}(t_m),2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)+\left(\aleph^m,2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right),\\
&\left(D_\tau e_{1\mathbf{b}}^m,2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right)+B_b\left(\mu^m(\widetilde{\theta}^m),e_{1\mathbf{b}}^m,2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right)\\
=&\left(\partial_t\mathbf{b}(t_m)-D_\tau F_h\mathbf{b}(t_m),2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right)+\left(\varkappa^m,2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right),\\
&\left(D_\tau e_{1\theta}^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)+B_\theta\left(\kappa^m(\widetilde{\theta}^m),e_{1\theta}^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\\
=&\left(\partial_t\theta(t_m)-D_\tau T_h\theta(t_m),2e_{1\theta}^m-e_{1\theta}^{m-1}\right)+\left(\wp^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right).
\end{aligned}$$

Moreover, the Cauchy-Schwarz inequality, the triangle inequality, and (2.7), (2.8) give

$$\begin{aligned}
&\left(\partial_t\mathbf{u}(t_m)-D_\tau R_h\mathbf{u}(t_m),2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\\
=&\left(\partial_t\mathbf{u}(t_m)-D_\tau(\mathbf{u}(t_m)),2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)+\left(D_\tau(\mathbf{u}(t_m))-D_\tau(R_h\mathbf{u}(t_m)),2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\\
\le&\, c\Big(\tau^{5/2}\left\|\partial_t^4\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\\
&\quad+\frac{h^{r+1}}{\sqrt{\tau}}\|(\partial_t\mathbf{u},\partial_t p)\|_{L^2(t_{m-3},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}\Big)\left(\left\|e_{1\mathbf{u}}^m\right\|_{L^2}+\left\|e_{1\mathbf{u}}^{m-1}\right\|_{L^2}\right)
\end{aligned}$$

$$\leq c\epsilon\left(\tau^5\left\|\partial_t^4\mathbf{u}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\frac{h^{2r+2}}{\tau}\|(\partial_t\mathbf{u},\partial_t p)\|^2_{L^2(t_{m-3},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}\right)$$
$$+\epsilon^{-1}\left(\left\|e^m_{1\mathbf{u}}\right\|^2_{L^2}+\left\|e^{m-1}_{1\mathbf{u}}\right\|^2_{L^2}\right). \tag{5.1}$$

Similar arguments show that

$$\begin{aligned}&\left(\partial_t\mathbf{b}(t_m)-D_\tau F_h\mathbf{b}(t_m),2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\\ &\leq c\epsilon\left(\tau^5\left\|\partial_t^4\mathbf{b}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\frac{h^{2r+2}}{\tau}\|\partial_t\mathbf{b}\|^2_{L^2(t_{m-3},t_m;H^{r+1}(\Omega)^d)}\right)\\ &\quad+\epsilon^{-1}\left(\left\|e^m_{1\mathbf{b}}\right\|^2_{L^2}+\left\|e^{m-1}_{1\mathbf{b}}\right\|^2_{L^2}\right),\\ &\left(\partial_t\theta(t_m)-D_\tau T_h\theta(t_m),2e^m_{1\theta}-e^{m-1}_{1\theta}\right)\\ &\leq c\epsilon\left(\tau^5\left\|\partial_t^4\theta\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\frac{h^{2r+2}}{\tau}\|\partial_t\theta\|^2_{L^2(t_{m-3},t_m;H^{r+1}(\Omega))}\right)\\ &\quad+\epsilon^{-1}\left(\left\|e^m_{1\theta}\right\|^2_{L^2}+\left\|e^{m-1}_{1\theta}\right\|^2_{L^2}\right).\end{aligned} \tag{5.2}$$

Adding and subtracting suitable terms, we rewrite the nonlinear term $(\aleph^m,\mathbf{v}_h)$ as follows:

$$\begin{aligned}(\aleph^m,\mathbf{v}_h)=&B_u\left(\nu^m(\theta(t_m))-\nu^m(\widetilde{\theta}(t_m)),\mathbf{u}(t_m),\mathbf{v}_h\right)+B_u\left(\nu^m(\widetilde{\theta}(t_m))-\nu^m(\widetilde{\theta}^m_h),\mathbf{u}(t_m),\mathbf{v}_h\right)\\ &+T_u\left(\mathbf{u}(t_m),\mathbf{e}^m_{2\mathbf{u}},\mathbf{v}_h\right)+T_u\left(\mathbf{u}(t_m)-\widetilde{\mathbf{u}}(t_m),R_h\mathbf{u}(t_m),\mathbf{v}_h\right)\\ &+T_u\left(\widetilde{e}^m_{2\mathbf{u}},R_h\mathbf{u}(t_m),\mathbf{v}_h\right)-T_u\left(\widetilde{e}^m_{1\mathbf{u}},R_h\mathbf{u}(t_m),\mathbf{v}_h\right)-T_u\left(\widetilde{\mathbf{u}}^m_h,e^m_{1\mathbf{u}},\mathbf{v}_h\right)\\ &+sT_{b1}\left(\mathbf{b}(t_m),\mathbf{e}^m_{2\mathbf{b}},\mathbf{v}_h\right)+sT_{b1}\left(\mathbf{b}(t_m)-\widetilde{\mathbf{b}}(t_m),F_h\mathbf{b}(t_m),\mathbf{v}_h\right)\\ &+sT_{b1}\left(\widetilde{e}^m_{2\mathbf{b}},F_h\mathbf{b}(t_m),\mathbf{v}_h\right)-sT_{b1}\left(\widetilde{e}^m_{1\mathbf{b}},F_h\mathbf{b}(t_m),\mathbf{v}_h\right)-sT_{b1}\left(\widetilde{\mathbf{b}}^m_h,e^m_{1\mathbf{b}},\mathbf{v}_h\right)\\ &+D_f\left(\beta^m(\widetilde{\theta}^m_h),\theta^m_h-\theta(t_m),\mathbf{v}_h\right)+D_f\left(\beta^m(\widetilde{\theta}^m_h)-\beta^m(\widetilde{\theta}(t_m)),\theta(t_m),\mathbf{v}_h\right)\\ &-D_f\left(\beta^m(\theta(t_m))-\beta^m(\widetilde{\theta}(t_m)),\theta(t_m),\mathbf{v}_h\right)\\ =:&\left(\sum_{i=1}^{15}\aleph^m_i,\mathbf{v}_h\right).\end{aligned}$$

