INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 22, Number 2, Pages 202–225

doi: 10.4208/ijnam2025-1010

HYBRIDIZABLE DISCONTINUOUS GALERKIN METHOD FOR LINEAR HYPERBOLIC INTEGRO-DIFFERENTIAL EQUATIONS

RIYA JAIN AND SANGITA YADAV*

Abstract. This article introduces the hybridizable discontinuous Galerkin (HDG) approach to numerically approximate the solution of a linear hyperbolic integro-differential equation. A priori error estimates for semi-discrete and fully discrete schemes are developed. It is shown that the optimal order of convergence is achieved for both scalar and flux variables. To achieve that, an intermediate projection is introduced for the semi-discrete error analysis, and it also shows that this projection achieves convergence of order $h^{k+3/2}$ for $k \geq 1$. Next, superconvergence is achieved for the scalar variable using element-by-element post-processing. For the fully discrete error analysis, the central difference scheme and the mid-point rule approximate the derivative and the integral term, respectively. Hence, the second order of convergence is achieved in the temporal direction. Finally, numerical experiments have been performed to validate the theory developed in this article.

Key words. Hyperbolic integro-differential equation, hybridizable discontinuous Galerkin method, Ritz-Volterra projection, a priori error bounds, post-processing.

1. Introduction

Throughout this paper, we have discussed HDG method for the following model problem:

(1a)
$$u_{tt}(x,t) - \nabla \cdot \left(a(x)\nabla u(x,t) + \int_0^t b(x,t,s)\nabla u(x,s)ds \right) = f(x,t) \quad \text{in } \Omega \times (0,T],$$

(1b)
$$u(x,t) = 0 \quad \text{on } \partial\Omega \times (0,T],$$

(1c)
$$u(x,0) = u_0(x) \ \forall\ x \in \Omega,$$

(1d)
$$u_t(x,0) = u_1(x) \ \forall\ x \in \Omega,$$

where $u_{tt}(x,t) = \dfrac{\partial^2 u}{\partial t^2}(x,t)$ and $u : \Omega\times(0,T] \to \mathbb{R}$. The functions $f : \Omega\times(0,T] \to \mathbb{R}$, $u_0 : \Omega \to \mathbb{R}$ and $u_1 : \Omega \to \mathbb{R}$ are known. We have assumed the following properties to be true on the domain Ω, it is convex, polygonal and bounded in $\mathbb{R}^2$ with smooth boundary $\partial\Omega$. The known functions $a : \Omega \to \mathbb{R}$ and $b : \Omega\times(0,T]\times(0,T] \to \mathbb{R}$ satisfy the following properties: function a is positive and bounded. There exists $\alpha_0 > 0$, $M > 0$ such that $0 < \alpha_0 \leq a \leq M$, whereas, b is smooth and twice differentiable with bounded derivatives and $|b| \leq M$. Such classes of problems and nonlinear versions, thereof arise naturally in many applications, such as in viscoelasticity, see [28] and references therein.

In the literature, Pani *et al.* [2] have analyzed fully discrete schemes for time-dependent partial integro-differential equations, using energy methods, paying attention to the storage required during time-stepping. Further, errors are estimated in L^2 and H^1 norms. In [13], Saedpanah has formulated a continuous space-time finite element method of degree one for an integro-differential equation of hyperbolic

Received by the editors on October 3, 2023 and, accepted on October 8, 2024.
2000 *Mathematics Subject Classification.* 65M60.
*Corresponding author.

type with mixed boundary conditions. Further, a posteriori error estimates are also established. Then, in [14], a first-order continuous space-time finite element method for a hyperbolic integro-differential equation has been formulated. Moreover, the theory is illustrated with the help of an example. In [12], Karaa *et al.* have applied DG method to (1). A priori error estimates are derived for both scalar as well as for vector variables, and the optimal rate of convergence is derived for the scalar variable and suboptimal rate of convergence for vector variables. In [3], Karaa *et al.* have discussed mixed finite element methods for the model problem (1). They have derived $L^\infty(L^2)$ and $L^\infty(L^\infty)$ error estimates and have shown to achieve optimal and quasi-optimal order of convergence, respectively, with minimal smoothness on the initial data. Later, in [15], Merad *et al.* proposed a Galerkin method based on least squares for a two-dimensional hyperbolic integro-differential equation with purely integral conditions. They have also discussed the existence and uniqueness of the solution of the model problem under specific conditions. In [4], Chen *et al.* have proposed a two-grid finite element scheme for a semi-linear hyperbolic integro-differential equation, which uses two grids to deal with the semi-linearity of the problem and achieves the same order of accuracy as that of the ordinary finite element method. Recently, Tan *et al.* [16] have applied a fully discrete two-grid finite element method on a hyperbolic integro-differential equation and achieved optimal order of convergence. The scheme has reduced the computational cost while maintaining numerical accuracy.

HDG method is a numerical technique for solving partial differential equations (PDEs) that combines the accuracy of the discontinuous Galerkin (DG) method with the computational efficiency of other finite element methods. HDG method was first introduced by Cockburn [6, 5, 7, 8], and has since been applied to a wide range of problems. In HDG method, the solution is approximated using piecewise polynomial functions, similar to the DG method, but with additional degrees of freedom that are defined at the element interfaces. These additional degrees of freedom are used to enforce the continuity of the solution across the element interfaces, which leads to a more efficient and accurate method than the standard DG method. HDG method has several advantages over other finite element methods, including the ability to handle complex geometries and nonlinear equations and the ability to achieve high-order accuracy with fewer degrees of freedom. HDG method has been successfully applied to a variety of PDE, including the Navier-Stokes equations [17, 18], the Maxwell equations [19, 20], and the advection-diffusion equation [21, 22], etc. In addition, the method has been extended to include time-dependent problems, such as the heat equation [11, 23], the wave equation [24, 25] and parabolic integro-differential equation [26]. In [11], Chabaud *et al.* have extended the analysis of HDG method and applied to second-order elliptic equations for the heat equation. They have shown that the superconvergence results hold for the heat equation when the HDG method is used to semi-discretize the equation. Further, in [23], Moon *et al.* have analyzed the method for the heat equation with nonlinear coefficients, which satisfy the Lipschitz condition. As far as the wave equations are considered, in [24], Cockburn *et al.* have analyzed the error estimates of the acoustic equation and have achieved optimal order of convergence for velocity as well as its gradient. They have also discussed the superconvergence result for the same. Stanglmeier *et al.* [25] has developed an explicit HDG method for acoustic wave equation that yields optimal convergence rates for the approximation of all the unknown variables and discussed some superconvergence properties. Recently, Jain *et al.* [26] have analyzed the HDG method for linear parabolic integro-differential equations, along with superconvergence results. Overall, HDG method is a powerful and versatile numerical method for solving PDEs, combining the accuracy of the DG method with the computational efficiency of other finite element methods. Its ability to handle complex geometries and nonlinear equations, and to achieve high-order accuracy with fewer degrees of freedom, make it a promising approach for a wide range of applications.

This paper analyzes HDG method for the model problem (1) and discusses *a priori* error estimates. The most significant points of this article are as follows:

- In contrast to DG methods, optimal convergence rates have been obtained for scalar as well as vector variables.
- For the semi-discrete error analysis, two intermediate projections, namely, HDG projection and the Ritz-Volterra projection, have been used for the model problem.
- For the scalar variable, a new post-processed approximation has been defined, which achieves the superconvergence.
- Mid-point rule and central difference scheme are used to approximate the integral and the derivative term, respectively, to achieve second order of convergence in the temporal direction.
- The theoretical results are verified by implementing HDG method for problems in the 2-dimensional domain.

The organization of the article is as follows: In Section 2, a few important assumptions and results are stated, which will be used throughout. Section 3 defines HDG formulation for the model problem (1). Section 4 provides the highlights of the paper by stating all the essential results of the article. In Section 5, a priori error estimates are derived using a few crucial steps. In Section 6, the superconvergence results for the scalar variable are analyzed. Section 7 is about discretizing the scheme in temporal direction. Finally, in Section 8, some numerical results are given to verify the theoretical findings. A few concluding remarks are made in Section 9.

2. Preliminaries

Let us denote a finite element subset of the domain Ω with the shape regularity by $\mathcal{T}_h$. For any $K \in \mathcal{T}_h$, K is a triangle or rectangle. Let the radius of the biggest ball in K be ρ_K and the diameter of K is h_K and $\rho = \min\limits_{K\in\mathcal{T}_h} \rho_K$, $h = \max\limits_{K\in\mathcal{T}_h} h_K$. The set of interior edges is denoted by Γ_I, the set of boundary edges by Γ_∂, and the set $\partial K : K \in \mathcal{T}_h$ is denoted by $\partial\mathcal{T}_h$. Lastly, we write $\Gamma = \Gamma_I \cup \Gamma_\partial$.
We take into account the following sets of finite elements:

$$\begin{aligned} V_h &= \{v \in L^2(\Omega) : v|_K \in P_k(K), \forall K \in \mathcal{T}_h\},\\ \boldsymbol{W}_h &= \{\boldsymbol{w} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{w}|_K \in \boldsymbol{P}_k(K), \forall K \in \mathcal{T}_h\},\\ M_h &= \{\mu \in L^2(\Gamma) : \mu|_F \in P_k(F), \forall F \in \Gamma\}.\end{aligned}$$

In this case, $\boldsymbol{P}_k(K) = [P_k(K)]^2$, and the space of polynomials defined on K with a maximum degree k is denoted by $P_k(K)$.
Next, let $u, v \in L^2(D)$, define $(u, v)_D = \displaystyle\int_D uv$, when the domain D is a subset of $\mathbb{R}^n$. If ∂D is in $\mathbb{R}^{n-1}$, define $\langle u, v\rangle_{\partial D} = \displaystyle\int_{\partial D} uvds$. Then, we introduce the following

notations:

$$(u,v) = \sum_{K\in\mathcal{T}_h} (u,v)_K \quad \text{with norm} \quad \|v\|^2 = \sum_{K\in\mathcal{T}_h} \|v\|^2_{L^2(K)},$$

$$\langle u,v\rangle_{\partial\mathcal{T}_h} = \sum_{K\in\mathcal{T}_h} \langle u,v\rangle_{\partial K} \quad \text{with norm} \quad \|\mu\|^2_\tau = \sum_{K\in\mathcal{T}_h} \tau\|\mu\|^2_{L^2(\partial K)}.$$

In addition, we use the following fragmented Sobolev spaces:

$$H^r(\mathcal{T}_h) = \big\{v \in L^2(\Omega) : \sum_{K\in\mathcal{T}_h} \|v\|^2_{H^r(K)} < \infty\big\},$$

with norm

$$\|v\|_{H^r(\mathcal{T}_h)} = \Big(\sum_{K\in\mathcal{T}_h} \|v\|^2_{H^r(K)}\Big)^{\frac{1}{2}},$$

where $H^r(K)$ is the Sobolev space of order r defined on K.

Lemma 2.1. *(L^2-projection) Let $\boldsymbol{I}_h^k$ denote the L^2-projection. If $\boldsymbol{w} \in \boldsymbol{H^{r+1}}(K)$ and $\boldsymbol{I}_h^k\boldsymbol{w} \in \boldsymbol{P_k}(K)$, the subsequent approximation property holds:*

$$\|\boldsymbol{w} - \boldsymbol{I}_h^k\boldsymbol{w}\|_{\boldsymbol{L^2}(K)} + h^{\frac{1}{2}}\|\boldsymbol{w} - \boldsymbol{I}_h^k\boldsymbol{w}\|_{\boldsymbol{L^2}(\partial K)} \le Ch^{min(r,k)+1}\|\boldsymbol{w}\|_{\boldsymbol{H^{r+1}}(K)}.$$

HDG Projection and related estimates.
The projection Π_h into $V_h \times \boldsymbol{W}_h$, which was first introduced in [9], is defined as follows:
Given $(u,\boldsymbol{z}) \in H^1(\mathcal{T}_h)\times\boldsymbol{H}_{div}(\mathcal{T}_h)$, for $k \ge 0$ and $\tau|_{\partial K}$ non negative, $\tau = \max\tau|_{\partial K} > 0$, the function $\Pi_h(u,\boldsymbol{z}) = (\Pi_V u, \boldsymbol{\Pi_W z})$ on an arbitrary simplex $K \in \mathbb{T}_h$ is the element of $V_h \times \boldsymbol{W}_h$ which uniquely solves

$$(\Pi_V u, v)_K = (u,v)_K, \qquad \forall\, v \in P_{k-1}(K) \tag{2a}$$

$$(\boldsymbol{\Pi_W z}, \boldsymbol{w})_K = (\boldsymbol{z},\boldsymbol{w})_K, \qquad \forall\, \boldsymbol{w} \in \boldsymbol{P}_{k-1}(K) \tag{2b}$$

$$\langle \boldsymbol{\Pi_W z}\cdot\boldsymbol{\nu} + \tau\Pi_V u, \mu\rangle_F = \langle \boldsymbol{z}\cdot\boldsymbol{\nu} + \tau u, \mu\rangle_F, \qquad \forall\, \mu \in P_k(F), \tag{2c}$$

for all faces F of the simplex K.
In addition, there exists a constant C that does not rely on K and τ, $\forall\, 1 \le \alpha,\beta \le k+1$, then:

$$\|\boldsymbol{\Pi_W z} - \boldsymbol{z}\|_K \le Ch^\alpha\|\boldsymbol{z}\|_{\boldsymbol{H^\alpha}(K)} + C\tau_K^* h^\beta\|u\|_{H^\beta(K)}, \tag{3a}$$

$$\|\Pi_V u - u\|_K \le Ch^\beta\|u\|_{H^\beta(K)} + C\frac{h^\alpha}{\tau_K^{\max}}\|\nabla\cdot\boldsymbol{z}\|_{\boldsymbol{H^{\alpha-1}}(K)}. \tag{3b}$$

Here, $\tau_K^* := \max\tau_{|\partial K\setminus F^*}$, where F^* is a face of K at which $\tau_{|\partial K}$ is maximum and $\tau_K^{\max} := \max\tau_{|\partial K} > 0$. For more details, please refer to [9].

