ESAIM: M2AN 57 (2023) 2097–2129
https://doi.org/10.1051/m2an/2023048

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A PRIORI ERROR ANALYSIS OF NEW SEMIDISCRETE, HAMILTONIAN HDG METHODS FOR THE TIME-DEPENDENT MAXWELL'S EQUATIONS

Bernardo Cockburn[1], Shukai Du[2,*] and Manuel A. Sánchez[3]

Abstract. We present the first *a priori* error analysis of a class of space-discretizations by Hybridizable Discontinuous Galerkin (HDG) methods for the time-dependent Maxwell's equations introduced in Sánchez *et al.* [*Comput. Methods Appl. Mech. Eng.* **396** (2022) 114969]. The distinctive feature of these discretizations is that they display a discrete version of the Hamiltonian structure of the original Maxwell's equations. This is why they are called "Hamiltonian" HDG methods. Because of this, when combined with symplectic time-marching methods, the resulting methods display an energy that does *not drift* in time. We provide a single analysis for several of these methods by exploiting the fact that they only differ by the choice of the approximation spaces and the stabilization functions. We also introduce a new way of discretizing the static Maxwell's equations in order to define the initial condition in a manner consistent with our technique of analysis. Finally, we present numerical tests to validate our theoretical orders of convergence and to explore the convergence properties of the method in situations not covered by our analysis.

Mathematics Subject Classification. 65M60, 74H15, 74J05, 74S05.

Received November 4, 2022. Accepted May 23, 2023.

1. Introduction

The goal of this paper is to present the first *a priori* error analysis of a class Hybridizable Discontinuous Galerkin (HDG) methods for the time-dependent Maxwell's equations. This class of HDG methods is based on the formulation of the magnetic field and its vector potential. It is well-suited for symplectic time-integration and has the advantage of providing time-invariant, non-drifting approximations to the total electric and magnetic charges, and to the total energy, see [31].

To describe our results, let us begin by considering the classical formulation of the Maxwell's equations based on the electric and magnetic field:

$$\epsilon\,\dot{\boldsymbol{E}} = \nabla\times\boldsymbol{H} - \sigma\boldsymbol{E} - \boldsymbol{J} \qquad \text{in } \Omega\times(0,T], \tag{1a}$$

Keywords and phrases. Time-dependent Maxwell's equations, symplectic Hamiltonian finite element methods, discontinuous Gakerkin, hybridizable discontinuous Galerkin, error analysis.

[1] School of mathematics, University of Minnesota, Minneapolis, MN 55455, USA.

[2] Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA.

[3] Instituto de Ingeniería Matemática y Computacional, Facultad de Matemáticas y Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile.

*Corresponding author: sdu49@wisc.edu

© The authors. Published by EDP Sciences, SMAI 2023

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

$$\mu\,\dot{\boldsymbol{H}} = -\nabla \times \boldsymbol{E} \qquad \text{in } \Omega \times (0,T], \tag{1b}$$
$$\nabla \cdot (\epsilon\,\boldsymbol{E}) = \rho \qquad \text{in } \Omega \times (0,T], \tag{1c}$$
$$\nabla \cdot (\mu\,\boldsymbol{H}) = 0 \qquad \text{in } \Omega \times (0,T], \tag{1d}$$

complete with the following initial and boundary conditions:

$$\boldsymbol{n} \times \boldsymbol{E} = \boldsymbol{g}_E \qquad \text{on } \Gamma \times (0,T], \quad \Gamma := \partial\Omega, \tag{1e}$$
$$\boldsymbol{E} = \boldsymbol{E}_0,\ \boldsymbol{H} = \boldsymbol{H}_0 \qquad \text{on } \Omega \times \{t=0\}. \tag{1f}$$

Here we assume $\Omega \subset \mathbb{R}^d$ is a polyhedral domain with Lipschtz boundary. In addition, we assume that ϵ and μ, both lying in $L^\infty(\Omega)$, are bounded from below by $\epsilon_0 > 0$ and $\mu_0 > 0$, respectively, and $\sigma, \rho \in L^\infty(\Omega)$ are non-negative functions.

We are interested in finite element spatial discretizations of (1). A natural way is to use Nédélec's $H(\text{curl})$-conforming edge elements [27, 28]. Compared to H^1-conforming nodal elements, they provide more robust convergence. For instance, it is well known that nodal elements can give spurious solution when the domain has re-entrant corners [12, 13], or spurious modes when the domain has closed cavities [3]. Edge elements are free from these issues. For error analyses of edge elements for the time-dependent Maxwell's equations, we refer to [23, 25, 26].

A key reason for the above-mentioned lack of convergence by H^1-conforming elements is their heavy enforcement of the inter-element conformity along the normal direction, since this renders the discrete space not dense in the solution space. Therefore, it is not surprising to see that Discontinuous Galerkin (DG) methods are free from spurious solutions [5]. In addition, compared to edge elements, DG methods are *more straightforward to implement* for *hp*-adaptivty and meshes with hanging-nodes. Numerical studies also suggest DG methods reach high scaliability due to their small parallel communication footprint [24]. These advantages make DG an appealing class of methods to solve the Maxwell's equations. However, since no conformity is enforced in the discrete space, DG methods display more globally-coupled degrees of freedom for the same mesh, as compared with edge methods. To overcome this problem, the Hybridizable Discontinuous Galerkin (HDG) methods were introduced. In the framework of electromagnetic problems, they first appeared in [30] for solving the time-harmonic Maxwell's equations. Error analyses are given in [6–9, 15, 22] for either the time-harmonic or the static Maxwell's equations.

Relatively fewer papers exist on HDG methods for the time-dependent Maxwell's equations; see [10, 19, 21, 29]. Most existing methods are based on a direct application of the HDG discretization of (1), namely, by using approximations of $\boldsymbol{E}$ and $\boldsymbol{H}$ on each element, as well as approximations of the tangential traces of $\boldsymbol{E}$ and $\boldsymbol{H}$ on the skeleton of the mesh. In [31], we show that the methods obtained in this way, that is, by using HDG space discretiations, are *not* a Hamiltonian dynamical system. As expected, we numerically observe a drifting of the total energy, the electric and magnetic charges, and the linear and angular momenta. This makes $\boldsymbol{E}$-$\boldsymbol{H}$ based methods unsuitable for long-time simulation. To overcome this problem, we proposed a new class of methods based on a rewriting of the original problem (1) in terms of the electric field $\boldsymbol{E}$ and the magnetic potential $\boldsymbol{A}$. The corresponding $\boldsymbol{E}$-$\boldsymbol{A}$ formulation [31] reads as follows:

$$-\,\epsilon\,\ddot{\boldsymbol{A}} = \nabla \times \boldsymbol{H} + \sigma\,\dot{\boldsymbol{A}} - \boldsymbol{J} \qquad \text{in } \Omega \times (0,T], \tag{2a}$$
$$\mu\boldsymbol{H} = \nabla \times \boldsymbol{A} \qquad \text{in } \Omega \times (0,T], \tag{2b}$$
$$\nabla \cdot (\epsilon\boldsymbol{E}) = \rho \qquad \text{in } \Omega \times (0,T], \tag{2c}$$

and is complete with the following initial and boundary conditions:

$$\boldsymbol{n} \times \boldsymbol{A} = \boldsymbol{g}_A \qquad \text{on } \Gamma \times (0,T], \quad \Gamma := \partial\Omega, \tag{2d}$$
$$\boldsymbol{A} = \boldsymbol{A}_0,\ \dot{\boldsymbol{A}} = -\boldsymbol{E}_0 \qquad \text{on } \Omega \times \{t=0\}. \tag{2e}$$

Here $\boldsymbol{A}$ is the magnetic vector potential; see (2b). Note that the condition (2c) (or (1c)) only needs to be enforced in the initial condition (2e), and similarly for (1d); see [25]. Correspondingly, in the HDG methods we shall introduce later based on (2), the equation (2c) is not used in the discretization of the evolutionary equation (3), but is used in the discretization of the initial condition (5e). We proved that the HDG methods obtained based on (2) are Hamiltonian dynamical systems. Consequently, they provide non-drifting energy, charges and momenta; see Corollary 4.2 of [31] for the proof for these non-drifting properties.

In this paper, we present an *a priori* error analysis for the HDG methods based on the formulation (2). To the best of our knowledge, this is the first error analysis for HDG methods for the time-dependent Maxwell's equations. By adopting the techniques used in [15], we are able to analyze at once most existing spatial HDG discretizations, see Table 1. Moreover, since (2) is essentially a wave-form formulation based on the magnetic potential, we are able to reuse many existing analysis techniques from [11, 14] for acoustic and elastic waves. However, for the Maxwell's equations special attention needs to be paid to the discretization of the initial condition in order to ensure that the initial error does not pollute the convergence order for the rest of the time. To achieve this, a new class of HDG methods for discretizing the static Maxwell's equations is introduced.

We present our main result in Section 2. We describe the class of HDG semi-discretizations as well as the choice of the approximation to the initial condition. We then state and and discuss our error estimates. Then in Section 3, we present its detailed proof. Finally, in Section 5, we show numerical tests which validate our theoretical results.

2. The main results

2.1. Notation and preliminaries

We begin by introducing some notation. Let $\mathcal{T}_h$ be a collection of conforming triangulations of Ω by shape-regular, star-shaped polyhedral elements. Let $\mathcal{F}_h$ denote the set of all the faces of $\mathcal{T}_h$, and $\mathcal{F}_K$ the set of faces of an element $K \in \mathcal{T}_h$. For any element $K \in \mathcal{T}_h$, we denote its diameter by h_K, and set $h := \max_{K\in\mathcal{T}_h} h_K$ and $h_{min} := \min_{K\in\mathcal{T}_h} h_K$.

Let X be either an element $K \in \mathcal{T}_h$ or a face $F \in \mathcal{F}_h$. We denote by $\mathcal{P}_k(X)$ and $\boldsymbol{\mathcal{P}}_k(X)$ the polynomial spaces of degree k, supported on X, taking values in $\mathbb{R}$ and $\mathbb{R}^d$, respectively. We let $\boldsymbol{\Pi}_k$ and Π_k denote the L^2-projection into $\boldsymbol{\mathcal{P}}_k(K)$ and $\mathcal{P}_k(K)$, respectively. For any vector normal to a face F, $\boldsymbol{n}_F$, and any linear space of $\mathbb{R}^d$-valued functions defined on F, $\boldsymbol{X}(F)$, we set $\boldsymbol{X}^t(F) := \{\boldsymbol{u} \in \boldsymbol{X}(F) : \boldsymbol{u}\cdot\boldsymbol{n}_F = 0\}$, and $\boldsymbol{X}^n(F) := \{\boldsymbol{u} \in \boldsymbol{X}(F) : \boldsymbol{u}\times\boldsymbol{n}_F = 0\}$. For an $\mathbb{R}^d$-valued space $\boldsymbol{X}(K)$ defined on an element K, we shall denote $\boldsymbol{X}^t(\partial K) := \prod_{F\in\mathcal{F}_K} \boldsymbol{X}^t(F)$, where $\mathcal{F}_K$ is the collection of the faces of K, and $\boldsymbol{X}(F)$ is the restriction of the space $\boldsymbol{X}(K)$ on the face F.

For any subdomain X of $\mathbb{R}^d$, we set $(u, v)_X := \int_X uv$ for scalar-valued functions u and v, and $(\boldsymbol{u}, \boldsymbol{v})_X := \int_X \boldsymbol{u}\cdot\mathbf{v}$ for $\mathbb{R}^d$-valued functions $\boldsymbol{u}$ and $\boldsymbol{v}$. We then write

$$(\boldsymbol{u}, \boldsymbol{v})_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (\boldsymbol{u}, \boldsymbol{v})_K, \quad \langle \boldsymbol{u}, \boldsymbol{v}\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \langle \boldsymbol{u}, \boldsymbol{v}\rangle_{\partial K}, \quad \langle \boldsymbol{u}, \boldsymbol{v}\rangle_{\partial\mathcal{T}_h\setminus\Gamma} := \sum_{K\in\mathcal{T}_h} \langle \boldsymbol{u}, \boldsymbol{v}\rangle_{\partial K\setminus\Gamma}.$$

For any natural number s, we set

$$|u|^2_{s,\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} |u|^2_{s,K} \quad \text{and} \quad \|u\|^2_{s,\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \|u\|^2_{s,K},$$

where, as usual, $|u|^2_{s,X} := \sum_{|\alpha|=s} \|D^\alpha u\|^2_X$ and $\|u\|^2_{s,X} := \sum_{l\leqslant s} |u|^2_{l,X}$. Finally, for functions $\boldsymbol{f}$ depending on time and space, and for any natural numbers ℓ, s, we set

$$\|\boldsymbol{f}\|_{W^{\ell,p}(T,H^s(\mathcal{T}_h))} := \left(\int_0^T \sum_{m=0}^{\ell} \left\|\boldsymbol{f}^{(m)}(t)\right\|^p_{s,\mathcal{T}_h} \mathrm{d}t\right)^{1/p} \qquad \forall p \in [1,\infty),$$

$$\|\boldsymbol{f}\|_{W^{\ell,\infty}(T,H^s(\mathcal{T}_h))} := \sup_{t\in(0,T)} \max_{0\leqslant m\leqslant\ell} \left\|\boldsymbol{f}^{(m)}(t)\right\|_{s,\mathcal{T}_h}$$

TABLE 1. The HDG variants ($k \geqslant 0$) for the spatial discretization of the $\boldsymbol{E}$-$\boldsymbol{A}$ formulation of the time-dependent Maxwell's equations. They satisfy the Assumption 2.1 on the local spaces.

Method	$\boldsymbol{W}$	$\boldsymbol{V}$	$\boldsymbol{M}$	Ref.
HDG-S$_k$	$\boldsymbol{\mathcal{P}}_k$	$\boldsymbol{\mathcal{P}}_k$	$\boldsymbol{\mathcal{P}}_k^t$	[8, 30]
HDG-B$_k$	$\boldsymbol{\mathcal{P}}_k$	$\boldsymbol{\mathcal{P}}_{k+1}$	$\boldsymbol{\mathcal{P}}_{k+1}^t$	[9, 15]
HDG-B$_k$+	$\boldsymbol{\mathcal{P}}_k$	$\boldsymbol{\mathcal{P}}_{k+1}$	$\boldsymbol{\mathcal{P}}_k^t \oplus \nabla_F \tilde{\mathcal{P}}_{k+2}$	[7]

where $\boldsymbol{f}^{(m)}$ represents the m-th order time derivative of the function $\boldsymbol{f}$. We write L^∞ instead of $W^{0,\infty}$, and L^2 instead of H^0. Similar definitions hold when replacing $\mathcal{T}_h$ by $\partial\mathcal{T}_h$.

We now introduce the notation for the approximation spaces of our numerical methods. Given any element $K \in \mathcal{T}_h$, we let $\boldsymbol{V}(K)$ and $\boldsymbol{W}(K)$ represent the approximation spaces for the magnetic vector potential $\boldsymbol{A}$ and the magnetic field $\boldsymbol{H}$ on K, respectively. Since $\dot{\boldsymbol{A}} = -\boldsymbol{E}$, $\boldsymbol{V}(K)$ is also the space for the electric field $\boldsymbol{E}$. For any face $F \in \mathcal{F}_h$, we denote by $\boldsymbol{M}(F)$ the approximation space for the *tangential component* of the magnetic potential $\boldsymbol{A}$ on F. We also set $\boldsymbol{X}(\partial K) := \prod_{F\in\mathcal{F}_K} \boldsymbol{X}(F)$ where $\boldsymbol{X}(F)$ can be $\boldsymbol{M}(F)$, $\mathcal{P}_k(F)$, or $\boldsymbol{\mathcal{P}}_k(F)$. With the local spaces defined, we introduce the global approximation spaces:

$$\boldsymbol{V}_h = \prod_{K\in\mathcal{T}_h} \boldsymbol{V}(K), \quad \boldsymbol{W}_h = \prod_{K\in\mathcal{T}_h} \boldsymbol{W}(K), \quad \boldsymbol{M}_h = \prod_{F\in\mathcal{F}_h} \boldsymbol{M}(F),$$

and set

$$\boldsymbol{M}_h := \boldsymbol{M}_h^0 \oplus \boldsymbol{M}_h^\partial,$$

where $\boldsymbol{M}_h^0$ and $\boldsymbol{M}_h^\partial$ represent the subspaces of $\boldsymbol{M}_h$ taking zero values on the boundary faces and the interior faces, respectively.

Finally, we denote by k_X the maximal polynomial degree k such that $\boldsymbol{\mathcal{P}}_k(K) \subset \boldsymbol{X}(K)$ or $\boldsymbol{\mathcal{P}}_k^t(F) \subset \boldsymbol{X}(F)$ hold for all $K \in \mathcal{T}_h$ or $F \in \mathcal{F}_h$, where $\boldsymbol{X}$ can be $\boldsymbol{V}$, $\boldsymbol{W}$, or $\boldsymbol{M}$.

2.2. Spatial semi-discretization

We assume the approximation spaces $\boldsymbol{V}$, $\boldsymbol{W}$ and $\boldsymbol{M}$ satisfy the following assumption:

Assumption 2.1 (On the local spaces)**.** *For all $K \in \mathcal{T}_h$, we have that*

(1) $\nabla \times \boldsymbol{V}(K) \subset \boldsymbol{W}(K)$,
(2) $\nabla \times \boldsymbol{W}(K) \subset \boldsymbol{V}(K)$,
(3) $\boldsymbol{n}_{\partial K} \times \left(\boldsymbol{W}(K)\big|_{\partial K}\right) \subset \boldsymbol{M}(\partial K)$,
(4) $\{\boldsymbol{v} \in \boldsymbol{V}(K) : \nabla \times \boldsymbol{v} = 0\}^t \subset \boldsymbol{M}(\partial K)$.

Assumptions 2.1-(1), -(2) and -(3) were first used in [15] for a unified analysis of HDG methods for the static Maxwell's equations. Assumption 2.1-(4) is added in order to guarantee the existence of a class of projections needed for our error analysis; see Proposition 3.1.

Note that we do not require $\boldsymbol{V}^t(\partial K) \subset \boldsymbol{M}(\partial K)$. Thus, $\boldsymbol{V}^t(\partial K)$ can be larger than $\boldsymbol{M}(\partial K)$. To handle this case, we use the L^2-projection $\boldsymbol{P}_M : L^2(\partial K; \mathbb{R}^3) \to \boldsymbol{M}(\partial K)$. The appearance of projections of this type is a hallmark of the Lehrenfeld–Schöberl projection for HDG methods, which was first introduced in [20] and later applied to the static Maxwell's equations in [7]. A remarkable feature of these methods is that they achieve super-convergence (of their numerical traces) on general polyhedral meshes.

Our analysis holds for any HDG method satisfying the above Assumption 2.1 on the local spaces. Examples are the three HDG methods listed in Table 1.

Let us now define the HDG space-discretization methods for equations (2). They determine the approximation $(\boldsymbol{A}_h, \widehat{\boldsymbol{A}}_h, \boldsymbol{H}_h)$ in the space $\boldsymbol{V}_h \times \boldsymbol{M}_h \times \boldsymbol{W}_h$ as the solution of

$$\left(\epsilon \ddot{\boldsymbol{A}}_h, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma \dot{\boldsymbol{A}}_h, \boldsymbol{v}\right)_{\mathcal{T}_h} + (\nabla \times \boldsymbol{H}_h, \boldsymbol{v})_{\mathcal{T}_h} + \left\langle \tau_t \left(\boldsymbol{P}_M \boldsymbol{A}_h - \widehat{\boldsymbol{A}}_h\right), \boldsymbol{v}\right\rangle_{\partial \mathcal{T}_h} = (\boldsymbol{J}, \boldsymbol{v})_{\mathcal{T}_h}, \tag{3a}$$

$$(\mu \boldsymbol{H}_h, \boldsymbol{w})_{\mathcal{T}_h} - \left\langle \boldsymbol{n} \times \widehat{\boldsymbol{A}}_h, \boldsymbol{w}\right\rangle_{\partial \mathcal{T}_h} - (\boldsymbol{A}_h, \nabla \times \boldsymbol{w})_{\mathcal{T}_h} = 0, \tag{3b}$$

$$\left\langle \boldsymbol{n} \times \boldsymbol{H}_h + \tau_t (\boldsymbol{A}_h - \widehat{\boldsymbol{A}}_h), \boldsymbol{\eta}^0\right\rangle_{\partial \mathcal{T}_h} = 0, \tag{3c}$$

$$\left\langle \widehat{\boldsymbol{A}}_h, \boldsymbol{\eta}^\partial\right\rangle_{\partial \mathcal{T}_h} = \langle \boldsymbol{g}_A \times \boldsymbol{n}, \boldsymbol{\eta}^\partial\rangle_{\partial \mathcal{T}_h}, \tag{3d}$$

for all $(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{\eta}^0, \boldsymbol{\eta}^\partial) \in \boldsymbol{V}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h^0 \times \boldsymbol{M}_h^\partial$.

Clearly, the projection $\boldsymbol{P}_M$ becomes the tangential-restriction operator in equation (3a) if $\boldsymbol{V}^t(K) \subset \boldsymbol{M}(\partial K)$ for all $K \in \mathcal{T}_h$. Note also that this equation can be written as

$$\left(\epsilon \ddot{\boldsymbol{A}}_h, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma \dot{\boldsymbol{A}}_h, \boldsymbol{v}\right)_{\mathcal{T}_h} + (\boldsymbol{H}_h, \nabla \times \boldsymbol{v})_{\mathcal{T}_h} + \left\langle \boldsymbol{n} \times \widehat{\boldsymbol{H}}_h, \boldsymbol{v}\right\rangle_{\partial \mathcal{T}_h} = (\boldsymbol{J}, \boldsymbol{v})_{\mathcal{T}_h},$$

by introducing the numerical trace

$$\widehat{\boldsymbol{H}}_h := \boldsymbol{H}_h^t + \tau_t \left(\boldsymbol{P}_M \boldsymbol{A}_h - \widehat{\boldsymbol{A}}_h\right) \times \boldsymbol{n},$$

where τ_t is the stabilization function which we assume to be a positive face-wise constant function lying in $\prod_{K \in \mathcal{T}_h} \prod_{F \in \mathcal{F}_K} \mathcal{P}_0(F)$. In addition, we denote by $\tau_{t,K}^{\max}$ and $\tau_{t,K}^{\min}$ the maximum and the minimum value of τ_t on ∂K, respectively.

To complete the semi-discretization (3), we need to give an initial condition for $\boldsymbol{A}_h(0)$ and $\dot{\boldsymbol{A}}_h(0)$. Let us first consider $\dot{\boldsymbol{A}}_h(0)$. We define it as a projection of $\dot{\boldsymbol{A}}(0)$:

$$\dot{\boldsymbol{A}}_h(0) = \boldsymbol{\Pi}_V \dot{\boldsymbol{A}}(0) = \boldsymbol{\Pi}_V(-\boldsymbol{E}_0), \tag{4}$$

where $\boldsymbol{\Pi}_V$ is the L^2-projection onto $\boldsymbol{V}$. The last equation is motivated by the fact that $\dot{\boldsymbol{A}} = -\boldsymbol{E}$.

