SIAM J. NUMER. ANAL.
Vol. 57, No. 4, pp. 1697–1722

ANALYSIS OF AN HDG METHOD FOR LINEARIZED INCOMPRESSIBLE RESISTIVE MHD EQUATIONS*

JEONGHUN J. LEE†, STEPHEN J. SHANNON‡, TAN BUI-THANH§, AND JOHN N. SHADID¶

Abstract. We develop a hybridized discontinuous Galerkin (HDG) method for stationary linearized incompressible magnetohydrodynamics (MHD) equations. At the heart of the development is the introduction of an upwind HDG flux for the dual saddle-point form of the MHD equations that facilitates the hybridization of the discontinuous Galerkin (DG) method. We carry out the a priori error estimates for the proposed HDG method on simplicial meshes in both two and three dimensions. The analysis provides optimal convergence for the fluid velocity and the magnetic variables, and quasi-optimal convergence for the remaining quantities. Numerical examples are presented to verify the theoretical findings.

Key words. hybridized discontinuous Galerkin methods, resistive magnetohydrodynamics, a priori error analysis, Stokes equations, Maxwell equations

AMS subject classifications. 65N30, 65N12, 65N15, 76W05

DOI. 10.1137/18M1166729

1. Introduction.

An important base-level representation for continuum approximation of the dynamics of electrically conducting fluids in the presence of electromagnetic fields is the resistive magnetohydrodynamics (MHD) model. MHD models describe important physical phenomena in astrophysical systems (e.g., solar flares and planetary magnetic field generation) and in critical scientific and technological applications (e.g., magnetically confined fusion energy devices) [17].

There are a number of difficulties in discretizations of the MHD equations, for example, the dual saddle-point structure of the velocity-pressure $(\boldsymbol{u}, p)$, nonlinear coupling of unknowns, and the enforcement of the solenoidal involution/constraint on the magnetic induction $(\nabla \cdot \boldsymbol{b} = 0)$. In the context of finite volume and finite element methods, there are several popular approaches: discretizations with physics-compatible finite elements (see, e.g., [23, 21, 2, 27, 20]); methods that transform into potential-based formulations to eliminate saddle-point subsystems [6, 28]; exact and weighted-exact penalty formulations [18, 15, 12, 13]; stabilization methods [26, 11, 29]; and discontinuous Galerkin (DG) methods [19].

*Received by the editors January 24, 2018; accepted for publication (in revised form) April 5, 2019; published electronically July 25, 2019. The proposed HDG method was first presented at the Oberwolfach Computational Engineering Workshop, 2015.
https://doi.org/10.1137/18M1166729
Funding: This work was partially supported by U.S. Department of Energy grants DE-SC0010518 and DE-SC0011118, by National Science Foundation grant DMS-1620352, and by U.S. Department of Energy NNSA ASC Algorithms effort, the U.S. Department of Energy Office of Science AMR program at Sandia National Laboratory under contract DE-AC04-94AL85000.
†Department of Mathematics, Baylor University, Waco, TX 76798 (Jeonghun_Lee@baylor.edu).
‡Institute for Computational and Engineering Sciences (ICES), University of Texas at Austin, Austin, TX 78712 (shannon@ices.utexas.edu).
§Institute for Computational and Engineering Sciences (ICES), University of Texas at Austin, Austin, TX 78712, and Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX 78712 (tanbui@ices.utexas.edu).
¶Computational Mathematics Department, Sandia National Laboratories, Albuquerque, NM 98185, and Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131 (jnshadi@sandia.gov).

In this paper we propose a *hybridized discontinuous Galerkin* (HDG) formulation for a linearized version of the resistive MHD system. This formulation can serve as a stand-alone solver for linearized MHD equations, or as the solver for a substep in a fixed-point nonlinear solver, such as Picard or Newton iteration. The hybridization technique and postprocessing have been proposed to reduce computational costs of saddle-point problems [1]. This has inspired the development of HDG methods to mitigate the computational costs of DG methods [7].

In HDG methods, unknowns on elements of the mesh can be reduced to single-valued trace unknowns on the mesh skeleton, so the total number of globally coupled unknowns is substantially smaller than in classical DG methods. In addition, once the trace unknowns are solved for, the element unknowns can be obtained by element-wise local solve, which can be efficiently implemented in parallel, and then constraints from physics can be imposed by elementwise postprocessing. However, devising an HDG method for coupled PDE systems is challenging because the construction of a consistent numerical flux leading to a stable numerical scheme is nontrivial. As may become apparent in this paper, the difficulties in carrying out the projection-based error analysis, along with the duality argument, become compounded, relative to simpler PDEs, due to the large and coupled nature of the system of PDEs that describes MHD. In this paper we apply an upwind HDG framework [4] to derive a numerical flux for linearized MHD equations and show that we can obtain stable HDG methods for the MHD equations.

We organize this paper as follows. Notation and conventions are introduced in section 2, and the description of our HDG method, including well-posedness proof, is presented in section 3. The *a priori error estimation* for the HDG method is discussed in section 4, and numerical results illustrating our theoretical findings are presented in section 5. Section 6 concludes the paper and describes ongoing work. Finally, in three appendices we briefly discuss the definitions of projection operators, auxiliary estimates, and the well-posedness of the adjoint equation.

2. Notation.

In this section we introduce common notation and conventions to be used in the rest of the paper. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded domain such that it is simply connected, and its boundary $\partial\Omega$ is a Lipschitz manifold with only one component. Suppose that we have a triangulation of Ω, i.e., a partition of Ω into a finite number of nonoverlapping d-dimensional simplices. We assume that the triangulation is shape-regular; i.e., for all d-dimensional simplices in the triangulation, the ratio of the diameter of the simplex and the radius of an inscribed d-dimensional ball is uniformly bounded. We will use Ω_h and $\mathcal{E}_h$ to denote the sets of d- and $(d-1)$-dimensional simplices of the triangulation, and call $\mathcal{E}_h$ the mesh skeleton of the triangulation. The boundary and interior mesh skeletons are defined by $\mathcal{E}_h^\partial := \{e \in \mathcal{E}_h : e \subset \partial\Omega\}$ and $\mathcal{E}_h^o := \mathcal{E}_h \backslash \mathcal{E}_h^\partial$. We also define $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$. The mesh size of triangulations is $h := \max_{K\in\Omega_h} \mathrm{diam}(K)$.

We use $(\cdot,\cdot)_D$ (respectively, $\langle\cdot,\cdot\rangle_D$) to denote the L^2-inner product on D if D is a d- (respectively, $(d-1)$-) dimensional domain. The standard notation $W^{s,p}(D)$, $s \geq 0$, $1 \leq p \leq \infty$, is used for the Sobolev space on D based on the L^p-norm with differentiability s (see, e.g., [14]), and $\|\cdot\|_{W^{s,p}(D)}$ denotes the associated norm. In particular, if $p = 2$, we use $H^s(D) := W^{s,p}(D)$ and $\|\cdot\|_{s,D}$. $W^{s,p}(\Omega_h)$ denotes the space of functions whose restrictions on K reside in $W^{s,p}(K)$ for each $K \in \Omega_h$, and its norm is $\|u\|^p_{W^{s,p}(\Omega_h)} := \sum_{K\in\Omega_h} \|u|_K\|^p_{W^{s,p}(K)}$ if $1 \leq p < \infty$ and $\|u\|_{W^{s,\infty}(\Omega_h)} := \max_{K\in\Omega_h} \|u|_K\|_{W^{s,\infty}(K)}$. For simplicity, we use $(\cdot,\cdot)$, $\langle\cdot,\cdot\rangle$, $\|\cdot\|_s$, $\|\cdot\|_{\partial\Omega_h}$, and $\|\cdot\|_{W^{s,\infty}}$ for $(\cdot,\cdot)_\Omega$, $\langle\cdot,\cdot\rangle_{\partial\Omega_h}$, $\|\cdot\|_{s,\Omega}$, $\|\cdot\|_{0,\partial\Omega_h}$, and $\|\cdot\|_{W^{s,\infty}(\Omega_h)}$, respectively. We define $\|u, v\| :=$

$\|u\| + \|v\|$. Furthermore, we denote by $A \lesssim B$ the inequality $A \leq \lambda B$ with a constant $\lambda > 0$ independent of the mesh size and by $A \sim B$ the combination of $A \lesssim B$ and $B \lesssim A$.

For vector- or matrix-valued functions this notation is naturally extended with a componentwise inner product. We define similar spaces (respectively, inner products and norms) on a single element and a single skeleton face/edge by replacing Ω_h with K and $\mathcal{E}_h$ with e. We define the gradient of a vector, the divergence of a matrix, and the outer product symbol $\otimes$ as

$$(\nabla \boldsymbol{u})_{ij} = \frac{\partial u_i}{\partial x_j}, \quad (\nabla \cdot \boldsymbol{L})_i = \nabla \cdot \boldsymbol{L}(i,:) = \sum_{j=1}^{3} \frac{\partial \boldsymbol{L}_{ij}}{\partial x_j}, \quad (\boldsymbol{a} \otimes \boldsymbol{b})_{ij} = a_i b_j = \left(\boldsymbol{a}\boldsymbol{b}^T\right)_{ij}.$$

In this paper $\boldsymbol{n}$ denotes a unit outward normal vector field on faces/edges. If $\partial K^- \cap \partial K^+ \in \mathcal{E}_h$ for two distinct simplices K^-, K^+, then $\boldsymbol{n}^-$ and $\boldsymbol{n}^+$ denote the outward unit normal vector fields on ∂K^- and ∂K^+, respectively, and $\boldsymbol{n}^- = -\boldsymbol{n}^+$ on $\partial K^- \cap \partial K^+$. We simply use $\boldsymbol{n}$ to denote either $\boldsymbol{n}^-$ or $\boldsymbol{n}^+$ in an expression that is valid for both cases, and this convention is also used for other quantities (restricted) on a face/edge $e \in \mathcal{E}_h$. For a scalar quantity u which is double-valued on $e := \partial K^- \cap \partial K^+$, the jump term on e is defined by $[\![u\boldsymbol{n}]\!]|_e = u^+\boldsymbol{n}^+ + u^-\boldsymbol{n}^-$, where u^+ and u^- are the traces of u from K^+- and K^--sides, respectively. For double-valued vector quantity $\boldsymbol{u}$ and matrix quantity $\boldsymbol{L}$, jump terms are $[\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]|_e = \boldsymbol{u}^+ \cdot \boldsymbol{n}^+ + \boldsymbol{u}^- \cdot \boldsymbol{n}^-$ and $[\![\boldsymbol{L}\boldsymbol{n}]\!]|_e = \boldsymbol{L}^+\boldsymbol{n}^+ + \boldsymbol{L}^-\boldsymbol{n}^-$, where $\boldsymbol{L}\boldsymbol{n}$ denotes the matrix-vector product.

We define $\mathcal{P}_k(K)$ as the space of polynomials of degree at most k on K, with $k \geq 0$, and we define

$$\mathcal{P}_k(\Omega_h) = \left\{ u \in L^2(\Omega) \ : \ u|_K \in \mathcal{P}_k(K) \ \forall K \in \Omega_h \right\}.$$

The space of polynomials on the mesh skeleton $\mathcal{P}_k(\mathcal{E}_h)$ is similarly defined, and their extensions to vector- or matrix-valued polynomials $[\mathcal{P}_k(\Omega_h)]^d$, $[\mathcal{P}_k(\Omega_h)]^{d\times d}$, $[\mathcal{P}_k(\mathcal{E}_h)]^d$, etc., are straightforward.

3. HDG formulation.

We consider a linearized incompressible MHD system

$$-\frac{1}{\mathrm{Re}}\Delta \boldsymbol{u} + \nabla p + (\boldsymbol{w} \cdot \nabla)\boldsymbol{u} + \kappa \boldsymbol{d} \times (\nabla \times \boldsymbol{b}) = \boldsymbol{g}, \tag{3.1a}$$

$$\nabla \cdot \boldsymbol{u} = 0, \tag{3.1b}$$

$$\frac{\kappa}{\mathrm{Rm}} \nabla \times (\nabla \times \boldsymbol{b}) + \nabla r - \kappa \nabla \times (\boldsymbol{u} \times \boldsymbol{d}) = \boldsymbol{f}, \tag{3.1c}$$

$$\nabla \cdot \boldsymbol{b} = 0, \tag{3.1d}$$

where $\boldsymbol{u}$ is velocity of the fluid (plasma or liquid metal), $\boldsymbol{b}$ is the magnetic field, p is the fluid pressure, and r is a scalar potential. The following are constant parameters: a fluid Reynolds number $\mathrm{Re} > 0$; a magnetic Reynolds number $\mathrm{Rm} > 0$; and a coupling parameter $\kappa = \mathrm{Ha}^2/(\mathrm{Re}\mathrm{Rm})$, with the Hartmann number $\mathrm{Ha} > 0$. Here, $\boldsymbol{d}$ is a prescribed magnetic field and $\boldsymbol{w}$ is a prescribed velocity field. From this point forward, we assume (see, e.g., [5, 19] for similar assumptions) that $\boldsymbol{d} \in \left[W^{1,\infty}(\Omega)\right]^d$, $\boldsymbol{w} \in \left[W^{1,\infty}(\Omega_h)\right]^d \cap H(\mathrm{div}, \Omega)$, and $\nabla \cdot \boldsymbol{w} = 0$.

By introducing auxiliary variables $\boldsymbol{L}$ and $\boldsymbol{J}$, we cast (3.1) into a system

$$\mathrm{Re}\boldsymbol{L} - \nabla \boldsymbol{u} = \boldsymbol{0}, \tag{3.2a}$$

$$-\nabla \cdot \boldsymbol{L} + \nabla p + (\boldsymbol{w} \cdot \nabla)\boldsymbol{u} + \kappa \boldsymbol{d} \times (\nabla \times \boldsymbol{b}) = \boldsymbol{g}, \tag{3.2b}$$

$$\nabla \cdot \boldsymbol{u} = 0, \tag{3.2c}$$

$$\frac{\mathrm{Rm}}{\kappa}\boldsymbol{J} - \nabla \times \boldsymbol{b} = \boldsymbol{0}, \tag{3.2d}$$

$$\nabla \times \boldsymbol{J} + \nabla r - \kappa \nabla \times (\boldsymbol{u} \times \boldsymbol{d}) = \boldsymbol{f}, \tag{3.2e}$$

$$\nabla \cdot \boldsymbol{b} = 0, \tag{3.2f}$$

with (Dirichlet) boundary conditions

$$\boldsymbol{u} = \boldsymbol{u}_D, \quad \boldsymbol{b}^t := \boldsymbol{h}_D, \quad r = 0 \quad \text{on } \partial\Omega, \tag{3.3}$$

where $\boldsymbol{a}^t := -\boldsymbol{n} \times (\boldsymbol{n} \times \boldsymbol{a})$. In addition, we require the compatibility condition for $\boldsymbol{u}_D$ and the mean-value zero condition for p:

$$\langle \boldsymbol{u}_D \cdot \boldsymbol{n}, 1 \rangle_{\partial\Omega} = 0, \qquad (p, 1)_\Omega = 0. \tag{3.4}$$

Following the upwind HDG framework in [4] we define the HDG flux as

$$\begin{bmatrix} \hat{\boldsymbol{F}}^1 \cdot \boldsymbol{n} \\ \hat{\boldsymbol{F}}^2 \cdot \boldsymbol{n} \\ \hat{\boldsymbol{F}}^3 \cdot \boldsymbol{n} \\ \hat{\boldsymbol{F}}^4 \cdot \boldsymbol{n} \\ \hat{\boldsymbol{F}}^5 \cdot \boldsymbol{n} \\ \hat{\boldsymbol{F}}^6 \cdot \boldsymbol{n} \end{bmatrix} = \begin{bmatrix} -\hat{\boldsymbol{u}} \otimes \boldsymbol{n} \\ -\boldsymbol{L}\boldsymbol{n} + m\boldsymbol{u} + p\boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\boldsymbol{b}^t + \hat{\boldsymbol{b}}^t\right)\right) + \alpha_1 (\boldsymbol{u} - \hat{\boldsymbol{u}}) \\ \hat{\boldsymbol{u}} \cdot \boldsymbol{n} \\ -\boldsymbol{n} \times \hat{\boldsymbol{b}}^t \\ \boldsymbol{n} \times \boldsymbol{J} + \hat{r}\boldsymbol{n} - \frac{1}{2}\kappa \boldsymbol{n} \times ((\boldsymbol{u} + \hat{\boldsymbol{u}}) \times \boldsymbol{d}) + \alpha_2 \left(\boldsymbol{b}^t - \hat{\boldsymbol{b}}^t\right) \\ \boldsymbol{b} \cdot \boldsymbol{n} + \alpha_3 (r - \hat{r}) \end{bmatrix}, \tag{3.5}$$

where $\hat{\boldsymbol{b}}^t$, $\hat{\boldsymbol{u}}$, and $\hat{r}$ are the restrictions (or trace) of $\boldsymbol{b}^t$, $\boldsymbol{u}$, r on $\mathcal{E}_h$. These $\hat{\boldsymbol{b}}^t$, $\hat{\boldsymbol{u}}$, $\hat{r}$ will be regarded as unknowns in discretizations to obtain a hybridized DG method. Here, $m := \boldsymbol{w} \cdot \boldsymbol{n}$, and α_1, α_2, and α_3 are constant parameters. It will be shown that the conditions $\alpha_1 > \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty}$, $\alpha_2 > 0$, and $\alpha_3 > 0$ are sufficient for the well-posedness of our HDG formulation. Note that for simplicity all six components of the HDG flux, $\hat{\boldsymbol{F}}$, are denoted in the same fashion (by a bold italic symbol). However, it is clear from (3.2) that $\hat{\boldsymbol{F}}^1$ is a third order tensor, $\hat{\boldsymbol{F}}^2$ is a second order tensor, $\hat{\boldsymbol{F}}^3$ is a vector, etc., and that the normal HDG flux components, $\hat{\boldsymbol{F}}^i \cdot \boldsymbol{n}$ in (3.5), are tensors of one order lower.