3. HDG Scheme

Throughout this article, we have used the following auxiliary variable in $\Omega \times (0,T]$:

$$\boldsymbol{\sigma} = -\nabla u, \quad \boldsymbol{z} = a\boldsymbol{\sigma} + \int_0^t b(t,s)\boldsymbol{\sigma}(s)ds.$$

Using these variables, (1) is rewritten as follows:

$$\boldsymbol{\sigma} = -\nabla u \quad \text{in } \Omega\times(0,T], \tag{4a}$$

$$\boldsymbol{z} = a\boldsymbol{\sigma} + \int_0^t b(t,s)\boldsymbol{\sigma}(s)ds \quad \text{in } \Omega\times(0,T], \tag{4b}$$

$$u_{tt} + \nabla\cdot\boldsymbol{z} = f \quad \text{in } \Omega\times(0,T]. \tag{4c}$$

For each time t on the interval $(0,T]$, the method yields a scalar approximation $u_h(t)$ to $u(t)$, a vector approximation $\boldsymbol{\sigma}_h(t)$ to $\boldsymbol{\sigma}(t)$, $\boldsymbol{z}_h(t)$ to $\boldsymbol{z}(t)$, and a scalar approximation $\hat{u}_h(t)$ to the trace of $u(t)$ on element boundaries, in the spaces V_h, $\boldsymbol{W_h}$, $\boldsymbol{W_h}$ and M_h, respectively.

With these spaces, HDG formulation seeks approximation $(u_h,\boldsymbol{\sigma}_h,\boldsymbol{z}_h,\hat{u}_h)(t)\in V_h\times\boldsymbol{W}_h\times\boldsymbol{W}_h\times M_h$, for $t\in(0,T]$, that satisfy the following equations:

$$(\boldsymbol{\sigma}_h,\boldsymbol{w}_h) - (u_h,\nabla\cdot\boldsymbol{w}_h) + \langle\hat{u}_h,\boldsymbol{w}_h\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{5a}$$

$$(a\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h) - (\boldsymbol{z}_h,\boldsymbol{\tau}_h) + \int_0^t (b(t,s)\boldsymbol{\sigma}_h(s),\boldsymbol{\tau}_h)ds = 0, \tag{5b}$$

$$(u_{h_{tt}},v_h) - (\boldsymbol{z}_h,\nabla v_h) + \langle\hat{\boldsymbol{z}}_h\cdot\boldsymbol{\nu},v_h\rangle_{\partial\mathcal{T}_h} = (f,v_h), \tag{5c}$$

$$\langle\hat{u}_h,\mu_h\rangle_{\partial\Omega} = 0, \tag{5d}$$

$$\langle\hat{\boldsymbol{z}}_h\cdot\boldsymbol{\nu},m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{5e}$$

$$u_h(0) = \Pi_V u_0, \tag{5f}$$

$$u_{h_t}(0) = \Pi_V u_1, \tag{5g}$$

for any $(v_h,\boldsymbol{w}_h,\boldsymbol{\tau}_h,\mu_h,m_h)\in V_h\times\boldsymbol{W}_h\times\boldsymbol{W}_h\times M_h\times M_h$, along with the following relation:

$$\hat{\boldsymbol{z}}_h\cdot\boldsymbol{\nu} = \boldsymbol{z}_h\cdot\boldsymbol{\nu} + \tau(u_h-\hat{u}_h) \quad \text{on } \partial\mathcal{T}_h,$$

where, $\tau\geq 0$ on Γ and piece-wise constant on the faces. Now, with the help of (4) and (5), we have the following:

$$(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{w}_h) - (u-u_h,\nabla\cdot\boldsymbol{w}_h) + \langle u-\hat{u}_h,\boldsymbol{w}_h\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0 \quad \forall\boldsymbol{w}_h\in\boldsymbol{W}_h, \tag{6a}$$

(6b)

$$(a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\boldsymbol{\tau}_h) - (\boldsymbol{z}-\boldsymbol{z}_h,\boldsymbol{\tau}_h) + \int_0^t (b(t,s)(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)(s),\boldsymbol{\tau}_h)ds = 0 \quad \forall\boldsymbol{\tau}_h\in\boldsymbol{W}_h,$$

$$(u_{tt}-u_{h_{tt}},v_h) - (\boldsymbol{z}-\boldsymbol{z}_h,\nabla v_h) + \langle(\boldsymbol{z}-\hat{\boldsymbol{z}}_h)\cdot\boldsymbol{\nu},v_h\rangle_{\partial\mathcal{T}_h} = 0 \quad \forall v_h\in V_h, \tag{6c}$$

$$\langle u-\hat{u}_h,\mu_h\rangle_{\partial\Omega} = 0 \quad \forall\mu_h\in M_h, \tag{6d}$$

$$\langle(\boldsymbol{z}-\hat{\boldsymbol{z}}_h)\cdot\boldsymbol{\nu},m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0 \quad \forall m_h\in M_h. \tag{6e}$$

4. The Main Results

In this section, we state the main results of the paper in the form of the following theorems:

Theorem 4.1. *Let $(u,\boldsymbol{\sigma},\boldsymbol{z})$ be the solution of (4) with $u\in L^\infty(H^{k+2}(\mathcal{T}_h))$ and $u_t,u_{tt}\in L^2(H^{k+2}(\mathcal{T}_h))$, $u_0,u_1\in H^{k+2}(\mathcal{T}_h)$ for $k\geq 0$. Additionally, let $(u_h,\boldsymbol{\sigma}_h,\boldsymbol{z}_h,\hat{u}_h)\in V_h\times\boldsymbol{W}_h\times\boldsymbol{W}_h\times M_h$ be the solution of (5) along with $u_h(0)=\Pi_V u_0$, $u_{h_t}(0)=I_h^k u_1$, $\boldsymbol{\sigma}_h(0)=-\boldsymbol{I}_h^k\nabla u_0$, $\boldsymbol{z}_h(0)=\boldsymbol{\Pi_W}(a\nabla u_0)$ and $\hat{u}_h(0)=P_M u_0$, where P_M is the L^2 projection onto M_h. Consequently, the following estimations hold:*

$$\|(u-u_h)(t)\| + \|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)(t)\| + \|(\boldsymbol{z}-\boldsymbol{z}_h)(t)\| + \|(u-\hat{u}_h)(t)\|_\tau \leq Ch^{k+1},$$

$$\|(u_t-u_{h_t})(t)\| \leq Ch^{k+1}.$$

For the next result, we define the post- processed solution $u_h^* \in P_{k+1}(K)$ on the element K, as

$$u_h^* = u_h^p + \frac{1}{|K|}\int_K u_h, \quad u_h^p \in P_{k+1}^0 \tag{7}$$

where u_h^p satisfies

$$(a\nabla u_h^p, \nabla v) = -(a\boldsymbol{\sigma}_h, \nabla v), \quad \forall v \in P_{k+1}^0 \tag{8}$$

where $P_{k+1}^0(K)$ represents the collection of polynomials in $P_{k+1}(K)$ with zero average. The next theorem gives the L^2 estimates of u_h^*.

Theorem 4.2. *Under the conditions of Theorem 4.1, there exists a positive constant C independent of h and k such that*

$$\|u - u_h^*\| \leq Ch^{k+2}.$$

Theorem 4.3. *Let $\Delta t = \dfrac{T}{M}$, for some positive integer M, and $t_n = n\Delta t$, for $1 \leq n \leq M$. Let $(U^n, \boldsymbol{S}^n, \boldsymbol{Z}^n) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h$ be the fully discrete approximations of u, $\boldsymbol{\sigma}$ and $\boldsymbol{z}$ and $\hat{U}^n \in M_h$ be the approximation of u on Γ, along with $U^0 = \Pi_V u_0$, $\boldsymbol{S}^0 = -\boldsymbol{I}_h^k \nabla u_0$, $\boldsymbol{Z}^0 = \boldsymbol{\Pi_W}(a\nabla u_0)$ and $\hat{U}^0 = P_M u_0$, where P_M is the L^2 projection onto M_h. Then, we have the following estimates:*

$$\|\partial_t \Upsilon \zeta_u^n\| + \|\Upsilon \boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n\| + \|\Upsilon \boldsymbol{\zeta}_{\boldsymbol{z}}^n\| + \|\Upsilon \hat{\zeta}_u^n\|_\tau \leq O(h^{k+1} + \Delta t^2), \tag{9a}$$

$$\|\zeta_u^{n+1}\| \leq O(h^{k+1} + \Delta t^2). \tag{9b}$$

where, $\zeta's$ are defined as: $\zeta_u^n = u_h(t_n) - U^n$. Similarly, $\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n$, $\boldsymbol{\zeta}_{\boldsymbol{z}}^n$ and $\hat{\zeta}_u^n$. Also, $\Upsilon U^n = \dfrac{U^{n+1} + U^n}{2}$ and $\partial_t \Upsilon U^n = \dfrac{U^{n+1} - U^n}{\Delta t}$, see Section 7.

5. Semi-Discrete Error Analysis

In this section, we provide detailed proofs of Theorem 4.1.

STEP I: Extended type Ritz-Voterra projection. For each $t \in (0, T]$, $(\tilde{u}_h, \tilde{\boldsymbol{\sigma}}_h, \tilde{\boldsymbol{z}}_h, \hat{\tilde{u}}_h) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h$ is defined as the Ritz-Volterra projection, provided it satisfy the following equations:

$$(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h, \boldsymbol{w}_h) - (u - \tilde{u}_h, \nabla \cdot \boldsymbol{w}_h) + \langle u - \hat{\tilde{u}}_h, \boldsymbol{w}_h \cdot \boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{10a}$$

$$(a(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h), \boldsymbol{\tau}_h) - (\boldsymbol{z} - \tilde{\boldsymbol{z}}_h, \boldsymbol{\tau}_h) + \int_0^t (b(t,s)(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h)(s), \boldsymbol{\tau}_h) ds = 0, \tag{10b}$$

$$-(\boldsymbol{z} - \tilde{\boldsymbol{z}}_h, \nabla v_h) + \langle (\boldsymbol{z} - \hat{\tilde{\boldsymbol{z}}}_h) \cdot \boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = 0, \tag{10c}$$

$$\langle u - \hat{\tilde{u}}_h, \mu_h\rangle_{\partial\Omega} = 0, \tag{10d}$$

$$\langle (\boldsymbol{z} - \hat{\tilde{\boldsymbol{z}}}_h) \cdot \boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h \setminus \partial\Omega} = 0, \tag{10e}$$

$\forall\ v_h \in V_h$, $\boldsymbol{w}_h, \boldsymbol{\tau}_h \in \boldsymbol{W}_h$ and $\mu_h, m_h \in M_h$, where

$$\hat{\tilde{\boldsymbol{z}}}_h \cdot \boldsymbol{\nu} = \tilde{\boldsymbol{z}}_h \cdot \boldsymbol{\nu} + \tau(\tilde{u}_h - \hat{\tilde{u}}_h) \;\; \text{on } \partial\mathcal{T}_h.$$

In order to derive the estimates of the Ritz-Volterra projection, we disintegrate it as follows:

$$\begin{aligned}
\eta_u &:= u - \tilde{u}_h = (u - \Pi_V u) - (\tilde{u}_h - \Pi_V u) = \theta_u - \rho_u,\\
\boldsymbol{\eta_\sigma} &:= \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h = (\boldsymbol{\sigma} - \boldsymbol{I}_h^k\boldsymbol{\sigma}) - (\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{I}_h^k\boldsymbol{\sigma}) = \boldsymbol{\theta_\sigma} - \boldsymbol{\rho_\sigma},\\
\boldsymbol{\eta_z} &:= \boldsymbol{z} - \tilde{\boldsymbol{z}}_h = (\boldsymbol{z} - \boldsymbol{\Pi_W z}) - (\tilde{\boldsymbol{z}}_h - \boldsymbol{\Pi_W z}) = \boldsymbol{\theta_z} - \boldsymbol{\rho_z},\\
\hat{\eta}_u &:= u - \hat{\tilde{u}}_h = (u - P_M u) - (\hat{\tilde{u}}_h - P_M u) = \hat{\theta}_u - \hat{\rho}_u,\\
\hat{\boldsymbol{\eta}}_{\boldsymbol{z}} &:= \boldsymbol{z} - \hat{\tilde{\boldsymbol{z}}}_h = (\boldsymbol{z} - \boldsymbol{P_M z}) - (\hat{\tilde{\boldsymbol{z}}}_h - \boldsymbol{P_M z}) = \hat{\boldsymbol{\theta}}_{\boldsymbol{z}} - \hat{\boldsymbol{\rho}}_{\boldsymbol{z}},
\end{aligned}$$

where, P_M is the L^2 projection onto M_h.
Therefore, the system of equations (10) can be rewritten as

$$(\boldsymbol{\rho_\sigma}, \boldsymbol{w_h}) - (\rho_u, \nabla\cdot\boldsymbol{w_h}) + \langle\hat{\rho}_u, \boldsymbol{w_h}\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{11a}$$

$$(a\boldsymbol{\rho_\sigma}, \boldsymbol{\tau_h}) - (\boldsymbol{\rho_z}, \boldsymbol{\tau}_h) + \int_0^t (b(t,s)\boldsymbol{\rho_\sigma}(s), \boldsymbol{\tau}_h)ds = (a\boldsymbol{\theta_\sigma}, \boldsymbol{\tau}_h) - (\boldsymbol{\theta_z}, \boldsymbol{\tau}_h) + \int_0^t (b(t,s)\boldsymbol{\theta_\sigma}(s), \boldsymbol{\tau}_h)ds, \tag{11b}$$

$$-(\boldsymbol{\rho_z}, \nabla v_h) + \langle\hat{\boldsymbol{\rho}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = 0, \tag{11c}$$

$$\langle\hat{\rho}_u, \mu_h\rangle_{\partial\Omega} = 0, \tag{11d}$$

$$\langle\hat{\boldsymbol{\rho}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{11e}$$

for all $(v_h, \boldsymbol{w}_h, \boldsymbol{\tau}_h, \mu_h, m_h) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h$.