It remains to define $\boldsymbol{A}_h(0)$. Note that $\boldsymbol{A}_h(0)$ has to satisfy (3b)–(3d) at $t = 0$. However, these three equations are not enough to determine $\boldsymbol{A}_h(0)$. Therefore, we add equations to complete its definition. So, we seek

$$\left(\boldsymbol{H}_h(0), \boldsymbol{A}_h(0), \widehat{\boldsymbol{A}}_h(0), p_h\right) \in \boldsymbol{W}_h \times \boldsymbol{V}_h \times \boldsymbol{M}_h \times Q_h^{c,0},$$

such that

$$(\nabla \times \boldsymbol{H}_h(0), \boldsymbol{v})_{\mathcal{T}_h} + \left\langle \tau_t \left(\boldsymbol{P}_M \boldsymbol{A}_h(0) - \widehat{\boldsymbol{A}}_h(0)\right), \boldsymbol{v}\right\rangle_{\partial \mathcal{T}_h} + (\nabla p_h, \boldsymbol{v})_{\mathcal{T}_h} = (\nabla \times \boldsymbol{H}(0), \boldsymbol{v})_{\mathcal{T}_h}, \tag{5a}$$

$$(\mu \boldsymbol{H}_h(0), \boldsymbol{w})_{\mathcal{T}_h} - \left\langle \boldsymbol{n} \times \widehat{\boldsymbol{A}}_h(0), \boldsymbol{w}\right\rangle_{\partial \mathcal{T}_h} - (\boldsymbol{A}_h(0), \nabla \times \boldsymbol{w})_{\mathcal{T}_h} = 0, \tag{5b}$$

$$\left\langle \boldsymbol{n} \times \boldsymbol{H}_h(0) + \tau_t \left(\boldsymbol{A}_h(0) - \widehat{\boldsymbol{A}}_h(0)\right), \boldsymbol{\eta}^0\right\rangle_{\partial \mathcal{T}_h} = 0, \tag{5c}$$

$$\left\langle \widehat{\boldsymbol{A}}_h(0), \boldsymbol{\eta}^\partial\right\rangle_{\partial \mathcal{T}_h} = \langle \boldsymbol{g}_A(0) \times \boldsymbol{n}, \boldsymbol{\eta}^\partial\rangle_{\partial \mathcal{T}_h}, \tag{5d}$$

and

$$(\boldsymbol{A}_h(0), \nabla \phi)_{\mathcal{T}_h} = 0, \tag{5e}$$

for all $(\boldsymbol{v}, \phi, \boldsymbol{w}, \boldsymbol{\eta}^0, \boldsymbol{\eta}^\partial) \in \boldsymbol{V}_h \times Q_h^{c,0} \times \boldsymbol{W}_h \times \boldsymbol{M}_h^0 \times \boldsymbol{M}_h^\partial$. Note that (5b)–(5d) are nothing but (3b)–(3d) at $t = 0$. Here, we require $Q_h^{c,0}$ to satisfy the following assumption.

Assumption 2.2 (On the pressure space)**.** *The pressure space* $Q_h^{c,0}$ *satisfies*

(1) $\nabla Q_h^{c,0} \subset \boldsymbol{V}_h$,
(2) $\{\boldsymbol{v}_h \in \boldsymbol{V}_h : \nabla \times \boldsymbol{v}_h = 0 \text{ in } \Omega,\ \boldsymbol{n} \times \boldsymbol{v}_h = 0 \text{ on } \Gamma\} \subset \nabla Q_h^{c,0}$,
(3) $Q_h^{c,0} \subset H_0^1(\Omega)$.

For instance, for the methods in Table 1, we choose $Q_h^{c,0} := Q_h \cap H_0^1(\Omega)$ where $Q_h := \prod_{K\in\mathcal{T}_h} \mathcal{P}_{k_V+1}(K)$, with $k_V = k$ for HDG-S$_k$ and $k_V = k+1$ for HDG-B$_k$/B$_k$+.

The equations (5) define a discretization of the following static Maxwell's equations:

$$\begin{aligned}
\nabla \times \left(\mu^{-1}\nabla \times \boldsymbol{A}(0)\right) + \nabla p &= \nabla \times \boldsymbol{H}(0) && \text{in } \Omega,\\
\nabla \cdot \boldsymbol{A}(0) &= 0 && \text{in } \Omega,\\
\boldsymbol{n} \times \boldsymbol{A}(0) &= \boldsymbol{g}_A(0) && \text{on } \Gamma,
\end{aligned}$$

where the curl operator is discretized by an HDG method while the gradient operator is discretized by Lagrange elements. Here p is a pseudo-pressure. Note that $\nabla \times \boldsymbol{H}(0)$ is divergence-free, which leads to $p = 0$. This property is inherited by the discretization (5). In fact, we have the following Proposition 2.1. Its proof is included in Section 3.

Proposition 2.1. *If Assumption 2.2 holds, then the equations* (5) *are uniquely solvable. In addition,* $p_h = 0$.

In our previous work [31], to define the initial condition, we used discontinuous approximations for the pressure and for its trace over the skeleton, and weakly imposed the continuity of the numerical normal trace of $\boldsymbol{A}_h$. In contrast, here, we propose an alternative approach to discretize the initial condition for two reasons. The first is that, the method has less degrees of freedom since it uses *continuous* approximations to the pressure p. The second is that, we can guarantee that $p_h = 0$, see Proposition 2.1, which makes this initial discretization consistent with the time-marching scheme since the time-marching (3) does not involve the pressure.

2.3. Error estimates

Now we present the error estimates to the HDG methods defined by (3)–(5). We first present two assumptions which are *only* used in the duality arguments for estimating the error of $\boldsymbol{A}_h$.

Assumption 2.3 (On the data and elliptic regularity)**.** *We assume that:*

(1) *there exist constants* $s_{\mathrm{reg},A}$ *in* $(-\frac{1}{2}, 1]$ *and* $s_{\mathrm{reg},H} \in (\frac{1}{2}, 1]$ *such that the following elliptic regularity inequality holds:*

$$\|\boldsymbol{u}\|_{1+s_{\mathrm{reg},A},\Omega} + \left\|\mu^{-1}\nabla \times \boldsymbol{u}\right\|_{s_{\mathrm{reg},H},\Omega} \leqslant C_{\mathrm{reg}}\left\|\nabla \times \left(\mu^{-1}\nabla \times \boldsymbol{u}\right)\right\|_{\Omega},$$

for all $\boldsymbol{u} \in H(\mathrm{curl};\Omega)$ *such that* $\boldsymbol{n} \times \boldsymbol{u} = 0$ *and* $\nabla \cdot \boldsymbol{u} = 0$*;*
(2) *the functions* μ, ϵ, *and* σ *are piecewise* C^1, *namely, there exists a constant* $c_{\mu,\epsilon,\sigma} > 0$ *such that* $\sup_{K\in\mathcal{T}_h}(\|\mu\|_{C^1(\overline{K})} + \|\epsilon\|_{C^1(\overline{K})} + \|\sigma\|_{C^1(\overline{K})}) \leqslant c_{\mu,\epsilon,\sigma}$*;*
(3) *the domain* Ω *is simply connected with a connected boundary.*

When μ is a constant function, the elliptic regularity inequality of Assumption 2.3-(1) holds by Proposition 3.7 of [1]. If in addition, Ω is a convex polyhedron, then it holds with $s_{\mathrm{reg},A} = 1$ and $s_{\mathrm{reg},H} = 1$. For the cases when $\boldsymbol{u}$ has a lower regularity, we refer to [2, 16].

Assumption 2.4 (Minimal polynomial degree)**.** *For all* $K \in \mathcal{T}_h$, *we have*

(1) $\boldsymbol{\mathcal{P}}_0(K) \subset \boldsymbol{W}(K)$,
(2) $\boldsymbol{\mathcal{P}}_1(K) \subset \boldsymbol{V}(K)$,

TABLE 2. Orders of convergence for any HDG method satisfying the assumptions of Theorem 2.1. We assume that the solution is very smooth and the we have the full elliptic regularity, that is, $s_{\mathrm{reg},A} = s_{\mathrm{reg},H} = 1$.

Error	$\tau_t \sim h$	$\tau_t \sim 1$	$\tau_t \sim h^{-1}$
$\\|\boldsymbol{H}_h - \boldsymbol{H}\\| + \\|\boldsymbol{E}_h - \boldsymbol{E}\\|$	$\min\{k_W, k_V + 1\}$	$\min\{k_W, k_V\} + 1/2$	$\min\{k_W + 1, k_V\}$
$\\|\boldsymbol{A}_h - \boldsymbol{A}\\|$	$\min\{k_W, k_V + 1\}$	$\min\{k_W, k_V\} + 1$	$\min\{k_W + 1, k_V\} + 1$

(3) $\boldsymbol{n} \times \boldsymbol{\mathcal{P}}_1(K) \times \boldsymbol{n} \subset \boldsymbol{M}(\partial K)$.

We are now ready to state our main result. It provides estimates of the quality of the approximate magnetic field, $\boldsymbol{H}_h$, magnetic potential, $\boldsymbol{A}_h$, and the electric field $\boldsymbol{E}_h := -\dot{\boldsymbol{A}}_h$.

Theorem 2.1. *Consider any HDG method defined by the weak formulation (3), and the initial condition (4) and (5). Assume that its approximation spaces satisfy the Assumptions 2.1 and 2.2. Then we have*

$$\|\boldsymbol{H}_h(T) - \boldsymbol{H}(T)\|_{\mathcal{T}_h} + \|\boldsymbol{E}_h(T) - \boldsymbol{E}(T)\|_{\mathcal{T}_h} + \left\|\tau_t^{1/2}\Big(\boldsymbol{P}_M \boldsymbol{A}_h(T) - \widehat{\boldsymbol{A}}_h(T)\Big)\right\|_{\partial\mathcal{T}_h} \leqslant C_1(1+T)\Theta_{h,1},$$

where

$$\Theta_{h,M} := \theta_H \,\|\boldsymbol{H}\|_{W^{M,\infty}(T,H^{s_H}(\mathcal{T}_h))} + \theta_A \,\|\boldsymbol{A}\|_{W^{M+1,\infty}(T,H^{s_A}(\mathcal{T}_h))},$$

and

$$\begin{aligned} \theta_H &:= \max_{K\in\mathcal{T}_h}\Big\{h_K^{\min(k_W+1,s_H)}\Big(1 + \big(h_K\,\tau_{t,K}^{\min}\big)^{-1/2}\Big)\Big\},\\ \theta_A &:= \max_{K\in\mathcal{T}_h}\Big\{h_K^{\min(k_V,s_A-1)}\Big(h_K + \big(h_K\,\tau_{t,K}^{\max}\big)^{1/2}\Big)\Big\}. \end{aligned}$$

If, in addition, the data and regularity Assumption 2.3 and the minimal polynomial degree Assumption 2.4 hold, then

$$\|\boldsymbol{A}_h(T) - \boldsymbol{A}(T)\|_{\mathcal{T}_h} \leqslant C_2(1+T)^2 \max\{\theta_{\mathrm{reg},H}, \theta_{\mathrm{reg},A}\}\,\Theta_{h,2},$$

where

$$\begin{aligned} \theta_{\mathrm{reg},H} &:= \max_{K\in\mathcal{T}_h}\Big\{h_K^{s_{\mathrm{reg},H}}\Big(1 + \big(h_K\tau_{t,K}^{\min}\big)^{-1/2}\Big)\Big\},\\ \theta_{\mathrm{reg},A} &:= \max_{K\in\mathcal{T}_h}\Big\{h_K^{s_{\mathrm{reg},A}}\Big(1 + \big(h_K\,\tau_{t,K}^{\max}\big)^{1/2}\Big)\Big\}. \end{aligned}$$

Here the constant C_1 depends only on the polynomial degree k_V, k_W, k_M, the shape regularity of $\mathcal{T}_h$, and the parameters μ, ϵ, σ. The constant C_2 depends additionally on the domain Ω through the elliptic regularity constant C_{reg}.

Observe that we can have an estimate for the approximation error of the magnetic vector potential $\boldsymbol{A}_h$ without assuming the data regularity by simply using the estimate for the electric field. This would give a sub-optimal order of convergence, though. For the main choices of the stabilization parameter τ, this result gives the orders of convergence displayed in Table 2. The orders of convergence of the HDG methods in Table 1 can now be easily obtained, see Table 3. Let us briefly discuss them.

First, note that for $\tau = h^{-1}$, the three methods considered here provide approximations to $\boldsymbol{A}$ converging optimally. Next, note that we chose to consider the method HDG-S$_{k+1}$, and not the method HDG-S$_k$, because

TABLE 3. Orders of convergence for the HDG methods listed in Table 1. The conditions of Theorem 2.1 are assumed to hold. In boldface are indicated the optimal orders of convergence. We assume that the solution is very smooth and that we have the full elliptic regularity, that is, $s_{\mathrm{reg},A} = s_{\mathrm{reg},H} = 1$.

Method	$\|\boldsymbol{H}_h - \boldsymbol{H}\|$	$\|\boldsymbol{E}_h - \boldsymbol{E}\|$	$\|\boldsymbol{A}_h - \boldsymbol{A}\|$	$\left\|h_K^{1/2}\left(\boldsymbol{P}_M \boldsymbol{A}_h - \widehat{\boldsymbol{A}}_h\right)\right\|$ quasi-uniform meshes
$k \geqslant 0$				
HDG-S$_{k+1}(\tau_t = h)$	$k+1$	$k+1$	$k+1$	$k+1$
HDG-S$_{k+1}(\tau_t = 1)$	$k+3/2$	$k+3/2$	$\boldsymbol{k+2}$	$\boldsymbol{k+2}$
HDG-S$_{k+1}(\tau_t = h^{-1})$	$k+1$	$k+1$	$\boldsymbol{k+2}$	$\boldsymbol{k+2}$
$k \geqslant 0$				
HDG-B$_k(\tau_t = h)$	k	k	k	k
HDG-B$_k(\tau_t = 1)$	$k+1/2$	$k+1/2$	$k+1$	$k+1$
HDG-B$_k(\tau_t = h^{-1})$	$\boldsymbol{k+1}$	$k+1$	$\boldsymbol{k+2}$	$\boldsymbol{k+2}$
$k \geqslant 1$				
HDG-B$_k$+$(\tau_t = h)$	k	k	k	k
HDG-B$_k$+$(\tau_t = 1)$	$k+1/2$	$k+1/2$	$k+1$	$k+1$
HDG-B$_k$+$(\tau_t = h^{-1})$	$\boldsymbol{k+1}$	$k+1$	$\boldsymbol{k+2}$	$\boldsymbol{k+2}$

this method has the same spaces than HDG-B$_k$ for both the magnetic potential $\boldsymbol{A}_h$ and its trace $\widehat{\boldsymbol{A}}_h$. Only the local space for the magnetic field, $\boldsymbol{H}_h$, is *bigger* for HDG-S$_{k+1}$. We see that this difference pays off for $\tau = h$ and $\tau = 1$ as all the unknowns of the method HDG-S$_{k+1}$ converge with one *additional* order of convergence. The difference disappears for the larger stabilization $\tau = h^{-1}$ though. Finally, note that all the variables of the methods HDG-B$_k$ and HDG-B$_k$+ converge with the same orders, even though the local space for the numerical traces, $M(F)$, is smaller for the latter. In fact, the numerical trace $\widehat{\boldsymbol{A}}_h$ provided by the HDG-B$_k$+ method achieves super-convergence since it lies in a *proper* subspace of $\boldsymbol{P}^t_{k+1}$.

The estimates of Theorem 2.1 hold for general polyhedral meshes. Since using polyhedral meshes allows an automatic handling of hanging nodes, our analysis naturally extends to non-conforming meshes.

3. PROOFS

In this section, we prove the results of Section 2. We proceed in several steps.

Step 1: The initial condition

Let us begin by proving the unique solvability of the discretization of the initial condition (5) of Proposition 2.1.

We first prove the existence and uniqueness of the approximate solution. We only have to show that (5) admits the trivial solution when we set $\nabla \times \boldsymbol{H}(0) := 0$ and $\boldsymbol{g}_A(0) := 0$. To prove this, we take $\boldsymbol{v} := \boldsymbol{A}_h(0)$, $\phi := -p_h$, $\mathbf{w} := \boldsymbol{H}_h(0)$, and $\boldsymbol{\eta} := -\widehat{\boldsymbol{A}}_h(0)$ in the equations (5). Adding these equations, we obtain

$$(\mu \boldsymbol{H}_h(0), \boldsymbol{H}_h(0))_{\mathcal{T}_h} + \left\langle \tau_t \left(\boldsymbol{P}_M \boldsymbol{A}_h(0) - \widehat{\boldsymbol{A}}_h(0)\right), \boldsymbol{A}_h(0) - \widehat{\boldsymbol{A}}_h(0)\right\rangle_{\partial\mathcal{T}_h} = 0.$$

Therefore $\boldsymbol{H}_h(0) = 0$ and $\boldsymbol{P}_M \boldsymbol{A}_h(0) = \widehat{\boldsymbol{A}}_h(0)$ since $\tau_t > 0$ on $\partial\mathcal{T}_h$. As a consequence, the equation (5b) reads

$$-\left\langle \widehat{\boldsymbol{A}}_h(0) - \boldsymbol{A}_h(0), \boldsymbol{n} \times \boldsymbol{w}\right\rangle_{\partial\mathcal{T}_h} - (\nabla \times \boldsymbol{A}_h(0), \boldsymbol{w})_{\mathcal{T}_h} = 0 \qquad \forall \boldsymbol{w} \in \boldsymbol{W}_h,$$

and since $\boldsymbol{n}_{\partial K} \times \boldsymbol{W}(K) \subset \boldsymbol{M}(\partial K)$, by Assumption 2.1-(3), we obtain that

$$-(\nabla \times \boldsymbol{A}_h(0), \boldsymbol{w})_{\mathcal{T}_h} = 0 \qquad \forall \boldsymbol{w} \in \boldsymbol{W}_h.$$

By Assumption 2.1-(1), we can take $\boldsymbol{w} := \nabla \times \boldsymbol{A}(0)$ and immediately get that $\nabla \times \boldsymbol{A}_h(0) = 0$ on all $K \in \mathcal{T}_h$. We also have that the tangential trace on the boundaries of the elements is continuous. Indeed, $\boldsymbol{A}_h^t(0) = \boldsymbol{P}_M \boldsymbol{A}_h(0)$ by Assumption 2.1-(4) and so $\boldsymbol{A}_h^t(0) = \widehat{\boldsymbol{A}}_h(0)$. We can now conclude, that

$$\nabla \times \boldsymbol{A}_h(0) = 0 \text{ in } \Omega \quad \text{and} \quad \boldsymbol{n} \times \boldsymbol{A}_h(0) = 0 \qquad \text{on } \Gamma.$$

By Assumption 2.2-(2), $\boldsymbol{A}_h(0) \in \nabla Q_h^{c,0}$, and taking $\phi \in Q_h^{c,0}$ in equation (5e), we get that $\boldsymbol{A}_h(0) = \mathbf{0}$. Then $\widehat{\boldsymbol{A}}_h(0) = 0$ and equation (5a) becomes

$$(\nabla p_h, \boldsymbol{v})_{\mathcal{T}_h} = 0 \qquad \forall \boldsymbol{v} \in \boldsymbol{V}_h.$$

By Assumption 2.2-(1), we can set $\boldsymbol{v} := \nabla p_h$ and get that p_h is a constant. Since p_h is zero on the boundary of the domain Ω, by Assumption 2.2-(3), we obtain that $p_h = 0$. This completes the proof of existence and uniqueness.

It remains to prove that the approximate pressure p_h is identically zero. Again, by Assumption 2.2-(1), we can set $\boldsymbol{v} := \nabla p_h$ in equation (5a) to get

$$\begin{aligned}\|\nabla p_h\|_{\mathcal{T}_h}^2 &= -\left\langle \boldsymbol{n} \times \boldsymbol{H}_h(0) + \tau_t\left(\boldsymbol{P}_M \boldsymbol{A}_h(0) - \widehat{\boldsymbol{A}}_h(0)\right), (\nabla p_h)^t \right\rangle_{\partial \mathcal{T}_h} + (\nabla \times \boldsymbol{H}(0), \nabla p_h)_\Omega, \\ &= -\left\langle \boldsymbol{n} \times \boldsymbol{H}_h(0) + \tau_t\left(\boldsymbol{P}_M \boldsymbol{A}_h(0) - \widehat{\boldsymbol{A}}_h(0)\right), (\nabla p_h)^t \right\rangle_{\partial \mathcal{T}_h},\end{aligned}$$

because $p_h \in H_0^1(\Omega)$ by Assumption 2.2-(3). Finally, by Assumption 2.2-(1), $\nabla p_h \in \boldsymbol{V}_h$ and so, $(\nabla p_h)^t \in \boldsymbol{M}_h$ by Assumption 2.1-(4). Since $p_h = 0$ on the boundary of Ω, by Assumption 2.2-(3), we have that $(\nabla p_h)^t \in \boldsymbol{M}_h^0$. This means that we can take $\boldsymbol{\eta}^0 := (\nabla p_h)^t$ in equation (5c), to get that

$$\|\nabla p_h\|_{\mathcal{T}_h}^2 = 0.$$

As in the previous argument, we can now conclude that $p_h \equiv 0$. This completes the proof of Proposition 2.1.

Step 2: The compact form of the weak formulation

Here, we introduce a compact form of the weak formulation (3), which allows us to investigate its properties more concisely. So, we define

$$\begin{aligned}B_h\left(\boldsymbol{A}, \widehat{\boldsymbol{A}}, \boldsymbol{H}; \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}\right) := &\langle \tau_t \boldsymbol{P}_M \boldsymbol{A}, \boldsymbol{v}\rangle_{\partial \mathcal{T}_h} - \left\langle \tau_t \widehat{\boldsymbol{A}}, \boldsymbol{v}\right\rangle_{\partial \mathcal{T}_h} + (\nabla \times \boldsymbol{H}, \boldsymbol{v})_{\mathcal{T}_h} \\ &- \langle \tau_t \boldsymbol{A}, \boldsymbol{\eta}\rangle_{\partial \mathcal{T}_h} + \left\langle \tau_t \widehat{\boldsymbol{A}}, \boldsymbol{\eta}\right\rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{n} \times \boldsymbol{H}, \boldsymbol{\eta}\rangle_{\partial \mathcal{T}_h} \\ &- (\boldsymbol{A}, \nabla \times \boldsymbol{w})_{\mathcal{T}_h} + \left\langle \widehat{\boldsymbol{A}}, \boldsymbol{n} \times \boldsymbol{w}\right\rangle_{\partial \mathcal{T}_h} + (\mu \boldsymbol{H}, \boldsymbol{w})_{\mathcal{T}_h}.\end{aligned} \tag{6}$$

Note how the lines of the definition of the bilinear form B_h correspond to the equations (3a), (3c), and (3b), respectively.

Now, the original weak formulation (3) can be rewritten into the following more compact form: find $(\boldsymbol{A}_h, \widehat{\boldsymbol{A}}_h, \boldsymbol{H}_h) \in \boldsymbol{V}_h \times \boldsymbol{M}_h \times \boldsymbol{W}_h$ such that

$$\left(\epsilon \ddot{\boldsymbol{A}}_h, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma \dot{\boldsymbol{A}}_h, \boldsymbol{v}\right)_{\mathcal{T}_h} + B_h\left(\boldsymbol{A}_h, \widehat{\boldsymbol{A}}_h, \boldsymbol{H}_h; \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}\right) - (\boldsymbol{J}, \boldsymbol{v})_{\mathcal{T}_h} = 0, \tag{7a}$$

for all $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h^0 \times \boldsymbol{W}_h$, with the boundary condition enforced by

$$\left\langle \widehat{\boldsymbol{A}}_h, \boldsymbol{\eta}\right\rangle_{\partial \mathcal{T}_h} = \langle \boldsymbol{g}_A \times \boldsymbol{n}, \boldsymbol{\eta}\rangle_{\partial \mathcal{T}_h}, \tag{7b}$$

for all $\boldsymbol{\eta} \in \boldsymbol{M}_h^\partial$.

The bilinear form B_h has a sub-skew-symmetry structure, as we see in the next result. This structure will be frequently exploited in our error analysis of the method.