Further, using the Hölder inequality and the relations (2.2)-(2.3), (2.8), (2.9), we obtain

$$\begin{aligned}\left|\left(\aleph^m_1,2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right|&\leq c|\nu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\|\theta(t_m)-\widetilde{\theta}(t_m)\|_{L^2}\|\nabla\mathbf{u}(t_m)\|_{L^\infty}\left\|\nabla\left(2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right\|_{L^2}\\ &\leq c\epsilon^{-1}\tau^5\left\|\partial_t^4\theta\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\epsilon\left(\left\|e^m_{1\mathbf{u}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{u}}\right\|^2_1\right),\\ \left|\left(\aleph^m_2,2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right|&\leq c|\nu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left\|\widetilde{\theta}(t_m)-\widetilde{\theta}^m_h\right\|_{L^2}\|\nabla\mathbf{u}(t_m)\|_{L^\infty}\left\|\nabla\left(2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right\|_{L^2}\\ &\leq c\epsilon^{-1}\left(h^{2r+2}\|\widetilde{\theta}\|_{C(t_{m-3},t_m;H^{r+1}(\Omega))}+\left\|\widetilde{e}^m_{1\theta}\right\|^2_{L^2}\right)+\epsilon\left(\left\|e^m_{1\mathbf{u}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{u}}\right\|^2_1\right).\end{aligned}$$

The expressions containing the terms $\aleph^m_j$, $j=3,\ldots,15$ can be handled by using the Hölder and Young inequalities along with the relations (2.2)-(2.5), (2.8), (2.9), (4.1)-(4.3), so that

$$\left|\left(\aleph^m_3,2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right|\leq c\|\mathbf{u}(t_m)\|_{L^\infty}\left\|\mathbf{e}^m_{2\mathbf{u}}\right\|_{L^2}\left\|2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right\|_1$$

$$
\begin{aligned}
&\le ch^{r+1}\|(\mathbf{u},p)\|_{C(t_{m-1},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}h^{2r+2}\|(\mathbf{u},p)\|^2_{C(t_{m-1},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_4^m,2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le c\|\mathbf{u}(t_m)-\widetilde{\mathbf{u}}(t_m)\|_{L^2}\big(\|\nabla R_h\mathbf{u}(t_m)\|_{L^3}+\|R_h\mathbf{u}(t_m)\|_{L^\infty}\big)\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\tau^{5/2}\left\|\partial_t^3\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}\tau^5\left\|\partial_t^3\mathbf{u}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_5^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le c\left\|\widetilde{e}_{2\mathbf{u}}^m\right\|_{L^2}\big(\|\nabla R_h\mathbf{u}(t_m)\|_{L^3}+\|R_h\mathbf{u}(t_m)\|_{L^\infty}\big)\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le ch^{r+1}\|(\mathbf{u},p)\|_{C(t_{m-3},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}h^{2r+2}\|(\mathbf{u},p)\|^2_{C(t_{m-3},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_6^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le c\left\|\widetilde{e}_{1\mathbf{u}}^m\right\|_{L^2}\big(\|\nabla R_h\mathbf{u}(t_m)\|_{L^3}+\|R_h\mathbf{u}(t_m)\|_{L^\infty}\big)\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}\left\|\widetilde{e}_{1\mathbf{u}}^m\right\|_{L^2}^2+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_7^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le \left|T_u\big(\widetilde{\mathbf{u}}_h^m,e_{1\mathbf{u}}^m,2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\big)\right|\\
&\le c\left(\left\|\widetilde{\mathbf{u}}_h^m\right\|_{L^\infty}+\left\|\nabla\widetilde{\mathbf{u}}_h^m\right\|_{L^3}\right)\left\|e_{1\mathbf{u}}^m\right\|_1\left\|e_{1\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\le c\epsilon^{-1}\left\|e_{1\mathbf{u}}^{m-1}\right\|_{L^2}^2+\epsilon\left\|e_{1\mathbf{u}}^m\right\|_1^2,
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_8^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le s\|\mathbf{b}(t_m)\|_1\left\|\mathbf{e}_{2\mathbf{b}}^m\right\|_1\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le ch^{r+1}\|\mathbf{b}\|_{C(t_{m-1},t_m;H^{r+2}(\Omega)^d)}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}h^{2r+2}\|\mathbf{b}\|^2_{C(t_{m-1},t_m;H^{r+2}(\Omega)^d)}+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_9^m,2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le cs\|\mathbf{b}(t_m)-\widetilde{\mathbf{b}}(t_m)\|_{L^2}\|\mathrm{curl}\,F_h\mathbf{b}(t_m)\|_{L^3}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\tau^{5/2}\left\|\partial_t^3\mathbf{b}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}\tau^5\left\|\partial_t^3\mathbf{b}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_{10}^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le cs\left\|\widetilde{e}_{2\mathbf{b}}^m\right\|_{L^2}\|\mathrm{curl}\,F_h\mathbf{b}(t_m)\|_{L^3}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le ch^{r+1}\|\mathbf{b}\|_{C(t_{m-3},t_m;H^{r+1}(\Omega)^d)}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}h^{2r+2}\|\mathbf{b}\|^2_{C(t_{m-3},t_m;H^{r+1}(\Omega)^d)}+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_{11}^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le cs\left\|\widetilde{e}_{1\mathbf{b}}^m\right\|_{L^2}\|\mathrm{curl}\,F_h\mathbf{b}(t_m)\|_{L^3}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_1\\
&\le c\epsilon^{-1}s\left\|\widetilde{e}_{1\mathbf{b}}^m\right\|_{L^2}^2+\epsilon\left(\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\left|\left(\aleph_{12}^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| &\le cs\left\|\widetilde{\mathbf{b}}_h^m\right\|_{L^\infty}\left\|\mathrm{curl}\,e_{1\mathbf{b}}^m\right\|_{L^2}\left\|e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_{L^2}\\
&\le c\epsilon^{-1}\left(\left\|e_{1\mathbf{u}}^m\right\|_{L^2}^2+\left\|e_{1\mathbf{u}}^{m-1}\right\|_{L^2}^2\right)+\epsilon\left\|\mathrm{curl}\,e_{1\mathbf{b}}^m\right\|_{L^2}^2,
\end{aligned}
$$

$$
\left|\left(\aleph_{13}^m,2e_{2\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)\right| \le c\beta_1\left\|\theta_h^m-\theta(t_m)\right\|_{L^2}\left\|2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right\|_{L^2}
$$

$$\leq c\epsilon^{-1}\left(h^{2r+2}\|\theta\|^2_{C(t_{m-1},t_m;H^{r+1}(\Omega))}+\left\|e^m_{1\theta}\right\|^2_{L^2}\right)$$
$$+\epsilon\left(\left\|e^m_{1\mathbf{u}}\right\|^2_{L^2}+\left\|e^{m-1}_{1\mathbf{u}}\right\|^2_{L^2}\right),$$