STEP II: Estimates of $\left\|\frac{\partial^l \boldsymbol{\rho_z}}{\partial t^l}\right\|$, $\left\|\frac{\partial^l \boldsymbol{\rho_\sigma}}{\partial t^l}\right\|$ and $\left\|\frac{\partial^l \rho_u}{\partial t^l}\right\|$ for $l = 0, 1, 2$.

Lemma 5.1. *There exists a positive constant C which does not rely on h and k such that $\forall t \in (0, T]$, the inequality below is valid for $l = 0, 1, 2$*

$$\left\|\frac{\partial^l \rho_u}{\partial t^l}\right\| + \left\|\frac{\partial^l \boldsymbol{\rho_\sigma}}{\partial t^l}\right\| + \left\|\frac{\partial^l \boldsymbol{\rho_z}}{\partial t^l}\right\| \leq Ch^{k+1}. \tag{12}$$

Proof. For $l = 0, 1$, see [26]. For $l = 2$, we begin by differentiating (11a)-(11e) twice w.r.t. t, to obtain

$$(\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{w_h}) - (\rho_{u_{tt}}, \nabla\cdot\boldsymbol{w_h}) + \langle\hat{\rho}_{u_{tt}}, \boldsymbol{w_h}\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{13a}$$

$$(a\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\tau_h}) - (\boldsymbol{\rho_{z_{tt}}}, \boldsymbol{\tau}_h) + \frac{\partial^2}{\partial t^2}\left(\int_0^t (b(t,s)\boldsymbol{\rho_\sigma}(s), \boldsymbol{\tau}_h)ds\right) = (a\boldsymbol{\theta_{\sigma_{tt}}}, \boldsymbol{\tau}_h) - (\boldsymbol{\theta_{z_{tt}}}, \boldsymbol{\tau}_h) + \frac{\partial^2}{\partial t^2}\left(\int_0^t (b(t,s)\boldsymbol{\theta_\sigma}(s), \boldsymbol{\tau}_h)ds\right), \tag{13b}$$

$$-(\boldsymbol{\rho_{z_{tt}}}, \nabla v_h) + \langle\hat{\boldsymbol{\rho}}_{\boldsymbol{z_{tt}}}\cdot\boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = 0, \tag{13c}$$

$$\langle\hat{\rho}_{u_{tt}}, \mu_h\rangle_{\partial\Omega} = 0, \tag{13d}$$

$$\langle\hat{\boldsymbol{\rho}}_{\boldsymbol{z_{tt}}}\cdot\boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\mathcal{T}_h} = 0, \tag{13e}$$

for all $(v_h, \boldsymbol{w}_h, \boldsymbol{\tau}_h, \mu_h, m_h) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h \times M_h$.

Using the Leibniz integral rule, (13b) can be rewritten as

$$(a\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\tau_h}) - (\boldsymbol{\rho_{z_{tt}}}, \boldsymbol{\tau}_h) + 2(b_t(t,t)\boldsymbol{\rho_{\sigma}}(t), \boldsymbol{\tau}_h) + (b(t,t)\boldsymbol{\rho_{\sigma_t}}(t), \boldsymbol{\tau}_h) + \int_0^t (b_{tt}(t,s)\boldsymbol{\rho_{\sigma}}(s), \boldsymbol{\tau}_h)ds = (a\boldsymbol{\theta_{\sigma_{tt}}}, \boldsymbol{\tau}_h) - (\boldsymbol{\theta_{z_{tt}}}, \boldsymbol{\tau}_h) + 2(b_t(t,t)\boldsymbol{\theta_{\sigma}}(t), \boldsymbol{\tau}_h) + (b(t,t)\boldsymbol{\theta_{\sigma_t}}(t), \boldsymbol{\tau}_h) + \int_0^t (b_{tt}(t,s)\boldsymbol{\theta_{\sigma}}(s), \boldsymbol{\tau}_h)ds. \tag{14}$$

Now, adding (13a), (14), (13c)-(13e) after taking $\boldsymbol{w}_h = \boldsymbol{\rho_{z_{tt}}}$, $\boldsymbol{\tau}_h = \boldsymbol{\rho_{\sigma_{tt}}}$, $v_h = \rho_{u_{tt}}$, $\mu_h = -\hat{\boldsymbol{\rho}}_{\boldsymbol{z}_{tt}} \cdot \boldsymbol{\nu}$ and $m_h = -\hat{\rho}_{u_{tt}}$ will yield the following inequality:

$$(a\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\rho_{\sigma_{tt}}}) + \|\hat{\rho}_{u_{tt}} - \rho_{u_{tt}}\|_\tau^2 = -2(b_t(t,t)\boldsymbol{\rho_{\sigma}}(t), \boldsymbol{\rho_{\sigma_{tt}}}) - (b(t,t)\boldsymbol{\rho_{\sigma_t}}(t), \boldsymbol{\rho_{\sigma_{tt}}}) - \int_0^t (b_{tt}(t,s)\boldsymbol{\rho_{\sigma}}(s), \boldsymbol{\rho_{\sigma_{tt}}})ds + (a\boldsymbol{\theta_{\sigma_{tt}}}, \boldsymbol{\rho_{\sigma_{tt}}}) - (\boldsymbol{\theta_{z_{tt}}}, \boldsymbol{\rho_{\sigma_{tt}}}) + 2(b_t(t,t)\boldsymbol{\theta_{\sigma}}(t), \boldsymbol{\rho_{\sigma_{tt}}}) + (b(t,t)\boldsymbol{\theta_{\sigma_t}}(t), \boldsymbol{\rho_{\sigma_{tt}}}) + \int_0^t (b_{tt}(t,s)\boldsymbol{\theta_{\sigma}}(s), \boldsymbol{\rho_{\sigma_{tt}}})ds.$$

and simplifying using Cauchy Schwarz inequality and using the estimates of $\|\boldsymbol{\rho_{\sigma_t}}\|$ and $\|\boldsymbol{\rho_{\sigma}}\|$ along with the boundedness properties of a, b and its derivatives, will yield the estimate of $\|\boldsymbol{\rho_{\sigma_{tt}}}\|$.

Next, we take $\boldsymbol{\tau}_h = \boldsymbol{\rho_{z_{tt}}}$ in (14). A use of Cauchy Schwarz inequality along with the boundedness properties of a, b and its derivatives gives the following estimate of $\|\boldsymbol{\rho_{z_{tt}}}\|$:

$$\|\boldsymbol{\rho_{z_{tt}}}\|^2 = C\bigg(\|\boldsymbol{\rho_{\sigma_{tt}}}\| + \|\boldsymbol{\rho_{\sigma}}(t)\| + \|\boldsymbol{\rho_{\sigma_t}}(t)\|\|\boldsymbol{\theta_{\sigma_{tt}}}\| + \|\boldsymbol{\theta_{z_{tt}}}\| + \|\boldsymbol{\theta_{\sigma}}(t)\| + \|\boldsymbol{\theta_{\sigma_t}}(t)\| + \int_0^t \|\boldsymbol{\theta_{\sigma}}(s)\| + \|\boldsymbol{\rho_{\sigma}}(s)\|ds\bigg)\|\boldsymbol{\rho_{z_{tt}}}\|.$$

For the estimate of $\|\rho_{u_{tt}}\|$, we begin by taking into account the following dual problem in the domain Ω:

$$\boldsymbol{\phi} = -\nabla\psi, \tag{15a}$$

$$\boldsymbol{p} = a\boldsymbol{\phi}, \tag{15b}$$

$$\nabla \cdot \boldsymbol{p} = \rho_{u_{tt}}, \tag{15c}$$

$$\psi = 0 \qquad \text{on } \partial\Omega, \tag{15d}$$

along with:

$$\|\psi\|_{H^2(\Omega)} \leq \|\rho_{u_{tt}}\|. \tag{16}$$

Consider,

$$\|\rho_{u_{tt}}\|^2 = (\rho_{u_{tt}}, \nabla \cdot \boldsymbol{p}).$$

Use of the HDG projection yields

$$\|\rho_{u_{tt}}\|^2 = (\rho_{u_{tt}}, \nabla \cdot \boldsymbol{\Pi_W p}) + \langle \rho_{u_{tt}}, \tau(\Pi_V\psi - \psi)\rangle_{\partial\mathcal{T}_h}.$$

Next, a use of (13a),(13c)-(13e) and the definition of the HDG projection yield

$$\|\rho_{u_{tt}}\|^2 = (\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\Pi_W p} - \boldsymbol{p}) + (a\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\phi}) - (\boldsymbol{\rho_{z_{tt}}}, \boldsymbol{\phi} - \boldsymbol{I_h^k\phi}) - (\boldsymbol{\rho_{z_{tt}}}, \boldsymbol{I_h^k\phi}).$$

Now, a use of (14) will yield the following equation

$$\begin{aligned}\|\rho_{u_{tt}}\|^2 &= (\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\Pi_W p} - \boldsymbol{p}) + (a\boldsymbol{\rho_{\sigma_{tt}}}, \boldsymbol{\phi} - \boldsymbol{I_h^k\phi}) - (\boldsymbol{\rho_{z_{tt}}}, \boldsymbol{\phi} - \boldsymbol{I_h^k\phi})\\ &\quad - 2(b_t(t,t)\boldsymbol{\rho_\sigma}(t), \boldsymbol{I_h^k\phi}) - (b(t,t)\boldsymbol{\rho_{\sigma_t}}(t), \boldsymbol{I_h^k\phi}) - \int_0^t (b_{tt}(t,s)\boldsymbol{\rho_\sigma}(s), \boldsymbol{I_h^k\phi})ds\\ &\quad + (a\boldsymbol{\theta_{\sigma_{tt}}}, \boldsymbol{I_h^k\phi}) - (\boldsymbol{\theta_{z_{tt}}}, \boldsymbol{I_h^k\phi}) + 2(b_t(t,t)\boldsymbol{\theta_\sigma}(t), \boldsymbol{I_h^k\phi}) + (b(t,t)\boldsymbol{\theta_{\sigma_t}}(t), \boldsymbol{I_h^k\phi})\\ &\quad + \int_0^t (b_{tt}(t,s)\boldsymbol{\theta_\sigma}(s), \boldsymbol{I_h^k\phi})ds.\end{aligned}$$

A use of Cauchy Schwarz inequality along with the boundedness properties of a, b and its derivatives gives the following estimate

$$\begin{aligned}\|\rho_{u_{tt}}\|^2 &= C\Bigg[\|\boldsymbol{\rho_{\sigma_{tt}}}\|\|\boldsymbol{\Pi_W p} - \boldsymbol{p}\| + \|\boldsymbol{\rho_{\sigma_{tt}}}\|\|\boldsymbol{\phi} - \boldsymbol{I_h^k\phi}\| + \|\boldsymbol{\rho_{z_{tt}}}\|\|\boldsymbol{\phi} - \boldsymbol{I_h^k\phi}\| + \Bigg(\|\boldsymbol{\rho_\sigma}(t)\|\\ &\quad + \|\boldsymbol{\rho_{\sigma_t}}(t)\| + \|\boldsymbol{\theta_{\sigma_{tt}}}\| + \|\boldsymbol{\theta_{z_{tt}}}\| + \|\boldsymbol{\theta_\sigma}(t)\| + \|\boldsymbol{\theta_{\sigma_t}}(t)\|\\ &\quad + \int_0^t \big(\|\boldsymbol{\rho_\sigma}(s)\| + \|\boldsymbol{\theta_\sigma}(s)\|\big)ds\Bigg)\|\boldsymbol{I_h^k\phi}\|_{\boldsymbol{H^1}(\Omega)}\Bigg].\end{aligned}$$

Now, an application of the estimates of HDG projection and the fact that $\|\boldsymbol{\phi}\|_{\boldsymbol{H^1}(\Omega)} \leq M\|\boldsymbol{p}\|_{\boldsymbol{H^1}(\Omega)}$ and $\|\boldsymbol{p}\|_{\boldsymbol{H^1}(\Omega)} \leq \|\psi\|_{H^2(\Omega)}$ gives desired estimate. □

STEP III: Estimates of Ritz-Volterra projection.