Lemma 3.1 (Sub-skew-symmetry)**.** *The following identities hold:*

$$
\begin{aligned}
B_h(\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{b};\boldsymbol{c},\widehat{\boldsymbol{c}},\boldsymbol{d}) &= B_h(\boldsymbol{c},\widehat{\boldsymbol{c}},-\boldsymbol{d};\boldsymbol{a},\widehat{\boldsymbol{a}},-\boldsymbol{b}), \quad B_h(\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{0};\boldsymbol{0},\boldsymbol{0},\boldsymbol{b}) + B_h(\boldsymbol{0},\boldsymbol{0},\boldsymbol{b};\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{0}) = 0,\\
B_h(\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{b};\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{b}) &= B_h(\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{0};\boldsymbol{a},\widehat{\boldsymbol{a}},\boldsymbol{0}) + B_h(\boldsymbol{0},\boldsymbol{0},\boldsymbol{b};\boldsymbol{0},\boldsymbol{0},\boldsymbol{b}),\\
&= \langle \tau_t(\boldsymbol{P}_M\boldsymbol{a}-\widehat{\boldsymbol{a}}), \boldsymbol{P}_M\boldsymbol{a}-\widehat{\boldsymbol{a}}\rangle_{\partial\mathcal{T}_h} + (\mu\,\boldsymbol{b},\boldsymbol{b})_{\mathcal{T}_h},
\end{aligned}
$$

provided $\boldsymbol{P}_M\widehat{\boldsymbol{a}} = \widehat{\boldsymbol{a}}$.

Proof. We obtain the first identity by examining the structure of (6). The next two identities follow from the first one. The last identity is a consequence of the definition of B_h. This completes the proof. □

Step 3: Projections

Here, we introduce some projections and study their properties. Our analysis consists in controlling approximation errors of these projections.

Proposition 3.1. *If Assumption 2.1 holds, there is a projection* $\boldsymbol{\Pi}_{W,V} := (\boldsymbol{\Pi}_W, \boldsymbol{\Pi}_V)$ *onto* $\boldsymbol{W}_h \times \boldsymbol{V}_h$ *such that*

$$(\boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{w},\nabla\times\boldsymbol{r})_K = \langle \boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w}),\boldsymbol{r}\rangle_{\partial K} \qquad \forall \boldsymbol{r}\in\boldsymbol{V}(K), \tag{8a}$$
$$(\boldsymbol{\Pi}_V\boldsymbol{v}-\boldsymbol{v},\nabla\times\boldsymbol{r})_K = 0 \qquad \forall \boldsymbol{r}\in\boldsymbol{W}(K). \tag{8b}$$

hold for all $K\in\mathcal{T}_h$. *In addition,*

$$h_K^{1/2}\|\boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{w}\|_{\partial K} + \|\boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{w}\|_K \leqslant C h_K^{\min(s,k_W+1)}|\boldsymbol{w}|_{s,K}, \tag{9a}$$
$$h_K^{1/2}\|\boldsymbol{\Pi}_V\boldsymbol{v}-\boldsymbol{v}\|_{\partial K} + \|\boldsymbol{\Pi}_V\boldsymbol{v}-\boldsymbol{v}\|_K \leqslant C h_K^{\min(s,k_V+1)}|\boldsymbol{v}|_{s,K} \tag{9b}$$

where C *depends only on the shape regularity of* K, *and the polynomial degree* k_W *and* k_V.

Proof. If we take $\boldsymbol{\Pi}_V$ as the $\boldsymbol{L}^2$-projection to $\boldsymbol{V}_h$, the equation (8b) holds by Assumption 2.1-(2), and the estimate (9b) holds by the well-known convergence properties of $\boldsymbol{L}^2$-projections; see [4].

If we define $\boldsymbol{\Pi}_W\boldsymbol{w}$, on each element $K\in\mathcal{T}_h$, as the element of $\boldsymbol{W}(K)$ which solves the equations

$$(\boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{w},\nabla\times\boldsymbol{v})_K = \langle \boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w}),\boldsymbol{v}\rangle_{\partial K} \qquad \forall \boldsymbol{v}\in\boldsymbol{V}(K), \tag{10a}$$
$$(\boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{w},\boldsymbol{r})_K = 0 \qquad \forall \boldsymbol{r}\in(\nabla\times\boldsymbol{V}(K))^{\perp_W}, \tag{10b}$$

where $\perp_W$ means taking the orthogonal complement in $\boldsymbol{W}(K)$, we will have that it satisfies the equation (8a), by construction. So, it remains to prove that it $\boldsymbol{\Pi}_W\boldsymbol{w}$ is well defined.

To do that, we only have to prove that the above set of equations defines a square system. The fact that $\boldsymbol{\Pi}_W\boldsymbol{w}$ is well defined will then follow from the inequality (9a).

Let us show that the system of equations is square. The number of unknowns is clearly $\dim\boldsymbol{W}(K)$. Let us count the number of linearly independent equations. By Assumption 2.1-(4), we know that, if $\nabla\times\boldsymbol{v}=0$, then $\boldsymbol{n}\times\boldsymbol{v}\in\boldsymbol{M}(\partial K)$ and $\langle \boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w}),\boldsymbol{v}\rangle_{\partial K}=0$. As a result, the enforcement by (10a) only has a rank of $\dim(\nabla\times\boldsymbol{V}(K))$. Considering that (10b) has a rank of $\dim(\boldsymbol{W}(K))-\dim(\nabla\times\boldsymbol{V}(K))$, we see that the equations (10) define a square system.

It remains to prove the inequality (9a). Let $\boldsymbol{P}_W$ be the $\boldsymbol{L}^2$-projection to $\boldsymbol{W}_h$. Now set $\boldsymbol{\varepsilon}_w := \boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{P}_W\boldsymbol{w}\in\boldsymbol{W}_h$. On each element $K\in\mathcal{T}_h$, we can write $\boldsymbol{\varepsilon}_w = \boldsymbol{\varepsilon}_w^1+\boldsymbol{\varepsilon}_w^2$ where $\boldsymbol{\varepsilon}_w^1\in\nabla\times\boldsymbol{V}(K)$ and $\boldsymbol{\varepsilon}_w^2\in(\nabla\times\boldsymbol{V}(K))^{\perp_W}$. By equation (10b), $\boldsymbol{\varepsilon}_w^2=0$ and so, $\boldsymbol{\varepsilon}_w=\boldsymbol{\varepsilon}_w^1=\nabla\times v$ for some $\boldsymbol{v}\in\boldsymbol{V}(K)$. Then, by equation (10a), we have

$$\|\boldsymbol{\varepsilon}_w\|_K^2 = (\boldsymbol{\varepsilon}_w,\nabla\times\boldsymbol{v})_K = (\boldsymbol{\varepsilon}_w,\nabla\times(\boldsymbol{v}+\boldsymbol{\phi}))_K = \langle \boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w}),\boldsymbol{v}+\boldsymbol{\phi}\rangle_{\partial K},$$

for any $\boldsymbol{\phi}\in\boldsymbol{V}(K)$ such that $\nabla\times\boldsymbol{\phi}=\boldsymbol{0}$. Applying the Cauchy–Schwarz inequality, we get

$$\|\boldsymbol{\varepsilon}_w\|_K^2 \leqslant h_K^{1/2}\|\boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w})\|_{\partial K}\, h_K^{-1/2} \inf_{\boldsymbol{\phi}\in\boldsymbol{V}(K),\,\nabla\times\boldsymbol{\phi}=0}\|\boldsymbol{v}+\boldsymbol{\phi}\|_{\partial K}$$

$$\leqslant Ch_K^{1/2}\|\boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w})\|_{\partial K}\|\nabla\times\boldsymbol{v}\|_K,$$

since the *curl* operator is a bijection from the orthogonal complement in $\boldsymbol{V}(K)$ of its kernel to $\nabla\times\boldsymbol{V}(K)$. Since $\nabla\times\boldsymbol{v}=\boldsymbol{\varepsilon}_w^1=\boldsymbol{\varepsilon}_w$, we get

$$\|\boldsymbol{\varepsilon}_w\|_K\leqslant Ch_K^{1/2}\|\boldsymbol{n}\times\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w})\|_{\partial K}\leqslant Ch_K^{1/2}\|\boldsymbol{w}-\boldsymbol{P}_W\boldsymbol{w}\|_{\partial K}.$$

This completes the proof of (9a). □

Step 4: The equations of the projections of the errors for $t>0$

Now that the projection $\boldsymbol{\Pi}_{W,V}$ is defined, we can obtain the equations satisfied by the projection of the errors. To do this, we first introduce an associated boundary remainder operator:

$$\begin{aligned}\boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}:\boldsymbol{W}(K)\times\boldsymbol{V}(K)&\to\boldsymbol{M}(\partial K)\\(\boldsymbol{w},\boldsymbol{v})&\mapsto\boldsymbol{n}\times\boldsymbol{\Pi}_W\boldsymbol{w}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{w})+\tau_t(\boldsymbol{P}_M\boldsymbol{\Pi}_V\boldsymbol{v}-\boldsymbol{P}_M\boldsymbol{v}).\end{aligned}\tag{11}$$

The boundary remainder can be regarded as an interpolation error on the element boundary. It also serves as a quantification for the "non-conformity" of the projection operator $\boldsymbol{\Pi}_{W,V}$. In the following result, we place its most important property, namely, a generalized version of the weak-commutativity property.

Lemma 3.2 (Weak-commutativity)**.** *Suppose Assumption 2.1 holds. Set $\boldsymbol{\delta}:=\boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{H},\boldsymbol{A})$. Then*

$$(\nabla\times\boldsymbol{H},\boldsymbol{v})_K=(\nabla\times\boldsymbol{\Pi}_W\boldsymbol{H},\boldsymbol{v})_K+\langle\tau_t(\boldsymbol{P}_M\boldsymbol{\Pi}_V\boldsymbol{A}-\boldsymbol{P}_M\boldsymbol{A}),\boldsymbol{v}\rangle_{\partial K}-\langle\boldsymbol{\delta},\boldsymbol{v}\rangle_{\partial K},$$

for all $\boldsymbol{v}\in\boldsymbol{V}(K)$.

Proof. By the first equation defining the projection $\boldsymbol{\Pi}_{W,V}$ (8a), we have

$$\begin{aligned}(\nabla\times\boldsymbol{H},\boldsymbol{v})_K&=\langle\boldsymbol{n}\times\boldsymbol{H},\boldsymbol{v}\rangle_{\partial K}+(\boldsymbol{\Pi}_W\boldsymbol{H},\nabla\times\boldsymbol{v})_K-\langle\boldsymbol{n}\times\boldsymbol{H}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{H}),\boldsymbol{v}\rangle_{\partial K}\\&=(\nabla\times\boldsymbol{\Pi}_W\boldsymbol{H},\boldsymbol{v})_K-\langle\boldsymbol{n}\times\boldsymbol{\Pi}_W\boldsymbol{H}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{H}),\boldsymbol{v}\rangle_{\partial K}.\end{aligned}$$

The proof is complete. □

Next, we obtain an identity for the bilinear form B_h involving the projections of our unknowns.

Lemma 3.3 (An identity for the bilinear form B_h)**.** *Suppose that Assumption 2.1 holds. Set $\boldsymbol{\delta}:=\boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{H},\boldsymbol{A})$. Then*

$$B_h(\boldsymbol{\Pi}_V\boldsymbol{A},\boldsymbol{P}_M\boldsymbol{A},\boldsymbol{\Pi}_W\boldsymbol{H};\boldsymbol{v},\boldsymbol{\eta},\boldsymbol{w})=(\nabla\times\boldsymbol{H},\boldsymbol{v})_{\mathcal{T}_h}+(\mu(\boldsymbol{\Pi}_W\boldsymbol{H}-\boldsymbol{H}),\boldsymbol{w})_{\mathcal{T}_h}+\langle\boldsymbol{\delta},\boldsymbol{v}-\boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h},$$

for all $(\boldsymbol{v},\boldsymbol{\eta},\boldsymbol{w})\in\boldsymbol{V}_h\times\boldsymbol{M}_h^0\times\boldsymbol{W}_h$.

Proof. By the definition of B_h, the equation (6), for all $(\boldsymbol{v},\boldsymbol{\eta},\boldsymbol{w})\in\boldsymbol{V}_h\times\boldsymbol{M}_h^0\times\boldsymbol{W}_h$, we have

$$\begin{aligned}B_h(\boldsymbol{\Pi}_V\boldsymbol{A},\boldsymbol{P}_M\boldsymbol{A},\boldsymbol{\Pi}_W\boldsymbol{H};\boldsymbol{v},\boldsymbol{\eta},\boldsymbol{w})=&\langle\tau_t(\boldsymbol{P}_M\boldsymbol{\Pi}_V\boldsymbol{A}-\boldsymbol{P}_M\boldsymbol{A}),\boldsymbol{v}-\boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}\\&-(\boldsymbol{\Pi}_V\boldsymbol{A},\nabla\times\boldsymbol{w})_{\mathcal{T}_h}-\langle\boldsymbol{n}\times\boldsymbol{P}_M\boldsymbol{A},\boldsymbol{w}\rangle_{\partial\mathcal{T}_h}\\&-\big(-(\boldsymbol{v},\nabla\times\boldsymbol{\Pi}_W\boldsymbol{H})_{\mathcal{T}_h}-\langle\boldsymbol{n}\times\boldsymbol{\eta},\boldsymbol{\Pi}_W\boldsymbol{H}\rangle_{\partial\mathcal{T}_h}\big)\\&+(\mu\boldsymbol{\Pi}_W\boldsymbol{H},\boldsymbol{w})_{\mathcal{T}_h}.\end{aligned}$$

Now, by the weak-commutativity property of Lemma 3.2, we have

$$B_h(\boldsymbol{\Pi}_V\boldsymbol{A},\boldsymbol{P}_M\boldsymbol{A},\boldsymbol{\Pi}_W\boldsymbol{H};\boldsymbol{v},\boldsymbol{\eta},\boldsymbol{w})=(\nabla\times\boldsymbol{H},\boldsymbol{v})_{\mathcal{T}_h}+(\mu(\boldsymbol{\Pi}_W\boldsymbol{H}-\boldsymbol{H}),\boldsymbol{w})_{\mathcal{T}_h}+\langle\boldsymbol{\delta},\boldsymbol{v}\rangle_{\partial\mathcal{T}_h}-\langle\boldsymbol{\delta},\boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}$$

$$- (\boldsymbol{\Pi}_V \boldsymbol{A}, \nabla \times \boldsymbol{w})_{\mathcal{T}_h} - \langle \boldsymbol{n} \times \boldsymbol{P}_M \boldsymbol{A}, \boldsymbol{w} \rangle_{\partial \mathcal{T}_h} + (\mu \boldsymbol{H}, \boldsymbol{w})_{\mathcal{T}_h}.$$

By (8b) and Assumption 2.1-(3), we have $(\boldsymbol{\Pi}_V \boldsymbol{A}, \nabla \times \boldsymbol{w})_{\mathcal{T}_h} = (\boldsymbol{A}, \nabla \times \boldsymbol{w})_{\mathcal{T}_h}$ and $\langle \boldsymbol{n} \times \boldsymbol{P}_M \boldsymbol{A}, \boldsymbol{w} \rangle_{\partial \mathcal{T}_h} = \langle \boldsymbol{n} \times \boldsymbol{A}, \boldsymbol{w} \rangle_{\partial \mathcal{T}_h}$, respectively. Thus

$$-(\boldsymbol{\Pi}_V \boldsymbol{A}, \nabla \times \boldsymbol{w})_{\mathcal{T}_h} - \langle \boldsymbol{n} \times \boldsymbol{P}_M \boldsymbol{A}, \boldsymbol{w} \rangle_{\partial \mathcal{T}_h} + (\mu \boldsymbol{H}, \boldsymbol{w})_{\mathcal{T}_h} = (-\nabla \times \boldsymbol{A} + \mu \boldsymbol{H}, \boldsymbol{w})_{\mathcal{T}_h} = 0.$$

This completes the proof. □

We are now ready to obtain the equation satisfied by the projection of the errors. It is written in terms of the projection of the errors

$$\boldsymbol{\varepsilon}_h^A := \boldsymbol{\Pi}_V \boldsymbol{A} - \boldsymbol{A}_h, \quad \widehat{\boldsymbol{\varepsilon}}_h^A := \boldsymbol{P}_M \boldsymbol{A} - \widehat{\boldsymbol{A}}_h, \quad \boldsymbol{\varepsilon}_h^E := \boldsymbol{\Pi}_V \boldsymbol{E} - \boldsymbol{E}_h, \quad \boldsymbol{\varepsilon}_h^H := \boldsymbol{\Pi}_W \boldsymbol{H} - \boldsymbol{H}_h,$$

and the approximation errors

$$\boldsymbol{e}_A := \boldsymbol{\Pi}_V \boldsymbol{A} - \boldsymbol{A}, \quad \boldsymbol{e}_E := \boldsymbol{\Pi}_V \boldsymbol{E} - \boldsymbol{E}, \quad \boldsymbol{e}_H := \boldsymbol{\Pi}_W \boldsymbol{W} - \boldsymbol{W}, \quad \boldsymbol{\delta} := \boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{H}, \boldsymbol{A}).$$

As we see in the next result, the projections of the errors satisfy a set of equations which share a similar structure to the HDG discretization defined by (7).

Lemma 3.4 (Equations of the projections of the errors)**.** *Suppose that Assumption 2.1 holds. Then*

$$\left(\epsilon\, \ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma\, \dot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\right)_{\mathcal{T}_h} + B_h\big(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}\big) = (\mu \boldsymbol{e}_H, \boldsymbol{w})_{\mathcal{T}_h} + \langle \boldsymbol{\delta}, \boldsymbol{v} - \boldsymbol{\eta} \rangle_{\partial \mathcal{T}_h} + (\epsilon\, \ddot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h} + (\sigma\, \dot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h},$$

for all $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h^0 \times \boldsymbol{W}_h$.

Proof. The equation can be directly obtained by taking the difference between the equation of the projections of Lemma 3.3 and the compact form of the HDG semi-discretization (7). This completes the proof. □

Step 5: The energy argument

With the error equations just obtained, we can obtain the following energy identities.

Proposition 3.2 (Energy identities)**.** *Suppose that Assumption 2.1 holds. Then*

$$\begin{aligned}
&\left(\epsilon \ddot{\boldsymbol{\varepsilon}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\sigma \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\mu \boldsymbol{\varepsilon}_h^H, \dot{\boldsymbol{\varepsilon}}_h^H\right)_{\mathcal{T}_h} + \left\langle \tau_t\big(\boldsymbol{P}_M \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A\big), \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A \right\rangle_{\partial \mathcal{T}_h} \\
&\quad = \left(\mu \dot{\boldsymbol{e}}_H, \boldsymbol{\varepsilon}_h^H\right)_{\mathcal{T}_h} + \left\langle \boldsymbol{\delta}, \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A \right\rangle_{\partial \mathcal{T}_h} + \left(\epsilon \ddot{\boldsymbol{e}}_A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\sigma \dot{\boldsymbol{e}}_A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h}, \\
&\left(\epsilon \dddot{\boldsymbol{\varepsilon}}_h^A, \ddot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\sigma \ddot{\boldsymbol{\varepsilon}}_h^A, \ddot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\mu \dot{\boldsymbol{\varepsilon}}_h^H, \ddot{\boldsymbol{\varepsilon}}_h^H\right)_{\mathcal{T}_h} + \left\langle \tau_t\Big(\boldsymbol{P}_M \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\Big), \ddot{\boldsymbol{\varepsilon}}_h^A - \ddot{\widehat{\boldsymbol{\varepsilon}}}_h^A \right\rangle_{\partial \mathcal{T}_h} \\
&\quad = \left(\mu \ddot{\boldsymbol{e}}_H, \dot{\boldsymbol{\varepsilon}}_h^H\right)_{\mathcal{T}_h} + \left\langle \dot{\boldsymbol{\delta}}, \ddot{\boldsymbol{\varepsilon}}_h^A - \ddot{\widehat{\boldsymbol{\varepsilon}}}_h^A \right\rangle_{\partial \mathcal{T}_h} + \left(\epsilon \dddot{\boldsymbol{e}}_A, \ddot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\sigma \ddot{\boldsymbol{e}}_A, \ddot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h}.
\end{aligned}$$

Proof. By the error equations of Lemma 3.4, we have

$$\begin{aligned}
\left(\epsilon\, \ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma\, \dot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\right)_{\mathcal{T}_h} + B_h\big(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{0}\big) &= \langle \boldsymbol{\delta}, \boldsymbol{v} - \boldsymbol{\eta} \rangle_{\partial \mathcal{T}_h} + (\epsilon\, \ddot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h} + (\sigma\, \dot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h}, \\
B_h\Big(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{w}\Big) &= (\mu \dot{\boldsymbol{e}}_H, \boldsymbol{w})_{\mathcal{T}_h}.
\end{aligned}$$

If we add the equations, we get

$$\left(\epsilon\, \ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma\, \dot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\right)_{\mathcal{T}_h} + B_h\big(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{0}\big) + B_h\Big(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{w}\Big)$$

$$= (\mu \dot{\boldsymbol{e}}_H, \boldsymbol{w})_{\mathcal{T}_h} + \langle \boldsymbol{\delta}, \boldsymbol{v} - \boldsymbol{\eta} \rangle_{\partial \mathcal{T}_h} + (\epsilon\, \ddot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h} + (\sigma\, \dot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h}.$$

By (2d) and (7b) we have that $\widehat{\boldsymbol{\varepsilon}}_h^A \in \boldsymbol{M}_h^0$. Thus, we can take $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) := (\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{\varepsilon}_h^H)$ and obtain

$$\begin{aligned} &\left(\epsilon\, \ddot{\boldsymbol{\varepsilon}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\sigma\, \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}\right) + B_h\left(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H\right) \\ &\quad = \left(\mu \dot{\boldsymbol{e}}_H, \boldsymbol{\varepsilon}_h^H\right)_{\mathcal{T}_h} + \left\langle \boldsymbol{\delta}, \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A \right\rangle_{\partial \mathcal{T}_h} + \left(\epsilon \ddot{\boldsymbol{e}}_A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\sigma \dot{\boldsymbol{e}}_A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h}. \end{aligned}$$

But

$$\begin{aligned} \Theta &:= B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}\right) + B_h\left(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H\right) \\ &= B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{0}; \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}\right) + B_h\left(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H; \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}\right) + B_h\left(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H\right) \\ &= B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{0}; \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}\right) - B_h\left(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H\right) + B_h\left(\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H\right), \end{aligned}$$

by the second identity of Lemma 3.1. Then

$$\begin{aligned} \Theta &= B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{0}; \dot{\boldsymbol{\varepsilon}}_h^A, \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A, \boldsymbol{0}\right) + B_h\left(\boldsymbol{0}, \boldsymbol{0}, \dot{\boldsymbol{\varepsilon}}_h^H; \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varepsilon}_h^H\right) \\ &= \left\langle \tau_t\left(\boldsymbol{P}_M \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A\right), \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A \right\rangle_{\partial \mathcal{T}_h} + \left(\mu \boldsymbol{\varepsilon}_h^H, \dot{\boldsymbol{\varepsilon}}_h^H\right)_{\mathcal{T}_h}, \end{aligned}$$

by the last identity of Lemma 3.1. This completes the proof of the first energy identity of Proposition 3.2.