$$\begin{aligned}\left|\left(\aleph^m_{14},2e^m_{2\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right| &\leq c|\beta|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left\|\widetilde{\theta}^m_h-\widetilde{\theta}(t_m)\right\|_{L^2}\|\theta(t_m)\|_{L^\infty}\left\|2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right\|_{L^2}\\ &\leq c\epsilon^{-1}\left(h^{2r+2}\|\theta\|^2_{C(t_{m-3},t_m;H^{r+1}(\Omega))}+\left\|\widetilde{e}^m_{1\theta}\right\|^2_{L^2}\right)\\ &\quad+\epsilon\left(\left\|e^m_{1\mathbf{u}}\right\|^2_{L^2}+\left\|e^{m-1}_{1\mathbf{u}}\right\|^2_{L^2}\right),\end{aligned}$$

$$\begin{aligned}\left|\left(\aleph^m_{15},2e^m_{2\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right)\right| &\leq c|\beta|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\|\theta(t_m)-\widetilde{\theta}(t_m)\|_{L^2}\|\theta(t_m)\|_{L^\infty}\left\|2e^m_{1\mathbf{u}}-e^{m-1}_{1\mathbf{u}}\right\|_{L^2}\\ &\leq c\epsilon^{-1}\tau^5\left\|\partial^3_t\theta\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\epsilon\left(\left\|e^m_{1\mathbf{u}}\right\|^2_{L^2}+\left\|e^{m-1}_{1\mathbf{u}}\right\|^2_{L^2}\right).\end{aligned}$$

Additionally, we write $(\varkappa^m,\mathbf{w}_h)$ as

$$\begin{aligned}(\varkappa^m,\mathbf{w}_h)=\;&B_b\left(\mu^m(\theta(t_m))-\mu^m(\widetilde{\theta}(t_m)),\mathbf{b}(t_m),\mathbf{w}_h\right)+B_b\left(\mu^m(\widetilde{\theta}(t_m))-\mu^m(\widetilde{\theta}^m_h),\mathbf{b}(t_m),\mathbf{w}_h\right)\\ &-T_{b2}\left(e^m_{2\mathbf{u}},\mathbf{b}(t_m),\mathbf{w}_h\right)-T_{b2}\left(F_h\mathbf{u}(t_m),\mathbf{b}(t_m)-\widetilde{\mathbf{b}}(t_m),\mathbf{w}_h\right)\\ &-T_{b2}\left(R_h\mathbf{u}(t_m),\widetilde{e}^m_{2\mathbf{b}},\mathbf{w}_h\right)+T_{b2}\left(R_h\mathbf{u}(t_m),\widetilde{e}^m_{1\mathbf{b}},\mathbf{w}_h\right)+T_{b2}\left(e^m_{1\mathbf{u}},\widetilde{\mathbf{b}}(t_m),\mathbf{w}_h\right)\\ =:&\left(\sum_{i=1}^7\varkappa^m_i,\mathbf{w}_h\right).\end{aligned}$$

Similar approach to $\varkappa^m_1,\varkappa^m_2$ shows that

$$\begin{aligned}\left|\left(\varkappa^m_1,2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right| &\leq c|\mu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\|\theta(t_m)-\widetilde{\theta}(t_m)\|_{L^2}\|\mathrm{curl}\,\mathbf{b}(t_m)\|_{W^{0,\infty}}\left\|\left(2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right\|_1\\ &\leq c\epsilon^{-1}\tau^5\left\|\partial^4_t\theta\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega))}+\epsilon\left(\left\|e^m_{1\mathbf{b}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{b}}\right\|^2_1\right),\end{aligned}$$

$$\begin{aligned}\left|\left(\varkappa^m_2,2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right| &\leq c|\mu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left\|\widetilde{\theta}(t_m)-\widetilde{\theta}^m_h\right\|_{L^2}\|\mathrm{curl}\,\mathbf{b}(t_m)\|_{L^\infty}\left\|\left(2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right\|_1\\ &\leq c\epsilon^{-1}\left(h^{2r+2}\|\widetilde{\theta}\|_{C(t_{m-3},t_m;H^{r+1}(\Omega))}+\left\|\widetilde{e}^m_{1\theta}\right\|^2_{L^2}\right)+\epsilon\left(\left\|e^m_{1\mathbf{b}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{b}}\right\|^2_1\right),\end{aligned}$$

and we estimate the terms containing $\varkappa^m_j$, $j=3,\ldots,7$ as

$$\begin{aligned}\left|\left(\varkappa^m_3,2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right| &\leq \left\|e^m_{2\mathbf{u}}\right\|_{L^2}\|\mathbf{b}(t_m)\|_{L^\infty}\left\|\mathrm{curl}\left(2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right\|_{L^2}\\ &\leq ch^{2r+2}\|(\mathbf{u},p)\|^2_{C(t_{m-1},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}+\epsilon\left(\left\|e^m_{1\mathbf{b}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{b}}\right\|^2_1\right),\end{aligned}$$

$$\begin{aligned}\left|\left(\varkappa^m_4,2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right| &\leq \|R_h\mathbf{u}(t_m)\|_{L^\infty}\|\mathbf{b}(t_m)-\widetilde{\mathbf{b}}(t_m)\|_{L^2}\left\|\mathrm{curl}\left(2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right\|_{L^2}\\ &\leq c\tau^5\left\|\partial^3_t\mathbf{b}\right\|^2_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}+\epsilon\left(\left\|e^m_{1\mathbf{b}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{b}}\right\|^2_1\right),\end{aligned}$$

$$\begin{aligned}\left|\left(\varkappa^m_5,2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right| &\leq \|R_h\mathbf{u}(t_m)\|_{L^\infty}\left\|\widetilde{e}^m_{2\mathbf{b}}\right\|_{L^2}\left\|\mathrm{curl}\left(2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right\|_{L^2}\\ &\leq ch^{2r+2}\|\mathbf{b}\|^2_{C(t_{m-3},t_m;H^{r+1}(\Omega)^d)}+\epsilon\left(\left\|e^m_{1\mathbf{b}}\right\|^2_1+\left\|e^{m-1}_{1\mathbf{b}}\right\|^2_1\right),\end{aligned}$$

$$\left|\left(\varkappa^m_6,2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right| \leq c\|R_h\mathbf{u}(t_m)\|_{L^\infty}\left\|\widetilde{e}^m_{1\mathbf{b}}\right\|_{L^2}\left\|\mathrm{curl}\left(2e^m_{1\mathbf{b}}-e^{m-1}_{1\mathbf{b}}\right)\right\|_{L^2}$$

$$
\begin{aligned}
&\leq c\epsilon^{-1}\left\|\widetilde{e}_{1\mathbf{b}}^m\right\|_{L^2}^2+\epsilon\left(\left\|e_{1\mathbf{b}}^m\right\|_1^2+\left\|e_{1\mathbf{b}}^{m-1}\right\|_1^2\right),\\
\left|\left(\varkappa_7^m,2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right)\right| &\leq c\left\|e_{1\mathbf{u}}^m\right\|_{L^2}\left\|\widetilde{\mathbf{b}}_h^m\right\|_{L^\infty}\left\|\operatorname{curl}\left(2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right)\right\|_{L^2}\\
&\leq c\epsilon^{-1}\left\|e_{1\mathbf{u}}^m\right\|_{L^2}^2+\epsilon\left(\left\|e_{1\mathbf{b}}^m\right\|_1^2+\left\|e_{1\mathbf{b}}^{m-1}\right\|_1^2\right).
\end{aligned}
$$