Theorem 5.1. *For $t \in (0,T]$, if $u \in L^\infty(H^{k+2}(\mathcal{T}_h))$, $u_t, u_{tt} \in L^1(H^{k+2}(\mathcal{T}_h))$ and $l = 0,1,2$ then there exists a positive constant C independent of h and k such that*

$$\left\|\frac{\partial^l \eta_u}{\partial t^l}\right\| + \left\|\frac{\partial^l \boldsymbol{\eta_\sigma}}{\partial t^l}\right\| + \left\|\frac{\partial^l \boldsymbol{\eta_z}}{\partial t^l}\right\| \leq Ch^{k+1}, \tag{17}$$

$$\left\|I_h^{k-1}\left(\frac{\partial^l \eta_u}{\partial t^l}\right)\right\| \leq Ch^{k+2}. \tag{18}$$

Proof. The inequality (17) can be obtained with the help of (3), Lemma 5.1 and the triangle inequality.
For the estimates of $\|I_h^{k-1}\eta_u\|$, the following dual problem is considered in $\Omega \times (0,T]$

$$\begin{aligned}\boldsymbol{\phi} &= -\nabla\psi,\\ \boldsymbol{p} &= a\boldsymbol{\phi},\\ \nabla \cdot \boldsymbol{p} &= \theta.\end{aligned}$$

which satisfies the elliptic regularity

$$\|\psi\|_{H^2(\Omega)} \leq \|\theta\|.$$

Now, using (10a) and proceeding as in [27], concludes the proof.

$$\begin{aligned}(I_h^{k-1}\eta_u, \theta) &\leq (I_h^{k-1}\eta_u, \nabla\cdot\boldsymbol{p})\\ &\leq (\eta_u, \nabla\cdot\boldsymbol{\Pi}_{k-1}^{RT}\boldsymbol{p})\\ &\leq (\boldsymbol{\eta_\sigma}, \nabla\cdot\boldsymbol{\Pi}_{k-1}^{RT}\boldsymbol{p} - \boldsymbol{p}) - (a\boldsymbol{\eta_\sigma}, \nabla\psi)\\ &\leq Ch^{k+2}\|\theta\|.\end{aligned}$$

A similar procedure can be followed for $l = 1,2$. □

Remark: The order of convergence of $\|\rho_u\|$ can be further increased to $k + 3/2$, using dual norm estimates. This additional result is stated in the form of the following lemma:

Lemma 5.2. *For $t \in (0,T]$, a positive constant C that is unaffected by the values of h and $k \geq 1$ exists, such that it ensures the validity of the following inequality:*

$$\|\rho_u(t)\| \leq Ch^{k+3/2}\left[\|u(t)\|_{H^{k+2}(\Omega)} + \int_0^t \|u(s)\|_{H^{k+2}(\Omega)} ds\right].$$

Proof. We begin by defining the following dual norm:

$$\|\boldsymbol{v}\|_{(\boldsymbol{H}^1(\Omega))^*} = \sup_{\boldsymbol{w} \in \boldsymbol{H}^1(\Omega), w \neq 0} \frac{(\boldsymbol{v}, \boldsymbol{w})}{\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}}.$$

Next, we consider a similar dual problem as (15) after replacing $\rho_{u_{tt}}$ by ρ_u, which is written as follows

$$\begin{aligned} \boldsymbol{\phi} &= -\nabla \psi, \\ \boldsymbol{p} &= a\boldsymbol{\phi}, \\ \nabla \cdot \boldsymbol{p} &= \rho_u, \\ \psi &= 0, \end{aligned}$$

along with:

$$\|\psi\|_{H^2(\Omega)} \leq \|\rho_u\|.$$

Next, we proceed as done in Theorem 3.3 of [26], to achieve the following inequality:

$$\begin{aligned} \|\rho_u\|^2 \leq{}& \|\boldsymbol{\rho_\sigma}\|\|\boldsymbol{\Pi_W p} - \boldsymbol{p}\| + C\|\boldsymbol{\rho_\sigma}\|\|\boldsymbol{\phi} - \boldsymbol{I}_h^k\boldsymbol{\phi}\| + C\|\boldsymbol{\theta_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*}\|\boldsymbol{I}_h^k\boldsymbol{\phi}\|_{\boldsymbol{H}^1(\Omega)} \\ & + \|\boldsymbol{\theta_z}\|_{\boldsymbol{H}^1(\Omega)^*}\|\boldsymbol{I}_h^k\boldsymbol{\phi}\|_{\boldsymbol{H}^1(\Omega)} + C\int_0^t \Big(\|\boldsymbol{\theta_\sigma}(s)\|_{\boldsymbol{H}^1(\Omega)^*} + \|\boldsymbol{\rho_\sigma}(s)\|_{\boldsymbol{H}^1(\Omega)^*}\Big)\|\boldsymbol{I}_h^k\boldsymbol{\phi}\|_{\boldsymbol{H}^1(\Omega)}. \end{aligned} \tag{19}$$

Hence, we require the estimates of $\|\boldsymbol{\theta_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*}$, $\|\boldsymbol{\theta_z}\|_{\boldsymbol{H}^1(\Omega)^*}$ and $\|\boldsymbol{\rho_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*}$. For the estimates of $\|\boldsymbol{\theta_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*}$, we will proceed as follows:

$$\begin{aligned} (\boldsymbol{\theta_\sigma}, \boldsymbol{w}) &= (\boldsymbol{\theta_\sigma}, \boldsymbol{w} - \boldsymbol{I}_h^k\boldsymbol{w}) + (\boldsymbol{\theta_\sigma}, \boldsymbol{I}_h^k\boldsymbol{w}) \\ &\leq \|\boldsymbol{\theta_\sigma}\|\|\boldsymbol{w} - \boldsymbol{I}_h^k\boldsymbol{w}\| \\ &\leq Ch\|\boldsymbol{\theta_\sigma}\|\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}. \end{aligned}$$

Therefore, we have

$$\|\boldsymbol{\theta_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*} \leq Ch\|\boldsymbol{\theta_\sigma}\|. \tag{20}$$

Now, for $\|\boldsymbol{\theta_z}\|_{\boldsymbol{H}^1(\Omega)^*}$, we have for $k \geq 1$

$$\begin{aligned} (\boldsymbol{\theta_z}, \boldsymbol{w}) &= (\boldsymbol{\theta_z}, \boldsymbol{w} - \boldsymbol{I}_h^{k-1}\boldsymbol{w}) + (\boldsymbol{\theta_z}, \boldsymbol{I}_h^{k-1}\boldsymbol{w}) \\ &\leq \|\boldsymbol{\theta_z}\|\|\boldsymbol{w} - \boldsymbol{I}_h^{k-1}\boldsymbol{w}\| \\ &\leq Ch\|\boldsymbol{\theta_z}\|\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}. \end{aligned}$$

Therefore, we have

$$\|\boldsymbol{\theta_z}\|_{\boldsymbol{H}^1(\Omega)^*} \leq Ch\|\boldsymbol{\theta_z}\|. \tag{21}$$

Finally, for the estimates of $\|\boldsymbol{\rho_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*}$, we have

$$\begin{aligned} (\boldsymbol{\rho_\sigma}, \boldsymbol{w}) &= (\boldsymbol{\rho_\sigma}, \boldsymbol{w} - \boldsymbol{I}_h^k\boldsymbol{w}) + (\boldsymbol{\rho_\sigma}, \boldsymbol{I}_h^k\boldsymbol{w}) \\ &= (\rho_u, \nabla \cdot \boldsymbol{w_h}) - \langle \hat{\rho}_u, \boldsymbol{w_h} \cdot \boldsymbol{\nu} \rangle \quad \text{by (11a)} \\ &= (\rho_u, \boldsymbol{\nabla \cdot w}) + \langle \rho_u - \hat{\rho}_u, (\boldsymbol{I}_h^k\boldsymbol{w} - \boldsymbol{w}) \cdot \boldsymbol{\nu} \rangle \\ &\leq C(\|\rho_u\| + h^{1/2}\|\rho_u - \hat{\rho}_u\|)\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}. \end{aligned}$$

Therefore, we have

$$\|\boldsymbol{\rho_\sigma}\|_{\boldsymbol{H}^1(\Omega)^*} \leq C(\|\rho_u\| + h^{1/2}\|\rho_u - \hat{\rho}_u\|). \tag{22}$$

Use of (20), (21) and (22) in (19), will give the desired improved estimates of $\|\rho_u\|$, and hence, conclude the lemma. □

STEP IV: Estimates of $\|\xi_u\|$, $\|\boldsymbol{\xi_\sigma}\|$ and $\|\boldsymbol{\xi_z}\|$

In order to derive the error estimates, we disintegrate them in the following form

$$e_u = u - u_h = (u - \tilde{u}_h) - (u_h - \tilde{u}_h) = \eta_u - \xi_u.$$

Similarly for $\boldsymbol{e_\sigma}$, $\boldsymbol{e_z}$, $\hat{e}_u$ and $\boldsymbol{\hat{e}_z}$. Hence, (6) can be written as

$$(\boldsymbol{\xi_\sigma}, \boldsymbol{w_h}) - (\xi_u, \nabla\cdot\boldsymbol{w_h}) + \langle\hat{\xi}_u, \boldsymbol{w_h}\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0 \qquad \forall \boldsymbol{w}_h \in \boldsymbol{W}_h, \tag{23a}$$

$$(a\boldsymbol{\xi_\sigma}, \boldsymbol{\tau_h}) - (\boldsymbol{\xi_z}, \boldsymbol{\tau}_h) + \int_0^t (b(t,s)\boldsymbol{\xi_\sigma}(s), \boldsymbol{\tau}_h)ds = 0 \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{W}_h, \tag{23b}$$

$$(\xi_{u_{tt}}, v_h) - (\boldsymbol{\xi_z}, \nabla v_h) + \langle\boldsymbol{\hat{\xi}_z}\cdot\boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = (\eta_{u_{tt}}, v_h) \qquad \forall v_h \in V_h, \tag{23c}$$

$$\langle\hat{\xi}_u, \mu_h\rangle_{\partial\Omega} = 0 \qquad \forall \mu_h \in M_h, \tag{23d}$$

$$\langle\boldsymbol{\hat{\xi}_z}\cdot\boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0 \qquad \forall m_h \in M_h. \tag{23e}$$

For any function w in $[0,t]$, let us define $\bar{w}$ as:

$$\bar{w}(t) = \int_0^t w(s)ds.$$

Clearly, $\bar{w}_t = w$ and $\bar{w}(0) = 0$.

Lemma 5.3. *There exists a positive constant C which does not rely on h and k such that $\forall t \in (0,T]$, the inequality below is valid*

$$\|\xi_u(t)\|^2 + \|\bar{\boldsymbol{\xi}}_\sigma(t)\|^2 + \|(\bar{\hat{\xi}}_u - \bar{\xi}_u)(t)\|_\tau^2 \leq C\bigg(\|\xi_u(0)\|^2 + \|a^{1/2}\bar{\boldsymbol{\xi}}_\sigma(0)\|^2 + \|\bar{\hat{\xi}}_u(0) - \bar{\xi}_u(0)\|_\tau^2 + \int_0^T \|\eta_{u_t}(t)\|^2 dt\bigg).$$

Proof. We integrate (23b), (23c) and (23e) from 0 to t to get the following equations:

$$(a\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}, \boldsymbol{\tau_h}) - (\bar{\boldsymbol{\xi}}_{\boldsymbol{z}}, \boldsymbol{\tau}_h) + \int_0^t\int_0^s (b(s,\gamma)\boldsymbol{\xi_\sigma}(\gamma), \boldsymbol{\tau}_h)d\gamma ds = 0 \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{W}_h, \tag{24a}$$

$$(\xi_{u_t}, v_h) - (\bar{\boldsymbol{\xi}}_{\boldsymbol{z}}, \nabla v_h) + \langle\bar{\hat{\boldsymbol{\xi}}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = (\eta_{u_t}, v_h) - (e_{u_t}(0), v_h) \qquad \forall v_h \in V_h, \tag{24b}$$

$$\langle\bar{\hat{\boldsymbol{\xi}}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0 \qquad \forall m_h \in M_h. \tag{24c}$$

Note that with $u_{h_t}(0) = I_h^k u_1$, we have

$$(e_{u_t}(0), v_h) = 0 \ \ \forall\ v_h \in V_h$$

Next, we choose $\boldsymbol{w}_h = \bar{\boldsymbol{\xi}}_{\boldsymbol{z}}$ in (23a), $\boldsymbol{\tau}_h = \boldsymbol{\xi_\sigma}$ in (24a), $v_h = \xi_u$ in (24b), $\mu_h = -\bar{\hat{\boldsymbol{\xi}}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}$ in (23d) and $m_h = -\hat{\xi}_u$ in (24c) and add them, to obtain

$$\frac{1}{2}\frac{d}{dt}\left(\|a^{1/2}\bar{\boldsymbol{\xi}}_\sigma\|^2 + \|\xi_u\|^2 + \|\bar{\hat{\xi}}_u - \bar{\xi}_u\|_\tau^2\right) = (\eta_{u_t}, \xi_u) - \int_0^t\int_0^s (b(s,\gamma)\boldsymbol{\xi_\sigma}(\gamma), \boldsymbol{\xi_\sigma}(s))d\gamma ds.$$