The second identity can be obtained in a similar manner by taking temporal derivatives of the error equations of Lemma 3.4. This completes the proof of Proposition 3.2. □

With the above energy identities, we can obtain estimates for the following quantities:

$$\mathrm{Er}_0(t) := \left(\left\| \epsilon^{1/2} \boldsymbol{\varepsilon}_h^E(t) \right\|_{\mathcal{T}_h}^2 + \left\| \mu^{1/2} \boldsymbol{\varepsilon}_h^H(t) \right\|_{\mathcal{T}_h}^2 + \left\| \tau_t^{1/2} \left(\boldsymbol{P}_M \boldsymbol{\varepsilon}_h^A(t) - \widehat{\boldsymbol{\varepsilon}}_h^A(t) \right) \right\|_{\partial \mathcal{T}_h}^2 \right)^{1/2}, \tag{12a}$$

$$\mathrm{Er}_1(t) := \left(\left\| \epsilon^{1/2} \dot{\boldsymbol{\varepsilon}}_h^E(t) \right\|_{\mathcal{T}_h}^2 + \left\| \mu^{1/2} \dot{\boldsymbol{\varepsilon}}_h^H(t) \right\|_{\mathcal{T}_h}^2 + \left\| \tau_t^{1/2} \left(\boldsymbol{P}_M \dot{\boldsymbol{\varepsilon}}_h^A(t) - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(t) \right) \right\|_{\partial \mathcal{T}_h}^2 \right)^{1/2}. \tag{12b}$$

Note that $\boldsymbol{\varepsilon}_h^E = -\dot{\boldsymbol{\varepsilon}}_h^A$. The estimates are contained in the following result.

Proposition 3.3 (Error estimates by energy arguments)**.** *Suppose that Assumption 2.1 holds. Then, we have*

$$\mathrm{Er}_0(t) \lesssim \mathrm{Er}_0(0) + \left\| \tau_t^{-1/2} \boldsymbol{\delta} \right\|_{L^\infty(t, L^2(\partial \mathcal{T}_h))} + \int_0^t \left(\left\| \epsilon^{1/2} \ddot{\boldsymbol{e}}_A \right\|_{\mathcal{T}_h}^2 + \left\| \sigma \epsilon^{-1/2} \dot{\boldsymbol{e}}_A \right\|_{\mathcal{T}_h}^2 + \left\| \mu^{1/2} \dot{\boldsymbol{e}}_H \right\|_{\mathcal{T}_h}^2 + \left\| \tau_t^{-1/2} \dot{\boldsymbol{\delta}} \right\|_{\partial \mathcal{T}_h}^2 \right)^{1/2},$$

$$\mathrm{Er}_1(t) \lesssim \mathrm{Er}_1(0) + \left\| \tau_t^{-1/2} \dot{\boldsymbol{\delta}} \right\|_{L^\infty(t, L^2(\partial \mathcal{T}_h))} + \int_0^t \left(\left\| \epsilon^{1/2} \dddot{\boldsymbol{e}}_A \right\|_{\mathcal{T}_h}^2 + \left\| \sigma \epsilon^{-1/2} \ddot{\boldsymbol{e}}_A \right\|_{\mathcal{T}_h}^2 + \left\| \mu^{1/2} \ddot{\boldsymbol{e}}_H \right\|_{\mathcal{T}_h}^2 + \left\| \tau_t^{-1/2} \ddot{\boldsymbol{\delta}} \right\|_{\partial \mathcal{T}_h}^2 \right)^{1/2}.$$

To prove these estimates, we are going to rely on the following inequality.

Lemma 3.5 (Grönwall-type inequality)**.** *Suppose ϕ, β, and l are continuous and positive functions defined on $[0, \infty)$ and r is a constant. If*

$$\phi^2(t) \leqslant r + 2 \int_0^t \phi(s) \beta(s)\, \mathrm{d}s + \phi(t) l(t) \qquad \forall t \in [0, \infty),$$

then

$$\phi^2(t) \leqslant 2r + \left(2 \int_0^t \beta(s)\, \mathrm{d}s + \sup_{s \in [0,t]} l(s) \right)^2.$$

Proof. We provide a proof for the sake of completeness. Consider the interval $[0,t]$ and let $\phi(t^*)$ maximize ϕ in the interval. Then we have

$$\phi^2(t^*) \leqslant r + \phi(t^*)\left(2\int_0^t \beta(s)\,\mathrm{d}s + \sup_{s\in[0,t]} l(s)\right) \leqslant r + \frac{1}{2}\left(\phi^2(t^*) + \left(2\int_0^t \beta(s)\,\mathrm{d}s + \sup_{s\in[0,t]} l(s)\right)^2\right).$$

This completes the proof. □

We are now ready to prove Proposition 3.3.

Proof. Set

$$\begin{aligned}
\phi^2(t) &:= \left(\epsilon\dot{\boldsymbol{\varepsilon}}_h^A, \dot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\mu\boldsymbol{\varepsilon}_h^H, \boldsymbol{\varepsilon}_h^H\right)_{\mathcal{T}_h} + \left\langle \tau_t\left(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A\right), \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A\right\rangle_{\partial\mathcal{T}_h},\\
\beta^2(t) &:= (\epsilon\ddot{\boldsymbol{e}}_A, \ddot{\boldsymbol{e}}_A)_{\mathcal{T}_h} + (\mu\dot{\boldsymbol{e}}_H, \dot{\boldsymbol{e}}_H)_{\mathcal{T}_h} + \left(\sigma^2\epsilon^{-1}\dot{\boldsymbol{e}}_A, \dot{\boldsymbol{e}}_A\right)_{\mathcal{T}_h} + \left\langle \tau_t^{-1}\dot{\boldsymbol{\delta}}, \dot{\boldsymbol{\delta}}\right\rangle_{\partial\mathcal{T}_h},\\
l^2(t) &:= 4\left\langle \tau_t^{-1}\boldsymbol{\delta}, \boldsymbol{\delta}\right\rangle_{\partial\mathcal{T}_h},\\
r &:= \phi^2(0) + \phi(0)l(0).
\end{aligned}$$

Then by the first energy identity of Proposition 3.2, we know ϕ, β, l, and r satisfy the condition of Lemma 3.5. This proves the first estimate. Similarly, by letting

$$\begin{aligned}
\phi^2(t) &:= \left(\epsilon\ddot{\boldsymbol{\varepsilon}}_h^A, \ddot{\boldsymbol{\varepsilon}}_h^A\right)_{\mathcal{T}_h} + \left(\mu\dot{\boldsymbol{\varepsilon}}_h^H, \dot{\boldsymbol{\varepsilon}}_h^H\right)_{\mathcal{T}_h} + \left\langle \tau_t\left(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\right), \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\right\rangle_{\partial\mathcal{T}_h},\\
\beta^2(t) &:= (\epsilon\dddot{\boldsymbol{e}}_A, \dddot{\boldsymbol{e}}_A)_{\mathcal{T}_h} + (\mu\ddot{\boldsymbol{e}}_H, \ddot{\boldsymbol{e}}_H)_{\mathcal{T}_h} + \left\langle \sigma^2\epsilon^{-1}\ddot{\boldsymbol{e}}_A, \ddot{\boldsymbol{e}}_A\right\rangle_{\partial\mathcal{T}_h} + \left\langle \tau_t^{-1}\ddot{\boldsymbol{\delta}}, \ddot{\boldsymbol{\delta}}\right\rangle_{\partial\mathcal{T}_h},\\
l^2(t) &:= 4\left\langle \tau_t^{-1}\dot{\boldsymbol{\delta}}, \dot{\boldsymbol{\delta}}\right\rangle_{\partial\mathcal{T}_h},\\
r &:= \phi^2(0) + \phi(0)l(0).
\end{aligned}$$

Then we know, from the second energy identity of Proposition 3.2, that ϕ, β, l, and r satisfy condition of Lemma 3.5. This completes the proof of Proposition 3.3. □

Step 6: The equations of the projection of the errors at the initial time

Lemma 3.6 (Error equations at the initial time)**.** *Suppose that Assumptions 2.1 and 2.2 hold. Then*

$$B_h\left(\boldsymbol{\varepsilon}_h^A(0), \widehat{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{\varepsilon}_h^H(0); \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}\right) = (\mu\boldsymbol{e}_H(0), \boldsymbol{w})_{\mathcal{T}_h} + \langle \boldsymbol{\delta}(0), \boldsymbol{v} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h},$$

for all $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h^0 \times \boldsymbol{W}_h$.

Proof. This result can be proved in the same way as Lemma 3.4 by exploiting the similarity between the equations defining the weak formulation of the method for $t > 0$ (3), and the equations defining its initial condition (5). We also must use the fact that the approximate pressure p_h is zero, by Proposition 2.1. This completes the proof. □

Step 7: Estimates of the errors at the initial time

The following result gives an estimate of the initial errors $\mathrm{Er}_0(0)$ and $\mathrm{Er}_1(0)$.

Proposition 3.4 (Estimates of the initial error)**.** *Suppose that Assumptions 2.1 and 2.2 hold. Then*

$$\begin{aligned}
\mathrm{Er}_0(0) &\leqslant \left(\left\|\mu^{1/2}\boldsymbol{e}_H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\right\|_{\partial\mathcal{T}_h}^2\right)^{1/2},\\
\mathrm{Er}_1(0) &\leqslant \left(\left(\left\|\epsilon^{1/2}\ddot{\boldsymbol{e}}_A(0)\right\|_{\mathcal{T}_h} + \left\|\epsilon^{-1/2}\sigma\dot{\boldsymbol{e}}_A(0)\right\|_{\mathcal{T}_h}\right)^2 + \left\|\mu^{1/2}\dot{\boldsymbol{e}}_H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}(0)\right\|_{\partial\mathcal{T}_h}^2\right)^{1/2}.
\end{aligned}$$

Proof. Let us prove the first estimate. By definition of $\mathrm{Er}_0(0)$ (12a), we have that

$$\begin{aligned}\mathrm{Er}_0(0)^2 &:= \left\|\epsilon^{1/2}\boldsymbol{\varepsilon}_h^E(0)\right\|_{\mathcal{T}_h}^2 + \left\|\mu^{1/2}\boldsymbol{\varepsilon}_h^H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{1/2}\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big)\right\|_{\partial\mathcal{T}_h}^2 \\ &= \left\|\mu^{1/2}\boldsymbol{\varepsilon}_h^H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{1/2}\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big)\right\|_{\partial\mathcal{T}_h}^2\end{aligned}$$

because $\boldsymbol{\varepsilon}_h^E(0) = -\dot{\boldsymbol{\varepsilon}}_h^A(0) = \boldsymbol{0}$ by the definition of the initial condition in equation (4). Now, by the last identity of Lemma 3.1, we see that

$$\begin{aligned}\mathrm{Er}_0(0)^2 &= B_h\big(\boldsymbol{\varepsilon}_h^A(0), \widehat{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{\varepsilon}_h^H(0); \boldsymbol{\varepsilon}_h^A(0), \widehat{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{\varepsilon}_h^H(0)\big) \\ &= \big(\mu\boldsymbol{e}_H(0), \boldsymbol{\varepsilon}_h^H(0)\big)_{\mathcal{T}_h} + \big\langle\boldsymbol{\delta}(0), \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big\rangle_{\partial\mathcal{T}_h},\end{aligned}$$

by Lemma 3.6. The result now follows after a simple application of Young's inequality. This proves the first estimate.

It remains to prove the second estimate. By definition of $\mathrm{Er}_1(0)$ (12b), we have that $\mathrm{Er}_1(0)^2 := \mathrm{Er}_{11}(0)^2 + \mathrm{Er}_{12}(0)^2$, where

$$\begin{aligned}\mathrm{Er}_{11}(0)^2 &:= \left\|\epsilon^{1/2}\dot{\boldsymbol{\varepsilon}}_h^E(0)\right\|_{\mathcal{T}_h}^2, \\ \mathrm{Er}_{12}(0)^2 &:= \left\|\mu^{1/2}\dot{\boldsymbol{\varepsilon}}_h^H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{1/2}\Big(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A(0) - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(0)\Big)\right\|_{\partial\mathcal{T}_h}^2.\end{aligned}$$

The second term can be estimated as in the proof of the first inequality. Indeed, we have that

$$\begin{aligned}\mathrm{Er}_{12}(0)^2 &= \left\|\mu^{1/2}\dot{\boldsymbol{\varepsilon}}_h^H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{1/2}\Big(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A(0) - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(0)\Big)\right\|_{\partial\mathcal{T}_h}^2 \\ &= B_h\Big(\dot{\boldsymbol{\varepsilon}}_h^A(0), \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(0), \dot{\boldsymbol{\varepsilon}}_h^H(0); \dot{\boldsymbol{\varepsilon}}_h^A(0), \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(0), \dot{\boldsymbol{\varepsilon}}_h^H(0)\Big),\end{aligned}$$

by the fourth identity of Lemma 3.1. Then, by Lemma 3.4 with $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) := (\dot{\boldsymbol{\varepsilon}}_h^A(0), \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(0), \dot{\boldsymbol{\varepsilon}}_h^H(0))$, we get

$$\mathrm{Er}_{12}(0)^2 = \big(\mu\dot{\boldsymbol{e}}_H(0), \dot{\boldsymbol{\varepsilon}}_h^H(0)\big)_{\mathcal{T}_h} + \Big\langle\dot{\boldsymbol{\delta}}(0), \boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A(0) - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(0)\Big\rangle_{\partial\mathcal{T}_h},$$

because $\dot{\boldsymbol{\varepsilon}}_h^A(0) = 0$. An application of Young's inequality gives us the estimate

$$\mathrm{Er}_{12}(0)^2 \leqslant \left\|\mu^{1/2}\dot{\boldsymbol{e}}_H(0)\right\|_{\mathcal{T}_h}^2 + \left\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}(0)\right\|_{\partial\mathcal{T}_h}^2.$$

It remains to estimate the first term. But, the error equations of Lemma 3.4 at time $t = 0$ are

$$\big(\epsilon\,\ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\big)_{\mathcal{T}_h} + \big(\sigma\,\dot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{v}\big)_{\mathcal{T}_h} + B_h\big(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}\big) = (\mu\boldsymbol{e}_H, \boldsymbol{w})_{\mathcal{T}_h} + \langle\boldsymbol{\delta}, \boldsymbol{v} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} + (\epsilon\,\ddot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h} + (\sigma\,\dot{\boldsymbol{e}}_A, \boldsymbol{v})_{\mathcal{T}_h},$$

and, by Lemma 3.6, we have that

$$\big(\epsilon\,\ddot{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{v}\big)_{\mathcal{T}_h} + \big(\sigma\,\dot{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{v}\big)_{\mathcal{T}_h} = (\epsilon\,\ddot{\boldsymbol{e}}_A(0), \boldsymbol{v})_{\mathcal{T}_h} + (\sigma\,\dot{\boldsymbol{e}}_A(0), \boldsymbol{v})_{\mathcal{T}_h},$$

Since $\dot{\boldsymbol{\varepsilon}}_h^A(0) = 0$, we obtain

$$\big(\epsilon\ddot{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{v}\big)_{\mathcal{T}_h} = (\epsilon\ddot{\boldsymbol{e}}_A(0), \boldsymbol{v})_{\mathcal{T}_h} + (\sigma\dot{\boldsymbol{e}}_A(0), \boldsymbol{v})_{\mathcal{T}_h}, \tag{13}$$

we get

$$\mathrm{Er}_{11}(0) = \left\|\epsilon^{1/2}\ddot{\boldsymbol{\varepsilon}}_h^A(0)\right\|_{\mathcal{T}_h} \leqslant \left\|\epsilon^{1/2}\ddot{\boldsymbol{e}}_A(0)\right\|_{\mathcal{T}_h} + \left\|\epsilon^{-1/2}\sigma\dot{\boldsymbol{e}}_A(0)\right\|_{\mathcal{T}_h}.$$

The proof of the second estimate is complete. □

Note that, by (12b), the error term $\mathrm{Er}_1(0)$ depends on the estimate of $\|\ddot{\boldsymbol{\varepsilon}}_h^A(0)\|$. If a full HDG discretization were used to define the initial condition, then instead of (13), we would obtain

$$\left(\epsilon\ddot{\boldsymbol{\varepsilon}}_h^A(0),\boldsymbol{v}\right)_{\mathcal{T}_h}=(\epsilon\ddot{\boldsymbol{e}}_A(0),\boldsymbol{v})_{\mathcal{T}_h}+(\sigma\dot{\boldsymbol{e}}_A(0),\boldsymbol{v})_{\mathcal{T}_h}+(p_h,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h}-\langle\widehat{p}_h,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

The above identity suggests that p_h can degrade the order of convergence for $\mathrm{Er}_1(0)$, and then consequently also $\|\dot{\boldsymbol{\varepsilon}}_h^E(t)\|$, $\|\dot{\boldsymbol{\varepsilon}}_h^H(t)\|$, and $\|\tau_t^{1/2}(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A(t)-\dot{\widehat{\boldsymbol{\varepsilon}}}_h^A(t))\|_{\partial\mathcal{T}_h}$. Since these term are used for the duality estimate of $\boldsymbol{A}_h(t)$, which we shall present in the next subsection, the error pollution from p_h may also have an effect on the convergence of $\|\boldsymbol{\varepsilon}_h^A(t)\|$. Our choice of the initial condition by (5) is free from these problems since $p_h=0$ by Proposition 2.1.

Step 8: Proof of the first estimate of Theorem 2.1

Combining the estimates obtained in Propositions 3.3 and 3.4, we can bound the projection of the errors

$$\mathrm{Er}_0(T)=\left(\left\|\epsilon^{1/2}\boldsymbol{\varepsilon}_h^E(T)\right\|_{\mathcal{T}_h}^2+\left\|\mu^{1/2}\boldsymbol{\varepsilon}_h^H(T)\right\|_{\mathcal{T}_h}^2+\left\|\tau_t^{1/2}\left(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(T)-\widehat{\boldsymbol{\varepsilon}}_h^A(T)\right)\right\|_{\partial\mathcal{T}_h}^2\right)^{1/2},$$

as we get that

$$\begin{aligned}\mathrm{Er}_0(T)&\lesssim\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h}+\left\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\right\|_{\partial\mathcal{T}_h}+\left\|\tau_t^{-1/2}\boldsymbol{\delta}\right\|_{L^\infty(T,L^2(\partial\mathcal{T}_h))}\\&\quad+\int_0^T\left(\|\ddot{\boldsymbol{e}}_A\|_{\mathcal{T}_h}^2+\|\dot{\boldsymbol{e}}_A\|_{\mathcal{T}_h}^2+\|\dot{\boldsymbol{e}}_H\|_{\mathcal{T}_h}^2+\left\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}\right\|_{\partial\mathcal{T}_h}^2\right)^{1/2}\\&\lesssim(1+T)\left(\|\boldsymbol{e}_H\|_{W^{1,\infty}(T,L^2(\mathcal{T}_h))}+\|\boldsymbol{e}_A\|_{W^{2,\infty}(T,L^2(\mathcal{T}_h))}+\left\|\tau_t^{-1/2}\boldsymbol{\delta}\right\|_{W^{1,\infty}(T,L^2(\partial\mathcal{T}_h))}\right),\end{aligned}$$

where $\boldsymbol{\varepsilon}_h^E=\boldsymbol{\Pi}_V\boldsymbol{E}-\boldsymbol{E}_h$, $\boldsymbol{\varepsilon}_h^H=\boldsymbol{\Pi}_W\boldsymbol{H}-\boldsymbol{H}_h$, and

$$\tau_t^{1/2}\left(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A-\widehat{\boldsymbol{\varepsilon}}_h^A\right)=\tau_t^{1/2}\left(\boldsymbol{P}_M(\boldsymbol{\Pi}_V\boldsymbol{A}-\boldsymbol{A}_h)-\left(\boldsymbol{P}_M\boldsymbol{A}-\widehat{\boldsymbol{A}}_h\right)\right).$$

Since $\tau_t^{-1/2}\boldsymbol{\delta}^{(m)}=\tau_t^{-1/2}(\boldsymbol{n}\times\boldsymbol{\Pi}_W\boldsymbol{H}^{(m)}-\boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{H}^{(m)}))+\tau_t^{1/2}(\boldsymbol{P}_M\boldsymbol{e}_A^{(m)})$, we have

$$\left\|\tau_t^{-1/2}\boldsymbol{\delta}^{(m)}\right\|_{\partial\mathcal{T}_h}\lesssim\left\|\tau_t^{-1/2}\boldsymbol{e}_H^{(m)}\right\|_{\partial\mathcal{T}_h}+\left\|\tau_t^{1/2}\boldsymbol{e}_A^{(m)}\right\|_{\partial\mathcal{T}_h},$$

and, by the triangle inequality, we get that

$$\begin{aligned}&\|\boldsymbol{E}(T)-\boldsymbol{E}_h(T)\|_{\mathcal{T}_h}+\|\boldsymbol{H}(T)-\boldsymbol{H}_h(T)\|_{\mathcal{T}_h}+\left\|\tau_t^{1/2}\left(\boldsymbol{P}_M\boldsymbol{A}_h(T)-\widehat{\boldsymbol{A}}_h(T)\right)\right\|_{\partial\mathcal{T}_h}\\&\qquad\lesssim\mathrm{Er}_0(T)+\|\boldsymbol{e}_H(T)\|_{\mathcal{T}_h}+\|\boldsymbol{e}_E(T)\|_{\mathcal{T}_h}+\left\|\tau_t^{1/2}\boldsymbol{e}_A(T)\right\|_{\partial\mathcal{T}_h}\\&\qquad\lesssim(1+T)\Bigg(\|\boldsymbol{e}_H\|_{W^{1,\infty}(T,L^2(\mathcal{T}_h))}+\|\boldsymbol{e}_A\|_{W^{2,\infty}(T,L^2(\mathcal{T}_h))}\\&\qquad\quad+\left\|\tau_t^{-1/2}\boldsymbol{e}_H\right\|_{W^{1,\infty}(T,L^2(\partial\mathcal{T}_h))}+\left\|\tau_t^{1/2}\boldsymbol{e}_A\right\|_{W^{1,\infty}(T,L^2(\partial\mathcal{T}_h))}\Bigg).\end{aligned}$$

The wanted estimate now follows from the approximation estimates (9):

$$\begin{aligned}\|\boldsymbol{e}_H\|_{W^{1,\infty}(T,L^2(\mathcal{T}_h))}&\lesssim\max_{K\in\mathcal{T}_h}\left(h_K^{\min(k_W+1,s_H)}\right)\|\boldsymbol{H}\|_{W^{1,\infty}(T,H^{s_H}(\mathcal{T}_h))},\\\|\boldsymbol{e}_A\|_{W^{2,\infty}(T,L^2(\mathcal{T}_h))}&\lesssim\max_{K\in\mathcal{T}_h}\left(h_K^{\min(k_V+1,s_A)}\right)\|\boldsymbol{A}\|_{W^{2,\infty}(T,H^{s_A}(\mathcal{T}_h))},\end{aligned}$$

$$\left\|\tau_t^{-1/2}\boldsymbol{e}_H\right\|_{W^{1,\infty}(T,L^2(\partial\mathcal{T}_h))} \lesssim \max_{K\in\mathcal{T}_h}\left(\left(h_K\,\tau_{t,K}^{\min}\right)^{-1/2}h_K^{\min(k_W+1,s_H)}\right)\|\boldsymbol{H}\|_{W^{1,\infty}(T,H^{s_H}(\mathcal{T}_h))},$$
$$\left\|\tau_t^{1/2}\boldsymbol{e}_A\right\|_{W^{1,\infty}(T,L^2(\partial\mathcal{T}_h))} \lesssim \max_{K\in\mathcal{T}_h}\left(\left(h_K\,\tau_{t,K}^{\max}\right)^{1/2}h_K^{\min(k_V,s_A-1)}\right)\|\boldsymbol{A}\|_{W^{1,\infty}(T,H^{s_A}(\mathcal{T}_h))}.$$

This completes the proof.