For discretization we introduce the discontinuous piecewise polynomial spaces

$$\mathbf{G}_h := [\mathcal{P}_k(\Omega_h)]^{d\times d}, \qquad \mathbf{V}_h := [\mathcal{P}_k(\Omega_h)]^d, \qquad \mathbf{Q}_h := \mathcal{P}_k(\Omega_h),$$
$$\mathbf{H}_h := [\mathcal{P}_k(\Omega_h)]^{\tilde{d}}, \qquad \mathbf{C}_h := [\mathcal{P}_k(\Omega_h)]^d, \qquad \mathbf{S}_h := \mathcal{P}_k(\Omega_h), \qquad \mathbf{M}_h := [\mathcal{P}_k(\mathcal{E}_h)]^d,$$
$$\boldsymbol{\Lambda}_h^t := \left\{\boldsymbol{\lambda} \in [\mathcal{P}_k(\mathcal{E}_h)]^d \ : \ \boldsymbol{\lambda} \cdot \boldsymbol{n}_e = 0 \ \forall e \in \mathcal{E}_h\right\}, \qquad \boldsymbol{\Gamma}_h := [\mathcal{P}_k(\mathcal{E}_h)]^d,$$

where $\tilde{d} = 3$ if $d = 3$, and $\tilde{d} = 1$ if $d = 2$.

Let us introduce the following two identities which are useful throughout the paper:

$$(\boldsymbol{u}, \boldsymbol{d} \times (\nabla \times \boldsymbol{b}))_K = (\boldsymbol{b}, \nabla \times (\boldsymbol{u} \times \boldsymbol{d}))_K + \langle \boldsymbol{d} \times (\boldsymbol{n} \times \boldsymbol{b}), \boldsymbol{u} \rangle_{\partial K}, \tag{3.6a}$$

$$[\boldsymbol{d} \times (\boldsymbol{n} \times \boldsymbol{b})] \cdot \boldsymbol{u} = -[\boldsymbol{n} \times (\boldsymbol{u} \times \boldsymbol{d})] \cdot \boldsymbol{b}. \tag{3.6b}$$

These identities follow from integration by parts and vector product identities.

Next, we multiply (3.2a)–(3.2f) by test functions $(\boldsymbol{G}, \boldsymbol{v}, q, \boldsymbol{H}, \boldsymbol{c}, s)$, integrate by parts all terms, and introduce the HDG flux (3.5) in the boundary terms. This results in a local discrete weak formulation,

$$\operatorname{Re}(\boldsymbol{L}_h, \boldsymbol{G})_K + (\boldsymbol{u}_h, \nabla\cdot\boldsymbol{G})_K + \left\langle \hat{\boldsymbol{F}}_h^1\cdot\boldsymbol{n}, \boldsymbol{G}\right\rangle_{\partial K} = 0, \tag{3.7a}$$

$$\begin{aligned}(\boldsymbol{L}_h, \nabla\boldsymbol{v})_K - (p_h, \nabla\cdot\boldsymbol{v})_K - (\boldsymbol{u}_h\otimes\boldsymbol{w}, \nabla\boldsymbol{v})_K \\ + \kappa(\boldsymbol{b}_h, \nabla\times(\boldsymbol{v}\times\boldsymbol{d}))_K + \left\langle \hat{\boldsymbol{F}}_h^2\cdot\boldsymbol{n}, \boldsymbol{v}\right\rangle_{\partial K} = (\boldsymbol{g}, \boldsymbol{v})_K,\end{aligned} \tag{3.7b}$$

$$-(\boldsymbol{u}_h, \nabla q)_K + \left\langle \hat{\boldsymbol{F}}_h^3\cdot\boldsymbol{n}, q\right\rangle_{\partial K} = 0, \tag{3.7c}$$

$$\frac{\mathrm{Rm}}{\kappa}(\boldsymbol{J}_h, \boldsymbol{H})_K - (\boldsymbol{b}_h, \nabla\times\boldsymbol{H})_K + \left\langle \hat{\boldsymbol{F}}_h^4\cdot\boldsymbol{n}, \boldsymbol{H}\right\rangle_{\partial K} = 0, \tag{3.7d}$$

$$\begin{aligned}(\boldsymbol{J}_h, \nabla\times\boldsymbol{c})_K - (r_h, \nabla\cdot\boldsymbol{c})_K - \kappa(\boldsymbol{u}_h, \boldsymbol{d}\times(\nabla\times\boldsymbol{c}))_K \\ + \left\langle \hat{\boldsymbol{F}}_h^5\cdot\boldsymbol{n}, \boldsymbol{c}\right\rangle_{\partial K} = (\boldsymbol{f}, \boldsymbol{c})_K,\end{aligned} \tag{3.7e}$$

$$-(\boldsymbol{b}_h, \nabla s)_K + \left\langle \hat{\boldsymbol{F}}_h^6\cdot\boldsymbol{n}, s\right\rangle_{\partial K} = 0, \tag{3.7f}$$

for all $(\boldsymbol{G}, \boldsymbol{v}, q, \boldsymbol{H}, \boldsymbol{c}, s) \in \mathbf{G}_h(K)\times\mathbf{V}_h(K)\times\mathbf{Q}_h(K)\times\mathbf{H}_h(K)\times\mathbf{C}_h(K)\times\mathbf{S}_h(K)$ and for all $K\in\Omega_h$, where $\boldsymbol{u}_h$, $\boldsymbol{L}_h, \dots$, are the discrete counterparts of $\boldsymbol{u}$, $\boldsymbol{L}, \dots$, and $\hat{\boldsymbol{F}}_h^i$ is the discrete counterpart of $\hat{\boldsymbol{F}}^i$ in (3.5) by replacing the unknowns $\boldsymbol{u}$, $\boldsymbol{L}, \dots$, with their discrete counterparts.

Since $\hat{\boldsymbol{b}}_h^t$, $\hat{\boldsymbol{u}}_h$, and $\hat{r}_h$ are (trace) unknowns, we need to equip extra equations to make the system (3.7) well-posed. To that end, we observe that an element K communicates with its neighbors only through the trace unknowns. For the HDG method to be conservative, we weakly enforce the continuity of the HDG flux (3.5) across each interior edge. Since $\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t$, and $\hat{r}_h$ are single-valued on $\mathcal{E}_h$, we have automatically that $[\![\hat{\boldsymbol{F}}_h^1\cdot\boldsymbol{n}]\!] = \boldsymbol{0}$, $[\![\hat{\boldsymbol{F}}_h^3\cdot\boldsymbol{n}]\!] = 0$, and $[\![\hat{\boldsymbol{F}}_h^4\cdot\boldsymbol{n}]\!] = \boldsymbol{0}$. The conservation constraints to be enforced reduce to

$$\left\langle [\![\hat{\boldsymbol{F}}_h^2\cdot\boldsymbol{n}]\!], \boldsymbol{\mu}\right\rangle_e = 0, \qquad \left\langle [\![\hat{\boldsymbol{F}}_h^5\cdot\boldsymbol{n}]\!], \boldsymbol{\lambda}^t\right\rangle_e = 0, \qquad \left\langle [\![\hat{\boldsymbol{F}}_h^6\cdot\boldsymbol{n}]\!], \gamma\right\rangle_e = 0 \tag{3.8}$$

for all $(\boldsymbol{\mu}, \boldsymbol{\lambda}^t, \gamma)\in\mathbf{M}_h(e)\times\boldsymbol{\Lambda}_h^t(e)\times\boldsymbol{\Gamma}_h(e)$ and for all e in $\mathcal{E}_h^o$. Finally, we enforce the Dirichlet boundary conditions through the trace unknowns,

$$\langle\hat{\boldsymbol{u}}_h, \boldsymbol{\mu}\rangle_e = \langle\boldsymbol{u}_D, \boldsymbol{\mu}\rangle_e, \qquad \left\langle\hat{\boldsymbol{b}}_h^t, \boldsymbol{\lambda}^t\right\rangle_e = \left\langle\boldsymbol{h}_D, \boldsymbol{\lambda}^t\right\rangle_e, \qquad \langle\hat{r}_h, \gamma\rangle_e = 0, \tag{3.9}$$

for all $(\boldsymbol{\mu}, \boldsymbol{\lambda}^t, \gamma)\in\boldsymbol{M}_h(e)\times\boldsymbol{\Lambda}_h^t(e)\times\Gamma_h(e)$ and for all e in $\mathcal{E}_h^\partial$.

In (3.7), (3.8), and (3.9), we seek $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)\in\mathbf{G}_h\times\mathbf{V}_h\times\mathbf{Q}_h\times\mathbf{H}_h\times\mathbf{C}_h\times\mathbf{S}_h$ and $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)\in\mathbf{M}_h\times\boldsymbol{\Lambda}_h^t\times\boldsymbol{\Gamma}_h$. For simplicity, we will not state explicitly that equations hold for all test functions, for all elements, or for all edges.

We will refer to $\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h$, and r_h as the *local variables*, and to equation (3.7) on each element as the *local solver*. This reflects the fact that we can solve for local variables element-by-element as functions of $\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t$, and $\hat{r}_h$. On the other hand, we will refer to $\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t$, and $\hat{r}_h$ as the *global variables*, which are governed by (3.8) and (3.9) on the mesh skeleton. Finally, for the uniqueness of the discrete pressure p_h, we enforce the discrete counterpart of (3.4),

$$(p_h, 1) = 0. \tag{3.10}$$

3.1. Well-posedness of the HDG formulation. In this subsection we discuss well-posedness of (3.7)–(3.10). Proofs with full details can be found in [22].

THEOREM 3.1. *Let Ω be simply connected with one component to $\partial\Omega$. Let $\alpha_1 > \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty(\Omega)}$, $\alpha_2 > 0$, and $\alpha_3 > 0$. The system* (3.7)–(3.10) *is well-posed, in that given $\boldsymbol{f}$, $\boldsymbol{g}$, $\boldsymbol{u}_D$, and $\boldsymbol{h}_D$, there exists a unique solution $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h, \hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)$.*

Proof. System (3.7)–(3.10) has the same number of equations and unknowns, so it is enough to show that $(\boldsymbol{g}, \boldsymbol{f}, \boldsymbol{u}_D, \boldsymbol{h}_D) = \boldsymbol{0}$ implies $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h, \hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h) = \boldsymbol{0}$. To begin, we take $(\boldsymbol{G}, \boldsymbol{v}, q, \boldsymbol{H}, \boldsymbol{c}, s) = (\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)$, integrate by parts the first four terms of (3.7b) and the first term of (3.7e), sum the resulting equations in (3.7), and sum over all elements to arrive at

$$\begin{aligned}&\mathrm{Re}\,\|\boldsymbol{L}_h\|_0^2 + \frac{\mathrm{Rm}}{\kappa}\|\boldsymbol{J}_h\|_0^2 - \langle \hat{\boldsymbol{u}}_h \otimes \boldsymbol{n}, \boldsymbol{L}_h\rangle + \left\langle \frac{m}{2}\boldsymbol{u}_h, \boldsymbol{u}_h\right\rangle + \langle \alpha_1(\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h), \boldsymbol{u}_h\rangle \\ &+ \left\langle \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \hat{\boldsymbol{b}}_h^t\right), \boldsymbol{u}_h\right\rangle + \langle \hat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, p_h\rangle - \left\langle \boldsymbol{n} \times \hat{\boldsymbol{b}}_h^t, \boldsymbol{J}_h\right\rangle + \langle \hat{r}_h \boldsymbol{n}, \boldsymbol{b}_h\rangle \\ &+ \left\langle \alpha_2(\boldsymbol{b}_h^t - \hat{\boldsymbol{b}}_h^t), \boldsymbol{b}_h^t\right\rangle - \left\langle \frac{1}{2}\kappa \boldsymbol{n} \times (\hat{\boldsymbol{u}}_h \times \boldsymbol{d}), \boldsymbol{b}_h\right\rangle + \langle \alpha_3 (r_h - \hat{r}_h), r_h\rangle = 0.\end{aligned} \tag{3.11}$$

Here, we used the following identity obtained from $\nabla \cdot \boldsymbol{w} = 0$ and the integration by parts:

$$-(\boldsymbol{u}_h, \boldsymbol{w} \cdot \nabla \boldsymbol{u}_h)_K = -\frac{1}{2}(\boldsymbol{w}, \nabla(\boldsymbol{u}_h \cdot \boldsymbol{u}_h))_K = -\left\langle \frac{m}{2}\boldsymbol{u}_h, \boldsymbol{u}_h\right\rangle_{\partial K}.$$

Next, we set $(\boldsymbol{\mu}, \boldsymbol{\lambda}^t, \gamma) = (\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)$ and sum (3.8) over all interior edges to obtain

$$\begin{aligned}&\left\langle -\boldsymbol{L}_h \boldsymbol{n} + m\boldsymbol{u}_h + p_h \boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \boldsymbol{b}_h^t\right) + \alpha_1 (\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h), \hat{\boldsymbol{u}}_h\right\rangle_{\partial\Omega_h \backslash \partial\Omega} \\ &+ \left\langle \boldsymbol{n} \times \boldsymbol{J}_h - \frac{1}{2}\kappa \boldsymbol{n} \times (\boldsymbol{u}_h \times \boldsymbol{d}) + \alpha_2 \left(\boldsymbol{b}_h^t - \hat{\boldsymbol{b}}_h^t\right), \hat{\boldsymbol{b}}_h^t\right\rangle_{\partial\Omega_h \backslash \partial\Omega} \\ &+ \langle \boldsymbol{b}_h \cdot \boldsymbol{n} + \alpha_3 (r_h - \hat{r}_h), \hat{r}_h\rangle_{\partial\Omega_h \backslash \partial\Omega} = 0,\end{aligned} \tag{3.12}$$

where we used the continuity of $\boldsymbol{d}$ to eliminate $\langle \boldsymbol{d} \times (\boldsymbol{n} \times \hat{\boldsymbol{b}}_h^t), \hat{\boldsymbol{u}}_h\rangle_{\partial\Omega_h\backslash\partial\Omega}$ and $\langle \boldsymbol{n} \times (\hat{\boldsymbol{u}}_h \times \boldsymbol{d}), \hat{\boldsymbol{b}}_h^t\rangle_{\partial\Omega_h\backslash\partial\Omega}$.