It follows by integrating the aforementioned equality

$$
\begin{aligned}
&\|a^{1/2}\bar{\boldsymbol{\xi}}_{\sigma}\|^2 + \|\xi_u\|^2 + \|\bar{\hat{\xi}}_u - \bar{\xi}_u\|_\tau^2 \\
&\le \|\xi_u(0)\|^2 + \|a^{1/2}\bar{\boldsymbol{\xi}}_{\sigma}(0)\|^2 + \|\bar{\hat{\xi}}_u(0) - \bar{\xi}_u(0)\|_\tau^2 + 2\int_0^t (\eta_{u_s}, \xi_u) ds \\
&- 2\int_0^t \int_0^s \int_0^\gamma (b(\gamma,\gamma^*)\boldsymbol{\xi}_{\boldsymbol{\sigma}}(\gamma^*), \boldsymbol{\xi}_{\boldsymbol{\sigma}}(\gamma)) d\gamma^* d\gamma ds.
\end{aligned}
$$

Let the last term on the right-hand side of the above equation be denoted by I, then we have

$$
\begin{aligned}
I &= 2\int_0^t \int_0^s (b(\gamma,\gamma)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(\gamma), \boldsymbol{\xi}_{\boldsymbol{\sigma}}(s)) d\gamma ds - 2\int_0^t \int_0^s \int_0^\gamma (b_{\gamma^*}(\gamma,\gamma^*)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(\gamma^*), \boldsymbol{\xi}_{\boldsymbol{\sigma}}(s)) d\gamma^* d\gamma ds \\
&= 2(I_1 - I_2).
\end{aligned}
$$

Here,

$$
I_1 = \int_0^t \int_0^s (b(\gamma,\gamma)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(\gamma), \boldsymbol{\xi}_{\boldsymbol{\sigma}}(s)) d\gamma ds.
$$

Using integration by parts, we will achieve the following:

$$
I_1 = \int_0^t (b(s,s)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s), \bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(t)) ds - \int_0^t \frac{d}{ds}\left(\int_0^s b(\gamma,\gamma)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(\gamma) d\gamma\right) \bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s) ds.
$$

Next, a use of Leibniz rule yields:

$$
I_1 = \int_0^t (b(s,s)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s), \bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(t)) ds - \int_0^t b(s,s)\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s), \bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s) ds).
$$

Finally, using of boundedness of b yields

$$
|I_1| \le C\left(\|\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(t)\| \int_0^t \|\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s)\| ds + \int_0^t \|\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s)\|^2 ds\right).
$$

Similarly, simplifying I_2, and combining the estimates of I_1 and I_2 will yield the following inequality

$$
|I| \le C\left(\|\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(t)\| \int_0^t \|\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s)\| ds + \int_0^t \|\bar{\boldsymbol{\xi}}_{\boldsymbol{\sigma}}(s)\|^2 ds\right)
$$

Lastly, we use Young's inequality and Gronwall's lemma to finish the proof. □

Lemma 5.4. *There exists a positive constant C which does not rely on h and k such that $\forall t \in (0,T]$, the inequality below is valid*

$$
\begin{aligned}
&\|\xi_{u_t}\|^2 + \|\boldsymbol{\xi}_{\boldsymbol{z}}(t)\|^2 + \|\boldsymbol{\xi}_{\boldsymbol{\sigma}}(t)\|^2 + \|\hat{\xi}_u - \xi_u\|_\tau^2 \\
&\le C\bigg(\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}(0)\|^2 + \|\xi_{u_t}(0)\|^2 + \|\hat{\xi}_u(0) - \xi_u(0)\|_\tau^2 + \int_0^T \|\eta_{u_{tt}}(t)\|^2 dt\bigg).
\end{aligned}
$$

Proof. Firstly, we differentiate (23a) with respect to t and then select $\boldsymbol{w_h} = \boldsymbol{\xi_z}$, $\boldsymbol{\tau_h} = \boldsymbol{\xi}_{\boldsymbol{\sigma}_t}$, $v_h = \xi_{u_t}$ in (23a), (23b), (23c), respectively. Then, we differentiate (23d) and select $\mu_h = -\hat{\boldsymbol{\xi}}_{\boldsymbol{z}} \cdot \boldsymbol{\nu}$ and $m_h = -\hat{\xi}_{u_t}$ in (23d) and (23e), respectively. Finally, by combining the ensuing equations, we have

$$
(a\boldsymbol{\xi}_{\boldsymbol{\sigma}}, \boldsymbol{\xi}_{\boldsymbol{\sigma}_t}) + (\xi_{u_{tt}}, \xi_{u_t}) + \frac{1}{2}\frac{d}{dt}\|\hat{\xi}_u - \xi_u\|_\tau^2 + \int_0^t (b(t,s)\boldsymbol{\xi}_{\boldsymbol{\sigma}}(s), \boldsymbol{\xi}_{\boldsymbol{\sigma}_t}(t)) ds = (\eta_{u_{tt}}, \xi_{u_t}).
$$

An application of Leibnitz's theorem shows,

$$\frac{1}{2}\frac{d}{dt}\left(\|a^{1/2}\boldsymbol{\xi_\sigma}\|^2+\|\xi_{u_t}\|^2+\|\hat{\xi}_u-\xi_u\|_\tau^2\right)$$
$$=(\eta_{u_{tt}},\xi_{u_t})+\frac{d}{dt}\int_0^t(b(t,s)\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(t))ds-(b(t,t)\boldsymbol{\xi_\sigma}(t),\boldsymbol{\xi_\sigma}(t))$$
$$-\int_0^t(b_t(t,s)\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(t))ds.$$

Integrating the aforementioned inequality from 0 to t yields

$$\|a^{1/2}\boldsymbol{\xi_\sigma}\|^2+\|\xi_{u_t}\|^2$$
$$\leq\|a^{1/2}\boldsymbol{\xi_\sigma}(0)\|+\|\xi_{u_t}(0)\|^2+\|\xi_{u_t}(0)\|^2+\|\hat{\xi}_u(0)-\xi_u(0)\|_\tau^2+\int_0^t(\eta_{u_{ss}},\xi_{u_s})ds$$
$$+\int_0^t(b(t,s)\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(t))ds-\int_0^t(b(s,s)\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(s))ds$$
$$-\int_0^t\int_0^s(b_s(s,\gamma)\boldsymbol{\xi_\sigma}(\gamma),\boldsymbol{\xi_\sigma}(s))d\gamma ds.$$

Using the boundedness condition of a, b and its derivative, we can rewrite the above inequality as follows:

$$\|\boldsymbol{\xi_\sigma}\|^2+\|\xi_{u_t}\|^2$$
$$\leq C\Bigg(\|\boldsymbol{\xi_\sigma}(0)\|+\|\xi_{u_t}(0)\|^2+\|\xi_{u_t}(0)\|^2+\|\hat{\xi}_u(0)-\xi_u(0)\|_\tau^2+\int_0^t(\eta_{u_{ss}},\xi_{u_s})ds$$
$$+\int_0^t(\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(t))ds-\int_0^t(\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(s))ds-\int_0^T\int_0^t(\boldsymbol{\xi_\sigma}(s),\boldsymbol{\xi_\sigma}(t))dsdt\Bigg).$$

which can be further written as

$$\|\boldsymbol{\xi_\sigma}\|^2+\|\xi_{u_t}\|^2$$
$$\leq C\Bigg(\|\boldsymbol{\xi_\sigma}(0)\|+\|\xi_{u_t}(0)\|^2+\|\xi_{u_t}(0)\|^2+\|\hat{\xi}_u(0)-\xi_u(0)\|_\tau^2+\int_0^T\big(\|\eta_{u_{tt}}(t)\|^2$$
$$+\|\xi_{u_s}\|^2+\|\boldsymbol{\xi_\sigma}(s)\|^2\big)dt\Bigg).$$

Finally, the following estimate is obtained using Gronwall's theorem:

(25)
$$\|\boldsymbol{\xi_\sigma}\|^2+\|\xi_{u_t}\|^2\leq C\left(\|\boldsymbol{\xi_\sigma}(0)\|^2+\|\xi_{u_t}(0)\|^2+\|\hat{\xi}_u(0)-\xi_u(0)\|_\tau^2+\int_0^T\|\eta_{u_{tt}}(t)\|^2dt\right).$$

Now, choosing $\boldsymbol{\tau_h}=\boldsymbol{\xi_z}$ in (23b) yields

$$\|\boldsymbol{\xi_z}\|\leq C\Bigg(\|\boldsymbol{\xi_\sigma}\|+\int_0^t\|\boldsymbol{\xi_\sigma}(s)\|ds\Bigg). \tag{26}$$

Combining (25) and (26) will finish the proof. □

Proof of Theorem 4.1: If we chose $u_h(0)=\tilde{u}_h(0)=\Pi_v u_0$, $\boldsymbol{\sigma}_h(0)=\tilde{\boldsymbol{\sigma}}_h(0)=-\boldsymbol{I}_h^k\nabla u_0$ and $\boldsymbol{z}_h(0)=\tilde{\boldsymbol{z}}_h(0)=\boldsymbol{\Pi_W}(a\nabla u_0)$, then triangle inequality, Lemma 5.1, Lemma 5.3 along with Lemma 5.4 yields the desired result. □

6. Post-processing

To begin with, we define the function $\psi(s) \in H^2(\Omega) \cap H_0^1(\Omega),\ s \leq t$ to be the solution of the following problem:

$$\psi_{ss} - \nabla \cdot \left(a(x)\nabla\psi + \int_s^t b(\gamma, s)\nabla\psi(\gamma)d\gamma \right) = 0, \tag{27}$$

with the following final and boundary conditions:

$$\begin{aligned} \psi(x,s) &= 0 && \text{on } \partial\Omega, s \leq t, \\ \psi(x,t) &= 0 && \text{in } x \in \Omega, \\ \psi_s(x,t) &= \lambda(x) && \text{in } x \in \Omega, \end{aligned}$$

Lemma 6.1. *(Regularity Results) There exists a constant C dependent on the data of the above problem, such that it satisfies the following inequality:*

$$\|\psi(s)\|_{L^\infty(H^1)} + \|\psi_s(s)\|_{L^\infty(L^2)} \leq C\|\lambda\|, \tag{28a}$$

$$\|\underset{\sim}{\psi}(s)\|_2 \leq C\|\lambda\|, \tag{28b}$$

where, $\underset{\sim}{\psi}(s) = \int_s^t \psi(\gamma)d\gamma$.

Proof. The first inequality is a direct consequence of the conservation of energy. To prove the second inequality, we begin by integrating (27) from s to t, noting that $-\psi_s(s) = \underset{\sim}{\psi}_{ss}(s)$ and using the boundary condition, to obtain

$$\underset{\sim}{\psi}_{ss}(s) - \nabla \cdot \left(a(x)\nabla\underset{\sim}{\psi} + \int_s^t \int_\gamma^t b(\gamma^*, \gamma)\nabla\psi(\gamma^*)d\gamma^* d\gamma \right) = -\lambda.$$

Next, we assume the following elliptic regularity on $\underset{\sim}{\psi}$ [24], and use (28a) to get

$$\begin{aligned} \|\underset{\sim}{\psi}\|_2 &\leq C \left\| \nabla \cdot \big(a(x)\nabla\underset{\sim}{\psi} \big) \right\| \\ &\leq C \left(\|\underset{\sim}{\psi}_{ss}(s)\| + \|\lambda\| + \left\| \int_s^t \int_\gamma^t b(\gamma^*, \gamma)\nabla\psi(\gamma^*)d\gamma^* d\gamma \right\| \right) \\ &\leq C \left(\|\psi_s(s)\| + \|\lambda\| + \left\| \int_s^t \int_\gamma^t b(\gamma^*, \gamma)\nabla\psi(\gamma^*)d\gamma^* d\gamma \right\| \right) \\ &\leq C \|\lambda\| . \end{aligned}$$

□

Lemma 6.2. *For the method of the form (5), There exists a positive constant C which does not rely on h and k such that $\forall t \in (0,T]$, the inequality below is valid*

$$\|I_h^{k-1} e_u\|_{L^2(K)} \leq Ch^{k+2}, \tag{29}$$

where, I_h^{k-1} is L^2-projection onto the space of polynomial for degree at most $k-1$.