Step 9: The dual problem

In what follows, we give an estimate of $\boldsymbol{\varepsilon}_h^A(T)$ by using a duality argument, also called the "Aubin-Nitsche trick" in second-order elliptic problems.

Since we are going to use the following characterization of the $\boldsymbol{L}^2$-norm of $\boldsymbol{\varepsilon}_h^A(T)$;

$$\left\|\epsilon\,\boldsymbol{\varepsilon}_h^A(T)\right\|_\Omega := \sup_{\Theta\in\boldsymbol{L}^2(\Omega)\backslash\{\boldsymbol{0}\}}\left(\boldsymbol{\Theta},\epsilon\,\boldsymbol{\varepsilon}_h^A(T)\right)_{\mathcal{T}_h}/\|\boldsymbol{\Theta}\|_{\boldsymbol{L}^2(\Omega)}.$$

We need to introduce the corresponding dual problem, namely,

$$-\epsilon\ddot{\boldsymbol{A}}^* + \sigma\dot{\boldsymbol{A}}^* - \nabla\times\boldsymbol{H}^* = 0 \qquad \text{in } \Omega\times[0,T], \tag{14a}$$
$$\mu\boldsymbol{H}^* - \nabla\times\boldsymbol{A}^* = 0 \qquad \text{in } \Omega\times[0,T], \tag{14b}$$
$$\nabla\cdot\boldsymbol{A}^* = 0 \qquad \text{in } \Omega\times[0,T], \tag{14c}$$
$$\boldsymbol{n}\times\boldsymbol{A}^* = 0 \qquad \text{on } \Gamma\times[0,T], \tag{14d}$$

with $\boldsymbol{A}^*(T) = 0$ and $\dot{\boldsymbol{A}}^*(T) = \boldsymbol{\Theta}\in L^2(\Omega)$.

We next define the projections of the dual solution, namely $\boldsymbol{\Pi}_W\boldsymbol{H}^*$ and $\boldsymbol{\Pi}_V\boldsymbol{A}^*$, and their corresponding boundary remainder:

$$\boldsymbol{\delta}^* := \boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{H}^*,\boldsymbol{A}^*). \tag{15}$$

For any time-dependent function $\boldsymbol{f}(t)$ with $t\in[0,T]$, we introduce the following notation for the time-integration from t to T:

$$\underline{\boldsymbol{f}}(t) = \int_t^T \boldsymbol{f}(s)\,\mathrm{d}s.$$

Note that, since $\boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}$ is linear, $\underline{\boldsymbol{\delta}}^* = \boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\underline{\boldsymbol{H}}^*,\underline{\boldsymbol{A}}^*)$.

Proposition 3.5. *Suppose Assumption 2.3 holds. We have*

$$\|\boldsymbol{A}^*\|_{W^{1,\infty}(T,L^2(\Omega))} + \|\nabla\times\boldsymbol{A}^*\|_{L^\infty(T,L^2(\Omega))} \leqslant C\|\boldsymbol{\Theta}\|_\Omega,$$
$$\left\|\underline{\boldsymbol{A}}^*\right\|_{L^\infty\left(T,H^{1+s_{\mathrm{reg},A}}(\Omega)\right)} + \left\|\underline{\boldsymbol{H}}^*\right\|_{L^\infty(T,H^{s_{\mathrm{reg},H}}(\Omega))} \leqslant C\|\boldsymbol{\Theta}\|_\Omega.$$

Proof. To obtain the first inequality, we use a simple energy argument. Indeed, if we set $E(t) := (\epsilon\dot{\boldsymbol{A}}^*(t),\dot{\boldsymbol{A}}^*(t))_\Omega + (\mu^{-1}\nabla\times\boldsymbol{A}^*(t),\nabla\times\boldsymbol{A}^*(t))_\Omega$, we get, by (14), that $\dot{E}(t) = 2(\sigma\dot{\boldsymbol{A}}^*(t),\dot{\boldsymbol{A}}^*(t))_\Omega \geqslant 0$. So, $E(t)\leqslant E(T) = \|\epsilon^{1/2}\boldsymbol{\Theta}\|_\Omega^2$ for all $t\in[0,T]$. It remains to estimate the term $\|\boldsymbol{A}^*\|_{L^\infty(T,L^2(\Omega))}$. To do that, we use the Poincaré-type inequality:

$$\|\boldsymbol{u}\|_\Omega \leqslant C(\|\nabla\times\boldsymbol{u}\|_\Omega + \|\nabla\cdot\boldsymbol{u}\|_\Omega) \qquad \boldsymbol{n}\times\boldsymbol{u}|_{\partial\Omega} = 0.$$

The above inequality holds for a bounded and simply connected domain Ω with a Lipschitz connected boundary, see Proposition 16.9 of [32]. This condition for Ω is satisfied here since we have assumed Assumption 2.3-(3). Now, since $\nabla\cdot\boldsymbol{A}^* = 0$ and $\boldsymbol{n}\times\boldsymbol{A}^* = 0$, we have $\|\boldsymbol{A}^*\|_\Omega \lesssim \|\nabla\times\boldsymbol{A}^*\|_\Omega$. This completes the proof of the first estimate.

To obtain the second inequality, we do use an elliptic regularity result. By integrating (14a) and (14b) from t to T, we obtain

$$\begin{aligned} \nabla \times \underline{\boldsymbol{H}}^*(t) &= -\epsilon \dot{\boldsymbol{A}}^*(T) + \epsilon \dot{\boldsymbol{A}}^*(t) - \sigma \boldsymbol{A}^*(t), \\ \mu \underline{\boldsymbol{H}}^*(t) - \nabla \times \underline{\boldsymbol{A}^*}(t) &= 0, \end{aligned}$$

hold for all $t \in [0, T]$. Then by the elliptic regularity inequality of Assumption 2.3-(1), we obtain

$$\|\underline{\boldsymbol{A}}^*(t)\|_{s_{\mathrm{reg},A}+1,\Omega} + \|\underline{\boldsymbol{H}}^*(t)\|_{s_{\mathrm{reg},H},\Omega} \lesssim \|\boldsymbol{A}^*\|_{W^{1,\infty}(T,L^2(\Omega))} \quad \forall t \in [0,T].$$

Now we use the first inequality. This completes the proof. □

Step 10: The duality argument

Next, we use a duality argument to find an expression for $\left(\epsilon \boldsymbol{\Theta}, \boldsymbol{\varepsilon}_h^A(T)\right)_{\mathcal{T}_h}$ suitable for the analysis.

Proposition 3.6. *We have that* $\left(\epsilon \boldsymbol{\Theta}, \boldsymbol{\varepsilon}_h^A(T)\right)_{\mathcal{T}_h} = R_0 + \sum_{i=1}^4 \int_0^T R_i(t)\,\mathrm{d}t$, *where*

$$\begin{aligned} R_0 &:= \left(\epsilon \dot{\boldsymbol{A}}^*(0) - \sigma \boldsymbol{A}^*(0), \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h}, \\ R_1 &:= \langle \boldsymbol{\delta}^*, \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{\delta}, \boldsymbol{\Pi}_V \boldsymbol{A}^* - \boldsymbol{P}_M \boldsymbol{A}^* \rangle_{\partial \mathcal{T}_h}, \\ R_2 &:= (\mu(\boldsymbol{\Pi}_W \boldsymbol{H}^* - \boldsymbol{H}^*), \boldsymbol{H}_h - \boldsymbol{H})_{\mathcal{T}_h} + (\mu(\boldsymbol{\Pi}_W \boldsymbol{H} - \boldsymbol{H}), \boldsymbol{H}^*)_{\mathcal{T}_h}, \\ R_3 &:= \left(\epsilon(\boldsymbol{\Pi}_V \boldsymbol{A}^* - \boldsymbol{A}^*), \ddot{\boldsymbol{A}} - \ddot{\boldsymbol{A}}_h\right)_{\mathcal{T}_h} + \left(\ddot{\boldsymbol{A}} - \boldsymbol{\Pi}_V \ddot{\boldsymbol{A}}, \epsilon \boldsymbol{A}^*\right)_{\mathcal{T}_h}, \\ R_4 &:= \left(\sigma(\boldsymbol{\Pi}_V \boldsymbol{A}^* - \boldsymbol{A}^*), \dot{\boldsymbol{A}} - \dot{\boldsymbol{A}}_h\right)_{\mathcal{T}_h} + \left(\dot{\boldsymbol{A}} - \boldsymbol{\Pi}_V \dot{\boldsymbol{A}}, \sigma \boldsymbol{A}^*\right)_{\mathcal{T}_h}. \end{aligned}$$

Proof. If we set $\boldsymbol{\eta} := (\epsilon \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A)_{\mathcal{T}_h} - (\epsilon \boldsymbol{A}^*, \dot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h} - (\sigma \boldsymbol{A}^*, \boldsymbol{\varepsilon}_h^A)_{\mathcal{T}_h}$, we have that

$$\boldsymbol{\eta}(T) = \boldsymbol{\eta}(0) + \int_0^T \left(\left(\epsilon \ddot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - (\epsilon \boldsymbol{A}^*, \ddot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h} - \left(\sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - (\sigma \boldsymbol{A}^*, \dot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h} \right).$$

Since $\boldsymbol{A}^*(T) = 0$, we have that $\boldsymbol{\eta}(T) = (\epsilon \dot{\boldsymbol{A}}^*(T), \boldsymbol{\varepsilon}_h^A(T))_{\mathcal{T}_h} = (\epsilon\, \boldsymbol{\Theta}, \boldsymbol{\varepsilon}_h^A(T))_{\mathcal{T}_h}$ because $\dot{\boldsymbol{A}}^*(T) = \boldsymbol{\Theta}$. Moreover, $\boldsymbol{\eta}(0) = (\epsilon \dot{\boldsymbol{A}}^*(0) - \sigma \boldsymbol{A}^*(0), \boldsymbol{\varepsilon}_h^A(0))_{\mathcal{T}_h} = R_0$ because $\dot{\boldsymbol{\varepsilon}}_h^A(0) = 0$. Thus

$$\left(\epsilon \boldsymbol{\Theta}, \boldsymbol{\varepsilon}_h^A(T)\right)_{\mathcal{T}_h} = R_0 + \int_0^T R(t)\,\mathrm{d}t,$$

where $R := (\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A)_{\mathcal{T}_h} - (\epsilon \boldsymbol{A}^*, \ddot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h} - (\sigma \boldsymbol{A}^*, \dot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h}$.

To prove that $R = \sum_{i=1}^4 R_i$, we begin by carrying out a few simple algebraic manipulations. We have,

$$\begin{aligned} R &= \left(\epsilon \ddot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - (\epsilon \boldsymbol{A}^*, \ddot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h} - \left(\sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - (\sigma \boldsymbol{A}^*, \dot{\boldsymbol{\varepsilon}}_h^A)_{\mathcal{T}_h} \\ &= \left(\epsilon \ddot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - \left(\epsilon \boldsymbol{A}^*, \boldsymbol{\Pi}_V \ddot{\boldsymbol{A}} - \ddot{\boldsymbol{A}}\right)_{\mathcal{T}_h} - \left(\epsilon \boldsymbol{A}^*, \ddot{\boldsymbol{A}} - \ddot{\boldsymbol{A}}_h\right)_{\mathcal{T}_h} \\ &\quad - \left(\sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - \left(\sigma \boldsymbol{A}^*, \boldsymbol{\Pi}_V \dot{\boldsymbol{A}} - \dot{\boldsymbol{A}}\right)_{\mathcal{T}_h} - \left(\sigma \boldsymbol{A}^*, \dot{\boldsymbol{A}} - \dot{\boldsymbol{A}}_h\right)_{\mathcal{T}_h} \\ &= \left(\epsilon \ddot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + R_3 - \left(\epsilon \boldsymbol{\Pi}_V \boldsymbol{A}^*, \ddot{\boldsymbol{A}} - \ddot{\boldsymbol{A}}_h\right)_{\mathcal{T}_h} - \left(\sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + R_4 - \left(\sigma \boldsymbol{\Pi}_V \boldsymbol{A}^*, \dot{\boldsymbol{A}} - \dot{\boldsymbol{A}}_h\right)_{\mathcal{T}_h} \\ &= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - \left(\boldsymbol{\Pi}_V \boldsymbol{A}^*, \epsilon(\ddot{\boldsymbol{A}} - \ddot{\boldsymbol{A}}_h) + \sigma\left(\dot{\boldsymbol{A}} - \dot{\boldsymbol{A}}_h\right)\right)_{\mathcal{T}_h} + R_3 + R_4 \end{aligned}$$

$$
\begin{aligned}
&= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + \left(\boldsymbol{\Pi}_V \boldsymbol{A}^*, \epsilon(\ddot{\boldsymbol{e}}_h^A - \ddot{\boldsymbol{\varepsilon}}_h^A) + \sigma\left(\dot{\boldsymbol{e}}_h^A - \dot{\boldsymbol{\varepsilon}}_h^A\right)\right)_{\mathcal{T}_h} + R_3 + R_4 \\
&= \Xi + R_3 + R_4,
\end{aligned}
$$

where $\Xi := (\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A)_{\mathcal{T}_h} + (\boldsymbol{\Pi}_V \boldsymbol{A}^*, \epsilon(\ddot{\boldsymbol{e}}_h^A - \ddot{\boldsymbol{\varepsilon}}_h^A) + \sigma(\dot{\boldsymbol{e}}_h^A - \dot{\boldsymbol{\varepsilon}}_h^A))_{\mathcal{T}_h}$.

It remains to show that $\Xi = R_1 + R_2$. To do that, we begin by inserting information about the error equations. Using the equations of the projections of the errors, Lemma 3.4 with $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) := (\boldsymbol{\Pi}_V \boldsymbol{A}^*, \boldsymbol{P}_M \boldsymbol{A}^*, -\boldsymbol{\Pi}_W \boldsymbol{H}^*)$, we get that

$$
\begin{aligned}
\Xi &= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \boldsymbol{\Pi}_V \boldsymbol{A}^*, \boldsymbol{P}_M \boldsymbol{A}^*, -\boldsymbol{\Pi}_W \boldsymbol{H}^*\right) \\
&\quad + (\mu\, \boldsymbol{e}_H, \boldsymbol{\Pi}_W \boldsymbol{H}^*)_{\mathcal{T}_h} - \langle \boldsymbol{\delta}, \boldsymbol{\Pi}_V \boldsymbol{A}^* - \boldsymbol{P}_M \boldsymbol{A}^* \rangle_{\partial \mathcal{T}_h} \\
&= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + B_h\left(\boldsymbol{\Pi}_V \boldsymbol{A}^*, \boldsymbol{P}_M \boldsymbol{A}^*, \boldsymbol{\Pi}_W \boldsymbol{H}^*; \boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, -\boldsymbol{\varepsilon}_h^H\right) \\
&\quad + (\mu\, \boldsymbol{e}_H, \boldsymbol{\Pi}_W \boldsymbol{H}^*)_{\mathcal{T}_h} - \langle \boldsymbol{\delta}, \boldsymbol{\Pi}_V \boldsymbol{A}^* - \boldsymbol{P}_M \boldsymbol{A}^* \rangle_{\partial \mathcal{T}_h},
\end{aligned}
$$

by the first identity of Lemma 3.1. By the identity for the bilinear form B_h in Lemma 3.3, we have

$$
\begin{aligned}
\Xi &= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + \left(\nabla \times \boldsymbol{H}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - \left(\mu(\boldsymbol{\Pi}_W \boldsymbol{H}^* - \boldsymbol{H}^*), \boldsymbol{\varepsilon}_h^H\right)_{\mathcal{T}_h} + \left\langle \boldsymbol{\delta}^*, \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A \right\rangle_{\partial \mathcal{T}_h} \\
&\quad + (\mu \boldsymbol{e}_H, \boldsymbol{\Pi}_W \boldsymbol{H}^*)_{\mathcal{T}_h} - \langle \boldsymbol{\delta}, \boldsymbol{\Pi}_V \boldsymbol{A}^* - \boldsymbol{P}_M \boldsymbol{A}^* \rangle_{\partial \mathcal{T}_h} \\
&= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^* + \nabla \times \boldsymbol{H}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} - \left(\mu(\boldsymbol{\Pi}_W \boldsymbol{H}^* - \boldsymbol{H}^*), \boldsymbol{\varepsilon}_h^H\right)_{\mathcal{T}_h} + (\mu \boldsymbol{e}_H, \boldsymbol{\Pi}_W \boldsymbol{H}^*)_{\mathcal{T}_h} + R_1 \\
&= \left(\epsilon \ddot{\boldsymbol{A}}^* - \sigma \dot{\boldsymbol{A}}^* + \nabla \times \boldsymbol{H}^*, \boldsymbol{\varepsilon}_h^A\right)_{\mathcal{T}_h} + R_2 + R_1 \\
&= R_2 + R_1,
\end{aligned}
$$

since $\nabla \times \boldsymbol{H}^* = -\epsilon \ddot{\boldsymbol{A}}^* + \sigma \dot{\boldsymbol{A}}^*$. This completes the proof. ☐

All we have to do now is to estimate each of the terms $R_0, \int_0^T R_i, i = 1, 2, 3, 4$ and conclude. Before that, we introduce a lemma on the projection $\boldsymbol{\Pi}_W$ and $\boldsymbol{\Pi}_V$, which will be used frequently in the estimates of the R_i terms.

To do that,we are going to use several times the following auxiliary result.

Lemma 3.7. *If Assumption 2.4 holds, then we have*

$$
(\boldsymbol{\Pi}_W \boldsymbol{w} - \boldsymbol{w}, \boldsymbol{r})_K = 0 \quad \forall \boldsymbol{r} \in \boldsymbol{\mathcal{P}}_0(K), \quad (\boldsymbol{\Pi}_V \boldsymbol{v} - \boldsymbol{v}, \boldsymbol{r})_K = 0 \quad \forall \boldsymbol{r} \in \boldsymbol{\mathcal{P}}_1(K). \tag{16}
$$

Proof. Note that by Assumption 2.4, we know $\boldsymbol{\mathcal{P}}_1(K) \subset \boldsymbol{V}(K)$. In addition, recall that we define $\boldsymbol{\Pi}_V$ as the L^2 projection to $\boldsymbol{V}(K)$. Thus, we have $(\boldsymbol{\Pi}_V \boldsymbol{v} - \boldsymbol{v}, \boldsymbol{\mathcal{P}}_1(K))_K = 0$.

To prove $(\boldsymbol{\Pi}_W \boldsymbol{w} - \boldsymbol{w}, \boldsymbol{r})_K = 0$, for all $\boldsymbol{r} \in \boldsymbol{\mathcal{P}}_0(K)$, first recall the definition of $\boldsymbol{\Pi}_W$, namely (10). By Assumption 2.4, we know $\boldsymbol{n} \times \boldsymbol{\mathcal{P}}_1(K) \times \boldsymbol{n} \subset \boldsymbol{M}(\partial K)$ and $\boldsymbol{\mathcal{P}}_1(K) \subset \boldsymbol{V}(K)$. Now consider (10a), we take $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}_1(K)$, then we have

$$
(\boldsymbol{\Pi}_W \boldsymbol{w} - \boldsymbol{w}, \nabla \times \boldsymbol{\mathcal{P}}_1(K))_K = 0.
$$

This with the fact that $\boldsymbol{\mathcal{P}}_0 \subset \nabla \times \boldsymbol{\mathcal{P}}_1$ completes the proof. ☐

Step 11: Estimate of R_0

Lemma 3.8. *Suppose Assumptions 2.1–2.4 hold. Then we have*

$$
\begin{aligned}
|R_0| \leqslant \max_{K \in \mathcal{T}_h} &\left(h_K^{s_{\mathrm{reg},H}}, (\tau_{t,K}^{\min} h_K)^{-1/2} h_K^{s_{\mathrm{reg},H}}, (\tau_{t,K}^{\max} h_K)^{1/2} h_K^{s_{\mathrm{reg},A}}\right) \\
&\left(\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h} + \|\tau_t^{-1/2} \boldsymbol{\delta}(0)\|_{\partial \mathcal{T}_h}\right) \|\boldsymbol{\Theta}\|_{\Omega}.
\end{aligned}
$$

Proof. By the definition of R_0, we have

$$|R_0| = \left|\left(\epsilon \dot{\boldsymbol{A}}^*(0) - \sigma \boldsymbol{A}^*(0), \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h}\right| \lesssim \|\boldsymbol{A}^*\|_{W^{1,\infty}(T,L^2(\Omega))} \left\|\boldsymbol{\varepsilon}_h^A(0)\right\|_{\mathcal{T}_h} \lesssim \|\boldsymbol{\Theta}\|_\Omega \left\|\boldsymbol{\varepsilon}_h^A(0)\right\|_{\mathcal{T}_h},$$

by the first inequality of Proposition 3.5 (which requires Assumption 2.3 to hold). It remains to estimate the quantity $\|\boldsymbol{\varepsilon}_h^A(0)\|_{\mathcal{T}_h}$. Since

$$\left\|\boldsymbol{\varepsilon}_h^A(0)\right\|_{\mathcal{T}_h} = \sup_{\boldsymbol{\theta}\in \boldsymbol{L}^2(\Omega)\setminus\{\boldsymbol{0}\}} \left(\boldsymbol{\theta}, \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h} / \|\boldsymbol{\theta}\|_\Omega,$$

we use a duality technique to get the estimate. Thus, we begin by introducing the dual problem

$$\mu\boldsymbol{\psi} - \nabla\times(-\boldsymbol{\phi}) = 0, \quad -\nabla\times\boldsymbol{\psi} = \boldsymbol{\theta} \quad \text{in } \Omega, \quad \boldsymbol{\phi} = 0 \quad \text{on } \partial\Omega,$$

and writing that

$$\begin{aligned}
\left(\boldsymbol{\theta}, \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h} &= -\left(\nabla\times\boldsymbol{\psi}, \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h} \\
&= B_h\left(-\boldsymbol{\Pi}_V\boldsymbol{\phi}, -\boldsymbol{P}_M\boldsymbol{\phi}, \boldsymbol{\Pi}_W\boldsymbol{\psi}; -\boldsymbol{\varepsilon}_h^A(0), -\widehat{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{\varepsilon}_h^H(0)\right) \\
&\quad - \left(\mu\boldsymbol{e}_\psi, \boldsymbol{\varepsilon}_h^H(0)\right) + \left\langle \boldsymbol{\delta}_{-\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{\psi},\boldsymbol{\phi}), \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\right\rangle_{\partial\mathcal{T}_h},
\end{aligned}$$

by Lemma 3.3. Next we use the first identity of Lemma 3.1 to get

$$\begin{aligned}
\left(\boldsymbol{\theta}, \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h} &= B_h\left(\boldsymbol{\varepsilon}_h^A(0), \widehat{\boldsymbol{\varepsilon}}_h^A(0), \boldsymbol{\varepsilon}_h^H(0); \boldsymbol{\Pi}_V\boldsymbol{\phi}, \boldsymbol{P}_M\boldsymbol{\phi}, \boldsymbol{\Pi}_W\boldsymbol{\psi}\right) \\
&\quad - \left(\mu\boldsymbol{e}_\psi, \boldsymbol{\varepsilon}_h^H(0)\right) + \left\langle \boldsymbol{\delta}_{-\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{\psi},\boldsymbol{\phi}), \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

By the error equation at the initial time, Lemma 3.6 (which requires Assumption 2.2 to hold), we obtain

$$\begin{aligned}
\left(\boldsymbol{\theta}, \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h} &= (\mu\boldsymbol{e}_H(0), \boldsymbol{\Pi}_W\boldsymbol{\psi})_{\mathcal{T}_h} + \left\langle \boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{H}(0), \boldsymbol{A}(0)), \boldsymbol{\Pi}_V\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\right\rangle_{\partial\mathcal{T}_h} \\
&\quad - \left(\mu\boldsymbol{e}_\psi, \boldsymbol{\varepsilon}_h^H(0)\right) + \left\langle \boldsymbol{\delta}_{-\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{\psi},\boldsymbol{\phi}), \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

Setting $\boldsymbol{\delta}(0) := \boldsymbol{\delta}_{\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{H}(0), \boldsymbol{A}(0))$ and $\boldsymbol{\delta}' := \boldsymbol{\delta}_{-\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{\psi},\boldsymbol{\phi})$, we arrive at

$$\begin{aligned}
\left(\boldsymbol{\theta}, \boldsymbol{\varepsilon}_h^A(0)\right)_{\mathcal{T}_h} &= -\left(\mu\boldsymbol{e}_\psi, \boldsymbol{\varepsilon}_h^H(0)\right) + \left\langle \boldsymbol{\delta}', \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\right\rangle_{\partial\mathcal{T}_h} + (\mu\boldsymbol{e}_H(0), \boldsymbol{\Pi}_W\boldsymbol{\psi})_{\mathcal{T}_h} + \left\langle \boldsymbol{\delta}(0), \boldsymbol{\Pi}_V\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\right\rangle_{\partial\mathcal{T}_h} \\
&= (\mu\boldsymbol{e}_H(0), \boldsymbol{\psi})_{\mathcal{T}_h} + (\mu(\boldsymbol{\Pi}_W\boldsymbol{\psi} - \boldsymbol{\psi}), \boldsymbol{H}_h(0) - \boldsymbol{H}(0))_{\mathcal{T}_h} \\
&\quad + \left\langle \boldsymbol{\delta}(0), \boldsymbol{\Pi}_V\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\right\rangle_{\partial\mathcal{T}_h} + \left\langle \boldsymbol{\delta}', \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\right\rangle_{\partial\mathcal{T}_h} \\
&=: R_{01} + R_{02} + R_{03} + R_{04}.
\end{aligned}$$

By (16) (which requires Assumption 2.4 to hold), and Assumption 2.3-(2) saying that μ is piecewise C^1, we have

$$|R_{01}| = \left|(\boldsymbol{e}_H(0), \mu\boldsymbol{\psi} - \boldsymbol{\Pi}_0(\mu\boldsymbol{\psi}))_{\mathcal{T}_h}\right| \lesssim h_K^{s_{\mathrm{reg},H}} \|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h} \|\boldsymbol{\psi}\|_{s_{\mathrm{reg},H},\mathcal{T}_h} \lesssim h^{s_{\mathrm{reg},H}} \|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h} \|\boldsymbol{\theta}\|_{\mathcal{T}_h},$$

by the elliptic regularity estimate $\|\boldsymbol{\phi}\|_{1+s_{\mathrm{reg},A},\Omega} + \|\boldsymbol{\psi}\|_{s_{\mathrm{reg},H},\Omega} \leqslant C_{\mathrm{reg}}\|\boldsymbol{\theta}\|_\Omega$, which holds by the elliptic regularity inequality of Assumption 2.3-(1).