Since $\boldsymbol{u}_D = \boldsymbol{0}$ and $\boldsymbol{h}_D = \boldsymbol{0}$ by assumption, we conclude from the boundary conditions (3.9) that $\hat{\boldsymbol{u}}_h = \boldsymbol{0}$, $\hat{\boldsymbol{b}}_h^t = \boldsymbol{0}$, and $\hat{r}_h = 0$ on $\partial\Omega$. The integrals in (3.12) can then be written over $\partial\Omega_h$ since the contribution on the domain boundary, $\partial\Omega$, is zero. Subtracting (3.12) from (3.11), we arrive at

$$\begin{aligned}&\mathrm{Re}\,\|\boldsymbol{L}_h\|_0^2 + \frac{\mathrm{Rm}}{\kappa}\|\boldsymbol{J}_h\|_0^2 + \langle \alpha_1(\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h), (\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h)\rangle + \left\langle \frac{m}{2}\boldsymbol{u}_h, \boldsymbol{u}_h\right\rangle \\ &- \langle m\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h\rangle + \alpha_2 \left\|\boldsymbol{b}_h^t - \hat{\boldsymbol{b}}_h^t\right\|_{\partial\Omega_h}^2 + \alpha_3 \|r_h - \hat{r}_h\|_{\partial\Omega_h}^2 = 0.\end{aligned} \tag{3.13}$$

Finally, using the facts that $\boldsymbol{w} \in H(\mathrm{div}, \Omega)$ and $\hat{\boldsymbol{u}}_h = \boldsymbol{0}$ on $\partial\Omega$, we can freely add $0 = \left\langle \frac{m}{2}\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{u}}_h\right\rangle$ to rewrite (3.13) as

$$\begin{aligned}&\mathrm{Re}\,\|\boldsymbol{L}_h\|_0^2 + \frac{\mathrm{Rm}}{\kappa}\|\boldsymbol{J}_h\|_0^2 + \left\langle \left(\alpha_1 + \frac{m}{2}\right)(\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h), (\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h)\right\rangle \\ &+ \alpha_2 \left\|\boldsymbol{b}_h^t - \hat{\boldsymbol{b}}_h^t\right\|_{\partial\Omega_h}^2 + \alpha_3 \|r_h - \hat{r}_h\|_{\partial\Omega_h}^2 = 0.\end{aligned} \tag{3.14}$$

Recalling $\alpha_1 > \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty}$ and $\alpha_2, \alpha_3 > 0$, we can conclude that $\boldsymbol{L}_h = \boldsymbol{0}$, $\boldsymbol{J}_h = \boldsymbol{0}$; that $\boldsymbol{u}_h = \hat{\boldsymbol{u}}_h$, $\boldsymbol{b}_h^t = \hat{\boldsymbol{b}}_h^t$, and $r_h = \hat{r}_h$ on $\mathcal{E}_h^o$; and that $\boldsymbol{u}_h = \boldsymbol{b}_h^t = \boldsymbol{0}$ and $r_h = 0$ on $\partial\Omega$. Now, we integrate (3.7a) by parts to obtain $\nabla \boldsymbol{u}_h = \boldsymbol{0}$ in K, which implies that $\boldsymbol{u}_h$ is elementwise constant. The fact that $\boldsymbol{u}_h = \hat{\boldsymbol{u}}_h$ on $\mathcal{E}_h^o$ means $\boldsymbol{u}_h$ is continuous on $\mathcal{E}_h$. Since $\boldsymbol{u}_h = \boldsymbol{0}$ on $\partial\Omega$, we conclude that $\boldsymbol{u}_h = \boldsymbol{0}$, and therefore $\hat{\boldsymbol{u}}_h = \boldsymbol{0}$.

Since $\boldsymbol{b}_h^t = \hat{\boldsymbol{b}}_h^t$ on $\mathcal{E}_h^o$, $\boldsymbol{b}_h^t$ is continuous on Ω. Furthermore, the third conservation constraint in (3.8) implies that $\boldsymbol{b}_h \cdot \boldsymbol{n}$ is continuous on Ω. Integrating both (3.7d) and (3.7f) by parts, we have $\nabla \times \boldsymbol{b}_h = \boldsymbol{0}$ and $\nabla \cdot \boldsymbol{b}_h = 0$ on Ω. When $\boldsymbol{b}_h \in H(\mathrm{div}, \Omega) \cap H(\mathrm{curl}, \Omega)$ and $\boldsymbol{b}_h^t = \boldsymbol{0}$ on $\partial\Omega$, and recalling that Ω is simply connected with one component to the boundary, we see that there is a constant $C > 0$ such that $\|\boldsymbol{b}_h\|_0 \leq C(\|\nabla \cdot \boldsymbol{b}_h\|_0 + \|\nabla \times \boldsymbol{b}_h\|_0)$ [16, Lemma 3.4]. This implies that $\boldsymbol{b}_h = \boldsymbol{0}$, and hence $\hat{\boldsymbol{b}}_h^t = \boldsymbol{0}$.

Considering the vanishing unknowns above, integrating by parts reduces (3.7b) and (3.7e) to $(\nabla p_h, \boldsymbol{v})_K = 0$ and $(\nabla r_h, \boldsymbol{c})_K = 0$, respectively. Thus, p_h and r_h are elementwise constants. Since $r_h = \hat{r}_h$ on $\mathcal{E}_h^o$, r_h is continuous on Ω, and since $r_h = 0$ on $\partial\Omega$, we can conclude that $r_h = 0$, and hence $\hat{r}_h = 0$. Finally, we use the first conservation constraint in (3.8) to conclude that p_h is continuous and hence constant on Ω. Using the zero-average condition (3.10) yields $p_h = 0$. ◻

3.2. Well-posedness of the local solver. A key advantage of HDG methods is their ability to separate the computation of the volume unknowns $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)$ and the trace unknowns $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)$. In our HDG scheme, we first solve (3.7) for local unknowns $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)$ as a function of $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)$ (local solver), then these are substituted into (3.8) on the mesh skeleton to solve for the unknowns $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)$ (global solver). Finally, $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)$ are computed with the local solver using $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h)$, so well-posedness of the local solver is essential.

As in HDG methods for the Stokes equations [24, 9, 4], the local solver is not well-posed unless extra conditions are imposed on the pressure. Here, we introduce the elementwise pressure integral as a global unknown and require their sum to vanish. Toward this goal, we introduce $\mathbf{X}_h := \mathcal{P}_0(\Omega_h)$ and augment (3.7c) to read

$$-(\boldsymbol{u}_h, \nabla q)_K + \langle \hat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, q \rangle_{\partial K} + (p_h, \bar{q})_K = |K|^{-1}(\rho_h, \bar{q})_K, \tag{3.15}$$

with $\rho_h \in \mathbf{X}_h$, $\bar{q}|_K := |K|^{-1}(q, 1)_K$ the average of q in K, and $|K|$ the volume of element K. Next we augment the global solver with

$$\langle \hat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, \xi \rangle_{\partial K} + \sum_K \rho_h|_K = 0 \tag{3.16}$$

for all ξ in $\mathbf{X}_h$ and remove the constraint (3.10), which is satisfied by this construction.

To justify (3.15) and (3.16) we make the following observations. First, summing (3.16) over all elements and using the compatibility condition on $\boldsymbol{u}_D$, (3.4), we conclude that

$$\sum_{K \in \Omega_h} \rho_h|_K = 0 \quad \text{and} \quad \langle \hat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, \xi \rangle_{\partial K} = 0 \qquad \forall K \in \Omega_h. \tag{3.17}$$

Next, setting $q = 1$ on K in (3.15) and using the second condition in (3.17), we have

$$(p_h, 1)_K = \rho_h|_K, \tag{3.18}$$

and therefore (3.7c) holds for each K. In addition, (3.18) and the first condition in (3.17) imply (3.10). Finally, note that the number of the new unknowns ρ_h and the number of equations in (3.16) are the same.

For this modified HDG scheme we claim well-posedness of the local solver.

THEOREM 3.2. *Let* $\alpha_1 > \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty(\Omega)}$, $\alpha_2 > 0$, *and* $\alpha_3 > 0$. *The local solver given by* (3.7) *with* (3.7c) *replaced by* (3.15) *is well-posed. In other words, given* $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h, \boldsymbol{g}, \boldsymbol{f}, \rho_h)$, *there exists a unique solution* $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)$ *of the system.*

Proof. We show that $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h, \boldsymbol{g}, \boldsymbol{f}, \rho_h) = \boldsymbol{0}$ implies $(\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h) = \boldsymbol{0}$. To begin, set $(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{b}}_h^t, \hat{r}_h, \boldsymbol{g}, \boldsymbol{f}, \rho_h) = \boldsymbol{0}$. Then (3.15) reduces to $-(\boldsymbol{u}_h, \nabla q)_K + (p_h, \bar{q})_K = 0$, and taking q to be constant gives $(p_h, \bar{q})_K = 0$, and hence $-(\boldsymbol{u}_h, \nabla q)_K = 0$.

Take $(\boldsymbol{G}, \boldsymbol{v}, q, \boldsymbol{H}, \boldsymbol{c}, s) = (\boldsymbol{L}_h, \boldsymbol{u}_h, p_h, \boldsymbol{J}_h, \boldsymbol{b}_h, r_h)$, integrate by parts the first four terms in (3.7b) and the first term in (3.7e), and sum the resulting equations to get

$$\begin{aligned}(3.19)\qquad \mathrm{Re}\,\|\boldsymbol{L}_h\|_{0,K}^2 &+ \left\langle \left(\alpha_1 + \frac{m}{2}\right)\boldsymbol{u}_h, \boldsymbol{u}_h \right\rangle_{\partial K} \\ &+ \frac{\mathrm{Rm}}{\kappa}\|\boldsymbol{J}_h\|_{0,K}^2 + \alpha_2 \left\|\boldsymbol{b}_h^t\right\|_{0,\partial K}^2 + \alpha_3\|r_h\|_{0,\partial K}^2 = 0.\end{aligned}$$

Recalling $\alpha_1 > \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty}$, $\alpha_2 > 0$, and $\alpha_3 > 0$, we obtain

$$\boldsymbol{L}_h = \boldsymbol{0}, \quad \boldsymbol{J}_h = \boldsymbol{0} \quad \text{in } K, \qquad \boldsymbol{u}_h = \boldsymbol{0}, \quad \boldsymbol{b}_h^t = \boldsymbol{0}, \quad r_h = 0 \quad \text{on } \partial K.$$

Using an argument similar to that in subsection 3.1, we can conclude that $\boldsymbol{u}_h = \boldsymbol{b}_h = \boldsymbol{0}$ in K. From (3.7b) and (3.7e), $(\nabla p_h, \boldsymbol{v})_K = 0$ and $(\nabla r_h, \boldsymbol{c})_K = 0$, respectively. Thus, p_h and r_h must be constant, and since $r_h = 0$ on ∂K, r_h is identically zero in K. Now since $(p_h, \bar{q})_K = 0$, we have $p_h = 0$ in K. ∎

Remark 3.3. Introducing ρ_h and (3.15) and (3.16) does not alter the solution of the original HDG scheme. We refer the reader to [22] for details.

4. Error analysis. For an unknown σ we use ε_σ to denote the error between the exact solution σ and its numerical solution σ_h. For example, $\varepsilon_{\boldsymbol{L}} := \boldsymbol{L} - \boldsymbol{L}_h$ and $\varepsilon_{\hat{\boldsymbol{u}}} := \hat{\boldsymbol{u}} - \hat{\boldsymbol{u}}_h$, where $\hat{\boldsymbol{u}}$ is the trace of the exact solution $\boldsymbol{u}$ on the mesh skeleton. We use $\Pi\sigma$ to denote some interpolation (which will be defined later) of the unknown σ into its associated finite element space and decompose ε_σ into $\varepsilon_\sigma^I + \varepsilon_\sigma^h$, where

$$(4.1)\qquad \varepsilon_\sigma^I := \sigma - \Pi\sigma \quad \text{and} \quad \varepsilon_\sigma^h := \Pi\sigma - \sigma_h.$$

We will call the ε^I and ε^h error terms interpolation and approximation errors, respectively. We define a collective projection $\boldsymbol{\Pi}(\boldsymbol{L}, \boldsymbol{u}, p, \boldsymbol{J}, \boldsymbol{b}, r, \hat{\boldsymbol{u}}, \hat{\boldsymbol{b}}^t, \hat{r})$ in Appendix B, and each component of $\boldsymbol{\Pi}$ may depend on other unknowns; i.e., the $\boldsymbol{L}$-component of $\boldsymbol{\Pi}(\boldsymbol{L}, \boldsymbol{u}, p)$ also depends on $\boldsymbol{u}$ and p. Nonetheless, for simplicity of presentation we use $\Pi\boldsymbol{L}$ to denote the $\boldsymbol{L}$-component of $\boldsymbol{\Pi}$ for example. The properties of $\boldsymbol{\Pi}$ are summarized in Appendix B.

LEMMA 4.1. *Assume that the exact solution* $(\boldsymbol{L}, \boldsymbol{u}, p, \boldsymbol{J}, \boldsymbol{b}, r)$ *of* (3.2)–(3.3) *is sufficiently regular. Then the exact solution satisfies* (3.7)–(3.9).

Proof. The assertion follows from the sufficient regularity assumption of the exact solution and single-valuedness of $\boldsymbol{w}$ and $\boldsymbol{d}$. See [22] for details. ∎

LEMMA 4.2 (error equation). *The approximation errors satisfy*

$$E_h^2 := \mathrm{Re}\left\|\varepsilon_{\boldsymbol{L}}^h\right\|_0^2 + \frac{\mathrm{Rm}}{\kappa}\left\|\varepsilon_{\boldsymbol{J}}^h\right\|_0^2 + \left\langle\left(\alpha_1 + \frac{m}{2}\right)(\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h), (\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h)\right\rangle \tag{4.2}$$
$$+ \alpha_2\left\|\varepsilon_{\boldsymbol{b}^t}^h - \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right\|_{\partial\Omega_h}^2 + \alpha_3\left\|\varepsilon_r^h - \varepsilon_{\hat{r}}^h\right\|_{\partial\Omega_h}^2$$
$$= -\mathrm{Re}\left(\varepsilon_{\boldsymbol{L}}^I, \varepsilon_{\boldsymbol{L}}^h\right) - \kappa\left(\varepsilon_{\boldsymbol{b}}^I, \nabla\times\left(\varepsilon_{\boldsymbol{u}}^h\times\boldsymbol{d}\right)\right) + \kappa\left(\varepsilon_{\boldsymbol{u}}^I, \boldsymbol{d}\times\left(\nabla\times\varepsilon_{\boldsymbol{b}}^h\right)\right)$$
$$- \left\langle \boldsymbol{n}\times\varepsilon_{\boldsymbol{J}}^I - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon_{\boldsymbol{u}}^I + \varepsilon_{\hat{\boldsymbol{u}}}^I\right)\times\boldsymbol{d}\right) + \alpha_2\varepsilon_{\boldsymbol{b}^t}^I, \varepsilon_{\boldsymbol{b}^t}^h - \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right\rangle.$$

Proof. Since the numerical and exact solutions satisfy (3.7) (Lemma 4.1), the linearity of the operators leads to the following error equations:

$$\mathrm{Re}\left(\varepsilon_{\boldsymbol{L}}, \boldsymbol{G}\right) + \left(\varepsilon_{\boldsymbol{u}}, \nabla\cdot\boldsymbol{G}\right) - \langle\varepsilon_{\hat{\boldsymbol{u}}}\otimes\boldsymbol{n}, \boldsymbol{G}\rangle = 0, \tag{4.3a}$$
$$\left(\varepsilon_{\boldsymbol{L}}, \nabla\boldsymbol{v}\right) - \left(\varepsilon_p, \nabla\cdot\boldsymbol{v}\right) - \left(\varepsilon_{\boldsymbol{u}}\otimes\boldsymbol{w}, \nabla\boldsymbol{v}\right) + \kappa\left(\varepsilon_{\boldsymbol{b}}, \nabla\times(\boldsymbol{v}\times\boldsymbol{d})\right) \tag{4.3b}$$
$$+ \left\langle -\varepsilon_{\boldsymbol{L}}\boldsymbol{n} + m\varepsilon_{\boldsymbol{u}} + \varepsilon_p\boldsymbol{n} + \frac{1}{2}\kappa\boldsymbol{d}\times\left(\boldsymbol{n}\times\left(\varepsilon_{\boldsymbol{b}^t} + \varepsilon_{\hat{\boldsymbol{b}}^t}\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}} - \varepsilon_{\hat{\boldsymbol{u}}}\right), \boldsymbol{v}\right\rangle = 0,$$
$$-\left(\varepsilon_{\boldsymbol{u}}, \nabla q\right) + \langle\varepsilon_{\hat{\boldsymbol{u}}}\cdot\boldsymbol{n}, q\rangle = 0, \tag{4.3c}$$
$$\frac{\mathrm{Rm}}{\kappa}\left(\varepsilon_{\boldsymbol{J}}, \boldsymbol{H}\right) - \left(\varepsilon_{\boldsymbol{b}}, \nabla\times\boldsymbol{H}\right) - \left\langle\boldsymbol{n}\times\varepsilon_{\hat{\boldsymbol{b}}^t}, \boldsymbol{H}\right\rangle = 0, \tag{4.3d}$$
$$\left(\varepsilon_{\boldsymbol{J}}, \nabla\times\boldsymbol{c}\right) - \left(\varepsilon_r, \nabla\cdot\boldsymbol{c}\right) - \kappa\left(\varepsilon_{\boldsymbol{u}}, \boldsymbol{d}\times(\nabla\times\boldsymbol{c})\right) \tag{4.3e}$$
$$+ \left\langle \boldsymbol{n}\times\varepsilon_{\boldsymbol{J}} + \varepsilon_{\hat{r}}\boldsymbol{n} - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon_{\boldsymbol{u}} + \varepsilon_{\hat{\boldsymbol{u}}}\right)\times\boldsymbol{d}\right) + \alpha_2\left(\varepsilon_{\boldsymbol{b}^t} - \varepsilon_{\hat{\boldsymbol{b}}^t}\right), \boldsymbol{c}\right\rangle = 0,$$
$$-\left(\varepsilon_{\boldsymbol{b}}, \nabla s\right) + \langle\varepsilon_{\boldsymbol{b}}\cdot\boldsymbol{n} + \alpha_3\left(\varepsilon_r - \varepsilon_{\hat{r}}\right), s\rangle = 0. \tag{4.3f}$$