Furthermore, adding and subtracting suitable terms, we write $(\wp^m,\mathbf{v}_h)$ in the form

$$
\begin{aligned}
(\wp^m,\mathbf{v}_h)=&\,B_\theta\left(\kappa^m(\theta(t_m))-\kappa^m(\widetilde{\theta}(t_m)),\theta(t_m),\phi_h\right)+B_\theta\left(\kappa^m(\widetilde{\theta}(t_m))-\kappa^m(\widetilde{\theta}_h^m),\theta(t_m),\phi_h\right)\\
&+T_\theta\left(\mathbf{u}(t_m),\mathbf{e}_{2\theta}^m,\phi_h\right)+T_\theta\left(\mathbf{u}(t_m)-\widetilde{\mathbf{u}}(t_m),T_h\theta(t_m),\phi_h\right)\\
&+T_\theta\left(\widetilde{e}_{2\mathbf{u}}^m,T_h\theta(t_m),\phi_h\right)-T_\theta\left(\widetilde{e}_{1\mathbf{u}}^m,T_h\theta(t_m),\phi_h\right)-T_\theta\left(\widetilde{\mathbf{u}}_h^m,e_{1\theta}^m,\phi_h\right)\\
=:&\left(\sum_{i=1}^7\wp_i^m,\phi_h\right).
\end{aligned}
$$

Bounding $\wp_1^m$ and $\wp_2^m$ yields

$$
\begin{aligned}
\left|\left(\wp_1^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq c|\nu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\|\theta(t_m)-\widetilde{\theta}(t_m)\|_{L^2}\|\nabla\theta(t_m)\|_{L^\infty}\left\|\nabla\left(2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right\|_{L^2}\\
&\leq c\epsilon^{-1}\tau^5\left\|\partial_t^4\theta\right\|_{L^2(t_{m-3},t_m;L^2(\Omega))}^2+\epsilon\left(\left\|e_{1\theta}^m\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2\right),\\
\left|\left(\wp_2^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq c|\nu|_{C^{0,1}(\overline{\Omega}\times R;R^+)}\left\|\widetilde{\theta}(t_m)-\widetilde{\theta}_h^m\right\|_{L^2}\|\nabla\theta(t_m)\|_{L^\infty}\left\|\nabla\left(2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right\|_{L^2}\\
&\leq c\epsilon^{-1}\left(h^{2r+2}\|\widetilde{\theta}\|_{C(t_{m-3},t_m;H^{r+1}(\Omega))}+\left\|\widetilde{e}_{1\theta}^m\right\|_{L^2}^2\right)+\epsilon\left(\left\|e_{1\theta}^m\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2\right).
\end{aligned}
$$

The terms $\wp_3^m-\wp_7^m$ can be estimated by

$$
\begin{aligned}
\left|\left(\wp_3^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq c\|\mathbf{u}(t_m)\|_{L^\infty}\left\|\mathbf{e}_{2\theta}^m\right\|_{L^2}\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq ch^{r+1}\|\theta\|_{C(t_{m-1},t_m;H^{r+1}(\Omega))}\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq c\epsilon^{-1}h^{2r+2}\|\theta\|_{C(t_{m-1},t_m;H^{r+1}(\Omega))}^2+\epsilon\left(\left\|e_{1\theta}^m\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2\right),\\
\left|\left(\wp_4^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq c\|\mathbf{u}(t_m)-\widetilde{\mathbf{u}}(t_m)\|_{L^2}\left(\|\nabla T_h\theta(t_m)\|_{L^3}+\|T_h\theta(t_m)\|_{L^\infty}\right)\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq c\tau^{5/2}\left\|\partial_t^3\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq c\epsilon^{-1}\tau^5\left\|\partial_t^3\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2+\epsilon\left(\left\|e_{1\theta}^m\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2\right),\\
\left|\left(\wp_5^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq c\left\|\widetilde{e}_{2\mathbf{u}}^m\right\|_{L^2}\left(\|\nabla T_h\theta(t_m)\|_{L^3}+\|T_h\theta(t_m)\|_{L^\infty}\right)\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq ch^{r+1}\|\theta\|_{C(t_{m-3},t_m;H^{r+1}(\Omega))}\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq c\epsilon^{-1}h^{2r+2}\|\theta\|_{C(t_{m-3},t_m;H^{r+1}(\Omega))}^2+\epsilon\left(\left\|e_{1\theta}^m\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2\right),\\
\left|\left(\wp_6^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq c\left\|\widetilde{e}_{1\mathbf{u}}^m\right\|_{L^2}\left(\|\nabla T_h\theta(t_m)\|_{L^3}+\|T_h\theta(t_m)\|_{L^\infty}\right)\left\|2e_{1\theta}^m-e_{1\theta}^{m-1}\right\|_1\\
&\leq c\epsilon^{-1}\left\|\widetilde{e}_{1\mathbf{u}}^m\right\|_{L^2}^2+\epsilon\left(\left\|e_{1\theta}^m\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2\right),\\
\left|\left(\wp_7^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right| &\leq \left|T_\theta\left(\widetilde{\mathbf{u}}_h^m,e_{1\theta}^m,2e_{1\theta}^m-e_{1\theta}^{m-1}\right)\right|
\end{aligned}
$$

$$
\begin{aligned}
&\le c\left(\left\|\nabla\widetilde{\mathbf{u}}_h^m\right\|_{L^3}+\left\|\widetilde{\mathbf{u}}_h^m\right\|_{L^\infty}\right)\left\|e_{1\theta}^m\right\|_1\left\|e_{1\theta}^{m-1}\right\|_{L^2}\\
&\le c\epsilon^{-1}\left\|e_{1\theta}^{m-1}\right\|_{L^2}^2+\epsilon\left\|e_{1\theta}^m\right\|_1^2.
\end{aligned}
$$