Similarly, we have

$$\begin{aligned}
|R_{02}| &\lesssim h^{s_{\mathrm{reg},H}} \left(\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h} + \left\|\boldsymbol{\varepsilon}_h^H(0)\right\|_{\mathcal{T}_h}\right) \|\boldsymbol{\theta}\|_{\mathcal{T}_h}, \\
|R_{03}| &\lesssim \max_{K\in\mathcal{T}_h} \left(\left(\tau_{t,K}^{\max} h_K\right)^{1/2} h_K^{s_{\mathrm{reg},A}}\right) \left\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\right\|_{\partial\mathcal{T}_h} \|\boldsymbol{\theta}\|_{\mathcal{T}_h}.
\end{aligned}$$

To estimate R_{04}, first note that

$$\boldsymbol{\delta}' = \boldsymbol{\delta}_{-\tau_t}^{\boldsymbol{\Pi}_{W,V}}(\boldsymbol{\psi},\boldsymbol{\phi}) = \boldsymbol{n}\times\boldsymbol{\Pi}_W\boldsymbol{\psi} - \boldsymbol{P}_M(\boldsymbol{n}\times\boldsymbol{\psi}) - \tau_t(\boldsymbol{P}_M\boldsymbol{\Pi}_V\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}).$$

So we have

$$\begin{aligned}|R_{04}| &\lesssim \sum_{K\in\mathcal{T}_h}\Big(\big(\tau_{t,K}^{\min} h_K\big)^{-1/2} h_K^{s_{\mathrm{reg},H}}|\boldsymbol{\psi}|_{s_{\mathrm{reg},H},K} + \big(\tau_{t,K}^{\max} h_K\big)^{1/2} h_K^{s_{\mathrm{reg},A}}|\boldsymbol{\phi}|_{1+s_{\mathrm{reg},A},K}\Big) \\ &\qquad \Big\|\tau_t^{1/2}\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big)\Big\|_{\partial K} \\ &\lesssim \max_{K\in\mathcal{T}_h}\Big(\big(\tau_{t,K}^{\min} h_K\big)^{-1/2} h_K^{s_{\mathrm{reg},H}}, \big(\tau_{t,K}^{\max} h_K\big)^{1/2} h_K^{s_{\mathrm{reg},A}}\Big)\Big\|\tau_t^{1/2}\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big)\Big\|_{\partial\mathcal{T}_h}\|\boldsymbol{\theta}\|_{\mathcal{T}_h}.\end{aligned}$$

Finally, we combine the estimates for R_{0i} with $i = 1,2,3,4$, and then use the first estimate of Proposition 3.4 to control the terms $\|\boldsymbol{\varepsilon}_h^H(0)\|_{\mathcal{T}_h}$ and $\|\tau_t^{1/2}(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0))\|_{\partial\mathcal{T}_h}$. This completes the proof. ☐

Step 12: Estimate of $\int_0^T R_1$

Lemma 3.9. *Suppose Assumptions 2.1 and 2.3 hold. Then we have*

$$\begin{aligned}\left|\int_0^T R_1\right| &\leqslant \max_{K\in\mathcal{T}_h}\Big(\big(\tau_{t,K}^{\min} h_K\big)^{-1/2} h_K^{s_{\mathrm{reg},H}}, \big(\tau_{t,K}^{\max} h_K\big)^{1/2} h_K^{s_{\mathrm{reg},A}}\Big) \\ &\qquad \left(\mathrm{Er}_0(0) + T\sup_{s\in[0,T]}\mathrm{Er}_1(s) + \Big\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\Big\|_{\partial\mathcal{T}_h} + T\Big\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}\Big\|_{L^\infty(T,L^2(\partial\mathcal{T}_h))}\right)\|\boldsymbol{\Theta}\|_\Omega.\end{aligned}$$

Proof. Recall that

$$R_1 = \big\langle\boldsymbol{\delta}^*, \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A\big\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{\delta}, \boldsymbol{\Pi}_V\boldsymbol{A}^* - \boldsymbol{P}_M\boldsymbol{A}^*\rangle_{\partial\mathcal{T}_h} =: R_{11} + R_{12}.$$

So we have

$$\begin{aligned}\left|\int_0^T R_{11}\right| &= \left|\big\langle\underline{\boldsymbol{\delta}}^*(0), \boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big\rangle_{\partial\mathcal{T}_h} + \int_0^T\Big\langle\underline{\boldsymbol{\delta}}^*, \dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\Big\rangle_{\partial\mathcal{T}_h}\right| \\ &\leqslant \sum_{K\in\mathcal{T}_h}\Big\|\tau_t^{-1/2}\underline{\boldsymbol{\delta}}^*(0)\Big\|_{\partial K}\Big\|\tau_t^{1/2}\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big)\Big\|_{\partial K} + \int_0^T\sum_{K\in\mathcal{T}_h}\Big\|\tau_t^{-1/2}\underline{\boldsymbol{\delta}}^*\Big\|_{\partial K}\Big\|\tau_t^{1/2}\Big(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\Big)\Big\|_{\partial K}.\end{aligned}$$

Then we use the definition of $\boldsymbol{\delta}^*$, namely, equations (15) and (11), and the projection convergence properties (9) to estimate $\|\tau_t^{-1/2}\boldsymbol{\delta}^*\|_{\partial K}$. Note that Assumption 2.1 is needed for $\boldsymbol{\delta}^*$ to be well-defined. This gives us

$$\begin{aligned}\left|\int_0^T R_{11}\right| &\lesssim \sum_{K\in\mathcal{T}_h}\Big(\big(\tau_{t,K}^{\min} h_K\big)^{-1/2} h_K^{s_{\mathrm{reg},H}}\|\underline{\boldsymbol{H}}^*(0)\|_{s_{\mathrm{reg},H},K} + \big(\tau_{t,K}^{\max} h_K\big)^{1/2} h_K^{s_{\mathrm{reg},A}}\|\underline{\boldsymbol{A}}^*(0)\|_{1+s_{\mathrm{reg},A},K}\Big) \\ &\qquad \Big\|\tau_t^{1/2}\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0) - \widehat{\boldsymbol{\varepsilon}}_h^A(0)\big)\Big\|_{\partial K} \\ &\quad + \int_0^T\sum_{K\in\mathcal{T}_h}\Big(\big(\tau_{t,K}^{\min} h_K\big)^{-1/2} h_K^{s_{\mathrm{reg},H}}\|\underline{\boldsymbol{H}}^*\|_{s_{\mathrm{reg},H},K} + \big(\tau_{t,K}^{\max} h_K\big)^{1/2} h_K^{s_{\mathrm{reg},A}}\|\underline{\boldsymbol{A}}^*\|_{1+s_{\mathrm{reg},A},K}\Big) \\ &\qquad \Big\|\tau_t^{1/2}\Big(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A - \dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\Big)\Big\|_{\partial K}.\end{aligned}$$

Next, we use Proposition 3.5 (which needs Assumption 2.3 to hold) to estimate $(\|\underline{\boldsymbol{H}}^*\|_{s_{\text{reg},H},\mathcal{T}_h}+\|\underline{\boldsymbol{A}}^*\|_{s_{\text{reg},A}+1,\mathcal{T}_h})$. Then we have

$$\left|\int_0^T R_{11}\right| \lesssim \max_{K\in\mathcal{T}_h}\left(\left(\tau_{t,K}^{\min} h_K\right)^{-1/2} h_K^{s_{\text{reg},H}}, \left(\tau_{t,K}^{\max} h_K\right)^{1/2} h_K^{s_{\text{reg},A}}\right)\|\boldsymbol{\Theta}\|_\Omega\left(\mathrm{Er}_0(0)+T\sup_{s\in[0,T]}\mathrm{Er}_1(s)\right).$$

Similarly, we carry out the estimate for $|\int_0^T R_{12}|$ as follows:

$$\begin{aligned}\left|\int_0^T R_{12}\right| &\leqslant \left|\left\langle\boldsymbol{\delta}(0), \boldsymbol{P}_M\boldsymbol{\Pi}_V\underline{\boldsymbol{A}}^*(0)-\boldsymbol{P}_M\underline{\boldsymbol{A}}^*(0)\right\rangle_{\partial\mathcal{T}_h}+\int_0^T\left\langle\dot{\boldsymbol{\delta}}, \boldsymbol{P}_M\boldsymbol{\Pi}_V\underline{\boldsymbol{A}}^*-\boldsymbol{P}_M\underline{\boldsymbol{A}}^*\right\rangle_{\partial\mathcal{T}_h}\right|\\ &\leqslant \sum_{K\in\mathcal{T}_h}\left(\tau_{t,K}^{\max}\right)^{1/2}\left\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\right\|_{\partial K}\left\|\boldsymbol{\Pi}_V\underline{\boldsymbol{A}}^*(0)-\underline{\boldsymbol{A}}^*(0)\right\|_{\partial K}\\ &\quad+\int_0^T\sum_{K\in\mathcal{T}_h}\left(\tau_{t,K}^{\max}\right)^{1/2}\left\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}\right\|_{\partial K}\left\|\boldsymbol{\Pi}_V\underline{\boldsymbol{A}}^*-\underline{\boldsymbol{A}}^*\right\|_{\partial K}\\ &\lesssim \max_{K\in\mathcal{T}_h}\left(\left(\tau_{t,K}^{\max} h_K\right)^{1/2} h_K^{s_{\text{reg},A}}\right)\|\boldsymbol{\Theta}\|_\Omega\left(\left\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\right\|_{\partial\mathcal{T}_h}+T\left\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}\right\|_{L^\infty(T,L^2(\partial\mathcal{T}_h))}\right).\end{aligned}$$

This completes the proof. □

Step 13: Estimate of $\int_0^T R_2$

Lemma 3.10. *Suppose Assumptions 2.1, 2.3 and 2.4 hold. Then we have*

$$\left|\int_0^T R_2\right| \leqslant h^{s_{\text{reg},H}}\left(\mathrm{Er}_0(0)+\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h}+T\left(\sup_{s\in[0,T]}\mathrm{Er}_1(s)+\|\dot{\boldsymbol{e}}_H\|_{L^\infty(T,L^2(\mathcal{T}_h))}\right)\right)\|\boldsymbol{\Theta}\|_\Omega.$$

Proof. Recall that

$$R_2 := (\mu(\boldsymbol{\Pi}_W\boldsymbol{H}^*-\boldsymbol{H}^*), \boldsymbol{H}_h-\boldsymbol{H})_{\mathcal{T}_h}+(\mu(\boldsymbol{\Pi}_W\boldsymbol{H}-\boldsymbol{H}), \boldsymbol{H}^*)_{\mathcal{T}_h} =: R_{21}+R_{22}.$$

Then, we have

$$\begin{aligned}\left|\int_0^T R_{21}\right| &\leqslant \left|\left(\mu(\boldsymbol{H}_h(0)-\boldsymbol{H}(0)), \boldsymbol{\Pi}_W\underline{\boldsymbol{H}}^*(0)-\underline{\boldsymbol{H}}^*(0)\right)_{\mathcal{T}_h}+\int_0^T\left(\mu\left(\dot{\boldsymbol{H}}_h-\dot{\boldsymbol{H}}\right), \boldsymbol{\Pi}_W\underline{\boldsymbol{H}}^*-\underline{\boldsymbol{H}}^*\right)_{\mathcal{T}_h}\right|\\ &\leqslant \|\boldsymbol{\Pi}_W\underline{\boldsymbol{H}}^*-\underline{\boldsymbol{H}}^*\|_{L^\infty(T,L^2(\mathcal{T}_h))}\left(\|\boldsymbol{H}_h(0)-\boldsymbol{H}(0)\|_{\mathcal{T}_h}+T\left\|\dot{\boldsymbol{H}}_h-\dot{\boldsymbol{H}}\right\|_{L^\infty(T,L^2(\mathcal{T}_h))}\right)\\ &\lesssim h^{s_{\text{reg},H}}\|\boldsymbol{\Theta}\|_\Omega\left(\mathrm{Er}_0(0)+\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h}+T\left(\sup_{s\in[0,T]}\mathrm{Er}_1(s)+\|\dot{\boldsymbol{e}}_H\|_{L^\infty(T,L^2(\mathcal{T}_h))}\right)\right),\end{aligned}$$

where (9a) and Proposition 3.5 are used for the last inequality sign.

To estimate $|\int_0^T R_{22}|$, we first use (16) (which requires Assumption 2.4 to hold) and obtain

$$\left|\int_0^T R_{22}\right| = \left|\left(\boldsymbol{\Pi}_W\boldsymbol{H}(0)-\boldsymbol{H}(0), \mu\underline{\boldsymbol{H}}^*(0)-\boldsymbol{\Pi}_0\left(\mu\underline{\boldsymbol{H}}^*(0)\right)\right)_{\mathcal{T}_h}+\int_0^T\left(\boldsymbol{\Pi}_W\dot{\boldsymbol{H}}-\dot{\boldsymbol{H}}, \mu\underline{\boldsymbol{H}}^*-\boldsymbol{\Pi}_0\left(\mu\underline{\boldsymbol{H}}^*\right)\right)_{\mathcal{T}_h}\right|.$$

We then use Assumption 2.3-(2) which says that μ is piecewise C^1. So we have

$$\left|\int_0^T R_{22}\right| \lesssim \|\mu\underline{\boldsymbol{H}}^*-\boldsymbol{\Pi}_0\left(\mu\underline{\boldsymbol{H}}^*\right)\|_{L^\infty(T,L^2(\mathcal{T}_h))}\left(\|\boldsymbol{\Pi}_W\boldsymbol{H}(0)-\boldsymbol{H}(0)\|_{\mathcal{T}_h}+T\left\|\boldsymbol{\Pi}_W\dot{\boldsymbol{H}}-\dot{\boldsymbol{H}}\right\|_{L^\infty(T,L^2(\mathcal{T}_h))}\right)$$

$$
\begin{aligned}
&\lesssim h^{s_{\mathrm{reg},H}} \|\underline{\boldsymbol{H}}^*\|_{L^\infty(T,H^{s_{\mathrm{reg},H}}(\mathcal{T}_h))} \Big(\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h} + T\|\dot{\boldsymbol{e}}_H\|_{L^\infty(T,L^2(\mathcal{T}_h))} \Big) \\
&\lesssim h^{s_{\mathrm{reg},H}} \|\boldsymbol{\Theta}\|_\Omega \Big(\|\boldsymbol{e}_H(0)\|_{\mathcal{T}_h} + T\|\dot{\boldsymbol{e}}_H\|_{L^\infty(T,L^2(\mathcal{T}_h))} \Big),
\end{aligned}
$$

where Proposition 3.5 is used in the last step. This completes the proof. ☐

Step 14: Estimate of $\int_0^T R_3$

Lemma 3.11. *Suppose Assumptions 2.1, 2.3, and 2.4 hold. Then we have*

$$
\begin{aligned}
\left| \int_0^T R_3 \right| \lesssim \max_{K\in\mathcal{T}_h} \left(h_K^{s_{\mathrm{reg},A}}, \left(\tau_{t,K}^{\max}\right)^{1/2} h_K^{s_{\mathrm{reg},A}+1/2} \right) \|\boldsymbol{\Theta}\|_\Omega \\
\Bigg(\|\dot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h} + \|\ddot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h} + \left\| \tau_t^{-1/2} \boldsymbol{\delta}(0) \right\|_{\partial\mathcal{T}_h} + \mathrm{Er}_0(0) \\
+ T \Bigg(\|\ddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))} + \|\dddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))} + \left\| \tau_t^{-1/2} \dot{\boldsymbol{\delta}} \right\|_{L^\infty(T,L^2(\mathcal{T}_h))} + \sup_{s\in[0,T]} \mathrm{Er}_1(s) \Bigg) \Bigg).
\end{aligned}
$$

Proof. Recall that

$$
R_3 := \left(\epsilon \ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{e}_{A*}\right)_{\mathcal{T}_h}, -\left(\epsilon \ddot{\boldsymbol{e}}_A, \boldsymbol{e}_{A*}\right)_{\mathcal{T}_h}, -\left(\ddot{\boldsymbol{e}}_A, \epsilon \boldsymbol{A}^*\right)_{\mathcal{T}_h} =: R_{31} + R_{32} + R_{33}.
$$

We first estimate R_{32}. It follows that

$$
\begin{aligned}
\left| \int_0^T R_{32} \right| &= \left| (\epsilon \ddot{\boldsymbol{e}}_A(0), \underline{\boldsymbol{e}}_{A*}(0))_{\mathcal{T}_h} + \int_0^T (\epsilon \dddot{\boldsymbol{e}}_A, \underline{\boldsymbol{e}}_{A*})_{\mathcal{T}_h} \right| \\
&\lesssim h^{1+s_{\mathrm{reg},A}} \|\underline{\boldsymbol{A}}^*\|_{L^\infty\left(T,H^{1+s_{\mathrm{reg},A}}(\Omega)\right)} \Big(\|\ddot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h} + T\|\dddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))} \Big),
\end{aligned}
$$

by using the estimate (9b) to control $\underline{\boldsymbol{e}}_{A*}$, and the minimal polynomial degree Assumption 2.4-(2), *i.e.*, $\boldsymbol{\mathcal{P}}_1(K) \subset \boldsymbol{V}(K)$.

Next, we estimate R_{33}. By (16) and Assumption 2.3-(2) saying that ϵ is piecewise-C^1, we have

$$
\begin{aligned}
\left| \int_0^T R_{33} \right| &= \left| \int_0^T (\ddot{\boldsymbol{e}}_A, \epsilon \boldsymbol{A}^* - \boldsymbol{\Pi}_0(\epsilon \boldsymbol{A}^*))_{\mathcal{T}_h} \right| \\
&= \left| \left(\ddot{\boldsymbol{e}}_A(0), \epsilon \underline{\boldsymbol{A}}^*(0) - \boldsymbol{\Pi}_0\left(\epsilon \underline{\boldsymbol{A}}^*(0)\right)\right)_{\mathcal{T}_h} + \int_0^T \left(\dddot{\boldsymbol{e}}_A, \epsilon \underline{\boldsymbol{A}}^* - \boldsymbol{\Pi}_0\left(\epsilon \underline{\boldsymbol{A}}^*\right)\right)_{\mathcal{T}_h} \right| \\
&\lesssim h^{\min\{1,1+s_{\mathrm{reg},A}\}} \|\underline{\boldsymbol{A}}^*\|_{L^\infty\left(T,H^{1+s_{\mathrm{reg},A}}(\mathcal{T}_h)\right)} \Big(\|\ddot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h} + T\|\dddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))} \Big).
\end{aligned}
$$