Proof. Since, $e_u = \eta_u - \xi_u$, therefore, $\|I_h^{k-1} e_u\| \leq \|I_h^{k-1}\eta_u\| + \|I_h^{k-1}\xi_u\|$.
For the estimates of $\|I_h^{k-1}\xi_u\|$, we start by rewriting (27) in the following mixed

form:

$$\boldsymbol{\phi}(s) = \nabla\psi(s), \tag{30a}$$

$$\boldsymbol{p}(s) = a\boldsymbol{\phi}(s) + \int_s^t b(\gamma, s)\boldsymbol{\phi}(\gamma)d\gamma, \tag{30b}$$

$$\psi_{ss}(s) - \nabla\cdot\boldsymbol{p}(s) = 0, \tag{30c}$$

$$\psi(s) = 0 \quad \text{on } \partial\Omega, \tag{30d}$$

$$\psi(t) = 0, \tag{30e}$$

$$\psi_s(t) = I_h^{k-1}\xi_u(t). \tag{30f}$$

we begin by taking the inner product of (30c) with $I_h^{k-1}\xi_u(s)$, to obtain

$$(\psi_{ss}(s), I_h^{k-1}\xi_u(s)) - (\nabla\cdot\boldsymbol{p}(s), I_h^{k-1}\xi_u(s)) = 0.$$

Now,

$$\begin{aligned}
&\frac{d}{ds}\left[(\psi_s(s), I_h^{k-1}\xi_u(s)) - (\psi(s), I_h^{k-1}\xi_{u_s}(s))\right]\\
&= (\psi_{ss}(s), I_h^{k-1}\xi_u(s)) - (\psi(s), I_h^{k-1}\xi_{u_{ss}}(s))\\
&= -(\psi(s), I_h^{k-1}\xi_{u_{ss}}(s)) + (\nabla\cdot\boldsymbol{p}(s), I_h^{k-1}\xi_u(s)).
\end{aligned}$$

Use of (23) and intermediate projections, see [10], yields the following equality

$$\begin{aligned}
&\frac{d}{ds}\left[(\psi_s(s), I_h^{k-1}\xi_u(s)) - (\psi(s), I_h^{k-1}\xi_{u_s}(s))\right] = (\xi_{u_{ss}}(s), I_h^k\psi(s) - \psi(s))\\
&- (\xi_{u_{ss}}(s), I_h^{k-1}\psi(s) - \psi(s)) + (\boldsymbol{\xi_\sigma}(s), \boldsymbol{\Pi}_{k-1}^{RT}\boldsymbol{p}(s) - \boldsymbol{p}(s)) + (a\boldsymbol{\xi_\sigma}(s), \boldsymbol{\phi}(s) - \boldsymbol{I}_h^k\boldsymbol{\phi}(s))\\
&+ (\boldsymbol{\xi_z}(s), \boldsymbol{I}_h^k\boldsymbol{\phi}(s) - \boldsymbol{\phi}(s)) + (\boldsymbol{\xi_z}(s), \nabla(\psi - I_h^k\psi)(s)) + \langle\hat{\boldsymbol{\xi}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}, I_h^k\psi\rangle - (\eta_{u_{ss}}, I_h^k\psi)\\
&- \int_0^s (b(s,\gamma)\boldsymbol{\xi_\sigma}(\gamma), \boldsymbol{I}_h^k\boldsymbol{\phi}(s))d\gamma + \int_s^t (b(\gamma, s)\boldsymbol{\phi}(\gamma), \boldsymbol{\xi_\sigma}(s))d\gamma.
\end{aligned}$$

Taking $\xi_u(0) = \xi_{u_s}(0) = 0$ and integrating the equation from 0 to t followed by a change of order of integration of the last term, we obtain

$$\begin{aligned}
\|I_h^{k-1}\xi_u\|^2 &= \int_0^t \Big[(\xi_{u_{ss}}(s), I_h^k\psi(s) - \psi(s)) - (\xi_{u_{ss}}(s), I_h^{k-1}\psi(s) - \psi(s)) + (\boldsymbol{\xi_\sigma}(s),\\
&\quad \boldsymbol{\Pi}_{k-1}^{RT}\boldsymbol{p}(s) - \boldsymbol{p}(s)) + (a\boldsymbol{\xi_\sigma}(s), \boldsymbol{\phi}(s) - \boldsymbol{I}_h^k\boldsymbol{\phi}(s)) + (\boldsymbol{\xi_z}(s), \boldsymbol{I}_h^k\boldsymbol{\phi}(s) - \boldsymbol{\phi}(s))\\
&\quad + (\boldsymbol{\xi_z}(s), \nabla(\psi - I_h^k\psi)(s)) + \langle\hat{\boldsymbol{\xi}}_{\boldsymbol{z}}\cdot\boldsymbol{\nu}, I_h^k\psi\rangle - (\eta_{u_{ss}}, I_h^k\psi)\Big]ds\\
&\quad - \int_0^t\int_s^t (b(\gamma, s)\boldsymbol{\xi_\sigma}(s), \boldsymbol{\phi}(\gamma) - \boldsymbol{I}_h^k\boldsymbol{\phi}(\gamma))d\gamma ds\\
&= \int_0^t \left[E_1 + E_2 + E_3 + E_4 + E_5 + E_6 + E_7 + E_8\right]ds + E_9. \qquad (31)
\end{aligned}$$

Cauchy Schwarz's inequality and (28a) shows

$$E_1 + E_2 \leq Ch^{k+2}\|I_h^{k-1}\xi_u(s)\|$$

Next, a use of identity $\int_0^t f(r)g(r)dr = f(0)\bar{g}(0) + \int_0^t f_r(r)\bar{g}(r)dr$ along with (28b), yields

$$|E_3 + E_4 + E_5 + E_6| \leq Ch^{k+2}\|I_h^{k-1}\xi_u\|$$

Use of (6e), properties of the projection I_h and (28a) gives

$$|E_7| \leq \|\hat{\boldsymbol{\xi}}_{\boldsymbol{z}} \cdot \boldsymbol{\nu}\|_{\partial K} \|I_h^k \psi - \psi\|_{\partial K} \leq Ch^{k+2} \|I_h^{k-1} \xi_u\|$$

We rewrite E_8 as follows

$$\begin{aligned}(\eta_{u_{ss}}, I_h^k \psi) &= (\eta_{u_{ss}}, I_h^k \psi - I_h^{k-1}\psi) + (\eta_{u_{ss}}, I_h^{k-1}\psi)\\ &= (\eta_{u_{ss}}, I_h^k \psi - I_h^{k-1}\psi) + (I_h^{k-1}\eta_{u_{ss}}, I_h^{k-1}\psi)\\ &\leq \|\eta_{u_{ss}}\| \|I_h^k \psi - I_h^{k-1}\psi\| + \|I_h^{k-1}\eta_{u_{ss}}\| \|I_h^{k-1}\psi\|\\ &\leq Ch^{k+2} \|I_h^{k-1}\xi_u\|.\end{aligned}$$

Finally, a use of boundedness of b shows

$$\begin{aligned}|E_9| &\leq M \left| \int_0^t \left(\int_s^t \boldsymbol{\phi}(\gamma) - \boldsymbol{I}_h^k \boldsymbol{\phi}(\gamma) \right) d\gamma, \boldsymbol{\xi}_{\boldsymbol{\sigma}}(s) ds \right|\\ &= M \left| \int_0^t \bar{\boldsymbol{\phi}}(\gamma) - \boldsymbol{I}_h^k \bar{\boldsymbol{\phi}}(\gamma), \boldsymbol{\xi}_{\boldsymbol{\sigma}}(s) ds \right|\\ &\leq Ch^{k+2} \|\bar{\psi}(s)\|_2\\ &\leq Ch^{k+2} \|I_h^{k-1}\xi_u\| && \text{(by(28b))}.\end{aligned}$$

Substituting in (31), we get

$$\|I_h^{k-1}\xi_u(t)\|^2 \leq Ch^{k+2} \int_0^t \|I_h^{k-1}\xi_u(s)\| ds.$$

Using Young's inequality and Gronwall's Lemma, the following estimate is obtained:

$$\|I_h^{k-1}\xi_u(t)\| \leq Ch^{k+2}. \tag{32}$$

Finally, (32) and (18) conclude the proof of the theorem. □

Lemma 6.3. *There exists a positive constant C which does not rely on h and k such that $\forall t \in (0, T]$, the inequality below is valid*

$$\|u^p - u_h^p\| \leq Ch^{k+2}, \tag{33}$$

where, $u^p = u - \frac{1}{|K|} \int_K u dx.$

Proof. See [26] (Lemma 5.2). □

Proof of Theorem 4.2: By the definition of u_h^* from (7), on any $K \in \mathcal{T}_h$, we obtain

$$\begin{aligned}\|u - u_h^*\|_{L^2(K)} &\leq \left\| u - u_h^p - \frac{1}{|K|} \int_K u_h dx \right\|_{L^2(K)}\\ &\leq \left\| u^p - u_h^p + \frac{1}{|K|} \int_K (u - u_h) dx \right\|_{L^2(K)}\\ &\leq \|I_h^0 (u - u_h)\|_{L^2(K)} + \|u^p - u_h^p\|_{L^2(K)}\\ &\leq \|I_h^{k-1} e_u\|_{L^2(K)} + \|u^p - u_h^p\|_{L^2(K)}, && (34)\end{aligned}$$

where I_h^0 is L^2-projection onto the space of polynomials of total degree 0. A substitution of (29) and (33) in (34) finishes the proof. □

7. Fully discrete scheme

Here, we present a completely discrete method for approximating the solution to (1). To accomplish this, we discretize equation (5) in the time direction using a central difference scheme and the midpoint rule. First, we split $[0,T]$ into M parts with equal spacing using the following points:

$$0 = t_0 < t_1 < ... < t_M = T,$$

with $t_n = n\Delta t$ where, $\Delta t = T/M$.

We begin by defining the following notations,

$$\Upsilon U^n = \frac{U^{n+1}+U^n}{2}, \quad \Phi U^n = \frac{U^{n+1}+2U^n+U^{n-1}}{4} = \frac{\Upsilon U^n + \Upsilon U^{n-1}}{2},$$

$$\partial_t \Upsilon U^n = \frac{U^{n+1}-U^n}{\Delta t},\ \partial_t^2 U^n = \frac{U^{n+1}-2U^n+U^{n-1}}{\Delta t^2},\ \delta_t U^n = \frac{\partial_t\Upsilon U^n + \partial_t \Upsilon U^{n-1}}{2},$$

$$E_h^n(\boldsymbol{S}) = \Delta t \sum_{j=0}^{n-1} b(t_n, t_{j+1/2})\Upsilon \boldsymbol{S}^j, \quad \Upsilon E_h^n(\boldsymbol{S}) = \frac{E_h^{n+1}(\boldsymbol{S}) + E_h^n(\boldsymbol{S})}{2}.$$

$(U^n, \boldsymbol{S}^n, \boldsymbol{Z}^n, \hat{U}^n) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h$ when $1 \le n \le M$, such that, for any $(v_h, \boldsymbol{w}_h, \boldsymbol{\tau}_h, \mu_h, m_h) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h \times M_h$, we require

(35a)
$$\frac{2}{\Delta t}(\partial_t \Upsilon U^0, v_h) - (\Upsilon \boldsymbol{Z}^0, \nabla v_h) + \langle \Upsilon \hat{\boldsymbol{Z}}^0 \cdot \boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = (\Upsilon f^0 + \frac{2}{\Delta t}u_1, v_h),$$

(35b) $$\langle \Upsilon \hat{\boldsymbol{Z}}^0 \cdot \boldsymbol{\nu}, \mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0,$$

(35c)
$$(\Upsilon \boldsymbol{S}^n, \boldsymbol{w}_h) - (\Upsilon U^n, \nabla\cdot \boldsymbol{w}_h) + \langle \Upsilon \hat{U}^n, \boldsymbol{w}_h\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0 \qquad n \ge 0,$$

(35d) $$(a\Upsilon \boldsymbol{S}^n, \boldsymbol{\tau}_h) - (\Upsilon \boldsymbol{Z}^n, \boldsymbol{\tau}_h) + (\Upsilon E_h^n(\boldsymbol{S}), \boldsymbol{\tau}_h) = 0 \qquad n \ge 0,$$

(35e) $$(\partial_t^2 U^n, v_h) - (\Phi \boldsymbol{Z}^n, \nabla v_h) + \langle \Phi \hat{\boldsymbol{Z}}^n \cdot \boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = (\Phi f^n, v_h) \qquad n \ge 1,$$

(35f) $$\langle \Upsilon \hat{U}^n, \mu_h\rangle_{\partial\Omega} = 0 \qquad n \ge 0,$$

(35g) $$\langle \Phi \hat{\boldsymbol{Z}}^n \cdot \boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0 \qquad n \ge 1,$$

Proof of Theorem 4.3: We begin by writing $\|u(t_n) - U^n\| \le \|u(t_n) - u_h(t_n)\| + \|u_h(t_n) - U^n\|$. We only need to derive the estimate $\|u_h(t_n) - U^n\|$. We will use ζ_u^n to denote $u_h(t_n) - U^n$. Similarly, $\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n$, $\boldsymbol{\zeta}_{\boldsymbol{z}}^n$ and $\hat{\zeta}_u^n$.