Combining all the estimates and using (5.1)-(5.2) gives

$$
\begin{aligned}
&\left(D_\tau e_{1\mathbf{u}}^m, 2e_{1\mathbf{u}}^m-e_{1\mathbf{u}}^{m-1}\right)+c_1\nu_1\left\|e_{1\mathbf{u}}^m\right\|_1^2+\frac{\nu_1}{2}\left(\left\|\nabla e_{1\mathbf{u}}^m\right\|_{L^2}^2-\left\|\nabla e_{1\mathbf{u}}^{m-1}\right\|_{L^2}^2\right)\\
&\quad+\left(D_\tau e_{1\theta}^m, 2e_{1\theta}^m-e_{1\theta}^{m-1}\right)+c_1\kappa_1\left\|e_{1\theta}^m\right\|_1^2+\frac{\kappa_1}{2}\left(\left\|\nabla e_{1\theta}^m\right\|_{L^2}^2-\left\|\nabla e_{1\theta}^{m-1}\right\|_{L^2}^2\right)\\
&\quad+s\left(D_\tau e_{1\mathbf{b}}^m, 2e_{1\mathbf{b}}^m-e_{1\mathbf{b}}^{m-1}\right)+c_2\mu_1 s\left\|e_{1\mathbf{b}}^m\right\|_1^2\\
&\quad+\frac{s\mu_1}{2}\left(\left\|\operatorname{curl}e_{1\mathbf{b}}^m\right\|_{L^2}^2+\left\|\operatorname{div}e_{1\mathbf{b}}^m\right\|_{L^2}^2-\left\|\operatorname{curl}e_{1\mathbf{b}}^{m-1}\right\|_{L^2}^2-\left\|\operatorname{div}e_{1\mathbf{b}}^{m-1}\right\|_{L^2}^2\right)\\
&\le c\Bigg(h^{2r+2}+\tau^5\left(\left\|\partial_t^4\mathbf{u}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2+\left\|\partial_t^4\mathbf{b}\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2+\left\|\partial_t^4\theta\right\|_{L^2(t_{m-3},t_m;L^2(\Omega)^d)}^2\right)\\
&\qquad+\frac{h^{2r+2}}{\tau}\Big(\|(\partial_t\mathbf{u},\partial_t p)\|_{L^2(t_{m-3},t_m;H^{r+1}(\Omega)^d\times H^r(\Omega))}^2+\|\partial_t\mathbf{b}\|_{L^2(t_{m-3},t_m;H^{r+1}(\Omega)^d)}^2\\
&\qquad\qquad+\|\partial_t\theta\|_{L^2(t_{m-3},t_m;H^{r+1}(\Omega)^d)}^2\Big)\Bigg)\\
&\quad+c\epsilon^{-1}\left(\left\|\widetilde{e}_{1\mathbf{u}}^m\right\|_{L^2}^2+\left\|\widetilde{e}_{1\theta}^m\right\|_{L^2}^2+s\left\|\widetilde{e}_{1\mathbf{b}}^m\right\|_{L^2}^2+\left\|\mathbf{e}_{1\mathbf{u}}^m\right\|_{L^2}^2+\left\|\mathbf{e}_{1\theta}^m\right\|_{L^2}^2+s\left\|e_{1\mathbf{b}}^m\right\|_{L^2}^2\right)\\
&\quad+\epsilon\left(\left\|e_{1\mathbf{u}}^{m-1}\right\|_1^2+\left\|e_{1\theta}^{m-1}\right\|_1^2+s\left\|e_{1\mathbf{b}}^{m-1}\right\|_1^2+\left\|e_{1\mathbf{u}}^m\right\|_1^2+\left\|e_{1\theta}^m\right\|_1^2+s\left\|\mathbf{e}_{1\mathbf{b}}^m\right\|_1^2\right).
\end{aligned}
$$

Choosing a sufficiently small ϵ, summing in m from 3 to n and using Lemmas 2.1, 2.2, we show that if $\tau\le\tau_0$ for a $\tau_0>0$, then

$$
\begin{aligned}
&\left\|\mathbf{e}_{1\mathbf{u}}^n\right\|_{L^2}^2+\left\|\mathbf{e}_{1\theta}^n\right\|_{L^2}^2+s\left\|\mathbf{e}_{1\mathbf{b}}^n\right\|_{L^2}^2+\tau\sum_{i=3}^n\left(\nu_1\left\|\mathbf{e}_{1\mathbf{u}}^i\right\|_1^2+s\mu_1\left\|\mathbf{e}_{1\mathbf{b}}^i\right\|_1^2+\kappa_1\left\|\mathbf{e}_{1\theta}^i\right\|_1^2\right)\\
&\le\exp\left(\frac{Tc}{1-c\tau_0}\right)c\left(h^{2r+2}+\tau^6\right).
\end{aligned}\tag{5.3}
$$

The inequality (3.4) now follows from (4.1)-(4.3) and (5.3) if we choose $C_0>\exp(Tc/(1-c\tau_0))c$.

6. Numerical Tests

In this section, we present the results of numerical tests aimed to verify the optimal error estimates in Theorem 3.1. In all experiments, we employ the Taylor-Hood element to discretize the velocity $\mathbf{u}$ and pressure p and the quadratic element to discretize magnetic $\mathbf{b}$ and temperature θ. Choosing $\tau=h$ in Theorem 3.1 shows that the L^2-errors for $\mathbf{u},\mathbf{b},\theta$ are $\mathscr{O}(\tau^3+h^3)\sim\mathscr{O}(h^3)$, and the H^1-errors for $p,\mathbf{u},\mathbf{b},\theta$ are $\mathscr{O}(\tau^3+h^2)\sim\mathscr{O}(h^2)$.

The domain under consideration is $\Omega=[0,1]\times[0,1]$. We choose the right-hand side and boundary conditions such that the analytical solution to (1.1)-(1.2) has the following form:

$$\mathbf{u} = \left(y^5 + t^3, x^5 + t^3\right)^\top, \qquad \mathbf{b} = \left(\sin y + t^3, \sin x + t^3\right)^\top,$$
$$p = 10(2x-1)(2y-1)(1+t^3), \quad \theta = \left(\sin(\pi x y) + 1\right)e^{t/2}.$$

Moreover, for a two-dimensional system the differential operators in (1.1) are

$$\operatorname{curl}\mathbf{v} = \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}, \quad \mathbf{w} \times \mathbf{v} = w_1 v_2 - w_2 v_1.$$

We consider two cases of temperature-dependent functions ν, μ, κ, β, viz.

$$\text{Case 1.}\ \nu(\theta) = e^{-\theta}, \quad \mu(\theta) = \theta, \quad \kappa(\theta) = e^{\theta}, \quad \beta(\theta) = e^{\theta}.$$
$$\text{Case 2.}\ \nu(\theta) = \theta^2, \quad \mu(\theta) = e^{\theta}, \quad \kappa(\theta) = \theta^2, \quad \beta(\theta) = \theta^2.$$

Set the coupling parameter $s = 1$ and the terminal time $T = 1$. In order to test the convergence order of the proposed scheme, we take the decreasing mesh sizes $h = 1/4, 1/8, 1/16, 1/32, 1/64, 1/128$ and $\tau = h$ in these cases. Numerical results are displayed in Tables 1-4, which show that the L^2-errors of $\mathbf{u}, \mathbf{b}, \theta$ are of third-order accuracy, whereas the L^2-error of p and H^1-errors of $\mathbf{u}, \mathbf{b}, \theta$ are at least of the second-order accuracy. Thus the numerical results are consistent with the theoretical analysis — cf. Theorem 3.1.