Finally, we estimate R_{31}. Observe that, we can rewrite this term as $R_{31} = \left(\epsilon \ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*}\right)_{\mathcal{T}_h}$, where $\boldsymbol{P}_V^\epsilon$ is the ϵ-weighted L^2 projection to the space $\boldsymbol{V}_h$. We can now use the error equations of Lemma 3.4 with test function $(\boldsymbol{v}, \boldsymbol{\eta}, \boldsymbol{w}) := (\boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*}, \boldsymbol{0}, \boldsymbol{0})$ to get two new terms

$$
\begin{aligned}
R_{31} &= \left(\epsilon \ddot{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*}\right)_{\mathcal{T}_h} \\
&= (\epsilon(\ddot{\boldsymbol{e}}_A), \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*})_{\mathcal{T}_h} + \left(\sigma\left(\dot{\boldsymbol{e}}_A - \dot{\boldsymbol{\varepsilon}}_h^A\right) + \boldsymbol{\delta}, \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*}\right)_{\mathcal{T}_h} - B_h\left(\boldsymbol{\varepsilon}_h^A, \widehat{\boldsymbol{\varepsilon}}_h^A, \boldsymbol{\varepsilon}_h^H; \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*}, \boldsymbol{0}, \boldsymbol{0}\right) \\
&= \left(\epsilon \ddot{\boldsymbol{e}}_A + \sigma\left(\dot{\boldsymbol{e}}_A - \dot{\boldsymbol{\varepsilon}}_h^A\right) - \nabla \times \boldsymbol{\varepsilon}_h^H, \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*}\right)_{\mathcal{T}_h} + \left\langle \boldsymbol{\delta} - \tau_t\left(\boldsymbol{P}_M \boldsymbol{\varepsilon}_h^A - \widehat{\boldsymbol{\varepsilon}}_h^A\right), \boldsymbol{P}_V^\epsilon \boldsymbol{e}_{A*} \right\rangle_{\partial\mathcal{T}_h} \\
&=: R_{311} + R_{312},
\end{aligned}
$$

where we used the definition (6) of the bilinear form B_h. Now, we estimate the first of the terms above. By (12) (definition of Er_0 and Er_1) and (9b), we have

$$\begin{aligned}\left|\int_0^T R_{311}\right| &= \Big|\big(\epsilon\ddot{\boldsymbol{e}}_A(0)+\sigma\big(\dot{\boldsymbol{e}}_A(0)-\dot{\boldsymbol{\varepsilon}}_h^A(0)\big)-\nabla\times\boldsymbol{\varepsilon}_h^H(0),\boldsymbol{P}_V^\epsilon\underline{\boldsymbol{e}}_{A*}(0)\big)_{\mathcal{T}_h}\\&\quad+\int_0^T\big(\epsilon\dddot{\boldsymbol{e}}_A+\sigma\big(\ddot{\boldsymbol{e}}_A-\ddot{\boldsymbol{\varepsilon}}_h^A\big)-\nabla\times\dot{\boldsymbol{\varepsilon}}_h^H,\boldsymbol{P}_V^\epsilon\underline{\boldsymbol{e}}_{A*}\big)_{\mathcal{T}_h}\Big|\\&\lesssim \max_{K\in\mathcal{T}_h}\big(h_K^{s_{\mathrm{reg},A}}\big)\|\underline{\boldsymbol{A}}^*\|_{L^\infty(T,H^{1+s_{\mathrm{reg},A}}(\Omega))}\Bigg(h\big(\|\dot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h}+\|\ddot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h}\big)+\mathrm{Er}_0(0)\\&\quad+hT\Big(\|\ddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))}+\|\dddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))}\Big)+T\sup_{s\in[0,T]}\mathrm{Er}_1(s)\Bigg),\end{aligned}$$

where we have used the well-known inverse estimate $\|\nabla\times\boldsymbol{f}\|_K\lesssim h_K^{-1}\|\boldsymbol{f}\|_K$ for any $\boldsymbol{f}$ lying in a given polynomial space; see, for instance, [4]. Also note that $\dot{\boldsymbol{\varepsilon}}_h^A(0)=0$ by (4). It remains to estimate the term associated to R_{312}. We have

$$\begin{aligned}\left|\int_0^T R_{312}\right| &= \left|\big\langle\boldsymbol{\delta}(0)-\tau_t\big(\boldsymbol{P}_M\boldsymbol{\varepsilon}_h^A(0)-\widehat{\boldsymbol{\varepsilon}}_h^A(0)\big),\boldsymbol{P}_V^\epsilon\underline{\boldsymbol{e}}_{A*}(0)\big\rangle_{\partial\mathcal{T}_h}+\int_0^T\Big\langle\dot{\boldsymbol{\delta}}-\tau_t\Big(\boldsymbol{P}_M\dot{\boldsymbol{\varepsilon}}_h^A-\dot{\widehat{\boldsymbol{\varepsilon}}}_h^A\Big),\boldsymbol{P}_V^\epsilon\underline{\boldsymbol{e}}_{A*}\Big\rangle_{\partial\mathcal{T}_h}\right|\\&\lesssim\max_{K\in\mathcal{T}_h}\Big(\big(\tau_{t,K}^{\max}h_K\big)^{1/2}h_K^{s_{\mathrm{reg},A}}\Big)\|\underline{\boldsymbol{A}}^*\|_{L^\infty(T,H^{1+s_{\mathrm{reg},A}}(\mathcal{T}_h))}\Bigg(\Big\|\tau_t^{-1/2}\boldsymbol{\delta}(0)\Big\|_{\partial\mathcal{T}_h}+\mathrm{Er}_0(0)\\&\quad+T\Bigg(\Big\|\tau_t^{-1/2}\dot{\boldsymbol{\delta}}\Big\|_{L^\infty(T,L^2(\partial\mathcal{T}_h))}+\sup_{s\in[0,T]}\mathrm{Er}_1(s)\Bigg)\Bigg).\end{aligned}$$

Finally we use Proposition 3.5 to control $\|\underline{\boldsymbol{A}}^*\|_{L^\infty(T,H^{1+s_{\mathrm{reg},A}}(\mathcal{T}_h))}$. This completes the proof. ☐

Step 15: Estimate of $\int_0^T R_4$

Lemma 3.12. *Suppose Assumptions 2.1, 2.3, and 2.4 hold. Then we have*

$$\left|\int_0^T R_4\right|\leqslant h^{s_{\mathrm{reg},A}}\left(\|\dot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h}+T\|\ddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))}+T\sup_{t\in[0,T]}\mathrm{Er}_1(t)\right)\|\boldsymbol{\Theta}\|_\Omega.$$

Proof. Recall that

$$R_4:=\Big(\sigma(\boldsymbol{\Pi}_V\boldsymbol{A}^*-\boldsymbol{A}^*),\dot{\boldsymbol{A}}-\dot{\boldsymbol{A}}_h\Big)_{\mathcal{T}_h}+\Big(\dot{\boldsymbol{A}}-\boldsymbol{\Pi}_V\dot{\boldsymbol{A}},\sigma\boldsymbol{A}^*\Big)_{\mathcal{T}_h}=:R_{41}+R_{42}.$$

We begin by estimating the first term. Since $\dot{\boldsymbol{\varepsilon}}_h^A(0)=0$ by (4), we have

$$\begin{aligned}\left|\int_0^T R_{41}\right| &= \left|\Big(\sigma\big(\boldsymbol{\Pi}_V\underline{\boldsymbol{A}}^*(0)-\underline{\boldsymbol{A}}^*(0)\big),\dot{\boldsymbol{A}}(0)-\dot{\boldsymbol{A}}_h(0)\Big)_{\mathcal{T}_h}+\int_0^T\Big(\sigma\big(\boldsymbol{\Pi}_V\underline{\boldsymbol{A}}^*-\underline{\boldsymbol{A}}^*\big),\ddot{\boldsymbol{A}}-\ddot{\boldsymbol{A}}_h\Big)_{\mathcal{T}_h}\right|\\&\lesssim h^{1+s_{\mathrm{reg},A}}\|\underline{\boldsymbol{A}}^*\|_{L^\infty(T,H^{1+s_{\mathrm{reg},A}}(\mathcal{T}_h))}\left(\|\dot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h}+T\,\|\ddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))}+T\sup_{t\in[0,T]}\mathrm{Er}_1(t)\right).\end{aligned}$$

Now we estimate the second term. By (16) and Assumption 2.3-(2) saying that σ is piecewise-C^1, we get

$$\left|\int_0^T R_{42}\right|=\left|\Big(\dot{\boldsymbol{A}}(0)-\boldsymbol{\Pi}_V\dot{\boldsymbol{A}}(0),\sigma\underline{\boldsymbol{A}}^*(0)-\boldsymbol{\Pi}_0\big(\sigma\underline{\boldsymbol{A}}^*(0)\big)\Big)_{\mathcal{T}_h}+\int_0^T\Big(\ddot{\boldsymbol{A}}-\boldsymbol{\Pi}_V\ddot{\boldsymbol{A}},\sigma\underline{\boldsymbol{A}}^*-\boldsymbol{\Pi}_0\big(\sigma\underline{\boldsymbol{A}}^*\big)\Big)_{\mathcal{T}_h}\right|$$

$$\lesssim h^{\min\{1,1+s_{\mathrm{reg},A}\}}\|\underline{\boldsymbol{A}}^*\|_{L^\infty(T,H^{1+s_{\mathrm{reg},A}}(\mathcal{T}_h))}\Big(\|\dot{\boldsymbol{e}}_A(0)\|_{\mathcal{T}_h}+T\|\ddot{\boldsymbol{e}}_A\|_{L^\infty(T,L^2(\mathcal{T}_h))}\Big).$$

We finally use Proposition 3.5 to control the $\underline{\boldsymbol{A}}^*$ related terms. This completes the proof. □

Step 16: Conclusion

It only remains to prove the second inequality of Theorem 2.1. But, to do that, we simply have to gather the estimates of the terms $R_0, \int_0^T R_i, i = 1,2,3,4$, obtained in the previous steps. Then we use the second estimate of Proposition 3.3 to control $\sup_{s\in[0,T]} \mathrm{Er}_1(s)$ and Proposition 3.4 to control $\mathrm{Er}_0(0)$. This completes the proof.

4. Fully discrete scheme

In this section we present the fully discrete method. First, we cast the semidiscrete scheme (3) as a first order system of differential equations and then discretize using arbitrary-order, diagonally implicit Runge–Kutta methods. Alternative schemes such as explicit Runge–Kutta, partitioned Runge–Kutta, and fully implicit Runge Kutta methods can also be considered. See [18] for a comparison of explicit and implicit HDG methods for wave equation. Symplectic DIRK schemes are chosen here to showcase the energy conservation of the numerical scheme and corroborate the estimates in the error analysis.

4.1. First-order semidiscrete scheme

To obtain fully discrete schemes we first rewrite the second-order system of differential equations given by method (3) as a first-order system, by introducing the electric field $\boldsymbol{E}_h = -\dot{\boldsymbol{A}}_h$ into the equations and obtaining

$$\begin{aligned}
\Big(\epsilon\dot{\boldsymbol{A}}_h,\boldsymbol{v}\Big)_{\mathcal{T}_h}+(\epsilon\boldsymbol{E}_h,\boldsymbol{v})_{\mathcal{T}_h}&=0,\\
\Big(\epsilon\dot{\boldsymbol{E}}_h,\boldsymbol{v}\Big)_{\mathcal{T}_h}+(\sigma\boldsymbol{E}_h,\boldsymbol{v})_{\mathcal{T}_h}-(\nabla\times\boldsymbol{H}_h,\boldsymbol{v})_{\mathcal{T}_h}-\Big\langle\tau_t\Big(\boldsymbol{P}_M\boldsymbol{A}_h-\widehat{\boldsymbol{A}}_h\Big),\boldsymbol{v}\Big\rangle_{\partial\mathcal{T}_h}&=-(\boldsymbol{J},\boldsymbol{v})_{\mathcal{T}_h},\\
(\mu\boldsymbol{H}_h,\boldsymbol{w})_{\mathcal{T}_h}-\Big\langle\boldsymbol{n}\times\widehat{\boldsymbol{A}}_h,\boldsymbol{w}\Big\rangle_{\partial\mathcal{T}_h}-(\boldsymbol{A}_h,\nabla\times\boldsymbol{w})_{\mathcal{T}_h}&=0,\\
\Big\langle\boldsymbol{n}\times\boldsymbol{H}_h+\tau_t\Big(\boldsymbol{A}_h-\widehat{\boldsymbol{A}}_h\Big),\boldsymbol{\eta}^0\Big\rangle_{\partial\mathcal{T}_h}&=0,\\
\Big\langle\widehat{\boldsymbol{A}}_h,\boldsymbol{\eta}\Big\rangle_{\partial\mathcal{T}_h}&=\langle\boldsymbol{g}_A\times\boldsymbol{n},\boldsymbol{\eta}^\partial\rangle_{\partial\mathcal{T}_h},
\end{aligned}$$

for all $(\boldsymbol{v},\boldsymbol{w},\boldsymbol{\eta}^0,\boldsymbol{\eta}^\partial)\in\boldsymbol{V}_h\times\boldsymbol{W}_h\times\boldsymbol{M}_h^0\times\boldsymbol{M}_h^\partial$. In [31], we proposed the system above for the case $\sigma = 0$, obtaining a Hamiltonian system suitable for coupling with symplectic time integrators. We will demonstrate the energy-conserving property of the scheme with a numerical example in Section 5.

4.2. DIRK schemes

Before discretizing the system above we introduce the Diagonally Implicit Runge–Kutta (DIRK) schemes for a general setting. Consider the first order initial value problem $\dot{y} = f(t,y)$, $y(t_0) = y_0$. The s-stages DIRK schemes with coefficients a_{ij}, b_j, c_j, for $i,j = 1,\dots,s$, define an approximation $y^n \approx y(t^n)$, at $t^n = t_0 + n\Delta t$ where Δt is the time-step, by the method

$$\begin{aligned}
y^{n+1}&=y^n+\Delta t\sum_{i=1}^s b_i k_i,\quad k_i=f\big(t^{n,i},y^{n,i}\big),\\
y^{n,i}&=y^n+\Delta t\sum_{j=1}^i a_{ij}k_j,\quad t^{n,i}=t^n+c_i\Delta t,\qquad\text{for } i=1,\dots,s.
\end{aligned}$$

We now apply this setting to our semidiscrete HDG scheme. Consider an s-stage DIRK method with coefficients a_{ij}, b_j, c_j, for $i,j = 1,\dots,s$, then compute the following three steps for $i = 1,\dots,s$:

(1) $\boldsymbol{Z}_{\boldsymbol{Y}}^{n,i} = \boldsymbol{Y}_h^n + \Delta t \sum_{j=1}^{i-1} a_{ij} \boldsymbol{K}_j^{\boldsymbol{Y}}$, for $\boldsymbol{Y} = \boldsymbol{A}, \boldsymbol{E}$.

(2) Solution $(\boldsymbol{A}_h^{n,i}, \boldsymbol{E}_h^{n,i}, \boldsymbol{H}_h^{n,i}, \widehat{\boldsymbol{A}}_h^{n,i}) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times \boldsymbol{W}_h, \times \boldsymbol{M}_h$ of the system:

$$\begin{aligned}
\left(\epsilon \frac{\boldsymbol{A}_h^{n,i}}{a_{ii}\Delta t}, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\epsilon \boldsymbol{E}_h^{n,i}, \boldsymbol{v}\right)_{\mathcal{T}_h} &= \left(\frac{\boldsymbol{Z}_{\boldsymbol{A}}^{n,i}}{a_{ii}\Delta t}, \boldsymbol{v}\right)_{\mathcal{T}_h}, \\
\left(\epsilon \frac{\boldsymbol{E}_h^{n,i}}{a_{ii}\Delta t}, \boldsymbol{v}\right)_{\mathcal{T}_h} + \left(\sigma \boldsymbol{E}_h^{n,i}, \boldsymbol{v}\right)_{\mathcal{T}_h} & \\
-\left(\nabla \times \boldsymbol{H}_h^{n,i}, \boldsymbol{v}\right)_{\mathcal{T}_h} - \left\langle \tau_t \left(\boldsymbol{P}_M \boldsymbol{A}_h^{n,i} - \widehat{\boldsymbol{A}}_h^{n,i}\right), \boldsymbol{v} \right\rangle_{\partial\mathcal{T}_h} &= \left(\frac{\boldsymbol{Z}_{\boldsymbol{E}}^{n,i}}{a_{ii}\Delta t}, \boldsymbol{v}\right)_{\mathcal{T}_h} - \left(\boldsymbol{J}(t^{n,i}), \boldsymbol{v}\right)_{\mathcal{T}_h}, \\
\left(\mu \boldsymbol{H}_h^{n,i}, \boldsymbol{w}\right)_{\mathcal{T}_h} - \left\langle \boldsymbol{n} \times \widehat{\boldsymbol{A}}_h^{n,i}, \boldsymbol{w} \right\rangle_{\partial\mathcal{T}_h} - \left(\boldsymbol{A}_h^{n,i}, \nabla \times \boldsymbol{w}\right)_{\mathcal{T}_h} &= 0, \\
\left\langle \boldsymbol{n} \times \boldsymbol{H}_h^{n,i} + \tau_t \left(\boldsymbol{A}_h^{n,i} - \widehat{\boldsymbol{A}}_h^{n,i}\right), \boldsymbol{\eta}^0 \right\rangle_{\partial\mathcal{T}_h} &= 0, \\
\left\langle \widehat{\boldsymbol{A}}_h^{n,i}, \boldsymbol{\eta}^\partial \right\rangle_{\partial\mathcal{T}_h} &= \left\langle \boldsymbol{g}_A(t^{n,i}) \times \boldsymbol{n}, \boldsymbol{\eta}^\partial \right\rangle_{\partial\mathcal{T}_h},
\end{aligned}$$

for all $(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{\eta}^0, \boldsymbol{\eta}^\partial) \in \boldsymbol{V}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h^0 \times \boldsymbol{M}_h^\partial$.

(3) $\boldsymbol{K}_i^{\boldsymbol{Y}} := \frac{1}{a_{ii}\Delta t}(\boldsymbol{Y}_h^{n,i} - \boldsymbol{Y}_h^n) - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} \boldsymbol{K}_j^{\boldsymbol{Y}}$, for $\boldsymbol{Y} = \boldsymbol{A}, \boldsymbol{E}, \boldsymbol{H}, \widehat{\boldsymbol{A}}$.

Finally, update the approximate solution $\boldsymbol{Y}_h^{n+1} = \boldsymbol{Y}_h^n + \Delta t \sum_{j=1}^{s} b_j \boldsymbol{K}_j^{\boldsymbol{Y}}$, for $\boldsymbol{Y} = \boldsymbol{A}, \boldsymbol{E}, \boldsymbol{H}, \widehat{\boldsymbol{A}}$.

The fully discrete system is written for a general s-stages DIRK scheme. Here we can consider for instance singly DIRK methods, symplectic DIRK methods, symmetric DIRK methods, explicit first-stage, just to name a few. See [17] for a complete review of DIRK methods.

5. Numerical examples

In this section, we present several numerical examples demonstrating the properties of the semidiscrete scheme (3), implemented with the spaces HDG-S$_k$ and HDG-B$_k$ (see Tab. 1), and discretized in time using DIRK methods. In the first two examples, we study numerically the error convergence; in the third example, we provide numerical support for our choice of initial condition (5); and in the last example we show the energy-conserving property of the scheme. All experiments in this section were implemented using the open source finite element software NETGEN/NGSolve [33, 34].

Example 1: History of convergence test

Here we validate the predictions given in Theorem 2.1 for the methods HDG-S$_k$ of index k and HDG-B$_k$ of index k combined with with a DIRK scheme of order $k+2$. For each of the approximations $\boldsymbol{E}_h$, $\boldsymbol{A}_h$, and $\boldsymbol{H}_h$, we compute the maximum over the time steps t^n of the L^2-errors of the corresponding error, and then estimate their orders of convergence (e.o.c.). For instance, for the magnetic vector potential approximation we compute

$$\text{error}(h) = \max_{t^n} \|\boldsymbol{A}(t^n) - \boldsymbol{A}_h^n\|_{L^2(\Omega)^3}, \qquad \text{e.o.c}(h) = \frac{\log(\text{error}(h)/\text{error}(h'))}{\log(h/h')},$$

where h' correspond to the previous mesh size parameter used in the computations. The experiment is carried on the unit cubic domain $\Omega = (0,1)^3$ and using the solution of problem (2) given by

$$\boldsymbol{A}(x,y,z,t) = \begin{pmatrix} \cos(\pi x)\sin(\pi y)\sin(\pi z)\sin(\sqrt{3}\pi t) \\ 0 \\ -\sin(\pi x)\sin(\pi y)\cos(\pi z)\sin(\sqrt{3}\pi t) \end{pmatrix},$$

TABLE 4. History of convergence of numerical approximations of Maxwell's equations (2) by method DIRK order $k+2$ and HDG-S$_k$ order k, wtih $\tau_t = 1.0$, up to final time $t = 0.5$.

			$\boldsymbol{E}_h$		$\boldsymbol{H}_h$		$\boldsymbol{A}_h$	
k	h	N	Error	E.O.C.	Error	E.O.C.	Error	E.O.C.
0	7.94e−01	12	5.82e−01	–	3.70e−01	–	2.49e−01	–
	3.99e−01	96	4.41e−01	0.40	3.41e−01	0.12	1.23e−01	1.01
	1.98e−01	768	3.00e−01	0.55	2.39e−01	0.51	8.36e−02	0.57
	9.92e−02	6144	1.80e−01	0.74	1.49e−01	0.68	5.21e−02	0.68
	4.96e−02	49 152	9.98e−02	0.86	8.62e−02	0.79	2.87e−02	0.86
1	7.94e−01	12	2.56e−01	–	1.69e−01	–	5.26e−02	–
	3.99e−01	96	1.26e−01	1.02	9.10e−02	0.90	3.86e−02	0.44
	1.98e−01	768	5.26e−02	1.26	3.01e−02	1.59	1.40e−02	1.46
	9.92e−02	6144	1.95e−02	1.43	8.28e−03	1.86	3.62e−03	1.95
	4.96e−02	49 152	4.84e−03	2.01	2.07e−03	2.00	8.27e−04	2.13
2	7.94e−01	12	1.81e−01	–	9.41e−02	–	6.21e−02	–
	3.99e−01	96	4.31e−02	2.07	2.00e−02	2.23	1.29e−02	2.26
	1.98e−01	768	7.14e−03	2.59	2.67e−03	2.91	1.88e−03	2.78
	9.92e−02	6144	1.12e−03	2.66	3.44e−04	2.96	2.55e−04	2.88
3	7.94e−01	12	3.25e−02	–	1.91e−02	–	9.03e−03	–
	3.99e−01	96	4.36e−03	2.90	2.41e−03	2.98	1.13e−03	2.99
	1.98e−01	768	4.41e−04	3.30	1.92e−04	3.65	1.17e−04	3.28
	9.92e−02	6144	4.05e−05	3.45	1.21e−05	3.98	8.00e−06	3.87

with material parameters: permittivity $\epsilon = 1$, permeability $\mu = 1$, and conductivity $\sigma = 0.5$. The remaining data of the problem are computed accordingly to the exact solution.

We present in Tables 4 and 5 the errors and orders of convergence for the DIRK-(HDG-S$_k$) and DIRK-(HDG-B$_k$) methods for $k = 0, 1, 2, 3$, for a sequence of uniform triangulations (tetrahedra) with meshsize parameter h and number of elements N.

Optimal convergence of the L^2 errors is observed in Table 5 for method DIRK-(HDG-B$_k$) as predicted in Theorem 2.1 and Table 3. We observe convergence of order $k+2$ for the magnetic vector potential $\boldsymbol{A}_h$, and of order $k+1$ for the magnetic field $\boldsymbol{H}_h$. The convergence of the electric field $\boldsymbol{E}_h$ lies in between the convergence orders of $\boldsymbol{H}_h$ and $\boldsymbol{A}_h$, which is consistent with our theoretical prediction.

In Table 4, we observe a slightly faster convergence than the one predicted in Theorem 2.1 and Table 3, namely, that of $k+1/2$ for the electric and magnetic fields and of $k+1$ for the magnetic vector potential.

Example 2: Low-regularity test

We now consider the analytic solution of problem (2) with low regularity defined on the L-shaped domain $\Omega = ((-1,1)^2 \backslash ((-1,0)^2) \times (0,1)$ by

$$\boldsymbol{A}(x,y,z,t) = \nabla S(x,y)\cos(t), \qquad S(x,y) = r^{4/3}\sin\left(\frac{4}{3}\theta\right),$$

where S is given in cylindrical coordinates (r,θ). Observe that $\boldsymbol{A}$ and $\boldsymbol{E}$ have a singularity located at the z-axis, and $\boldsymbol{A}, \boldsymbol{E} \in [H^{4/3-\varepsilon}(\Omega)]^3$, $\varepsilon > 0$. Similar test are presented in [7, 15] for the static case. We compute the numerical approximations on a sequence of uniform triangulations with meshsize h and number of elements N by the methods DIRK-(HDG)-S and DIRK-(HDG-B$_k$) for $k = 0, 1$. Errors and rates of convergence are displayed in Tables 6 and 7.

Table 6 shows optimal order approximation of the errors using the HDG-S$_k$ method and with stabilization parameter $\tau_t = 1$. Indeed, it shows order 1 for $\boldsymbol{E}$ and $\boldsymbol{A}$ in the case $k = 0$, and (approximately) order 4/3 in the

TABLE 5. History of convergence of numerical approximations of Maxwell's equations (2) by method DIRK order $k+2$ and HDG-B$_k$, with $\tau_t = 1/h$, up to final time $t = 0.5$.