Next, we split the error terms into their interpolation and approximation components as in (4.1) using the projections Π defined in Appendix B. Due to the cancellation properties of Π in Appendix B, we obtain the following reduced error equations (see [22] for details):

$$\mathrm{Re}\left(\varepsilon_{\boldsymbol{L}}^h, \boldsymbol{G}\right) + \left(\varepsilon_{\boldsymbol{u}}^h, \nabla\cdot\boldsymbol{G}\right) - \left\langle\varepsilon_{\hat{\boldsymbol{u}}}^h\otimes\boldsymbol{n}, \boldsymbol{G}\right\rangle = -\mathrm{Re}\left(\varepsilon_{\boldsymbol{L}}^I, \boldsymbol{G}\right), \tag{4.4a}$$
$$\left(\varepsilon_{\boldsymbol{L}}^h, \nabla\boldsymbol{v}\right) - \left(\varepsilon_p^h, \nabla\cdot\boldsymbol{v}\right) - \left(\varepsilon_{\boldsymbol{u}}^h\otimes\boldsymbol{w}, \nabla\boldsymbol{v}\right) + \kappa\left(\varepsilon_{\boldsymbol{b}}^h, \nabla\times(\boldsymbol{v}\times\boldsymbol{d})\right) \tag{4.4b}$$
$$+ \left\langle -\varepsilon_{\boldsymbol{L}}^h\boldsymbol{n} + m\varepsilon_{\boldsymbol{u}}^h + \varepsilon_p^h\boldsymbol{n} + \frac{1}{2}\kappa\boldsymbol{d}\times\left(\boldsymbol{n}\times\left(\varepsilon_{\boldsymbol{b}^t}^h + \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h\right), \boldsymbol{v}\right\rangle$$
$$= -\kappa\left(\varepsilon_{\boldsymbol{b}}^I, \nabla\times(\boldsymbol{v}\times\boldsymbol{d})\right),$$
$$-\left(\varepsilon_{\boldsymbol{u}}^h, \nabla q\right) + \left\langle\varepsilon_{\hat{\boldsymbol{u}}}^h\cdot\boldsymbol{n}, q\right\rangle = 0, \tag{4.4c}$$
$$\frac{\mathrm{Rm}}{\kappa}\left(\varepsilon_{\boldsymbol{J}}^h, \boldsymbol{H}\right) - \left(\varepsilon_{\boldsymbol{b}}^h, \nabla\times\boldsymbol{H}\right) - \left\langle\boldsymbol{n}\times\varepsilon_{\hat{\boldsymbol{b}}^t}^h, \boldsymbol{H}\right\rangle = 0, \tag{4.4d}$$
$$\left(\varepsilon_{\boldsymbol{J}}^h, \nabla\times\boldsymbol{c}\right) - \left(\varepsilon_r^h, \nabla\cdot\boldsymbol{c}\right) - \kappa\left(\varepsilon_{\boldsymbol{u}}^h, \boldsymbol{d}\times(\nabla\times\boldsymbol{c})\right) \tag{4.4e}$$
$$+ \left\langle \boldsymbol{n}\times\varepsilon_{\boldsymbol{J}}^h + \varepsilon_{\hat{r}}^h\boldsymbol{n} - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon_{\boldsymbol{u}}^h + \varepsilon_{\hat{\boldsymbol{u}}}^h\right)\times\boldsymbol{d}\right) + \alpha_2\left(\varepsilon_{\boldsymbol{b}^t}^h - \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right), \boldsymbol{c}\right\rangle$$
$$= \kappa\left(\varepsilon_{\boldsymbol{u}}^I, \boldsymbol{d}\times(\nabla\times\boldsymbol{c})\right) - \left\langle \boldsymbol{n}\times\varepsilon_{\boldsymbol{J}}^I - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon_{\boldsymbol{u}}^I + \varepsilon_{\hat{\boldsymbol{u}}}^I\right)\times\boldsymbol{d}\right) + \alpha_2\varepsilon_{\boldsymbol{b}^t}^I, \boldsymbol{c}\right\rangle,$$
$$-\left(\varepsilon_{\boldsymbol{b}}^h, \nabla s\right) + \left\langle\varepsilon_{\boldsymbol{b}}^h\cdot\boldsymbol{n} + \alpha_3\left(\varepsilon_r^h - \varepsilon_{\hat{r}}^h\right), s\right\rangle = 0. \tag{4.4f}$$

Notice that (4.4) looks like (3.7) but with the approximation error replacing the finite element solution, and with some nonzero right-hand side terms. Since the approximation errors are in the finite element spaces, we can choose the test functions to be the approximation error terms. Similarly to the procedure for arriving at (3.11), we take $(\boldsymbol{G}, \boldsymbol{v}, q, \boldsymbol{H}, \boldsymbol{c}, s) = (\varepsilon^h_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{u}}, \varepsilon^h_p, \varepsilon^h_{\boldsymbol{J}}, \varepsilon^h_{\boldsymbol{b}}, \varepsilon^h_r)$, integrate by parts the first four terms of (4.4b) and the first term of (4.4e), and sum the resulting equations in (4.4) to arrive at

$$
\begin{aligned}
&\mathrm{Re}\left\|\varepsilon^h_{\boldsymbol{L}}\right\|_0^2 + \frac{\mathrm{Rm}}{\kappa}\left\|\varepsilon^h_{\boldsymbol{J}}\right\|_0^2 - \left\langle \varepsilon^h_{\hat{\boldsymbol{u}}} \otimes \boldsymbol{n}, \varepsilon^h_{\boldsymbol{L}}\right\rangle + \left\langle \frac{m}{2}\varepsilon^h_{\boldsymbol{u}}, \varepsilon^h_{\boldsymbol{u}}\right\rangle + \left\langle \alpha_1(\varepsilon^h_{\boldsymbol{u}} - \varepsilon^h_{\hat{\boldsymbol{u}}}), \varepsilon^h_{\boldsymbol{u}}\right\rangle \\
&+ \left\langle \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right), \varepsilon^h_{\boldsymbol{u}}\right\rangle + \left\langle \varepsilon^h_{\hat{\boldsymbol{u}}} \cdot \boldsymbol{n}, \varepsilon^h_p\right\rangle - \left\langle \boldsymbol{n} \times \varepsilon^h_{\hat{\boldsymbol{b}}^t}, \varepsilon^h_{\boldsymbol{J}}\right\rangle + \left\langle \varepsilon^h_{\hat{r}}\boldsymbol{n}, \varepsilon^h_{\boldsymbol{b}}\right\rangle \\
&+ \left\langle \alpha_2(\varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}), \varepsilon^h_{\boldsymbol{b}^t}\right\rangle - \left\langle \frac{1}{2}\kappa \boldsymbol{n} \times \left(\varepsilon^h_{\hat{\boldsymbol{u}}} \times \boldsymbol{d}\right), \varepsilon^h_{\boldsymbol{b}}\right\rangle + \left\langle \alpha_3\left(\varepsilon^h_r - \varepsilon^h_{\hat{r}}\right), \varepsilon^h_r\right\rangle \\
&= -\mathrm{Re}\left(\varepsilon^I_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{L}}\right) - \kappa\left(\varepsilon^I_{\boldsymbol{b}}, \nabla \times \left(\varepsilon^h_{\boldsymbol{u}} \times \boldsymbol{d}\right)\right) + \kappa\left(\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{d} \times \left(\nabla \times \varepsilon^h_{\boldsymbol{b}}\right)\right) \\
&\quad - \left\langle \boldsymbol{n} \times \varepsilon^I_{\boldsymbol{J}} - \frac{1}{2}\kappa \boldsymbol{n} \times \left(\left(\varepsilon^I_{\boldsymbol{u}} + \varepsilon^I_{\hat{\boldsymbol{u}}}\right) \times \boldsymbol{d}\right) + \alpha_2 \varepsilon^I_{\boldsymbol{b}^t}, \varepsilon^h_{\boldsymbol{b}}\right\rangle .
\end{aligned} \tag{4.5}
$$

For the boundary conditions and conservation conditions, since the exact solution satisfies (3.8)–(3.9), we have

$$
\left\langle -\varepsilon_{\boldsymbol{L}}\boldsymbol{n} + m\varepsilon_{\boldsymbol{u}} + \varepsilon_p \boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\varepsilon_{\boldsymbol{b}^t} + \varepsilon_{\hat{\boldsymbol{b}}^t}\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}} - \varepsilon_{\hat{\boldsymbol{u}}}\right), \boldsymbol{\mu}\right\rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.6a}
$$

$$
\left\langle \boldsymbol{n} \times \varepsilon_{\boldsymbol{J}} - \frac{1}{2}\kappa \boldsymbol{n} \times \left(\left(\varepsilon_{\boldsymbol{u}} + \varepsilon_{\hat{\boldsymbol{u}}}\right) \times \boldsymbol{d}\right) + \alpha_2\left(\varepsilon_{\boldsymbol{b}^t} - \varepsilon_{\hat{\boldsymbol{b}}^t}\right), \boldsymbol{\lambda}^t\right\rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.6b}
$$

$$
\left\langle \varepsilon_{\boldsymbol{b}} \cdot \boldsymbol{n} + \alpha_3\left(\varepsilon_r - \varepsilon_{\hat{r}}\right), \gamma\right\rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.6c}
$$

$$
\left\langle \varepsilon_{\hat{\boldsymbol{u}}}, \boldsymbol{\mu}\right\rangle_{\partial\Omega} = 0, \tag{4.6d}
$$

$$
\left\langle \varepsilon_{\hat{\boldsymbol{b}}^t}, \boldsymbol{\lambda}^t\right\rangle_{\partial\Omega} = 0, \tag{4.6e}
$$

$$
\left\langle \varepsilon_{\hat{r}}, \gamma\right\rangle_{\partial\Omega} = 0. \tag{4.6f}
$$

We split the errors into interpolation and approximation errors as before, and use the interpolations defined in Appendix B to cancel terms. We refer the reader to [22] for more details on cancellation of terms. Then we have

$$
\left\langle -\varepsilon^h_{\boldsymbol{L}}\boldsymbol{n} + m\varepsilon^h_{\boldsymbol{u}} + \varepsilon^h_p \boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \varepsilon^h_{\boldsymbol{b}^t}\right) + \alpha_1\left(\varepsilon^h_{\boldsymbol{u}} - \varepsilon^h_{\hat{\boldsymbol{u}}}\right), \boldsymbol{\mu}\right\rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.7a}
$$

$$
\begin{aligned}
&\left\langle \boldsymbol{n} \times \varepsilon^h_{\boldsymbol{J}} - \frac{1}{2}\kappa \boldsymbol{n} \times \left(\varepsilon^h_{\boldsymbol{u}} \times \boldsymbol{d}\right) + \alpha_2\left(\varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right), \boldsymbol{\lambda}^t\right\rangle_{\partial\Omega_h \setminus \partial\Omega} \\
&= -\left\langle \boldsymbol{n} \times \varepsilon^I_{\boldsymbol{J}} - \frac{1}{2}\kappa \boldsymbol{n} \times \left(\left(\varepsilon^I_{\boldsymbol{u}} + \varepsilon^I_{\hat{\boldsymbol{u}}}\right) \times \boldsymbol{d}\right) + \alpha_2 \varepsilon^I_{\boldsymbol{b}^t}, \boldsymbol{\lambda}^t\right\rangle_{\partial\Omega_h \setminus \partial\Omega},
\end{aligned} \tag{4.7b}
$$

$$
\left\langle \varepsilon^h_{\boldsymbol{b}} \cdot \boldsymbol{n} + \alpha_3\left(\varepsilon^h_r - \varepsilon^h_{\hat{r}}\right), \gamma\right\rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.7c}
$$

$$
\left\langle \varepsilon^h_{\hat{\boldsymbol{u}}}, \boldsymbol{\mu}\right\rangle_{\partial\Omega} = 0, \tag{4.7d}
$$

$$
\left\langle \varepsilon^h_{\hat{\boldsymbol{b}}^t}, \boldsymbol{\lambda}^t\right\rangle_{\partial\Omega} = 0, \tag{4.7e}
$$

$$
\left\langle \varepsilon^h_{\hat{r}}, \gamma\right\rangle_{\partial\Omega} = 0. \tag{4.7f}
$$

Equations (4.7d)–(4.7f) imply that on $\partial\Omega$, $\varepsilon^h_{\hat{\boldsymbol{u}}} = \boldsymbol{0}$, $\varepsilon^h_{\hat{\boldsymbol{b}}^t} = \boldsymbol{0}$, and $\varepsilon^h_{\hat{r}} = 0$. With this zero contribution on $\partial\Omega$, summing of the formulae (4.7a)–(4.7c) with $(\boldsymbol{\mu}, \boldsymbol{\lambda}^t, \gamma) =$

$(\varepsilon^h_{\hat{\boldsymbol{u}}}, \varepsilon^h_{\hat{\boldsymbol{b}}^t}, \varepsilon^h_{\hat{r}})$, including $\partial\Omega$, gives

$$\begin{aligned}&\left\langle -\varepsilon^h_{\boldsymbol{L}}\boldsymbol{n} + m\varepsilon^h_{\boldsymbol{u}} + \varepsilon^h_p\boldsymbol{n} + \frac{1}{2}\kappa\boldsymbol{d}\times\left(\boldsymbol{n}\times\varepsilon^h_{\boldsymbol{b}^t}\right) + \alpha_1\left(\varepsilon^h_{\boldsymbol{u}} - \varepsilon^h_{\hat{\boldsymbol{u}}}\right), \varepsilon^h_{\hat{\boldsymbol{u}}}\right\rangle \\ &\quad + \left\langle \boldsymbol{n}\times\varepsilon^h_{\boldsymbol{J}} - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\varepsilon^h_{\boldsymbol{u}}\times\boldsymbol{d}\right) + \alpha_2\left(\varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right), \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right\rangle \\ &\quad + \left\langle \varepsilon^h_{\boldsymbol{b}}\cdot\boldsymbol{n} + \alpha_3\left(\varepsilon^h_r - \varepsilon^h_{\hat{r}}\right), \varepsilon^h_{\hat{r}}\right\rangle \\ &= \left\langle \boldsymbol{n}\times\varepsilon^I_{\boldsymbol{J}} - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon^I_{\boldsymbol{u}} + \varepsilon^I_{\hat{\boldsymbol{u}}}\right)\times\boldsymbol{d}\right) + \alpha_2\varepsilon^I_{\boldsymbol{b}^t}, \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right\rangle. \end{aligned} \tag{4.8}$$

Subtracting (4.8) from (4.5), we arrive at

$$\begin{aligned}&\mathrm{Re}\left\|\varepsilon^h_{\boldsymbol{L}}\right\|_0^2 + \frac{\mathrm{Rm}}{\kappa}\left\|\varepsilon^h_{\boldsymbol{J}}\right\|_0^2 + \left\langle\alpha_1(\varepsilon^h_{\boldsymbol{u}} - \varepsilon^h_{\hat{\boldsymbol{u}}}), (\varepsilon^h_{\boldsymbol{u}} - \varepsilon^h_{\hat{\boldsymbol{u}}})\right\rangle + \left\langle\frac{m}{2}\varepsilon^h_{\boldsymbol{u}}, \varepsilon^h_{\boldsymbol{u}}\right\rangle \\ &\quad - \left\langle m\varepsilon^h_{\boldsymbol{u}}, \varepsilon^h_{\hat{\boldsymbol{u}}}\right\rangle + \alpha_2\left\|\varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right\|^2_{\partial\Omega_h} + \alpha_3\left\|\varepsilon^h_r - \varepsilon^h_{\hat{r}}\right\|^2_{\partial\Omega_h} \\ &\quad = -\mathrm{Re}\left(\varepsilon^I_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{L}}\right) - \kappa\left(\varepsilon^I_{\boldsymbol{b}}, \nabla\times\left(\varepsilon^h_{\boldsymbol{u}}\times\boldsymbol{d}\right)\right) + \kappa\left(\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{d}\times\left(\nabla\times\varepsilon^h_{\boldsymbol{b}}\right)\right) \\ &\quad\quad - \left\langle \boldsymbol{n}\times\varepsilon^I_{\boldsymbol{J}} - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon^I_{\boldsymbol{u}} + \varepsilon^I_{\hat{\boldsymbol{u}}}\right)\times\boldsymbol{d}\right) + \alpha_2\varepsilon^I_{\boldsymbol{b}^t}, \varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right\rangle. \end{aligned}$$