Now, using (5) and (35), we have the following

(36a)
$$\frac{2}{\Delta t}(\partial_t \Upsilon \zeta_u^0, v_h) - (\Upsilon \boldsymbol{\zeta}_{\boldsymbol{z}}^0, \nabla v_h) + \langle \Upsilon \hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^0 \cdot \boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = \left(\frac{2}{\Delta t}\left(\partial_t \Upsilon u_h^0 - u_1\right) - \Upsilon u_{h_{tt}}^0, v_h\right),$$

(36b) $$\langle \Upsilon \hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^0 \cdot \boldsymbol{\nu}, \mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0,$$

(36c) $$(\Upsilon \boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \boldsymbol{w}_h) - (\Upsilon \zeta_u^n, \nabla\cdot\boldsymbol{w}_h) + \langle \Upsilon \hat{\zeta}_u^n, \boldsymbol{w}_h\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0,$$

(36d) $$(a\Upsilon \boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \boldsymbol{\tau}_h) - (\Upsilon \boldsymbol{\zeta}_{\boldsymbol{z}}^n, \boldsymbol{\tau}_h) + (\Upsilon I^n(\boldsymbol{\sigma}_{\boldsymbol{h}}), \boldsymbol{\tau}_h) = (\Upsilon E_h^n(\boldsymbol{S}), \boldsymbol{\tau}_h),$$

(36e) $$(\partial_t^2 \zeta_u^n, v_h) - (\Phi \boldsymbol{\zeta}_{\boldsymbol{z}}^n, \nabla v_h) + \langle \Phi \hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^n \cdot \boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = \left(\partial_t^2 u_h^n - \Phi u_{h_{tt}}^n, v_h\right),$$

(36f) $$\langle \Upsilon \hat{\zeta}_u^n, \mu_h\rangle_{\partial\Omega} = 0,$$

(36g) $$\langle \Phi \hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^n \cdot \boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0,$$

for all $(v_h, \boldsymbol{w}_h, \boldsymbol{\tau}_h, \mu_h, m_h) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h \times M_h$. Here,

$$I^n(\boldsymbol{\sigma_h}) = \int_0^{t_n} b(t_n, s)\boldsymbol{\sigma}_h(s)ds,$$

We begin with the proof of (9a). Let $n \geq 1$; then, we start by subtracting (36c) from itself after replacing n by $n-1$ and then, dividing the resulting equation by $2\Delta t$. Secondly, we will perform the same operations in (36f). Next, in (36d), we will replace n by $n-1$ and take the average of the resulting equation with itself. Now, take $\boldsymbol{w}_h = \Phi\boldsymbol{\zeta}_{\boldsymbol{z}}^n$, $\boldsymbol{\tau}_h = \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n$, $v_h = \delta_t\zeta_u^n$, $\mu_h = -\Phi\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^n \cdot \boldsymbol{\nu}$ and $m_h = -\delta_t\hat{\zeta}_u^n$ in (36c), (36d), (36e), (36f) and (36g), respectively and then, add (36c)-(36e), (36f) and (36g) to obtain

$$(a\Phi\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) + \left(\partial_t^2\zeta_u^n, \delta_t\zeta_u^n\right) + \left\langle \Phi\hat{\zeta}_u^n - \Phi\zeta_u^n, \tau(\delta_t\hat{\zeta}_u^n - \delta_t\zeta_u^n)\right\rangle = (\Phi E_h^n(\boldsymbol{S}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n)$$
$$- (\Phi I^n(\boldsymbol{\sigma_h}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) + \left(\Phi u_{h_{tt}}^n - \partial_t^2 u_h^n, \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n\right).$$

Now, we can write $(a\Phi\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n)$ as

$$\begin{aligned}(a\Phi\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) &= \left(a\left(\frac{\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n + \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1}}{2}\right), \frac{\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n - \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1}}{\Delta t}\right)\\ &= \frac{1}{2\Delta t}\Big[(a\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) - (a\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1}, \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1})\Big].\end{aligned}$$

Using a similar approach for other terms, the equation can be further written as

$$\frac{1}{2\Delta t}\Big[\Upsilon\|\partial_t\Upsilon\zeta_u^n\|^2 - \|\partial_t\Upsilon\zeta_u^{n-1}\|^2 + (a\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) - (a\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1}, \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1})$$
$$+ \|\Upsilon\hat{\zeta}_u^n - \Upsilon\zeta_u^n\|_\tau^2 - \|\Upsilon\hat{\zeta}_u^{n-1} - \Upsilon\zeta_u^{n-1}\|_\tau^2\Big]$$
$$= (\Phi E_h^n(\boldsymbol{\sigma_h}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) - (\Phi I^n(\boldsymbol{\sigma_h}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) + (\Phi E_h^n(\boldsymbol{\zeta_\sigma}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) + \left(\Phi u_{h_{tt}}^n - \partial_t^2 u_h^n, \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n\right),$$

Now, multiplying the equation by $2\Delta t$ and adding from $n=1$ to $n=m$, we obtain the following inequality

$$\begin{aligned}\|\partial_t\Upsilon\zeta_u^m\|^2 + \|\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^m\|^2 + \|\Upsilon\hat{\zeta}_u^m - \Upsilon\zeta_u^m\|_\tau^2 &\leq \|\partial_t\Upsilon\zeta_u^0\|^2 + \|\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0\|^2 + \|\Upsilon\hat{\zeta}_u^0 - \Upsilon\zeta_u^0\|_\tau^2\\ &+ 2\Delta t\sum_{n=1}^m (J_1^n + J_2^n + J_3^n),\end{aligned} \tag{37}$$

where

$$J_1^n = (\Phi E_h^n(\boldsymbol{\sigma_h}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n) - (\Phi I^n(\boldsymbol{\sigma_h}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n),\ J_2^n = (\Phi E_h^n(\boldsymbol{\zeta_\sigma}), \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n),$$
$$J_3^n = \left(\Phi u_{h_{tt}}^n - \partial_t^2 u_h^n, \delta_t\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n\right).$$

For the estimates of $\|\partial_t\Upsilon\zeta_u^0\|^2 + \|\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0\|^2 + \|\Upsilon\hat{\zeta}_u^0 - \Upsilon\zeta_u^0\|_\tau^2$, we consider the following equations

(38a)
$$\frac{2}{\Delta t}(\partial_t\Upsilon\zeta_u^0, v_h) - (\Upsilon\boldsymbol{\zeta}_{\boldsymbol{z}}^0, \nabla v_h) + \langle\Upsilon\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^0 \cdot \boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = \left(\frac{2}{\Delta t}\left(\partial_t\Upsilon u_h^0 - u_1\right) - \Upsilon u_{h_{tt}}^0, v_h\right),$$

$$(\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0, \boldsymbol{w}_h) - (\Upsilon\zeta_u^0, \nabla\cdot\boldsymbol{w}_h) + \langle\Upsilon\hat{\zeta}_u^0, \boldsymbol{w}_h\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{38b}$$

$$(a\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0, \boldsymbol{\tau}_h) - (\Upsilon\boldsymbol{\zeta}_{\boldsymbol{z}}^0, \boldsymbol{\tau}_h) + \left(I_1^0, \boldsymbol{\tau}_h\right)ds = \left(\Upsilon E_h^0(\boldsymbol{S}), \boldsymbol{\tau}_h\right), \tag{38c}$$

$$\langle\Upsilon\hat{\zeta}_u^0, \mu_h\rangle_{\partial\Omega} = 0, \tag{38d}$$

$$\langle\Upsilon\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}^0 \cdot \boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{38e}$$

for all $(v_h, \boldsymbol{w}_h, \boldsymbol{\tau}_h, \mu_h, m_h) \in V_h \times \boldsymbol{W}_h \times \boldsymbol{W}_h \times M_h \times M_h$. We take $v_h = \Upsilon\zeta_u^0$, $\boldsymbol{w}_h = \Upsilon \boldsymbol{z}^0$, $\boldsymbol{\tau}_h = \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0$, $\mu_h = -\Upsilon\hat{\zeta}_{\boldsymbol{z}}^0 \cdot \boldsymbol{\nu}$ and $m_h = -\Upsilon\delta_t\hat{\zeta}_u^0$ in (38a), (38b), (38c), (38d) and (38e), respectively and add the resulting equations, to get the following inequality

$$\begin{aligned}
&\|\partial_t \Upsilon\zeta_u^0\|^2 + \|\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0\|^2 + \|\Upsilon\hat{\zeta}_u^0 - \Upsilon\zeta_u^0\|_\tau^2 \\
&\le \frac{1}{2}\left(\Upsilon E_h^0(\boldsymbol{S}), \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0\right) - \frac{1}{2}\int_0^{t_1}\left(b(t_1,s)\boldsymbol{\sigma}_h(s), \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0\right) ds \\
&\quad + \left(\frac{2}{\Delta t}\left(\partial_t \Upsilon u_h^0 - u_1\right) - \Upsilon u_{h_{tt}}^0, \Upsilon\zeta_u^0\right).
\end{aligned}$$

Now, proceeding in the similar way as to obtain (37) will prove that

$$\|\partial_t \Upsilon\zeta_u^0\|^2 + \|\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^0\|^2 + \|\Upsilon\hat{\zeta}_u^0 - \Upsilon\zeta_u^0\|_\tau^2 \le C(h^{2(k+1)} + \Delta t^4).$$

Next, for J_1^n, a use of theorem 4.1 along with quadrature error yields

$$\begin{aligned}
\|\Phi E_h^n(\boldsymbol{\sigma_h}) - \Phi I^n(\boldsymbol{\sigma_h})\| &\le \|\Phi E_h^n(\boldsymbol{\sigma}) - \Phi I^n(\boldsymbol{\sigma}) - \Phi E_h^n(\boldsymbol{e_\sigma}) + \Phi I^n(\boldsymbol{e_\sigma})\| \\
&\le C(h^{k+1} + \Delta t^2).
\end{aligned}$$

Further, use of Young's inequality yields

$$\Delta t \sum_{n=1}^{m} |J_1^n| \le C\left(h^{2(k+1)} + \Delta t^4\right) + \frac{\Delta t}{2}\sum_{n=1}^{m}\left\|\frac{\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n - \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^{n-1}}{\Delta t}\right\|^2. \tag{39}$$

Use of Taylor's series expansion, along with Young's inequality, yields

$$\Delta t \sum_{n=1}^{m} |J_3^n| \le C\left(h^{2(k+1)} + \Delta t^4\right) + \frac{1}{2}\sum_{n=1}^{m}\left\|\frac{\partial_t\Upsilon\zeta_u^n + \partial_t\Upsilon\zeta_u^{n-1}}{2}\right\|^2. \tag{40}$$

Use of (39) and (40) in (37) along with discrete Gronwall's lemma yields

$$\|\partial_t \Upsilon\zeta_u^m\|^2 + \|\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^m\|^2 + \|\Upsilon\hat{\zeta}_u^m - \Upsilon\zeta_u^m\|_\tau^2 \le C\left(h^{2(k+1)} + \Delta t^4\right).$$

Finally, use of Triangle inequality and theorem 4.1, finishes the proof of (9a).
Now, for the proof of (9b), we introduce the following notations:

$$\underset{\sim}{\phi}^0 = 0, \quad \underset{\sim}{\phi}^n = \Delta t\sum_{j=0}^{n-1}\Upsilon\phi^j, \quad \partial_t\Upsilon\underset{\sim}{\phi}^n = \Upsilon\phi^n, \quad \Delta t\sum_{j=0}^{n}\Phi\phi^j = \Upsilon\underset{\sim}{\phi}^n - \frac{\Delta t}{2}\Upsilon\phi^0.$$

Next, we multiply (36d), (36e) and (36g) by k, take summation over n and use (36a) and (36b) to get the following system of equation

(41a)

$$(\Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n, \boldsymbol{w}_h) - (\Upsilon\zeta_u^n, \nabla\cdot\boldsymbol{w}_h) + \langle\Upsilon\hat{\zeta}_u^n, \boldsymbol{w}_h\cdot\boldsymbol{\nu}\rangle_{\partial\mathcal{T}_h} = 0,$$

(41b)

$$(a\Upsilon\underset{\sim}{\boldsymbol{\zeta}_{\boldsymbol{\sigma}}}^n, \boldsymbol{\tau}_h) - (\Upsilon\underset{\sim}{\boldsymbol{\zeta}_{\boldsymbol{z}}}^n, \boldsymbol{\tau}_h) + \left(\Upsilon\underset{\sim}{E_h}^n(\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n), \boldsymbol{\tau}_h\right) = \left(\Upsilon\underset{\sim}{F_h}^n(\boldsymbol{\sigma_h}), \boldsymbol{\tau}_h\right),$$

(41c)

$$(\partial_t\Upsilon\zeta_u^n, v_h) - (\Upsilon\underset{\sim}{\boldsymbol{\zeta}_{\boldsymbol{z}}}^n, \nabla v_h) + \langle\Upsilon\underset{\sim}{\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}}^n\cdot\boldsymbol{\nu}, v_h\rangle_{\partial\mathcal{T}_h} = \left(\Delta t\sum_{j=0}^{n}\left(\partial_t^2 u_h^n - \Phi u_{h_{tt}}^n\right), v_h\right),$$

(41d) $$\langle\Upsilon\hat{\zeta}_u^n, \mu_h\rangle_{\partial\Omega} = 0,$$

(41e) $$\langle\Upsilon\underset{\sim}{\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}}^n\cdot\boldsymbol{\nu}, m_h\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0.$$

Choose $\boldsymbol{w}_h = \Upsilon\underset{\sim}{\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}}^n$, $\boldsymbol{\tau}_h = \Upsilon\boldsymbol{\zeta}_{\boldsymbol{\sigma}}^n$, $v_h = \Upsilon\zeta_u^n$, $\mu_h = -\Upsilon\underset{\sim}{\hat{\boldsymbol{\zeta}}_{\boldsymbol{z}}}^n\cdot\boldsymbol{\nu}$ and $m_h = -\Upsilon\hat{\zeta}_u^n$ in (41a), (41b), (41c), (41d) and (41e), respectively, and add the resulting equations. After simplifying as above, we attain the desired estimate. For further details, see [12].

□

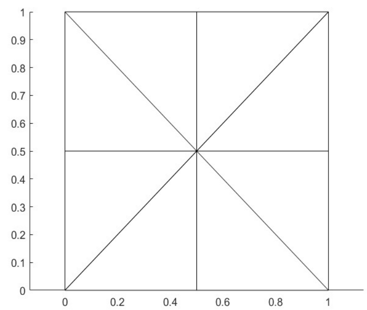
8. Numerical Results

The performance of the suggested HDG approach for the hyperbolic integro differential equations (1a)-(1c) is discussed in this section. Figure 1 shows the domain discretization used for different mesh sizes. The problem has been discretized using the central difference technique, and the integral term has been approximated using the mid-point rule. For the sake of simplicity, the function a is chosen to be 1 throughout, with the problem domain being $\Omega = (0,1)\times(0,1)$.