			$\boldsymbol{E}_h$		$\boldsymbol{H}_h$		$\boldsymbol{A}_h$	
k	h	N	Error	E.O.C.	Error	E.O.C.	Error	E.O.C.
0	7.94e−01	12	4.94e−01	–	3.78e−01	–	1.64e−01	–
	3.99e−01	96	3.93e−01	0.33	3.00e−01	0.34	1.03e−01	0.67
	1.98e−01	768	2.10e−01	0.86	1.90e−01	0.65	6.19e−02	0.74
	9.92e−02	6144	9.10e−02	1.25	9.76e−02	0.96	2.05e−02	1.59
	4.96e−02	49 152	2.38e−02	1.93	4.58e−02	1.09	5.04e−03	2.03
1	7.94e−01	12	2.79e−01	–	1.58e−01	–	9.69e−02	–
	3.99e−01	96	1.12e−01	1.31	8.04e−02	0.98	3.25e−02	1.58
	1.98e−01	768	2.84e−02	1.98	2.63e−02	1.61	7.12e−03	2.19
	9.92e−02	6144	4.64e−03	2.61	6.81e−03	1.95	7.42e−04	3.26
	4.96e−02	49 152	6.20e−04	2.90	1.69e−03	2.01	7.12e−05	3.38
2	7.94e−01	12	1.90e−01	–	8.58e−02	–	6.40e−02	–
	3.99e−01	96	3.28e−02	2.54	1.78e−02	2.27	9.77e−03	2.71
	1.98e−01	768	3.57e−03	3.20	2.38e−03	2.90	7.56e−04	3.69
	9.92e−02	6144	2.70e−04	3.72	2.88e−04	3.05	4.04e−05	4.23
3	7.94e−01	12	3.52e−02	–	1.84e−02	–	1.09e−02	–
	3.99e−01	96	4.03e−03	3.13	2.11e−03	3.13	1.18e−03	3.20
	1.98e−01	768	3.62e−04	3.48	1.92e−04	3.46	6.221-05	4.25
	9.92e−02	6144	1.17e−05	4.96	1.25e−05	3.94	1.56e−06	5.32

TABLE 6. History of convergence of numerical approximations to Maxwell's equations (2) by a DIRK method of order $k+2$ and the HDG-S$_k$ method of degree k, with $\tau = 1$, up to $t = 2$, CFL $= 0.05$.

			$\boldsymbol{E}_h$		$\boldsymbol{H}_h$		$\boldsymbol{A}_h$	
k	h	N	Error	E.O.C.	Error	E.O.C.	Error	E.O.C.
0	1.26e+00	17	3.61e−01	0.00	9.87e−02	0.00	1.30e−01	0.00
	6.30e−01	136	2.61e−01	0.47	8.19e−02	0.27	1.07e−01	0.27
	3.15e−01	1088	2.03e−01	0.36	5.43e−02	0.59	7.59e−02	0.50
	1.57e−01	8704	1.53e−01	0.41	2.63e−02	1.04	4.50e−02	0.76
	7.87e−02	69 632	8.61e−02	0.83	1.07e−02	1.30	2.37e−02	0.92
1	1.26e+00	17	2.89e−01	0.00	3.27e−01	0.00	8.47e−02	0.00
	6.30e−01	136	1.71e−01	0.76	1.09e−01	1.58	4.55e−02	0.90
	3.15e−01	1088	6.30e−02	1.44	4.07e−02	1.43	1.87e−02	1.29
	1.57e−01	8704	1.96e−02	1.68	1.54e−02	1.41	6.20e−03	1.59
	7.87e−02	69 632	6.83e−03	1.52	6.23e−03	1.30	1.96e−03	1.67

case $k = 1$ due to low regularity of the solution. Table 7 shows, for the case $k = 0, 1$, the optimal approximation order of the errors for the HDG-B$_k$ method with $\tau_t = 1/h$, close to $4/3$ in both cases. In Figure 1, we show the numerical approximation by the DIRK-(HDG-B$_k$) method for $k = 1$ and $h = \sqrt{2}/8$ of the magnitude of the magnetic field $\boldsymbol{E}_h$ and the magnitude of the magnetic vector potential $\boldsymbol{A}_h$.

TABLE 7. History of convergence of numerical approximations to Maxwell's equations (2) by a DIRK method of order $k+2$ and the HDG-B$_k$ method of degree k, with $\tau = 1/h$, up to $t = 2$, CFL = 0.05.

			$\boldsymbol{E}_h$		$\boldsymbol{H}_h$		$\boldsymbol{A}_h$	
k	h	N	Error	E.O.C.	Error	E.O.C.	Error	E.O.C.
	1.26e+00	17	2.08e−01	0.00	9.55e−02	0.00	7.58e−02	0.00
	6.30e−01	136	1.53e−01	0.45	4.39e−02	1.12	4.35e−02	0.80
0	3.15e−01	1088	5.11e−02	1.58	1.14e−02	1.94	1.56e−02	1.48
	1.57e−01	8704	1.41e−02	1.86	3.09e−03	1.89	4.29e−03	1.87
	7.87e−02	69 632	4.89e−03	1.52	1.51e−03	1.04	1.54e−03	1.47
	1.26e+00	17	1.38e−01	0.00	2.88e−01	0.00	2.62e−02	0.00
	6.30e−01	136	3.69e−02	1.91	1.03e−01	1.49	7.92e−03	1.72
1	3.15e−01	1088	1.42e−02	1.38	3.74e−02	1.46	2.67e−03	1.57
	1.57e−01	8704	5.23e−03	1.44	1.36e−02	1.46	7.15e−04	1.90
	7.87e−02	69 632	1.86e−03	1.49	4.90e−03	1.47	2.70e−04	1.41

Example 3: Initial condition test

The estimates obtained in our main result, Theorem 2.1, rely on the particular definition of the initial condition given by (5). In this numerical experiment, we provide evidence indicating that the use of this initial condition is indeed essential for the optimal convergence of the scheme. We compare the performance of the DIRK-(HDG-B$_k$) method initialized with: a) the standard L^2 projection of the initial data $\boldsymbol{A}_0$ and $\boldsymbol{E}_0$ onto the HDG spaces, and with b) the HDG-approximation proposed in (5).

As a test problem, we consider a solution on the domain $\Omega = (0,1)^3$, with parameters $\epsilon = \mu = 1$, and $\sigma = 0$, given by

$$\boldsymbol{A}(x,y,z,t) \;=\; \frac{1}{\sqrt{3}\pi}\begin{pmatrix} \cos(\pi x)\sin(\pi y)\sin(\pi z)\cos\big(\sqrt{3}\pi t\big) \\ 0 \\ -\sin(\pi x)\sin(\pi y)\cos(\pi z)\cos\big(\sqrt{3}\pi t\big)\end{pmatrix}.$$

Note that $\boldsymbol{A}_0 \neq \mathbf{0}$ and $\boldsymbol{E}_0 = \mathbf{0}$.

In Figure 2, we consider the DIRK-(HDG-B$_k$) method, and plot the L^2-errors, of the electric field and the magnetic vector potential, *vs.* the corresponding meshsizes h. In red, are the results obtained by using the standard L^2-projection as initial condition. In blue are the results obtained with our new initial condition (5). We observe sub-optimal rates of convergence corresponding to the scheme initialized with the L^2-projections. This supports our preference for our choice for the initial condition.

Example 4: Conservation test

In this example, we demonstrate the energy-conserving property of our numerical scheme. Consider the solution of equation of Maxwell's equations (2) in domain $\Omega = (-2,2)\times(0,1)^2$ with periodic boundary conditions, material parameters $\mu = 1$, $\sigma = 0$ and non-constant permittivity

$$\epsilon(x,y,z) = 1 + 2S(5(x+1)) - 2S(-5(x-1)), \qquad S(x) := \frac{1}{1+e^{-x}}\cdot$$

The system is initialized with

$$\boldsymbol{E}_0(x,y,z) = \begin{pmatrix} 0 \\ 0 \\ \sin(2\pi x)\end{pmatrix}, \quad \boldsymbol{A}_0 = \mathbf{0}.$$

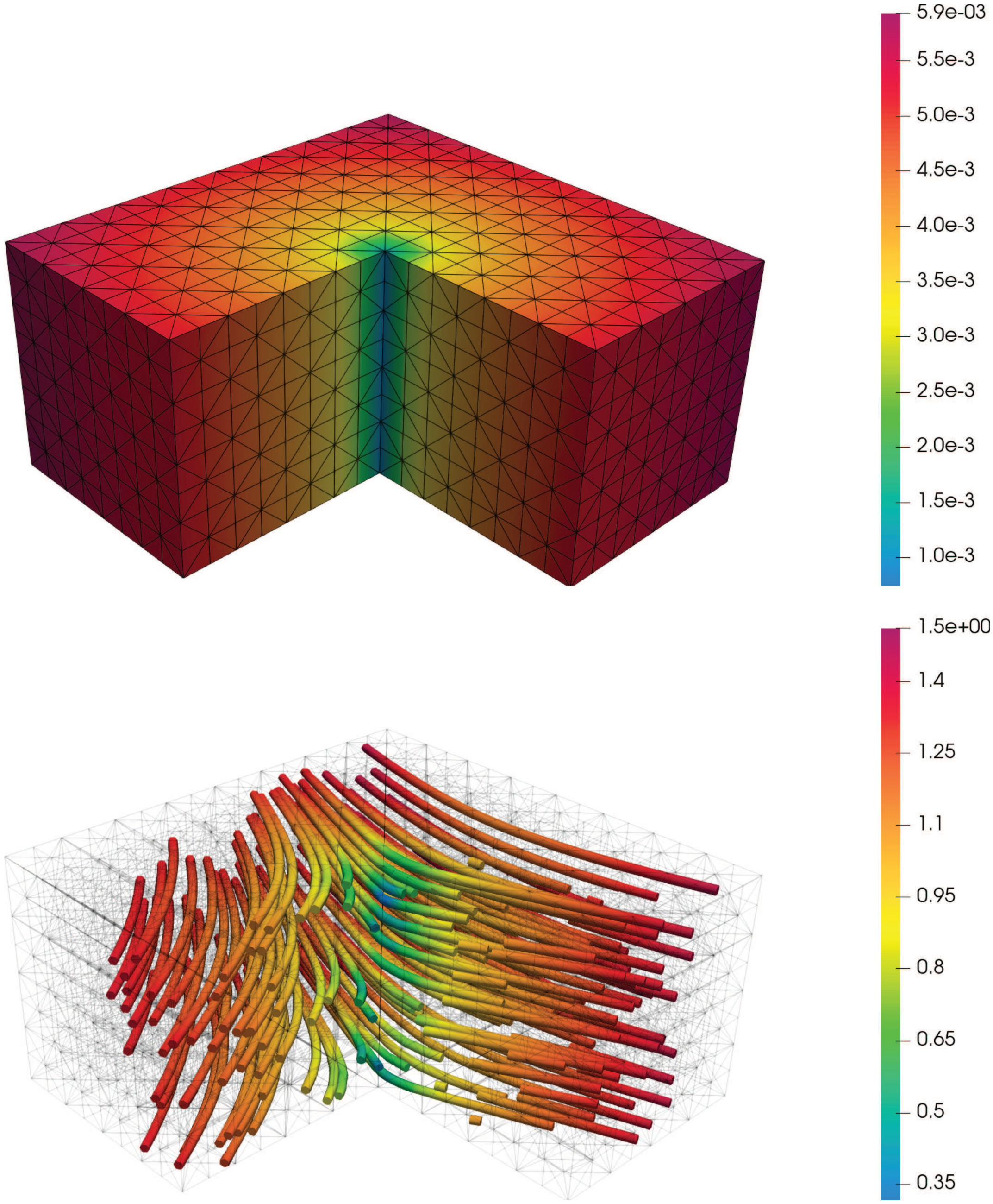

FIGURE 1. Numerical approximation by method the DIRK-(HDG-B_k) method for $k = 1$ and $h = \sqrt{2}/8$. *Top*: magnitude of $\boldsymbol{A}_h$. *Bottom*: streamlines and magnitude of $\boldsymbol{E}_h$.

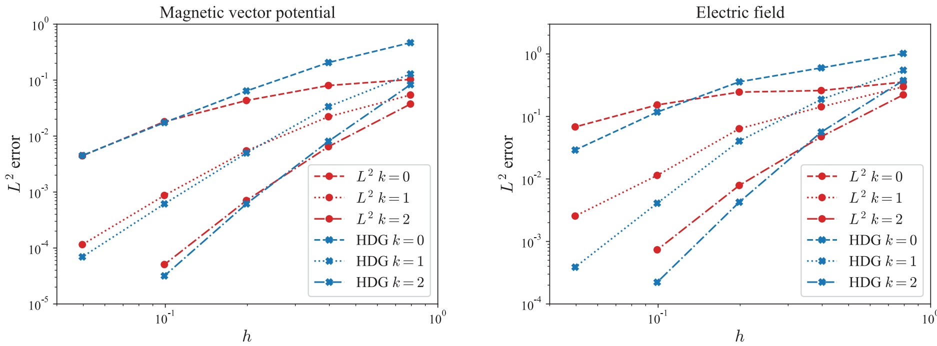

FIGURE 2. Plot the L^2 errors *vs.* meshsizes h of the DIRK-(HDG-B$_k$) method with initial condition L^2-projection (circle-dotted line), and HDG-approximation (cross-dashed line). *Left*: electric field.

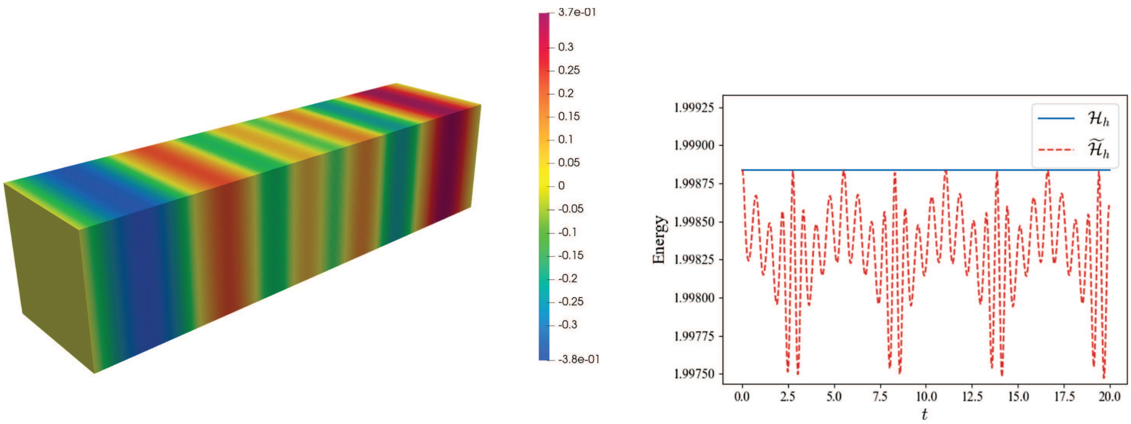

FIGURE 3. Magnitude of the approximate magnetic potential $\boldsymbol{A}_h$ (*left*) and the corresponding approximate, $\mathcal{H}_h$, and "exact" energies, $\widetilde{\mathcal{H}}_h$, (*right*) in a non-homogeneous medium. The numerical method is the DIRK-(HDG-B$_k$) method with $k = 1$.

We present, in Figure 3, a plot of the magnitude of the magnetic vector potential at time $t = 20$ and the evolution of the energy associated to the system, specifically we plot the quantities

$$\widetilde{\mathcal{H}}_h = \frac{1}{2}(\epsilon \boldsymbol{E}_h, \boldsymbol{E}_h)_{\mathcal{T}_h} + \frac{1}{2}(\mu \boldsymbol{H}_h, \boldsymbol{H}_h)_{\mathcal{T}_h}, \quad \mathcal{H}_h = \widetilde{\mathcal{H}}_h + \frac{1}{2}\Big\langle \tau_t\Big(P_{\boldsymbol{M}}\boldsymbol{A}_h - \widehat{\boldsymbol{A}}_h\Big), P_{\boldsymbol{M}}\boldsymbol{A}_h - \widehat{\boldsymbol{A}}_h\Big\rangle_{\partial\mathcal{T}_h}.$$

As it is shown in [31], the method DIRK-(HDG-B$_k$) method conserves the energy of the semidiscretization in space, $\mathcal{H}_h$, up to 10 digits of precision, and does not allow the "exact" energy, $\widetilde{\mathcal{H}}_h$, to drift in time.

6. Concluding remarks

In this paper, we provide the first *a priori* error analysis of a wide, new class of Symplectic-Hamiltonian finite element methods for the time-dependent Maxwell's equations, see [31]. We provide numerical results which validate our theoretical predictions and confirm that the schemes conserve *exactly* the energy of the discrete system and that the *exact* energy does not drift in time. Moreover, we propose a new way of defining the initial data which uses a continuous approximation for the "pressure". It behaves better than the simple L^2-projection, as our numerical experiments suggest.

Acknowledgements. B. Cockburn was supported in part by the NSF through DMS-1912646 grant. M.A. Sánchez was supported by FONDECYT Regular N. 1221189 and by Centro Nacional de Inteligencia Artificial CENIA, FB210017, Basal ANID Chile.

References

[1] C. Amrouche, C. Bernardi, M. Dauge and V. Girault, Vector potentials in three-dimensional non-smooth domains. *Math. Methods Appl. Sci.* **21** (1998) 823–864.

[2] A. Bonito, J.-L. Guermond and F. Luddens, Regularity of the Maxwell equations in heterogeneous media and Lipschitz domains. *J. Math. Anal. Appl.* **408** (2013) 498–512.

[3] A. Bossavit, Solving Maxwell equations in a closed cavity, and the question of 'spurious modes'. *IEEE Trans. Magn.* **26** (1990) 702–705.

[4] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods. Vol. 15 of *Texts in Applied Mathematics*, 3rd edition. Springer, New York (2008).

[5] A. Buffa and I. Perugia, Discontinuous Galerkin approximation of the Maxwell eigenproblem. *SIAM J. Numer. Anal.* **44** (2006) 2198–2226.

[6] L. Camargo, B. López-Rodríguez, M. Osorio and M. Solano, An HDG method for Maxwell's equations in heterogeneous media. *Comput. Methods Appl. Mech. Eng.* **368** (2020) 113178.

[7] H. Chen, W. Qiu, K. Shi and M. Solano, A superconvergent HDG method for the Maxwell equations. *J. Sci. Comput.* **70** (2017) 1010–1029.

[8] H. Chen, W. Qiu and K. Shi, A priori and computable a posteriori error estimates for an HDG method for the coercive Maxwell equations. *Comput. Methods Appl. Mech. Eng.* **333** (2018) 287–310.

[9] G. Chen, J. Cui and L. Xu, Analysis of a hybridizable discontinuous Galerkin method for the Maxwell operator. *ESAIM Math. Model. Numer. Anal.* **53** (2019) 301–324.

[10] A. Christophe, S. Descombes and S. Lanteri, An implicit hybridized discontinuous Galerkin method for the 3D time-domain Maxwell equations. *Appl. Math. Comput.* **319** (2018) 395–408.

[11] B. Cockburn, Z. Fu, A. Hungria, L. Ji, M.A. Sánchez and F.-J. Sayas, Stormer–Numerov HDG methods for acoustic waves. *J. Sci. Comput.* **75** (2018) 597–624.

[12] M. Costabel and M. Dauge, Maxwell and Lamé eigenvalues on polyhedra. *Math. Methods Appl. Sci.* **22** (1999) 243–258.

[13] M. Costabel and M. Dauge, Weighted regularization of Maxwell equations in polyhedral domains. A rehabilitation of nodal finite elements. *Numer. Math.* **93** (2002) 239–277.

[14] S. Du and F.-J. Sayas, New analytical tools for HDG in elasticity, with applications to elastodynamics. *Math. Comput.* **89** (2020) 1745–1782.

[15] S. Du and F.-J. Sayas, A unified error analysis of hybridizable discontinuous Galerkin methods for the static Maxwell equations. *SIAM J. Numer. Anal.* **58** (2020) 1367–1391.

[16] A. Ern and J.-L. Guermond, Analysis of the edge finite element approximation of the Maxwell equations with low regularity solutions. *Comput. Math. App.* **75** (2018) 918–932.

[17] C.A. Kennedy and M.H. Carpenter, *Diagonally implicit Runge–Kutta methods for ordinary differential equations. a review.* Technical report (2016).

[18] M. Kronbichler, S. Schoeder, C. Müller and W.A. Wall, Comparison of implicit and explicit hybridizable discontinuous Galerkin methods for the acoustic wave equation. *Int. J. Numer. Methods Eng.* **106** (2016) 712–739.

[19] A. La Spina and J. Fish, Time- and frequency-domain hybridizable discontinuous Galerkin solvers for the calculation of the Cherenkov radiation. *Comput. Methods Appl. Mech. Eng.* **402** (2022) 115170.

[20] C. Lehrenfeld, Hybrid discontinuous Galerkin methods for solving incompressible flow problems. *Rheinisch-Westfalischen Technischen Hochschule Aachen* (2010).

[21] X. Li, L. Xu, Z.-H. Yang and B. Li, A new explicit-implicit Hybridizable Discontinuous Galerkin time-domain method for electromagnetics. *IEEE Microwave Wirel. Compon. Lett.* **32** (2022) 1263–1266.

[22] P. Lu, H. Chen and W. Qiu, An absolutely stable *hp*-HDG method for the time-harmonic Maxwell equations with high wave number. *Math. Comput.* **86** (2017) 1553–1577.

[23] C.G. Makridakis and P. Monk, Time-discrete finite element schemes for Maxwell's equations. *RAIRO Modél. Math. Anal. Numér.* **29** (1995) 171–197.

[24] S. Marras, J.F. Kelly, M. Moragues, A. Müller, M.A. Kopera, M. Vázquez, F.X. Giraldo, G. Houzeaux and O.L. Jorba, A review of element-based Galerkin methods for numerical weather prediction: finite elements, spectral elements, and discontinuous Galerkin. *Arch. Comput. Methods Eng.* **23** (2016) 673–722.

[25] P. Monk, A mixed method for approximating Maxwell's equations. *SIAM J. Numer. Anal.* **28** (1991) 1610–1634.

[26] P. Monk, A finite element method for approximating the time-harmonic Maxwell equations. *Numer. Math.* **63** (1992) 243–261.

[27] J.-C. Nédélec, Mixed finite elements in $\mathbf{R}^3$. *Numer. Math.* **35** (1980) 315–341.

[28] J.-C. Nédélec, A new family of mixed finite elements in $\mathbf{R}^3$. *Numer. Math.* **50** (1986) 57–81.

[29] G. Nehmetallah, S. Lanteri, S. Descombes and A. Christophe, An explicit hybridizable discontinuous Galerkin method for the 3D time-domain Maxwell equations, in Spectral and High Order Methods for Partial Differential Equations – ICOSAHOM 2018. Vol. 134 of *Lect. Notes Comput. Sci. Eng.* Springer, Cham (2020) 513–523.

[30] N.C. Nguyen, J. Peraire and B. Cockburn, Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations. *J. Comput. Phys.* **230** (2011) 7151–7175.

[31] M.A. Sánchez, S. Du, B. Cockburn, N.-C. Nguyen and J. Peraire, Symplectic Hamiltonian finite element methods for electromagnetics. *Comput. Methods Appl. Mech. Eng.* **396** (2022) 114969.

[32] F.-J. Sayas, T.S. Brown and M.E. Hassell, Variational Techniques for Elliptic Partial Differential Equation: Theoretical Tools and Advanced Applications, 1st edition. CRC Press (2019).

[33] J. Schöberl, NETGEN: an advancing front 2D/3D-mesh generator based on abstract rules. *Comput. Visual. Sci.* **1** (1997) 41–52.

[34] J. Schöberl, *C++ 11 implementation of finite elements in NGSolve*. Institute for Analysis and Scientific Computing, Vienna University of Technology Report No. 30/2014 (2014).

Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at https://edpsciences.org/en/subscribe-to-open-s2o.