Following the same procedure to get (3.14) from (3.13), we obtain the conclusion. ∎

LEMMA 4.3. *There holds*

$$\begin{aligned}E_h^2 &\lesssim \mathrm{Re}\left\|\varepsilon^I_{\boldsymbol{L}}\right\|_0\left\|\varepsilon^h_{\boldsymbol{L}}\right\|_0 + \kappa\left\|\boldsymbol{d}\right\|_{W^{1,\infty}}\left(\left\|\varepsilon^I_{\boldsymbol{b}}\right\|_0\left\|\varepsilon^h_{\boldsymbol{u}}\right\|_0 + \left\|\varepsilon^I_{\boldsymbol{u}}\right\|_0\left\|\varepsilon^h_{\boldsymbol{b}}\right\|_0\right) \\ &\quad + \left(\left\|\varepsilon^I_{\boldsymbol{J}}\right\|_{\partial\Omega_h} + \kappa\left\|\boldsymbol{d}\right\|_{L^\infty}\left\|\varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\hat{\boldsymbol{u}}}\right\|_{\partial\Omega_h} + \alpha_2\left\|\varepsilon^I_{\boldsymbol{b}^t}\right\|_{\partial\Omega_h}\right)\left\|\varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}\right\|_{\partial\Omega_h}. \end{aligned} \tag{4.9}$$

Proof. Bounding the energy is the same as bounding the right-hand side of (4.2). The estimate of $\mathrm{Re}\left(\varepsilon^I_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{L}}\right)$ is straightforward by the Cauchy–Schwarz inequality. To estimate $\kappa\left(\varepsilon^I_{\boldsymbol{b}}, \nabla\times\left(\varepsilon^h_{\boldsymbol{u}}\times\boldsymbol{d}\right)\right)$, note that an algebraic computation gives

$$\begin{aligned}\kappa\left(\varepsilon^I_{\boldsymbol{b}}, \nabla\times\left(\varepsilon^h_{\boldsymbol{u}}\times\boldsymbol{d}\right)\right) &= \kappa\left(\varepsilon^I_{\boldsymbol{b}}, \varepsilon^h_{\boldsymbol{u}}\left(\nabla\cdot\boldsymbol{d}\right) - \left(\varepsilon^h_{\boldsymbol{u}}\cdot\nabla\right)\boldsymbol{d}\right) \\ &\quad + \kappa\left(\varepsilon^I_{\boldsymbol{b}}, \left(\boldsymbol{d}\cdot\nabla\right)\varepsilon^h_{\boldsymbol{u}} - \boldsymbol{d}\left(\nabla\cdot\varepsilon^h_{\boldsymbol{u}}\right)\right). \end{aligned} \tag{4.10}$$

The boundedness of the left-hand side can be obtained by

$$\begin{aligned}&\kappa\left|\left(\varepsilon^I_{\boldsymbol{b}}, \varepsilon^h_{\boldsymbol{u}}\left(\nabla\cdot\boldsymbol{d}\right) - \left(\varepsilon^h_{\boldsymbol{u}}\cdot\nabla\right)\boldsymbol{d}\right)_K\right| \le \kappa\|\varepsilon^I_{\boldsymbol{b}}\|_{0,K}\|\varepsilon^h_{\boldsymbol{u}}\|_{0,K}\|\boldsymbol{d}\|_{W^{1,\infty}(K)}, \\ &\kappa\left|\left(\varepsilon^I_{\boldsymbol{b}}, \left(\boldsymbol{d}\cdot\nabla\right)\varepsilon^h_{\boldsymbol{u}} - \boldsymbol{d}\left(\nabla\cdot\varepsilon^h_{\boldsymbol{u}}\right)\right)_K\right| \\ &\quad = \kappa\left|\left(\varepsilon^I_{\boldsymbol{b}}, \left(\left(\boldsymbol{d} - \mathbb{P}_0\boldsymbol{d}\right)\cdot\nabla\right)\varepsilon^h_{\boldsymbol{u}} - \left(\boldsymbol{d} - \mathbb{P}_0\boldsymbol{d}\right)\left(\nabla\cdot\varepsilon^h_{\boldsymbol{u}}\right)\right)_K\right| \\ &\quad \lesssim \kappa h_K\|\varepsilon^I_{\boldsymbol{b}}\|_{0,K}\|\boldsymbol{d}\|_{W^{1,\infty}(K)}\|\nabla\varepsilon^h_{\boldsymbol{u}}\|_{0,K} \lesssim \kappa\|\varepsilon^I_{\boldsymbol{b}}\|_{0,K}\|\boldsymbol{d}\|_{W^{1,\infty}(K)}\|\varepsilon^h_{\boldsymbol{u}}\|_{0,K}, \end{aligned}$$

where $\mathbb{P}_0\boldsymbol{d}$ is the L^2 projection of $\boldsymbol{d}$ to the piecewise constant space on K, and here we used (B.1f), Hölder inequality $\|f_1f_2f_3\|_{L^1} \le \|f_1\|_{L^2}\|f_2\|_{L^\infty}\|f_3\|_{L^2}$, the Bramble–Hilbert lemma (see, e.g., [3]), and the inverse estimate in the last two inequalities.

For an estimate of $\kappa\left(\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{d}\times\left(\nabla\times\varepsilon^h_{\boldsymbol{b}}\right)\right)$, we first note that

$$\kappa\left(\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{d}\times\left(\nabla\times\varepsilon^h_{\boldsymbol{b}}\right)\right) = \kappa\left(\varepsilon^I_{\boldsymbol{u}}, \left(\boldsymbol{d} - \mathbb{P}_0\boldsymbol{d}\right)\times\left(\nabla\times\varepsilon^h_{\boldsymbol{b}}\right)\right)$$

due to (B.1j). An argument similar to the above gives

$$\kappa| \left(\varepsilon_{\boldsymbol{u}}^I, \boldsymbol{d} \times \left(\nabla \times \varepsilon_{\boldsymbol{b}}^h\right)\right) |_K \leq \kappa\|\varepsilon_{\boldsymbol{u}}^I\|_{0,K}\|\boldsymbol{d} - \mathbb{P}_0\boldsymbol{d}\|_{L^\infty(K)}\|\nabla \times \varepsilon_{\boldsymbol{b}}^h\|_{0,K} \lesssim \kappa\|\varepsilon_{\boldsymbol{u}}^I\|_{0,K}\|\boldsymbol{d}\|_{W^{1,\infty}(K)}\|\varepsilon_{\boldsymbol{b}}^h\|_{0,K}.$$

Finally, we use the Cauchy–Schwarz inequality for the last term in (4.2). □

COROLLARY 4.4 (energy estimate). *There holds*

$$E_h^2 \lesssim \mathrm{Re}\left\|\varepsilon_{\boldsymbol{L}}^I\right\|_0^2 + \kappa\left\|\boldsymbol{d}\right\|_{W^{1,\infty}}\left(\left\|\varepsilon_{\boldsymbol{b}}^I\right\|_0\left\|\varepsilon_{\boldsymbol{u}}^h\right\|_0 + \left\|\varepsilon_{\boldsymbol{u}}^I\right\|_0\left\|\varepsilon_{\boldsymbol{b}}^h\right\|_0\right) + \alpha_2^{-1}\left\|\varepsilon_{\boldsymbol{J}}^I\right\|_{\partial\Omega_h}^2 + \kappa^2\left\|\boldsymbol{d}\right\|_{L^\infty}^2\left\|\varepsilon_{\boldsymbol{u}}^I\right\|_{\partial\Omega_h}^2 + \alpha_2\left\|\varepsilon_{\boldsymbol{b}^t}^I\right\|_{\partial\Omega_h}^2. \tag{4.11}$$

Proof. Apply Young's inequality to each of the terms on the right-hand side of (4.9) involving $\left\|\varepsilon_{\boldsymbol{L}}^h\right\|_0$ and $\|\varepsilon_{\boldsymbol{b}^t}^h - \varepsilon_{\hat{\boldsymbol{b}}^t}^h\|_{\partial\Omega_h}$. Note also that $\Pi\hat{\boldsymbol{u}}$ is the best approximation of $\boldsymbol{u}$ on $\partial\Omega_h$, so $\left\|\varepsilon_{\hat{\boldsymbol{u}}}^I\right\|_{\partial\Omega_h}$ is bounded by $\left\|\varepsilon_{\boldsymbol{u}}^I\right\|_{\partial\Omega_h}$. □

In the energy estimate (4.11), we do not have direct control over $\|\varepsilon_{\boldsymbol{u}}^h\|_0$ and $\|\varepsilon_{\boldsymbol{b}}^h\|_0$, so we employ a duality argument to estimate these terms. A similar approach for the Oseen equation appeared in [5], but $\varepsilon_{\boldsymbol{u}}^h$ and $\varepsilon_{\boldsymbol{b}}^h$ are coupled in our MHD system, so there are nontrivial modifications to complete this duality argument.

First, we define a dual (adjoint) problem of the MHD system (3.2) as

$$\mathrm{Re}\boldsymbol{L}^* - \nabla\boldsymbol{u}^* = \boldsymbol{0}, \tag{4.12a}$$
$$-\nabla\cdot\boldsymbol{L}^* - \nabla p^* - (\boldsymbol{w}\cdot\nabla)\boldsymbol{u}^* - \kappa\boldsymbol{d}\times(\nabla\times\boldsymbol{b}^*) = \boldsymbol{\theta}, \tag{4.12b}$$
$$-\nabla\cdot\boldsymbol{u}^* = 0, \tag{4.12c}$$
$$\frac{\mathrm{Rm}}{\kappa}\boldsymbol{J}^* - \nabla\times\boldsymbol{b}^* = \boldsymbol{0}, \tag{4.12d}$$
$$\nabla\times\boldsymbol{J}^* - \nabla r^* + \kappa\nabla\times(\boldsymbol{u}^*\times\boldsymbol{d}) = \boldsymbol{\sigma}, \tag{4.12e}$$
$$-\nabla\cdot\boldsymbol{b}^* = 0, \tag{4.12f}$$

with homogeneous boundary conditions. Here, $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ are two given functions in $L^2(\Omega)$, and the superscript $*$ is used to denote the corresponding unknowns in the adjoint equation. We assume the following elliptic regularity assumption:

$$\|\boldsymbol{u}^*\|_2 + \|\boldsymbol{b}^*, \boldsymbol{L}^*, \boldsymbol{J}^*, p^*, r^*\|_1 \lesssim \|\boldsymbol{\theta}, \boldsymbol{\sigma}\|_0. \tag{4.13}$$

The well-posedness of (4.12) and the conditions under which the regularity estimate (4.13) holds are discussed in Appendix C in [22].

We use the interpolation operators Π^* defined in (B.3) and (B.4) below, and $\varepsilon_{\boldsymbol{L}^*}^I, \varepsilon_{p^*}^I, \ldots,$ will denote $\boldsymbol{L}^* - \Pi^*\boldsymbol{L}^*$, $p^* - \Pi^*p^*$, etc. Testing (4.12b) with $\varepsilon_{\boldsymbol{u}}^h$ and (4.12e) with $\varepsilon_{\boldsymbol{b}}^h$ we have

$$\begin{aligned} &\left(\varepsilon_{\boldsymbol{u}}^h, \boldsymbol{\theta}\right) + \left(\varepsilon_{\boldsymbol{b}}^h, \boldsymbol{\sigma}\right) \\ &= \left(\varepsilon_{\boldsymbol{u}}^h, -\nabla\cdot\boldsymbol{L}^* - \nabla p^* - (\boldsymbol{w}\cdot\nabla)\boldsymbol{u}^* - \kappa\boldsymbol{d}\times(\nabla\times\boldsymbol{b}^*)\right) \\ &\quad + \left(\varepsilon_{\boldsymbol{b}}^h, \nabla\times\boldsymbol{J}^* - \nabla r^* + \kappa\nabla\times(\boldsymbol{u}^*\times\boldsymbol{d})\right) \\ &= \left(\nabla\varepsilon_{\boldsymbol{u}}^h, \boldsymbol{L}^*\right) + \left(\nabla\cdot\varepsilon_{\boldsymbol{u}}^h, p^*\right) + \left((\boldsymbol{w}\cdot\nabla)\varepsilon_{\boldsymbol{u}}^h, \boldsymbol{u}^*\right) - \kappa\left(\varepsilon_{\boldsymbol{u}}^h, \boldsymbol{d}\times(\nabla\times\boldsymbol{b}^*)\right) \end{aligned} \tag{4.14}$$

$$
\begin{aligned}
&+ \langle \varepsilon^h_{\boldsymbol{u}}, -\boldsymbol{L}^*\boldsymbol{n} - p^*\boldsymbol{n} - m\boldsymbol{u}^* \rangle + (\nabla\times\varepsilon^h_{\boldsymbol{b}}, \boldsymbol{J}^*) + (\nabla\cdot\varepsilon^h_{\boldsymbol{b}}, r^*) \\
&+ \kappa\left(\varepsilon^h_{\boldsymbol{b}}, \nabla\times(\boldsymbol{u}^*\times\boldsymbol{d})\right) + \langle \varepsilon^h_{\boldsymbol{b}}, \boldsymbol{n}\times\boldsymbol{J}^* - r^*\boldsymbol{n}\rangle \\
=& \left(\nabla\varepsilon^h_{\boldsymbol{u}}, \Pi^*\boldsymbol{L}^*\right) + \left(\nabla\cdot\varepsilon^h_{\boldsymbol{u}}, \Pi^*p^*\right) + \left((\boldsymbol{w}\cdot\nabla)\varepsilon^h_{\boldsymbol{u}}, \Pi^*\boldsymbol{u}^*\right) \\
&- \kappa\left(\varepsilon^h_{\boldsymbol{u}}, \boldsymbol{d}\times(\nabla\times\boldsymbol{b}^*)\right) + \langle \varepsilon^h_{\boldsymbol{u}}, -\boldsymbol{L}^*\boldsymbol{n} - p^*\boldsymbol{n} - m\boldsymbol{u}^*\rangle + \left(\nabla\times\varepsilon^h_{\boldsymbol{b}}, \Pi^*\boldsymbol{J}^*\right) \\
&+ \left(\nabla\cdot\varepsilon^h_{\boldsymbol{b}}, \Pi^*r^*\right) + \kappa\left(\nabla\times(\boldsymbol{u}^*\times\boldsymbol{d}), \varepsilon^h_{\boldsymbol{b}}\right) + \langle \varepsilon^h_{\boldsymbol{b}}, \boldsymbol{n}\times\boldsymbol{J}^* - r^*\boldsymbol{n}\rangle \\
=& \left(\varepsilon^h_{\boldsymbol{u}}, -\nabla\cdot\Pi^*\boldsymbol{L}^* - \nabla\Pi^*p^* - (\boldsymbol{w}\cdot\nabla)\Pi^*\boldsymbol{u}^* - \kappa\boldsymbol{d}\times(\nabla\times\boldsymbol{b}^*)\right) \\
&+ \left(\varepsilon^h_{\boldsymbol{b}}, \nabla\times\Pi^*\boldsymbol{J}^* - \nabla\Pi^*r^* + \kappa\nabla\times(\boldsymbol{u}^*\times\boldsymbol{d})\right) \\
&+ \langle \varepsilon^h_{\boldsymbol{u}}, -\varepsilon^I_{\boldsymbol{L}^*}\boldsymbol{n} - \varepsilon^I_{p^*}\boldsymbol{n} - m\varepsilon^I_{\boldsymbol{u}^*}\rangle + \langle \varepsilon^h_{\boldsymbol{b}}, \boldsymbol{n}\times\varepsilon^I_{\boldsymbol{J}^*} - \varepsilon^I_{r^*}\boldsymbol{n}\rangle,
\end{aligned}
$$

where we have used integration by parts in the second equality, the properties of the Π^* operators (B.3) and (B.4) in the third equality, and integration by parts again in the last equality.