To illustrate almost unconditional convergence, we also test the proposed scheme with large time steps — viz. $\tau = 1/4, 1/6, 1/8$ and spatial meshes $h = 1/40, 1/80$. The L^2-errors

Table 1: Case 1. L^2-errors, $t_m = 1$.

h	$\frac{\lVert\mathbf{u}_h^m - \mathbf{u}(t_m)\rVert_{L^2}}{\lVert\mathbf{u}(t_m)\rVert_{L^2}}$	Rate	$\frac{\lVert\mathbf{b}_h^m - \mathbf{b}(t_m)\rVert_{L^2}}{\lVert\mathbf{b}(t_m)\rVert_{L^2}}$	Rate	$\frac{\lVert p_h^m - p(t_m)\rVert_{L^2}}{\lVert p(t_m)\rVert_{L^2}}$	Rate	$\frac{\lVert\theta_h^m - \theta(t_m)\rVert_{L^2}}{\lVert\theta(t_m)\rVert_{L^2}}$	Rate
1/4	7.899e-03	/	7.905e-04	/	4.952e-02	/	2.072e-02	/
1/8	8.041e-04	2.89	9.307e-05	3.09	1.214e-02	2.03	2.175e-03	3.25
1/16	8.104e-05	3.31	1.150e-05	3.02	3.027e-03	2.00	2.877e-05	2.92
1/32	9.816e-06	3.05	1.432e-06	3.01	7.565e-04	2.00	3.425e-06	3.07
1/64	1.205e-06	3.03	1.789e-07	3.00	1.891e-04	2.00	4.264e-07	3.01
1/128	1.495e-07	3.01	2.238e-08	3.00	4.738e-05	2.00	5.319e-08	3.00

Table 2: Case 1. H^1-errors, $t_m = 1$.

h	$\frac{\lVert\mathbf{u}_h^m - \mathbf{u}(t_m)\rVert_1}{\lVert\mathbf{u}(t_m)\rVert_1}$	Rate	$\frac{\lVert\mathbf{b}_h^m - \mathbf{b}(t_m)\rVert_1}{\lVert\mathbf{b}(t_m)\rVert_1}$	Rate	$\frac{\lVert\theta_h^m - \theta(t_m)\rVert_1}{\lVert\theta(t_m)\rVert_1}$	Rate
1/4	8.735e-02	/	8.36722e-03	/	8.735e-02	/
1/8	1.201e-02	2.86	1.06793e-03	2.97	1.201e-02	2.86
1/16	2.490e-03	2.27	1.83424e-04	2.54	2.490e-03	2.27
1/32	5.949e-04	2.07	3.89267e-05	2.24	5.949e-04	2.07
1/64	1.470e-04	2.02	9.26094e-06	2.07	1.470e-04	2.02
1/128	3.666e-05	2.00	2.28488e-06	2.02	3.666e-05	2.00

Table 3: Case 2. L^2-errors, $t_m = 1$.

h	$\frac{\|\mathbf{u}_h^m - \mathbf{u}(t_m)\|_{L^2}}{\|\mathbf{u}(t_m)\|_{L^2}}$	Rate	$\frac{\|\mathbf{b}_h^m - \mathbf{b}(t_m)\|_{L^2}}{\|\mathbf{b}(t_m)\|_{L^2}}$	Rate	$\frac{\|p_h^m - p(t_m)\|_{L^2}}{\|p(t_m)\|_{L^2}}$	Rate	$\frac{\|\theta_h^m - \theta(t_m)\|_{L^2}}{\|\theta(t_m)\|_{L^2}}$	Rate
1/4	7.899e-03	/	7.905e-04	/	5.054e-02	/	2.197e-03	/
1/8	8.041e-04	2.97	9.307e-05	3.79	1.233e-02	2.03	2.354e-04	3.22
1/16	8.104e-05	2.99	1.150e-05	2.84	3.032e-03	2.02	3.151e-05	2.90
1/32	9.816e-06	3.00	1.432e-06	3.12	7.567e-04	2.00	3.749e-06	3.07
1/64	1.205e-06	3.00	1.789e-07	3.01	1.891e-04	2.00	4.646e-07	3.01
1/128	1.495e-07	3.00	2.238e-08	3.00	4.740e-05	2.00	5.785e-08	3.01

Table 4: Case 2. H^1-errors, $t_m = 1$.

h	$\frac{\|\mathbf{u}_h^m - \mathbf{u}(t_m)\|_1}{\|\mathbf{u}(t_m)\|_1}$	Rate	$\frac{\|\mathbf{b}_h^m - \mathbf{b}(t_m)\|_1}{\|\mathbf{b}(t_m)\|_1}$	Rate	$\frac{\|\theta_h^m - \theta(t_m)\|_1}{\|\theta(t_m)\|_1}$	Rate
1/4	3.694e-02	/	2.744e-03	/	3.694e-02	/
1/8	9.354-e03	1.98	6.00e-04	2.19	9.354e-03	1.98
1/16	2.344e-03	2.00	1.470e-04	2.03	2.344e-03	2.00
1/32	5.860e-04	2.00	3.648e-05	2.01	5.860e-04	2.00
1/64	1.465e-04	2.00	9.104e-06	2.00	1.465e-04	2.00
1/128	3.662e-05	2.00	2.275e-06	2.00	3.662e-05	2.00

Table 5: Case 1. L^2-errors, $t_m = 1$.

	τ	$\frac{\|\mathbf{u}_h^m - \mathbf{u}(t_m)\|_{L^2}}{\|\mathbf{u}(t_m)\|_{L^2}}$	$\frac{\|\mathbf{b}_h^m - \mathbf{b}(t_m)\|_{L^2}}{\|\mathbf{b}(t_m)\|_{L^2}}$	$\frac{\|p_h^m - p(t_m)\|_{L^2}}{\|p(t_m)\|_{L^2}}$	$\frac{\|\theta_h^m - \theta(t_m)\|_{L^2}}{\|\theta(t_m)\|_{L^2}}$
$h = 1/40$	1/4	2.620e-03	1.024e-03	6.621e-03	5.406e-04
	1/6	1.043e-03	3.198e-04	1.601e-03	1.789e-04
	1/8	6.017e-04	1.321e-04	9.579e-04	4.947e-05
$h = 1/80$	1/4	2.620e-03	1.024e-03	6.604e-03	5.405e-04
	1/6	1.043e-03	3.198e-04	1.531e-03	1.789e-04
	1/8	6.015e-04	1.322e-04	8.356e-04	4.946e-05

Table 6: Case 2. L^2-errors, $t_m = 1$.