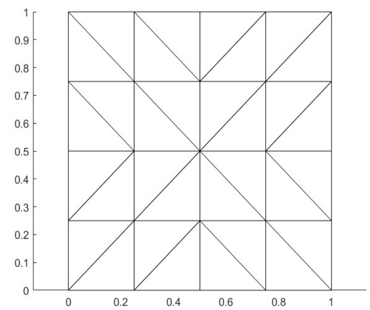
We demonstrate the order of convergence for the L^2-norm of the error in u, the gradient $\boldsymbol{\sigma} = -\nabla u$, and post-processed solution u_h^*. We see that the superconvergence for u_h^* and the optimal convergence for u and $\boldsymbol{\sigma}$ are realized as anticipated by our derived results.

Example 1. Let $u(x,y,t) = t^2e^tx(1-x)y(1-y)$ represent the precise solution with $b(x,t,s) = e^{t-s}$. Table 1 displays the computed order of convergence and L^2 error estimates for u and $\boldsymbol{\sigma}$, while Table 3 displays the computed order of convergence and L^2 error estimates for u_h^* at $t = \frac{1}{2}$ for $k=1$, $k=2$, and $k=3$ for a variety of h values. We observe that the convergence rates for $\|e_u\|$, $\|\boldsymbol{e}_{\boldsymbol{\sigma}}\|$ and $\|e_u^*\|$ are on the order of $O(h^{k+1})$, $O(h^{k+1})$ and $O(h^{k+2})$, respectively.

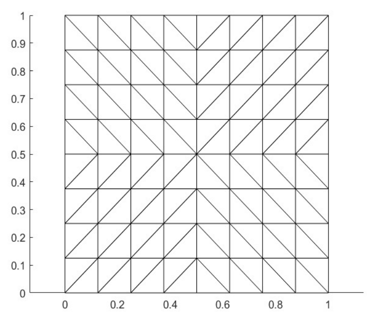
Example 2. Let $u(x,y,t) = t\sin(\pi t)\sin(\pi x)\sin(\pi y)$ represent the precise solution with $b(x,t,s) = \sin(\pi t)\cos(\pi s)$. Table 2 displays the computed order of convergence and L^2 error estimates for u and $\boldsymbol{\sigma}$, while Table 3 displays the computed order of convergence and L^2 error estimates for u_h^* at $t = \frac{1}{2}$ for $k=1$, $k=2$, and $k=3$ for a variety of h values. We observe that the convergence rates for $\|e_u\|$, $\|\boldsymbol{e}_{\boldsymbol{\sigma}}\|$ and $\|e_u^*\|$ are on the order of $O(h^{k+1})$, $O(h^{k+1})$ and $O(h^{k+2})$, respectively.



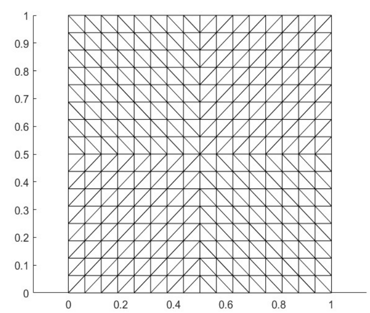
(A) $h = \frac{1}{2}$



(B) $h = \frac{1}{4}$



(C) $h = \frac{1}{8}$



(D) $h = \frac{1}{16}$

FIGURE 1. domain discretization for different values of h

TABLE 1. Order of convergence and L^2 error estimates for *Example* 1

		u_h		$\boldsymbol{\sigma}_h$	
k	h	$\|u^M - U^M\|$	**Order**	$\|\boldsymbol{\sigma}^M - \boldsymbol{S}^M\|$	**Order**
1	**1/2**	3.1116e-02	-	1.1484e-01	-
	1/4	6.9328e-03	2.1662	2.2498e-02	2.3518
	1/8	1.8259e-03	1.9248	5.5337e-03	2.0235
	1/16	4.7854e-04	1.9319	1.3897e-03	1.9934
2	**1/2**	7.1740e-03	-	2.5155e-02	-
	1/4	6.8389e-04	3.3909	1.9663e-03	3.6773
	1/8	8.3343e-05	3.0366	2.0802e-04	3.2407
	1/16	1.9478e-06	3.0635	3.6215e-06	3.3576
3	**1/2**	2.2952e-03	-	1.0145e-02	-
	1/4	7.1660e-05	5.0013	3.0543e-04	5.0538
	1/8	2.5553e-06	4.8096	9.6065e-06	4.9907
	1/16	1.3766e-08	4.5306	3.7858e-08	4.8432

9. Conclusion

This paper proposes and analyses HDG method for a hyperbolic integro-differential equation. Error estimates have been derived using HDG and by introducing Ritz-Volterra projections for the model problem. It is also shown that the Ritz-Volterra

TABLE 2. Order of convergence and L^2 error estimates for *Example* 2

		u_h		$\boldsymbol{\sigma}_h$	
k	h	$\|u^M - U^M\|$	**Order**	$\|\boldsymbol{\sigma}^M - \boldsymbol{S}^M\|$	**Order**
1	**1/2**	1.6315e-01	-	6.1879e-01	-
	1/4	2.2194e-02	2.8780	8.9694e-02	2.7864
	1/8	2.5502e-03	3.1215	1.1478e-02	2.9661
	1/16	2.9320e-04	3.1206	1.3000e-03	3.1424
2	**1/2**	1.6115e-02	-	6.6283e-02	-
	1/4	1.0174e-03	3.9855	4.4295e-03	3.9034
	1/8	6.4802e-05	3.9726	2.6761e-04	4.0489
	1/16	6.0993e-07	3.9564	2.3337e-06	4.0931
3	**1/2**	7.8359e-03	-	3.4452e-02	-
	1/4	2.2730e-04	5.1075	1.0040e-03	5.1007
	1/8	6.8771e-07	4.9264	4.3293e-06	4.8241
	1/16	2.5063e-09	4.9214	1.8006e-08	4.7892

TABLE 3. Order of convergence and L^2 error estimates for u_h^*

		Example 1		*Example* 2	
k	h	$\|u^M - u_h^{*M}\|$	**Order**	$\|u^M - u_h^{*M}\|$	**Order**
1	**1/2**	9.0267e-03	-	1.3336e-01	-
	1/4	8.3514e-04	3.4341	1.9980e-02	2.7387
	1/8	9.9647e-05	3.0671	2.5594e-03	2.9646
	1/16	1.2100e-05	3.0417	2.8749e-04	3.1543
2	**1/2**	4.6657e-03	-	1.4340e-02	-
	1/4	2.7192e-04	4.1009	9.8668e-04	3.8614
	1/8	1.6997e-05	3.9998	5.8242e-05	4.0825
	1/16	1.5014e-07	4.0361	4.9113e-07	4.1347
3	**1/2**	2.2604e-03	-	7.7002e-03	-
	1/4	6.7900e-05	5.0570	2.2518e-04	5.0957
	1/8	2.1192e-06	5.0018	6.5609e-07	5.1452
	1/16	7.1367e-09	5.0004	1.8317e-09	5.1879

projection achieves convergence of order $h^{k+3/2}$, for $k \geq 1$. In addition, the element-by-element post-processing of the numerical solution was accomplished by utilizing the dual of the problem. The findings demonstrate that all the three variables, namely, u, $\boldsymbol{\sigma}$ and $\boldsymbol{z}$ attain convergence of order $k+1$, for non-negative k in h, which is the discretizing parameter of the space domain. In contrast, the post-processed solution attains superconvergence; that is, it converges with order $k+2$, for $k \geq 1$. The analysis of this article provides better accuracy results compared to [12]. Finally, numerical results are reviewed. This study may be carried over to the three-dimensional domain by making the appropriate adjustments.

10. Acknowledgments

The second author acknowledges the support of DST-SERB (CRG/2020/001599).

References

[1] Pani, A.K. and Sinha, R.K., Error estimates for semidiscrete Galerkin approximation to a time dependent parabolic integro-differential equation with nonsmooth data, 2000.

[2] Pani, A.K. and Thomée, V. and Wahlbin, L.B., Numerical methods for hyperbolic and parabolic integro-differential equations, Journal of Integral Equations and Applications, 1992.

[3] Karaa, S. and Pani, A.K., Optimal error estimates of mixed FEMs for second order hyperbolic integro-differential equations with minimal smoothness on initial data, Journal of Computational and Applied Mathematics, 2015.
[4] Chen, C. and Zhang, X. and Zhang, G. and Zhang, Y., A two-grid finite element method for nonlinear parabolic integro-differential equations, International Journal of Computer Mathematics, 2019.
[5] Cockburn, B. and Dong, B. and Guzmán, J., A Superconvergent LDG-Hybridizable Galerkin Method for Second-Order Elliptic Problems, Mathematics of Computation, 2008.
[6] Cockburn, B. and Gopalakrishnan, J., A Characterization of Hybridized Mixed Methods for Second Order Elliptic Problems, SIAM J. Numerical Analysis, 2004.
[7] Cockburn, B. and Gopalakrishnan, J. and Lazarov, R., Unified Hybridization of Discontinuous Galerkin, Mixed, and Continuous Galerkin Methods for Second Order Elliptic Problems, SIAM J. Numer. Anal, 2009.
[8] Cockburn, B. and Guzmán, J. and Wang, H., Superconvergent discontinuous Galerkin methods for second-order elliptic problems, Math. Comput., 2009.
[9] Cockburn, B. and Gopalakrishnan, J. and Sayas, F.-J., A projection-based error analysis of HDG methods, Mathematics of Computation, 2010.
[10] Yadav, S. and Pani, A.K., Superconvergent discontinuous Galerkin methods for nonlinear parabolic initial and boundary value problems, Journal of Numerical Mathematics, 2019.
[11] Chabaud, B. and Cockburn, B., Uniform-in-time superconvergence of HDG methods for the heat equation, Mathematics of Computation, 2012.
[12] Karaa, S. and Pani, A.K. and Yadav, S., A priori *hp*-estimates for discontinuous Galerkin approximations to linear hyperbolic integro-differential equations, Applied Numerical Mathematics, 2015.
[13] Saedpanah, F., A posteriori error analysis for a continuous space-time finite element method for a hyperbolic integro-differential equation, BIT Numerical Mathematics, 2013.
[14] Saedpanah, F., Continuous Galerkin finite element methods for hyperbolic integro-differential equations, IMA Journal of Numerical Analysis, 2015.
[15] Merad, A. and Martin-Vaquero, J., A Galerkin method for two-dimensional hyperbolic integro-differential equation with purely integral conditions, Applied Mathematics and Computation, 2016.
[16] Tan, Z. and Li, K. and Chen, Y., A fully discrete two-grid finite element method for nonlinear hyperbolic integro-differential equation, Applied Mathematics and Computation, 2022.
[17] Nguyen, N.C. and Peraire, J. and Cockburn, B., An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier–Stokes equations, Journal of Computational Physics, 2011.
[18] Cesmelioglu, A. and Cockburn, B. and Qiu, W., Analysis of a hybridizable discontinuous Galerkin method for the steady-state incompressible Navier-Stokes equations, Mathematics of Computation, 2017.
[19] Nguyen, N.C. and Peraire, J. and Cockburn, B., Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations, Journal of Computational Physics, 2011.
[20] Du, S. and Sayas, F.-J., A unified error analysis of hybridizable discontinuous Galerkin methods for the static Maxwell equations, SIAM Journal on Numerical Analysis, 2020.
[21] Rhebergen, S. and Cockburn, B., Space-time hybridizable discontinuous Galerkin method for the advection–diffusion equation on moving and deforming meshes, The Courant–Friedrichs–Lewy (CFL) Condition: 80 Years After Its Discovery, 2013.
[22] Kirk, K.LA. and Horvath, T.L. and Cesmelioglu, A. and Rhebergen, S., Analysis of a space-time hybridizable discontinuous Galerkin method for the advection-diffusion problem on time-dependent domains, SIAM Journal on Numerical Analysis, 2019.
[23] Moon, M. and Jun, H.K. and Suh, T., Error estimates on hybridizable discontinuous Galerkin methods for parabolic equations with nonlinear coefficients, . Advances in Mathematical Physics, 2017.
[24] Cockburn, B. and Quenneville-Bélair, V., Uniform-in-time superconvergence of the HDG methods for the acoustic wave equation, Mathematics of Computation, 2014.
[25] Stanglmeier, M. and Nguyen, N.C. and Peraire, J. and Cockburn, B., An explicit hybridizable discontinuous Galerkin method for the acoustic wave equation, Computer Methods in Applied Mechanics and Engineering, 2016.
[26] Jain, R. and Pani, A.K. and Yadav, S., HDG method for linear parabolic integro-Differential equations, Applied Mathematics and Computation, 2023.

[27] Cockburn, Bernardo and Guzmán, Johnny and Wang, Haiying, Superconvergent discontinuous Galerkin methods for second-order elliptic problems, Mathematics of Computation, 2009.

[28] Renardy, Michael and Hrusa, William and Nohel, John A, Mathematical problems in viscoelasticity, 1987.

Department of Mathematics, Birla Institute of Technology and Science, Pilani, Pilani Campus, Vidya Vihar, Pilani Rajasthan 333031 India

E-mail: p20180422@pilani.bits-pilani.ac.in, sangita.yadav@pilani.bits-pilani.ac.in

URL: www.bits-pilani.ac.in/, universe.bits-pilani.ac.in/pilani/sangitayadav/profile