This can be reduced to (see [22] for full details)

$$
\begin{aligned}
(4.15)\quad & \left(\varepsilon^h_{\boldsymbol{u}}, \boldsymbol{\theta}\right) + \left(\varepsilon^h_{\boldsymbol{b}}, \boldsymbol{\sigma}\right) \\
&= \operatorname{Re}\underbrace{\left(\varepsilon^h_{\boldsymbol{L}}, -\varepsilon^I_{\boldsymbol{L}^*}\right)}_{=:I_1} + \operatorname{Re}\underbrace{\left(\varepsilon^I_{\boldsymbol{L}}, \Pi^*\boldsymbol{L}^*\right)}_{=:I_2} + \kappa\underbrace{\left(\varepsilon^h_{\boldsymbol{b}}, \nabla\times(\varepsilon^I_{\boldsymbol{u}^*}\times\boldsymbol{d})\right)}_{=:I_3} \\
&\quad - \kappa\underbrace{\left(\varepsilon^I_{\boldsymbol{b}}, \nabla\times(\Pi^*\boldsymbol{u}^*\times\boldsymbol{d})\right)}_{=:I_4} + \kappa\underbrace{\left(\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{d}\times(\nabla\times\Pi^*\boldsymbol{b}^*)\right)}_{=:I_5} - \kappa\underbrace{\left(\varepsilon^h_{\boldsymbol{u}}, \boldsymbol{d}\times(\nabla\times\varepsilon^I_{\boldsymbol{b}^*})\right)}_{=:I_6} \\
&\quad \underbrace{-\left\langle \varepsilon^h_{\boldsymbol{u}}, m\varepsilon^I_{\hat{\boldsymbol{u}}^*} + \kappa\boldsymbol{d}\times\left(\boldsymbol{n}\times\left(-\varepsilon^I_{\boldsymbol{b}^*} + \varepsilon^I_{(\hat{\boldsymbol{b}}^*)^t}\right)\right)\right\rangle}_{=:I_7} \\
&\quad + \underbrace{\left\langle \varepsilon^h_{\hat{\boldsymbol{b}}^t}, \frac{1}{2}\kappa\boldsymbol{n}\times\left((-\varepsilon^I_{\boldsymbol{u}^*} + \varepsilon^I_{\hat{\boldsymbol{u}}^*})\times\boldsymbol{d}\right) + \boldsymbol{n}\times(-\varepsilon^I_{\boldsymbol{J}^*}) - \alpha_2\varepsilon^I_{(\boldsymbol{b}^*)^t}\right\rangle}_{=:I_8} \\
&\quad + \underbrace{\left\langle \varepsilon^h_{\boldsymbol{b}^t}, \frac{1}{2}\kappa\boldsymbol{n}\times\left((-\varepsilon^I_{\boldsymbol{u}^*} + \varepsilon^I_{\hat{\boldsymbol{u}}^*})\times\boldsymbol{d}\right) + \boldsymbol{n}\times\varepsilon^I_{\boldsymbol{J}^*} + \alpha_2\varepsilon^I_{(\boldsymbol{b}^*)^t}\right\rangle}_{=:I_9} \\
&\quad - \underbrace{\left\langle \boldsymbol{n}\times\varepsilon^I_{\boldsymbol{J}} - \frac{1}{2}\kappa\boldsymbol{n}\times\left((\varepsilon^I_{\boldsymbol{u}} + \varepsilon^I_{\hat{\boldsymbol{u}}})\times\boldsymbol{d}\right) + \alpha_2\varepsilon^I_{\boldsymbol{b}^t}, \Pi^*\boldsymbol{b}^* - \mathbb{P}_e\boldsymbol{b}^*\right\rangle}_{=:I_{10}}.
\end{aligned}
$$

Estimation for I_1. Combining the estimate for $\varepsilon^I_{\boldsymbol{L}^*}$ and (B.6) gives

$$
(4.16)\qquad |\mathrm{Re} I_1| \le \operatorname{Re}\left\|\varepsilon^h_{\boldsymbol{L}}\right\|_0 \left\|\varepsilon^I_{\boldsymbol{L}^*}\right\|_0 \lesssim h\operatorname{Re}\left\|\varepsilon^h_{\boldsymbol{L}}\right\|_0 \|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0.
$$

Estimation for I_2. Using (4.12a), (B.1i), (B.1j), Lemma A.3, and the regularity of the adjoint solutions, we have

$$
\begin{aligned}
(4.17)\quad |\mathrm{Re} I_2| &\le |\operatorname{Re}\left(\varepsilon^I_{\boldsymbol{L}}, \nabla\boldsymbol{u}^*\right)| + |\operatorname{Re}\left(\varepsilon^I_{\boldsymbol{L}}, -\varepsilon^I_{\boldsymbol{L}^*}\right)| \\
&= |\operatorname{Re}\left(\varepsilon^I_{\boldsymbol{L}}, \nabla(\boldsymbol{u}^* - \mathbb{P}_1\boldsymbol{u}^*)\right) + \left(\varepsilon^I_{\boldsymbol{u}}\otimes\boldsymbol{w}, \nabla\mathbb{P}_1\boldsymbol{u}^*\right)| + |\operatorname{Re}\left(\varepsilon^I_{\boldsymbol{L}}, -\varepsilon^I_{\boldsymbol{L}^*}\right)| \\
&\lesssim \operatorname{Re}\left(h\left\|\varepsilon^I_{\boldsymbol{L}}\right\|_0\|\boldsymbol{u}^*\|_2 + \|\boldsymbol{w} - \mathbb{P}_0\boldsymbol{w}\|_{L^\infty}\left\|\varepsilon^I_{\boldsymbol{u}}\right\|_0\|\boldsymbol{u}^*\|_1 + \left\|\varepsilon^I_{\boldsymbol{L}}\right\|_0\left\|\varepsilon^I_{\boldsymbol{L}^*}\right\|_0\right) \\
&\lesssim h\operatorname{Re}\left(\left\|\varepsilon^I_{\boldsymbol{L}}\right\|_0 + \|\boldsymbol{w}\|_{W^{1,\infty}}\left\|\varepsilon^I_{\boldsymbol{u}}\right\|_0\right)\|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0.
\end{aligned}
$$

Estimation for I_4. By the identity (4.10), it suffices to estimate

$$\left(\varepsilon_{\boldsymbol{b}}^I, \Pi^* \boldsymbol{u}^* \left(\nabla \cdot \boldsymbol{d}\right) - \left(\Pi^* \boldsymbol{u}^* \cdot \nabla\right) \boldsymbol{d}\right),$$
$$\left(\varepsilon_{\boldsymbol{b}}^I, \left(\left(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d}\right) \cdot \nabla\right) \Pi^* \boldsymbol{u}^* - \left(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d}\right) \left(\nabla \cdot \Pi^* \boldsymbol{u}^*\right)\right).$$

By the triangle inequality, the inverse estimate, and (B.6), we have

$$\begin{aligned}
\left\|\nabla \Pi^* \boldsymbol{u}^*\right\|_0 &\le \left\|\nabla(\Pi^* \boldsymbol{u}^* - \mathbb{P}_1 \boldsymbol{u}^*)\right\|_0 + \left\|\nabla \mathbb{P}_1 \boldsymbol{u}^*\right\|_0 \\
&\lesssim h^{-1} \left\|\Pi^* \boldsymbol{u}^* - \mathbb{P}_1 \boldsymbol{u}^*\right\|_0 + \left\|\boldsymbol{u}^*\right\|_1 \\
&\le h^{-1}(\left\|\varepsilon_{\boldsymbol{u}^*}^I\right\|_0 + \left\|\boldsymbol{u}^* - \mathbb{P}_1 \boldsymbol{u}^*\right\|_0) + \left\|\boldsymbol{u}^*\right\|_1 \\
&\lesssim \left\|\boldsymbol{\theta}, \boldsymbol{\sigma}\right\|_0, \\
\left\|\Pi^* \boldsymbol{u}^*\right\|_0 &\le \left\|\varepsilon_{\boldsymbol{u}^*}^I\right\|_0 + \left\|\boldsymbol{u}^*\right\|_0 \lesssim \left\|\boldsymbol{\theta}, \boldsymbol{\sigma}\right\|_0,
\end{aligned}$$

and thus,

$$\begin{aligned}
\left|\kappa I_4\right| &\lesssim \kappa \left\|\boldsymbol{d}\right\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{b}}^I\right\|_0 \left(\left\|\Pi^* \boldsymbol{u}^*\right\|_0 + h \left\|\nabla \Pi^* \boldsymbol{u}^*\right\|_0\right) \\
&\lesssim \kappa \left\|\boldsymbol{d}\right\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{b}}^I\right\|_0 \left\|\boldsymbol{\theta}, \boldsymbol{\sigma}\right\|_0.
\end{aligned} \tag{4.18}$$

Estimation for I_5. By an argument similar to the estimate of $\left\|\nabla \Pi^* \boldsymbol{u}^*\right\|_0$ above, $\left\|\nabla \times \Pi^* \boldsymbol{b}^*\right\|_0 \lesssim \left\|\boldsymbol{\theta}, \boldsymbol{\sigma}\right\|_0$. Since $I_5 = \left(\varepsilon_{\boldsymbol{u}}^I, \left(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d}\right) \times \left(\nabla \times \Pi^* \boldsymbol{b}^*\right)\right)$,

$$\begin{aligned}
\left|\kappa I_5\right| &\lesssim h\kappa \left\|\boldsymbol{d}\right\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{u}}^I\right\|_0 \left\|\nabla \times \Pi^* \boldsymbol{b}^*\right\|_0 \\
&\lesssim h\kappa \left\|\boldsymbol{d}\right\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{u}}^I\right\|_0 \left\|\boldsymbol{\theta}, \boldsymbol{\sigma}\right\|_0.
\end{aligned} \tag{4.19}$$

Estimation for I_6 and I_7. Integrating I_6 by parts (see (3.6)), we have

$$-\kappa I_6 = -\kappa \left(\varepsilon_{\boldsymbol{b}^*}^I, \nabla \times \left(\varepsilon_{\boldsymbol{u}}^h \times \boldsymbol{d}\right)\right) - \kappa \left\langle \boldsymbol{d} \times \left(\boldsymbol{n} \times \varepsilon_{\boldsymbol{b}^*}^I\right), \varepsilon_{\boldsymbol{u}}^h \right\rangle.$$

Now we can write $-\kappa I_6 + I_7$ as

$$-\kappa I_6 + I_7 = -\kappa \left(\varepsilon_{\boldsymbol{b}^*}^I, \nabla \times \left(\varepsilon_{\boldsymbol{u}}^h \times \boldsymbol{d}\right)\right) + \left\langle \varepsilon_{\boldsymbol{u}}^h, -m \varepsilon_{\hat{\boldsymbol{u}}^*}^I - \kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \varepsilon_{(\hat{\boldsymbol{b}}^*)^t}^I\right) \right\rangle.$$

For the first term, as in the estimate of I_4, it suffices to estimate

$$\left(\varepsilon_{\boldsymbol{b}^*}^I, \varepsilon_{\boldsymbol{u}}^h \left(\nabla \cdot \boldsymbol{d}\right) - \left(\varepsilon_{\boldsymbol{u}}^h \cdot \nabla\right) \boldsymbol{d}\right) \text{ and } \left(\varepsilon_{\boldsymbol{b}^*}^I, \left(\left(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d}\right) \cdot \nabla\right) \varepsilon_{\boldsymbol{u}}^h - \left(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d}\right) \left(\nabla \cdot \varepsilon_{\boldsymbol{u}}^h\right)\right).$$

Invoking Hölder's inequality and an inverse estimate, we can bound the upper bounds of the first term as

$$\left\|\varepsilon_{\boldsymbol{b}^*}^I\right\|_0 \left\|\boldsymbol{d}\right\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_0 \lesssim h \left\|\boldsymbol{d}\right\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_0 \left\|\boldsymbol{\theta}, \boldsymbol{\sigma}\right\|_0.$$

For the second term, we first observe that

$$\begin{aligned}
&\left\langle \varepsilon_{\boldsymbol{u}}^h, -m \varepsilon_{\hat{\boldsymbol{u}}^*}^I - \kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \varepsilon_{(\hat{\boldsymbol{b}}^*)^t}^I\right) \right\rangle \\
&\quad = \left\langle \varepsilon_{\boldsymbol{u}}^h, -(\boldsymbol{w} - \mathbb{P}_0 \boldsymbol{w}) \cdot \boldsymbol{n} \varepsilon_{\hat{\boldsymbol{u}}^*}^I - \kappa(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d}) \times \left(\boldsymbol{n} \times \varepsilon_{(\hat{\boldsymbol{b}}^*)^t}^I\right) \right\rangle.
\end{aligned}$$

By the Hölder inequality,

$$\begin{aligned}
\left|\left\langle \varepsilon_{\boldsymbol{u}}^h, -m \varepsilon_{\hat{\boldsymbol{u}}^*}^I - \kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \varepsilon_{(\hat{\boldsymbol{b}}^*)^t}^I\right) \right\rangle\right| &\le \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_{\partial \Omega_h} \Big(\left\|\boldsymbol{w} - \mathbb{P}_0 \boldsymbol{w}\right\|_{L^\infty(\partial \Omega_h)} \left\|\boldsymbol{u}^* - \mathbb{P}_1 \boldsymbol{u}^*\right\|_{\partial \Omega_h} \\
&\quad + \kappa \left\|(\boldsymbol{d} - \mathbb{P}_0 \boldsymbol{d})\right\|_{L^\infty(\partial \Omega_h)} \left\|\boldsymbol{b}^* - \mathbb{P}_0 \boldsymbol{b}^*\right\|_{\partial \Omega_h}\Big),
\end{aligned}$$

where we used the fact that $\Pi^*\hat{\boldsymbol{b}^*}$ and $\Pi^*\hat{\boldsymbol{u}^*}$ are the best approximations on $\partial\Omega_h$. By Lemma A.2, this is estimated by $\left(h\kappa \|\boldsymbol{d}\|_{W^{1,\infty}} + h^2 \|\boldsymbol{w}\|_{W^{1,\infty}}\right) \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_0 \|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0$, so

$$|-\kappa I_6 + I_7| \lesssim \left(h\kappa \|\boldsymbol{d}\|_{W^{1,\infty}} + h^2 \|\boldsymbol{w}\|_{W^{1,\infty}}\right) \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_0 \|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0 . \tag{4.20}$$

Estimation for I_3, I_8, and I_9. Integrating I_3 by parts (see (3.6)) gives

$$\kappa I_3 = \kappa\left(\varepsilon_{\boldsymbol{u}^*}^I, \boldsymbol{d}\times\left(\nabla\times\varepsilon_{\boldsymbol{b}}^h\right)\right) + \kappa\left\langle \boldsymbol{n}\times\left(\varepsilon_{\boldsymbol{u}^*}^I\times\boldsymbol{d}\right), \varepsilon_{\boldsymbol{b}}^h\right\rangle .$$

Some algebraic manipulations give

$$\begin{aligned}\kappa I_3 + I_8 + I_9 = {}& \kappa\left(\varepsilon_{\boldsymbol{u}^*}^I, \boldsymbol{d}\times\left(\nabla\times\varepsilon_{\boldsymbol{b}}^h\right)\right) + \left\langle \varepsilon_{\boldsymbol{b}^t}^h, \kappa\boldsymbol{n}\times\left(\varepsilon_{\hat{\boldsymbol{u}}^*}^I\times\boldsymbol{d}\right)\right\rangle \\ &+ \left\langle \varepsilon_{\hat{\boldsymbol{b}}^t}^h - \varepsilon_{\boldsymbol{b}^t}^h, \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon_{\hat{\boldsymbol{u}}^*}^I - \varepsilon_{\boldsymbol{u}^*}^I\right)\times\boldsymbol{d}\right) - \boldsymbol{n}\times\varepsilon_{\boldsymbol{J}^*}^I - \alpha_2\varepsilon_{(\boldsymbol{b}^*)^t}^I\right\rangle .\end{aligned}$$

The first term is easily estimated by

$$\begin{aligned}\left|\left(\varepsilon_{\boldsymbol{u}^*}^I, \boldsymbol{d}\times\left(\nabla\times\varepsilon_{\boldsymbol{b}}^h\right)\right)\right| &= \left|\left(\varepsilon_{\boldsymbol{u}^*}^I, (\boldsymbol{d}-\mathbb{P}_0\boldsymbol{d})\times\left(\nabla\times\varepsilon_{\boldsymbol{b}}^h\right)\right)\right| \\ &\lesssim h\|\boldsymbol{d}\|_{W^{1,\infty}}\left\|\varepsilon_{\boldsymbol{b}}^h\right\|_0\|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0 .\end{aligned}$$

For the second term, we have

$$\begin{aligned}\left|\left\langle \varepsilon_{\boldsymbol{b}^t}^h, \kappa\boldsymbol{n}\times\left(\varepsilon_{\hat{\boldsymbol{u}}^*}^I\times\boldsymbol{d}\right)\right\rangle\right| &\le \kappa\left\|\varepsilon_{\boldsymbol{b}^t}^h\right\|_{\partial\Omega_h}\|\boldsymbol{u}^*-\mathbb{P}_k\boldsymbol{u}^*\|_{\partial\Omega_h}\|\boldsymbol{d}-\mathbb{P}_0\boldsymbol{d}\|_{L^\infty(\partial\Omega_h)} \\ &\lesssim h^2\kappa\|\boldsymbol{d}\|_{W^{1,\infty}}\left\|\varepsilon_{\boldsymbol{b}}^h\right\|_0\|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0 ,\end{aligned}$$

where we used Lemma A.2 and the discrete trace inequality. Using the Cauchy–Schwarz inequality, (B.6), and Lemma A.2, the third term is bounded by

$$h^{\frac{1}{2}}\left(\kappa\|\boldsymbol{d}\|_{L^\infty} + \alpha_2 + 1\right)\left\|\varepsilon_{\hat{\boldsymbol{b}}^t}^h - \varepsilon_{\boldsymbol{b}^t}^h\right\|_{\partial\Omega_h}\|\boldsymbol{\sigma},\boldsymbol{\theta}\|_0 .$$