	τ	$\frac{\|\mathbf{u}_h^m - \mathbf{u}(t_m)\|_{L^2}}{\|\mathbf{u}(t_m)\|_{L^2}}$	$\frac{\|\mathbf{b}_h^m - \mathbf{b}(t_m)\|_{L^2}}{\|\mathbf{b}(t_m)\|_{L^2}}$	$\frac{\|p_h^m - p(t_m)\|_{L^2}}{\|p(t_m)\|_{L^2}}$	$\frac{\|\theta_h^m - \theta(t_m)\|_{L^2}}{\|\theta(t_m)\|_{L^2}}$
$h = 1/40$	1/4	1.217e-04	9.964e-04	4.744e-03	8.980e-04
	1/6	2.284e-05	3.263e-04	8.782e-04	3.273e-04
	1/8	6.554e-06	1.279e-04	1.005e-03	8.353e-05
$h = 1/80$	1/4	1.217e-04	9.965e-04	4.720e-03	8.979e-04
	1/6	2.275e-05	3.263e-04	7.425e-04	3.272e-04
	1/8	6.285e-06	1.279e-04	8.884e-04	8.348e-05

showed in Tables 5 and 6 indicate that the time step size τ is independent of the spatial mesh size h. Namely, the proposed fully discretized numerical scheme performs well and has no restriction on mesh ratio.

7. Conclusions

In this paper, we studied a third-order linearized scheme for thermally coupled MHD equations with temperature-dependent coefficients. The scheme is based on the third-order BDF discretization in time and mixed finite element approximations in space. In addition, we obtain almost unconditional optimal L^2-errors of the considered fully discrete scheme for the velocity, magnetic, and temperature. The keys to success of the proof are the temporal-spatial error splitting technique and the telescope formula for the third-order BDF temporal discretization operator. Several numerical tests are presented to verify the convergence properties of the proposed scheme.

Acknowledgments

The authors would like to thank the editors and anonymous referees for their reviews, which improved the quality of this paper.

This work is supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (Grant Nos. 2021D01E11, 2023D14014) and by the Tianshan Talent Training Program of Xinjiang Uygur Autonomous Region (Grant No. 2023TSYCCX0103).

References

[1] R.A. Adams and J.F. Fournier, *Sobolev Spaces*, Academic Press (2003).

[2] D.N. Arnold, F. Brezzi and M. Fortin, *A stable finite element for the Stokes equations*, Calcolo **21**, 337–344 (1984).

[3] G.A. Baker, V.A. Dougalis and O.A. Karakashian, *On a higher order accurate fully discrete Galerkin approximation to the Navier-Stokes equations*, Math. Comp. **39**, 339–375 (1982).

[4] Y.-Z. Chen and L.-C. Wu, *Second Order Elliptic Equations and Elliptic Systems*, AMS (1998).

[5] G. Cimatti, *Existence and uniqueness for magnetohydrodynamic flows in pipes with viscosity dependent on the temperature*, Electron. J. Differ. Equ. **59**, 1–8 (2009).

[6] P. Davidson, *An Introduction to Magnetohydrodynamics*, Cambridge University Press (2001).

[7] Q. Ding, X. Long and S. Mao, *Convergence analysis of Crank-Nicolson extrapolated fully discrete scheme for thermally coupled incompressible magnetohydrodynamic system*, Appl. Numer. Math. **157**, 522–543 (2020).

[8] X. Dong and Y. He, *Optimal convergence analysis of Crank-Nicolson extrapolation scheme for the threedimensional incompressible magnetohydrodynamics*, Comput. Math. Appl. **76**, 2678–2700 (2018).

[9] X. Dong, Y. He and Y. Zhang, *Convergence analysis of three finite element iterative methods for the 2D/3D stationary incompressible magnetohydrodynamics*, Comput. Methods Appl. Mech. Engrg. **276**, 287–311 (2014).

[10] H. Gao, *Unconditional optimal error estimates of BDF-Galerkin FEMs for nonlinear thermistor equations*, J. Sci. Comput. **66**, 504–527 (2016).

[11] H. Gao and W. Qiu, *A semi-implicit energy conserving finite element method for the dynamical incompressible magnetohydrodynamics equations*, Comput. Methods Appl. Mech. Engrg. **346**, 982–1001 (2019).

[12] J.F. Gerbeau, *A stabilized finite element method for the incompressible magnetohydrodynamic equations*, Numer. Math. **87**, 83–111 (2000).

[13] J.F. Gerbeau, C. Le Bris and T. Lelièvre, *Mathematical Methods for the Magnetohydrodynamics of Liquid Metals*, Oxford University Press (2006).

[14] A.V. Getling, *Rayleigh-Bénard Convection: Structures and Dynamics*, World Scientific (1998).

[15] C. Greif, D. Li, D. Schötzau and X. Wei, *A mixed finite element method with exactly divergence-free velocities for incompressible magnetohydrodynamics*, Comput. Methods Appl. Mech. Engrg. **199**, 2840–2855 (2010).

[16] J.L. Guermond and P.D. Minev, *Mixed finite element approximation of an MHD problem involving conducting and insulating regions: The 3D case*, Numer. Meth. Part. Differ. Equs. **19**, 709–731 (2003).

[17] Y. He, *Unconditional convergence of the Euler semi-implicit scheme for the three-dimensional incompressible MHD equations*, Comput. IMA J. Numer. Anal. **35**, 767–801 (2015).

[18] J.G. Heywood and R. Rannacher, *Finite-element approximation of the nonstationary Navier-Stokes problem. Part IV: Error analysis for second-order time discretization*, SIAM J. Numer. Anal. **27**, 353–384 (1990).

[19] P.Z. Huang, *A finite element algorithm for the nonstationary incompressible magnetohydrodynamic system based on a correction method*, Mediterr. J. Math. **19**, 113 (2022).

[20] F. Kanbar, R. Touma and C. Klingenberg, *Well-balanced central scheme for the system of MHD equations with gravitational source term*, Commun. Comput. Phys. **32**, 878–898 (2022).

[21] A. Keram and P.Z. Huang, *The Arrow-Hurwicz iterative finite element method for the stationary thermally coupled incompressible magnetohydrodynamics flow*, J. Sci. Comput. **92**, 11 (2022).

[22] B. Li, *A bounded numerical solution with a small mesh size implies existence of a smooth solution to the Navier-Stokes equations*, Numer. Math. **147**, 283–304 (2021).