Combining the above estimates, we conclude that

$$\begin{aligned}|\kappa I_3 + I_8 + I_9| \lesssim \Bigg(& h\kappa\|\boldsymbol{d}\|_{W^{1,\infty}}\left\|\varepsilon_{\boldsymbol{b}}^h\right\|_0 \\ &+ h^{\frac{1}{2}}\left(\kappa\|\boldsymbol{d}\|_{L^\infty} + \alpha_2 + 1\right)\left\|\varepsilon_{\hat{\boldsymbol{b}}^t}^h - \varepsilon_{\boldsymbol{b}^t}^h\right\|_{\partial\Omega_h}\Bigg)\|\boldsymbol{\sigma},\boldsymbol{\theta}\|_0 .\end{aligned} \tag{4.21}$$

Estimation for I_{10}. Using the approximation capability of the projector $\boldsymbol{\Pi}$ (see Appendix B) we have

$$\begin{aligned}&\left|\left\langle \boldsymbol{n}\times\varepsilon_{\boldsymbol{J}}^I - \frac{1}{2}\kappa\boldsymbol{n}\times\left(\left(\varepsilon_{\boldsymbol{u}}^I + \varepsilon_{\hat{\boldsymbol{u}}}^I\right)\times\boldsymbol{d}\right) + \alpha_2\varepsilon_{\boldsymbol{b}^t}^I, \Pi^*\boldsymbol{b}^* - \mathbb{P}_e\boldsymbol{b}^*\right\rangle\right| \\ &\quad\lesssim h^{\frac{1}{2}}\left(\left\|\varepsilon_{\boldsymbol{J}}^I\right\|_{\partial\Omega_h} + \kappa\|\boldsymbol{d}\|_{L^\infty}\left\|\varepsilon_{\boldsymbol{u}}^I\right\|_{\partial\Omega_h} + \alpha_2\left\|\varepsilon_{\boldsymbol{b}}^I\right\|_{\partial\Omega_h}\right)\|\boldsymbol{\theta},\boldsymbol{\sigma}\|_0 .\end{aligned} \tag{4.22}$$

At this point we are ready to estimate the approximation errors for $\boldsymbol{L}$, $\boldsymbol{J}$, $\boldsymbol{u}$, and $\boldsymbol{b}$. For readability, let us absorb α_1, α_2, α_3, Re, Rm, κ, and the norms on $\boldsymbol{d}$ and $\boldsymbol{w}$ into the implicit constants. Note that the error estimates stated in the theorem below are optimal for $\boldsymbol{u}$ and $\boldsymbol{b}$, and suboptimal for $\boldsymbol{L}$, p, $\boldsymbol{H}$, and r. The results in section 5 indicate that it is possible for numerical results to exceed the suboptimal estimates proven in this work.

Theorem 4.5. *Suppose that* $\alpha_1 - \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty}$, α_2, *and* α_3 *are positive constants independent of* h, Re, Rm, κ, $\boldsymbol{d}$, *and* $\boldsymbol{w}$, *and suppose that* h *is sufficiently small, i.e.,*

$$h \leq C \ll 1, \tag{4.23}$$

with C *depending on the coefficients in estimates* (4.20) *and* (4.21). *Then it holds that*

$$E_h \lesssim h^{k+\frac{1}{2}} \|\boldsymbol{L}, \boldsymbol{J}, \boldsymbol{u}, \boldsymbol{b}, p, r\|_{k+1}, \tag{4.24}$$

and the following error estimates hold:

$$\|\boldsymbol{L} - \boldsymbol{L}_h, \boldsymbol{J} - \boldsymbol{J}_h\|_0 \lesssim h^{k+\frac{1}{2}} \|\boldsymbol{L}, \boldsymbol{J}, \boldsymbol{u}, \boldsymbol{b}, p, r\|_{k+1}, \tag{4.25}$$

$$\|\boldsymbol{b} - \boldsymbol{b}_h, \boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim h^{k+1} \|\boldsymbol{L}, \boldsymbol{J}, \boldsymbol{u}, \boldsymbol{b}, p, r\|_{k+1}. \tag{4.26}$$

Proof. We proceed by taking $\boldsymbol{\theta} = \varepsilon^h_{\boldsymbol{u}}$, $\boldsymbol{\sigma} = \varepsilon^h_{\boldsymbol{b}}$ in (4.12). If we use (4.15), the estimates (4.16)–(4.22), and Young's inequality, we can obtain

$$\left\|\varepsilon^h_{\boldsymbol{b}}, \varepsilon^h_{\boldsymbol{u}}\right\|_0 \lesssim \left\|\varepsilon^I_{\boldsymbol{b}}\right\|_0 + h\left\|\varepsilon^I_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{L}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^h_{\boldsymbol{u}}, \varepsilon^h_{\boldsymbol{b}}\right\|_0 + h^{\frac{1}{2}}\left\|\varepsilon^I_{\boldsymbol{J}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\boldsymbol{b}}, \varepsilon^h_{\hat{\boldsymbol{b}}^t} - \varepsilon^h_{\boldsymbol{b}^t}\right\|_{\partial\Omega_h}, \tag{4.27}$$

which can be further simplified to

$$\left\|\varepsilon^h_{\boldsymbol{b}}, \varepsilon^h_{\boldsymbol{u}}\right\|_0 \lesssim \left\|\varepsilon^I_{\boldsymbol{b}}\right\|_0 + h\left\|\varepsilon^I_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{L}}, \varepsilon^I_{\boldsymbol{u}}\right\|_0 + h^{\frac{1}{2}}\left\|\varepsilon^I_{\boldsymbol{J}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\boldsymbol{b}}, \varepsilon^h_{\hat{\boldsymbol{b}}^t} - \varepsilon^h_{\boldsymbol{b}^t}\right\|_{\partial\Omega_h} \tag{4.28}$$

if the constants that multiply $\left\|\varepsilon^h_{\boldsymbol{u}}\right\|_0$ and $\left\|\varepsilon^h_{\boldsymbol{b}}\right\|_0$ in (4.20) and (4.21) are sufficiently small, which is true given the assumptions we have made in the statement of this theorem. The approximation error terms on the right-hand side of (4.28) (i.e., terms with superscript h) are bounded by E_h (see the definition of E_h in (4.9)). This implies

$$\left\|\varepsilon^h_{\boldsymbol{b}}, \varepsilon^h_{\boldsymbol{u}}\right\|_0 \lesssim \left\|\varepsilon^I_{\boldsymbol{b}}\right\|_0 + h\left\|\varepsilon^I_{\boldsymbol{L}}, \varepsilon^I_{\boldsymbol{u}}\right\|_0 + h^{\frac{1}{2}}\left\|\varepsilon^I_{\boldsymbol{J}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\boldsymbol{b}}\right\|_{\partial\Omega_h} + h^{\frac{1}{2}} E_h. \tag{4.29}$$

Applying Young's inequality to the right-hand side of (4.11) for $\left\|\varepsilon^h_{\boldsymbol{u}}\right\|_0$, $\left\|\varepsilon^h_{\boldsymbol{b}}\right\|_0$, and using (4.29), we get

$$E_h^2 \lesssim \left\|\varepsilon^I_{\boldsymbol{J}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\boldsymbol{b}}\right\|^2_{\partial\Omega_h} + \left\|\varepsilon^I_{\boldsymbol{L}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\boldsymbol{b}}\right\|^2_0.$$

Then (4.24) follows from the approximation properties of $\varepsilon^I_{\boldsymbol{J}}, \varepsilon^I_{\boldsymbol{u}}, \varepsilon^I_{\boldsymbol{b}}, \varepsilon^I_{\boldsymbol{L}}$ with the trace inequality Lemma A.2 discussed in Appendices B and A, respectively. Further, it gives

$$\left\|\varepsilon^h_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{J}}\right\|_0 + \left\|\varepsilon^h_{\boldsymbol{u}} - \varepsilon^h_{\hat{\boldsymbol{u}}}, \varepsilon^h_{\boldsymbol{b}^t} - \varepsilon^h_{\hat{\boldsymbol{b}}^t}, \varepsilon^h_r - \varepsilon^h_{\hat{r}}\right\|_{\partial\Omega_h} \lesssim E_h \lesssim h^{k+\frac{1}{2}} \|\boldsymbol{L}, \boldsymbol{J}, \boldsymbol{u}, \boldsymbol{b}, p, r\|_{k+1},$$

where the first inequality is from the definition of E_h in (4.2). Then (4.25) follows from the triangle inequality. Finally, the above estimate, with (4.29), and the triangle inequality give (4.26). ∎

What remains is to estimate $\left\|\varepsilon^h_p\right\|_0$ and $\left\|\varepsilon^h_r\right\|_0$.

Theorem 4.6. *There holds*

$$\left\|\varepsilon^h_p\right\|_0 \lesssim \left\|\varepsilon^h_{\boldsymbol{L}}, \varepsilon^h_{\boldsymbol{u}}, \varepsilon^h_{\boldsymbol{b}}, \varepsilon^I_{\boldsymbol{b}}\right\|_0 + h^{\frac{1}{2}} E_h \lesssim h^{k+\frac{1}{2}} \|\boldsymbol{L}, \boldsymbol{J}, \boldsymbol{u}, \boldsymbol{b}, p, r\|_{k+1}, \quad k \geq 0.$$

Proof. Let $\tilde{\Pi} : \left[H^1(\Omega)\right]^d \to \left[\mathcal{P}_k(\Omega_h)\right]^d$ be defined by (see [10, Lemma 4.1])

$$\left(\tilde{\Pi}\boldsymbol{\vartheta} - \boldsymbol{\vartheta}, \boldsymbol{v}\right)_K = 0, \qquad \boldsymbol{v} \in \left[\mathcal{P}_{k-1}(K)\right]^d,$$
$$\left\langle (\tilde{\Pi}\boldsymbol{\vartheta} - \boldsymbol{\vartheta}) \cdot \boldsymbol{n}, \boldsymbol{\mu} \cdot \boldsymbol{n}\right\rangle_{\partial K} = 0, \qquad \boldsymbol{\mu} \in \left[\mathcal{P}_k^{\perp}(K)\right]^d,$$

for $\boldsymbol{\vartheta} \in \left[H^1(\Omega)\right]^d$ and $K \in \Omega_h$, where $\mathcal{P}_k^{\perp}(K)$ is the subspace of $\mathcal{P}_k(K)$ which is orthogonal to $\mathcal{P}_{k-1}(K)$ in $L^2(K)$.

Since Πp is the L^2-projection, we have $\left(\varepsilon_p^h, 1\right) = -(\varepsilon_p^I, 1) = 0$ from (3.10). It is known [16] that there exists $\boldsymbol{\vartheta} \in \left[H_0^1(\Omega)\right]^d$ such that $\nabla \cdot \boldsymbol{\vartheta} = \varepsilon_p^h$, $\|\boldsymbol{\vartheta}\|_1 \lesssim \left\|\varepsilon_p^h\right\|_0$. Then

$$\begin{aligned}\left\|\varepsilon_p^h\right\|_0^2 &= \left(\varepsilon_p^h, \nabla \cdot \boldsymbol{\vartheta}\right) = -\left(\nabla \varepsilon_p^h, \boldsymbol{\vartheta}\right) + \left\langle \varepsilon_p^h, \boldsymbol{\vartheta} \cdot \boldsymbol{n}\right\rangle \\ &= -\left(\nabla \varepsilon_p^h, \tilde{\Pi}\boldsymbol{\vartheta}\right) + \left\langle \varepsilon_p^h, \boldsymbol{\vartheta} \cdot \boldsymbol{n}\right\rangle \\ &= \left(\varepsilon_p^h, \nabla \cdot \tilde{\Pi}\boldsymbol{\vartheta}\right) + \left\langle \varepsilon_p^h \boldsymbol{n}, \boldsymbol{\vartheta} - \tilde{\Pi}\boldsymbol{\vartheta}\right\rangle,\end{aligned} \tag{4.30}$$

where we used integration by parts twice and used the definition $\tilde{\Pi}$. Since the exact solution $(\boldsymbol{L}, p, \boldsymbol{u}, \boldsymbol{b})$ and its trace also satisfy the HDG local (sub)equation (3.7b), we can add and subtract the corresponding projections in (B.1) to obtain

$$\begin{aligned}\left(\varepsilon_p^h, \nabla \cdot \tilde{\Pi}\boldsymbol{\vartheta}\right) &= \left(\varepsilon_{\boldsymbol{L}}^h, \nabla\tilde{\Pi}\boldsymbol{\vartheta}\right) - \left(\varepsilon_{\boldsymbol{u}}^h, (\boldsymbol{w} \cdot \nabla)\tilde{\Pi}\boldsymbol{\vartheta}\right) + \kappa\left(\varepsilon_{\boldsymbol{b}}^h, \nabla \times \left(\tilde{\Pi}\boldsymbol{\vartheta} \times \boldsymbol{d}\right)\right) \\ &\quad + \left\langle m\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\boldsymbol{L}}^h \boldsymbol{n} + \varepsilon_p^h \boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\varepsilon_{\boldsymbol{b}^t}^h + \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h\right), \tilde{\Pi}\boldsymbol{\vartheta}\right\rangle \\ &\quad + \kappa\left(\varepsilon_{\boldsymbol{b}}^I, \nabla \times \left(\tilde{\Pi}\boldsymbol{\vartheta} \times \boldsymbol{d}\right)\right),\end{aligned} \tag{4.31}$$

where we have taken $\tilde{\Pi}\boldsymbol{\vartheta}$ as the test function in (3.7b). Combining (4.30) and (4.31) yields

$$\begin{aligned}\left\|\varepsilon_p^h\right\|_0^2 &= \left(\varepsilon_{\boldsymbol{L}}^h, \nabla\tilde{\Pi}\boldsymbol{\vartheta}\right) - \left(\varepsilon_{\boldsymbol{u}}^h, (\boldsymbol{w} \cdot \nabla)\tilde{\Pi}\boldsymbol{\vartheta}\right) + \kappa\left(\varepsilon_{\boldsymbol{b}}^h, \nabla \times \left(\tilde{\Pi}\boldsymbol{\vartheta} \times \boldsymbol{d}\right)\right) \\ &\quad + \left\langle m\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\boldsymbol{L}}^h \boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\varepsilon_{\boldsymbol{b}^t}^h + \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h\right), \tilde{\Pi}\boldsymbol{\vartheta}\right\rangle \\ &\quad + \kappa\left(\varepsilon_{\boldsymbol{b}}^I, \nabla \times \left(\tilde{\Pi}\boldsymbol{\vartheta} \times \boldsymbol{d}\right)\right) + \left\langle \varepsilon_p^h \boldsymbol{n}, \boldsymbol{\vartheta}\right\rangle,\end{aligned}$$

which can be further simplified using the following two facts: first, integrating by parts twice and using the definition of $\tilde{\Pi}$ give $(\varepsilon_{\boldsymbol{L}}^h, \nabla\tilde{\Pi}\boldsymbol{\vartheta}) = (\varepsilon_{\boldsymbol{L}}^h, \nabla\boldsymbol{\vartheta}) - \langle \varepsilon_{\boldsymbol{L}}^h \boldsymbol{n}, \boldsymbol{\vartheta}\rangle + \langle \varepsilon_{\boldsymbol{L}}^h \boldsymbol{n}, \tilde{\Pi}\boldsymbol{\vartheta}\rangle$; and second, combining the first equation in (3.8) and (B.1k) gives

$$\begin{aligned}\left\langle -\varepsilon_{\boldsymbol{L}}^h \boldsymbol{n} + \varepsilon_p^h \boldsymbol{n}, \boldsymbol{\vartheta}\right\rangle &= \left\langle -\varepsilon_{\boldsymbol{L}}^h \boldsymbol{n} + \varepsilon_p^h \boldsymbol{n}, \mathbb{P}_e\boldsymbol{\vartheta}\right\rangle \\ &= -\left\langle m\varepsilon_{\boldsymbol{u}}^h + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\varepsilon_{\boldsymbol{b}^t}^h + \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h\right), \mathbb{P}_e\boldsymbol{\vartheta}\right\rangle.\end{aligned}$$