[23] B. Li, S. Ma and C. Ueda, *Analysis of fully discrete finite element methods for 2D Navier-Stokes equations with critical initial data*, ESAIM Math. Model. Numer. Anal. **56**, 2105–2139 (2022).

[24] B. Li and W. Sun, *Unconditional convergence and optimal error estimates of a Galerkin-mixed FEM for incompressible miscible flow in porous media*, SIAM J. Numer. Anal. **51**, 1959–1977 (2013).

[25] B. Li, J. Wang and L. Xu, *A convergent linearized Lagrange finite element method for the magneto-hydrodynamic equations in two-dimensional nonsmooth and nonconvex domains*, SIAM J. Numer. Anal. **58**, 430–459 (2020).

[26] J. Liu, *Simple and efficient ALE methods with provable temporal accuracy up to fifth order for the Stokes equations on time varying domains*, SIAM J. Numer. Anal. **51**, 743–772 (2013).

[27] S. Liu, P. Huang and Y. He, *An optimal error estimates of the BDF-Galerkin FEM for the incompressible MHD system*, J. Math. Anal. Appl. **515**, 126460 (2022).

[28] X. Lu and P.Z. Huang, *A modular grad-div stabilization for the 2D/3D nonstationary incompressible magnetohydrodynamic equations*, J. Sci. Comput. **82**, 3 (2020).

[29] C. Lubich, D. Mansour and C. Venkataraman, *Backward difference time discretization of parabolic differential equations on evolving surfaces*, IMA J. Numer. Anal. **33**, 1365–1385 (2013).

[30] H. Ma and P. Huang, *A vector penalty-projection approach for the time-dependent incompressible magnetohydrodynamics flows*, Comput. Math. Appl. **120**, 28–44 (2022).

[31] A.J. Meir, *Thermally coupled magnetohydrodynamics flow*, Appl. Math. Comput. **65**, 79–94 (1994).

[32] A.J. Meir, *Thermally coupled, stationary, incompressible MHD flow: Existence, uniqueness, and finite element approximation*, Numer. Meth. Part. Differ. Equs. **11**, 311–337 (1995).

[33] A.J. Meir and P.G. Schmidt, *On electromagnetically and thermally driven liquid-metal flows*, Nonlinear Anal.-Theory Methods Appl. **47**, 3281–3294 (2001).

[34] P. Monk, *Finite Element Methods for Maxwell's Equations*, Oxford University Press (2003).

[35] O. Nevanlinna and F. Odeh, *Multiplier techniques for linear multistep methods*, Funct. Anal. Optim. **3**, 377–423 (1981).

[36] L. Nirenberg, *An extended interpolation inequality*, Ann. Scuola Norm. Sup. Pisa-Cl. Sci. **20**, 733–737 (1966).

[37] E.R. Priest and A.W. Hood, *Advances in Solar System Magnetohydrodynamics*, Cambridge University Press (1991).

[38] A. Prohl, *Convergent finite element discretizations of the nonstationary incompressible magnetohydrodynamics system*, ESAIM Math. Model. Numer. Anal. **42**, 1065–1087 (2008).

[39] H. Qiu, *Error analysis of Euler semi-implicit scheme for the nonstationary magneto-hydrodynamics problem with temperature dependent parameters*, J. Sci. Comput. **85**, 47 (2020).

[40] H. Qiu, *Well-posedness and finite element approximation for the stationary magneto-hydrodynamics problem with temperature-dependent parameters*, J. Sci. Comput. **85**, 58 (2020).

[41] S.S. Ravindran, *An extrapolated second order backward difference time-stepping scheme for the magnetohydrodynamics system*, Numer. Func. Anal. Opt. **37**, 990–1020 (2016).

[42] S.S. Ravindran, *Partitioned time-stepping scheme for an MHD system with temperature-dependent coefficients*, IMA J. Numer. Anal. **39**, 1860–1887 (2019).

[43] D. Schötzau, *Mixed finite element methods for stationary incompressible magnetohydrodynamics*, Numer. Math. **96**, 771–800 (2004).

[44] R. Temam, *Navier-Stokes Equations, Theory and Numerical Analysis*, North-Holland (1983).

[45] V. Thomee, *Galerkin Finite Element Methods for Parabolic Problems*, Springer (2006).

[46] B. Wacker, D. Arndt and G. Lube, *Nodal-based finite element methods with local projection stabilization for linearized incompressible magnetohydrodynamics*, Comput. Methods Appl. Mech. Engrg. **302**, 170–192 (2016).

[47] C. Wang, J. Wang, Z. Xia and L. Xu, *Optimal error estimates of a Crank-Nicolson finite element projection method for magnetohydrodynamic equations*, ESAIM Math. Model. Numer. Anal. **56**, 767–789 (2022).

[48] D. Wang, *Large solutions to the initial-boundary value problem for planar magnetohydrodynamics*, SIAM J. Appl. Math. **63**, 1424–1441 (2003).

[49] J. Wang, *Superconvergence analysis for a semilinear parabolic equation with BDF-3 finite element method*, Appl. Anal. **101**, 1822–1832 (2020).

[50] J. Wang and Q. Li, *Superconvergence analysis of a linearized three-step backward differential formula finite element method for nonlinear Sobolev equation*, Math. Meth. Appl. Sci. **42**, 3359–3376 (2019).

[51] K. Wang and Y. He, *Convergence analysis for a higher order scheme for the time-dependent Navier-Stokes equations*, Appl. Math. Comput. **218**, 8269–8278 (2012).

[52] L. Wang, J. Li and P.Z. Huang, *An efficient two-level algorithm for the 2D/3D stationary incompressible magnetohydrodynamics based on the finite element method*, Int. Commun. Heat Mass Transf. **98**, 183–190 (2018).

[53] K. Wu, D. Xiu and X. Zhong, *A WENO-based stochastic Galerkin scheme for ideal MHD equations with random inputs*, Commun. Comput. Phys. **30**, 423–447 (2021).

[54] J. Yang and Y. He, *Stability and error analysis for the first-order Euler implicit/explicit scheme for the 3D MHD equations*, Int. J. Comput. Methods **15**, 1750077 (2018).

[55] G. Yuksel and R. Ingram, *Numerical analysis of a finite element, Crank-Nicolson discretiza-*

tion for MHD flow at small magnetic Reynolds number, Int. J. Numer. Anal. Model. **10**, 74–98 (2013).

[56] G. Yuksel and O.R. Isik, *Numerical analysis of Backward-Euler discretization for simplified magnetohydrodynamic flows*, Appl. Math. Model. **39**, 1889–1898 (2015).

[57] G. Zhang, J. Yang and C. Bi, *Second order unconditionally convergent and energy stable linearized scheme for MHD equations*, Adv. Comput. Math. **44**, 505–541 (2018).