In particular, we obtain

$$\begin{aligned}\left\|\varepsilon_p^h\right\|_0^2 &= \left(\varepsilon_{\boldsymbol{L}}^h, \nabla\boldsymbol{\vartheta}\right) - \left(\varepsilon_{\boldsymbol{u}}^h, (\boldsymbol{w} \cdot \nabla)\tilde{\Pi}\boldsymbol{\vartheta}\right) + \kappa\left(\varepsilon_{\boldsymbol{b}}^h, \nabla \times \left(\tilde{\Pi}\boldsymbol{\vartheta} \times \boldsymbol{d}\right)\right) \\ &\quad + \left\langle m\varepsilon_{\boldsymbol{u}}^h + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\varepsilon_{\boldsymbol{b}^t}^h + \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right)\right) + \alpha_1\left(\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h\right), \tilde{\Pi}\boldsymbol{\vartheta} - \mathbb{P}_e\boldsymbol{\vartheta}\right\rangle \\ &\quad + \kappa\left(\varepsilon_{\boldsymbol{b}}^I, \nabla \times \left(\tilde{\Pi}\boldsymbol{\vartheta} \times \boldsymbol{d}\right)\right).\end{aligned}$$

By the triangle and Hölder inequalities,

$$\left\|\varepsilon_p^h\right\|_0^2 \lesssim \left\|\varepsilon_{\boldsymbol{L}}^h\right\|_0 \|\nabla\boldsymbol{\vartheta}\|_0 + \|\boldsymbol{w}\|_{L^\infty} \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_0 \left\|\nabla\tilde{\Pi}\boldsymbol{\vartheta}\right\|_0 + \kappa \|\boldsymbol{d}\|_{W^{1,\infty}} \left\|\varepsilon_{\boldsymbol{b}}^h, \varepsilon_{\boldsymbol{b}}^I\right\|_0 \left\|\nabla\tilde{\Pi}\boldsymbol{\vartheta}\right\|_0$$
$$+ \left(\|\boldsymbol{w}\|_{L^\infty} \left\|\varepsilon_{\boldsymbol{u}}^h\right\|_{\partial\Omega_h} + \kappa \|\boldsymbol{d}\|_{L^\infty} \left\|\varepsilon_{\boldsymbol{b}^t}^h, \varepsilon_{\boldsymbol{b}^t}^h - \varepsilon_{\hat{\boldsymbol{b}}^t}^h\right\|_{\partial\Omega_h} + \alpha_1 \left\|\varepsilon_{\boldsymbol{u}}^h - \varepsilon_{\hat{\boldsymbol{u}}}^h\right\|_{\partial\Omega_h}\right) \left\|\tilde{\Pi}\boldsymbol{\vartheta} - \mathbb{P}_e\boldsymbol{\vartheta}\right\|_{\partial\Omega_h}$$
$$\lesssim \left(\left\|\varepsilon_{\boldsymbol{L}}^h, \varepsilon_{\boldsymbol{u}}^h, \varepsilon_{\boldsymbol{b}}^h, \varepsilon_{\boldsymbol{b}}^I\right\|_0 + h^{\frac{1}{2}} E_h\right) \left\|\varepsilon_p^h\right\|_0,$$

where we have used Lemma A.3, the approximation capability of both $\tilde{\Pi}$ and the L^2-projection, the definition of E_h in (4.2), and the property of $\boldsymbol{\vartheta}$, and we absorb all mesh independent parameters into the implicit constant in the final inequality. As a consequence, we have $\left\|\varepsilon_p^h\right\|_0 \lesssim \left\|\varepsilon_{\boldsymbol{L}}^h, \varepsilon_{\boldsymbol{u}}^h, \varepsilon_{\boldsymbol{b}}^h, \varepsilon_{\boldsymbol{b}}^I\right\|_0 + h^{\frac{1}{2}} E_h$. Then the conclusion follows from the triangle inequality and the estimates of $\left\|\varepsilon_{\boldsymbol{L}}^h, \varepsilon_{\boldsymbol{u}}^h, \varepsilon_{\boldsymbol{b}}^h, \varepsilon_{\boldsymbol{b}}^I\right\|_0$ and E_h. ∎

To obtain an analogous result for $\left\|\varepsilon_r^h\right\|_0$, we need the following result.

LEMMA 4.7. *There exists* $\left[\boldsymbol{\vartheta} \in H^1(\Omega)\right]^d$ *such that* $\nabla \cdot \boldsymbol{\vartheta} = \varepsilon_r^h$, $\boldsymbol{n} \times \boldsymbol{\vartheta} = \boldsymbol{0}$ *on* $\partial\Omega$, *and* $\|\boldsymbol{\vartheta}\|_1 \lesssim \left\|\varepsilon_r^h\right\|_0$.

Proof. We omit the proof and refer the reader to [22]. ∎

THEOREM 4.8. *There holds*

$$\|r - r_h\|_0 \lesssim \left\|\varepsilon_{\boldsymbol{J}}^h, \varepsilon_{\boldsymbol{u}}^I, \varepsilon_{\boldsymbol{u}}^h\right\|_0 + h^{\frac{1}{2}} E_h \lesssim h^{k+\frac{1}{2}} \|\boldsymbol{L}, \boldsymbol{J}, \boldsymbol{u}, \boldsymbol{b}, p, r\|_{k+1}, \quad k \geq 0.$$

Proof. The proof is similar to the estimate of $\|p - p_h\|_0$, so we refer the reader to [22] for details. ∎

5. Numerical results. In this section we start with numerical results of the proposed HDG scheme for 2D MHD problems on simplicial meshes. The first example is Hartmann flow, whose analytical solution is known. The second example is a manufactured solution on a nonconvex domain to demonstrate that our HDG methods have good approximation capability, though the elliptic regularity assumption does not hold on nonconvex domains. We also consider a problem with a low regularity solution to demonstrate the benefit of high order methods. For 3D problems, hexahedral meshes with tensor product polynomial spaces can be more advantageous from a large-scale computational point of view. Though our analysis is not applicable in this case, the proposed HDG discretization is still valid, and we close the section with a 3D example on a hexahedral mesh.

5.1. Hartmann flow. In this numerical study, we consider a conducting incompressible fluid (liquid metal, for example) in a domain $[-\infty, \infty] \times [-l_0, l_0] \times [-\infty, \infty]$ (bounded by infinite parallel plates in the x_2 direction [19, 29]). The fluid is subject to a uniform pressure gradient $G := -\frac{\partial p}{\partial x_1}$ in the x_1 direction, and a uniform external magnetic field b_0 in the x_2 direction. If we consider no-slip boundary conditions on the x_2 boundaries, the resulting flow pattern is known as *Hartmann flow,* which admits an analytical solution that is 1D in nature. We assume that the infinite parallel plates are perfectly insulating. Here, we consider the Hartmann flow in a 2D domain $\Omega = [0, 0.025] \times [-1, 1]$. If we define the characteristic velocity as $u_0 := \sqrt{Gl_0/\rho}$ and consider the driving pressure gradient G as a forcing term (incorporated in $\boldsymbol{g}$), the

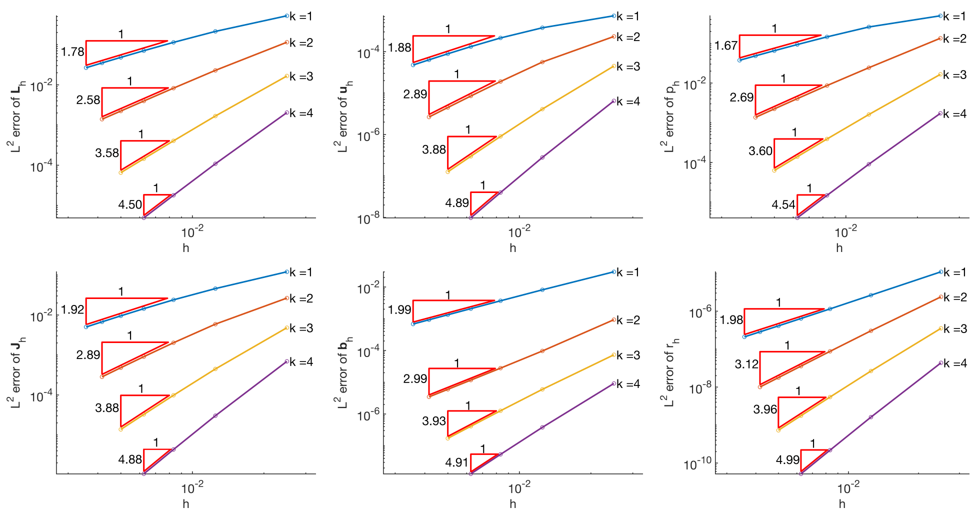

FIG. 1. *Hartmann flow problem: L^2 convergence.*

nondimensionalized solution with $\boldsymbol{g} = (1, 0)$, $\boldsymbol{f} = (0, 0)$ reads

$$\boldsymbol{u} = \left(\frac{\mathrm{Re}}{\mathrm{Ha}\tanh(\mathrm{Ha})} \left[1 - \frac{\cosh(\mathrm{Ha}\, x_2)}{\cosh(\mathrm{Ha})} \right], 0 \right), \quad p = -\frac{1}{2\kappa} \left[\frac{\sinh(\mathrm{Ha}\, x_2)}{\sinh(\mathrm{Ha})} - x_2 \right]^2 - p_0,$$
$$\boldsymbol{b} = \left(\frac{1}{\kappa} \left[\frac{\sinh(\mathrm{Ha}\, x_2)}{\sinh(\mathrm{Ha})} - x_2 \right], 1 \right), \qquad r = 0,$$

where $\mathrm{Ha} := \sqrt{\kappa \mathrm{Re} \mathrm{Rm}}$, and p_0 is a constant to enforce the zero mean-value condition to p. We set $\boldsymbol{w} = \boldsymbol{u}$ and $\boldsymbol{d} = \boldsymbol{b}$, and we enforce the boundary conditions on $\partial\Omega$ using the exact solution, i.e., $\boldsymbol{u}_D = \boldsymbol{u}$, $\boldsymbol{h}_D = \boldsymbol{b}^t$, and $r_D = 0$.

At refinement level l, the domain is divided into $l \times 80l$ squares, each of which is divided into two triangles from top right to bottom left. Figure 1 shows the convergence plots with $\mathrm{Re} = \mathrm{Rm} = 7.07$ and $\kappa = 200$. The convergence rates for $\boldsymbol{L}_h$, $\boldsymbol{u}_h$, p_h, $\boldsymbol{J}_h$, $\boldsymbol{b}_h$, and r_h are observed to be approximately $k + \frac{1}{2}$, $k+1$, $k + \frac{1}{2}$, $k+1$, $k+1$, and $k+1$, respectively. These observed rates approximately match or exceed their respective theoretical rates of $k + \frac{1}{2}$, $k+1$, $k+\frac{1}{2}$, $k+\frac{1}{2}$, $k+1$, and $k+\frac{1}{2}$, which were proven in section 4.

5.2. Nonconvex domain. This example illustrates the convergence of the HDG scheme applied to a problem posed on the nonconvex domain $\Omega = (-1,1) \times (-1,1) \backslash [0,1) \times (-1, 0]$ (similar to section 5.1.1 in [19]). We take $\mathrm{Re} = \mathrm{Rm} = \kappa = 1$, $\boldsymbol{w} = (2, 1)$, and $\boldsymbol{d} = (x_1, -x_2)$. We set $\boldsymbol{g}$ and $\boldsymbol{f}$ such that the manufactured solution for (3.1) is

$$\boldsymbol{u} = \left(-\left[x_2 \cos(x_2) + \sin(x_2) \right] e^{x_1}, x_2 \sin(x_2) e^{x_1} \right), \quad p = 2e^{x_1} \sin(x_2) - p_0,$$
$$\boldsymbol{b} = \left(-\left[x_2 \cos(x_2) + \sin(x_2) \right] e^{x_1}, x_2 \sin(x_2) e^{x_1} \right), \quad r = -\sin(\pi x_1) \sin(\pi x_2),$$

where p_0 is the constant chosen to enforce $(p, 1)_\Omega = 0$. We use the exact solution to enforce the boundary conditions $\partial\Omega$, i.e., $\boldsymbol{u}_D = \boldsymbol{u}$, $\boldsymbol{h}_D = \boldsymbol{b}^t$, and $r_D = r$.

At refinement level l, each quadrant of the domain is subdivided into $l \times l$ squares, each of which is divided into two triangles from top right to bottom left. Figure 2 shows the convergence plots. For this problem, we observe the optimal convergence

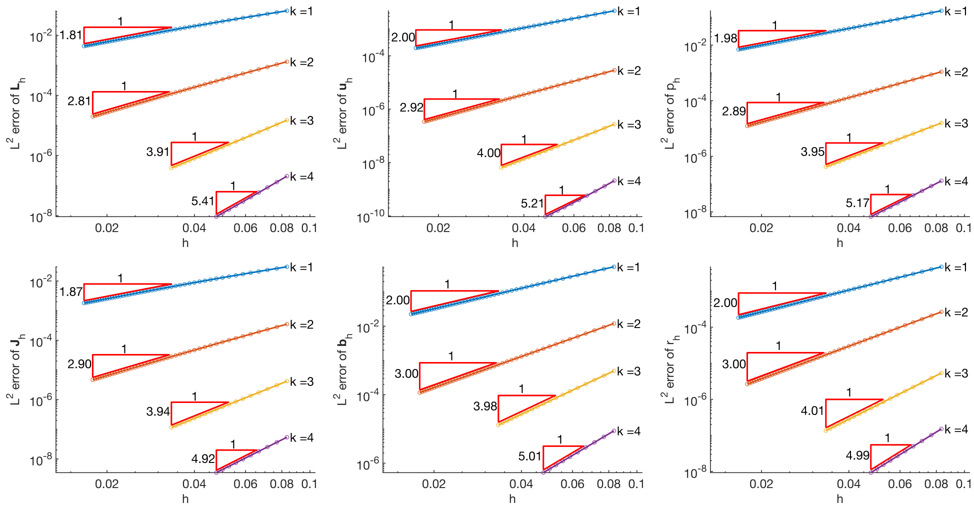

FIG. 2. *Nonconvex domain (L-shaped) problem: L^2 convergence.*

rates of $k+1$ for all of the local variables, which match or exceed the rates proven in section 4.

5.3. Singular solution. Although we do not discuss the implications of singular solutions on the theoretical convergence rates of the HDG scheme, applying the scheme to such a problem is instructive in assessing its robustness. This example illustrates the convergence of the HDG scheme using a manufactured solution with a singularity (similar to the example in section 5.2 of [19]). In particular, we consider the same nonconvex domain and mesh refinement as in the previous example. We take $\mathrm{Re} = \mathrm{Rm} = \kappa = 1$, $\boldsymbol{w} = \boldsymbol{0}$, and $\boldsymbol{d} = (-1, 1)$. We choose $\boldsymbol{g}$ and $\boldsymbol{f}$ such that the analytical solution of (3.1) has the form

$$\boldsymbol{u} = \begin{pmatrix} \rho^{\lambda}\left[(1+\lambda)\sin(\phi)\psi(\phi) + \cos(\phi)\psi'(\phi)\right], \\ \rho^{\lambda}\left[-(1+\lambda)\cos(\phi)\psi(\phi) + \sin(\phi)\psi'(\phi)\right] \end{pmatrix}, \qquad \boldsymbol{b} = \nabla\left(\rho^{2/3}\sin\left(\frac{2\phi}{3}\right)\right),$$

$$p = -\rho^{\lambda-1}\frac{(1+\lambda)^2\psi'(\phi) + \psi'''(\phi)}{1-\lambda}, \qquad r = 0,$$

$$\psi(\phi) = \cos\left(\frac{3\pi\lambda}{2}\right)\left[\frac{\sin((1+\lambda)\phi)}{1+\lambda} - \frac{\sin((1-\lambda)\phi)}{1-\lambda}\right] - \cos((1+\lambda)\phi) + \cos((1-\lambda)\phi),$$

where $\lambda \approx 0.54448373678246$. On $\partial\Omega$ we use the exact solution to set the boundary condition, i.e., $\boldsymbol{u}_D = \boldsymbol{u}$, $\boldsymbol{h}_D = \boldsymbol{b}^t$, and $r_D = r$. For this problem, it is known that $\boldsymbol{u} \in \left[H^{1+\lambda}(\Omega)\right]^2$, $p \in H^{\lambda}(\Omega)$, and $\boldsymbol{b} \in \left[H^{2/3}(\Omega)\right]^2$, and that the solution contains magnetic and hydrodynamic singularities that are among the strongest singularities [19].

Convergence results for this problem are shown in Figure 3. For the fluid variables $\boldsymbol{L}_h$, $\boldsymbol{u}_h$, and p_h, we observe convergence rates of approximately λ, 2λ, and λ, respectively. For the magnetic variables $\boldsymbol{J}_h$, $\boldsymbol{b}_h$, and r_h, we observe convergence rates of approximately $1/2$, $2/3$, and $1/3$, respectively.

5.4. 3D numerical experiments on structured hexahedral meshes. We now apply our HDG method with tensor product polynomial spaces to a 3D problem on structured hexahedral meshes. Note that our theoretical analysis is not valid for

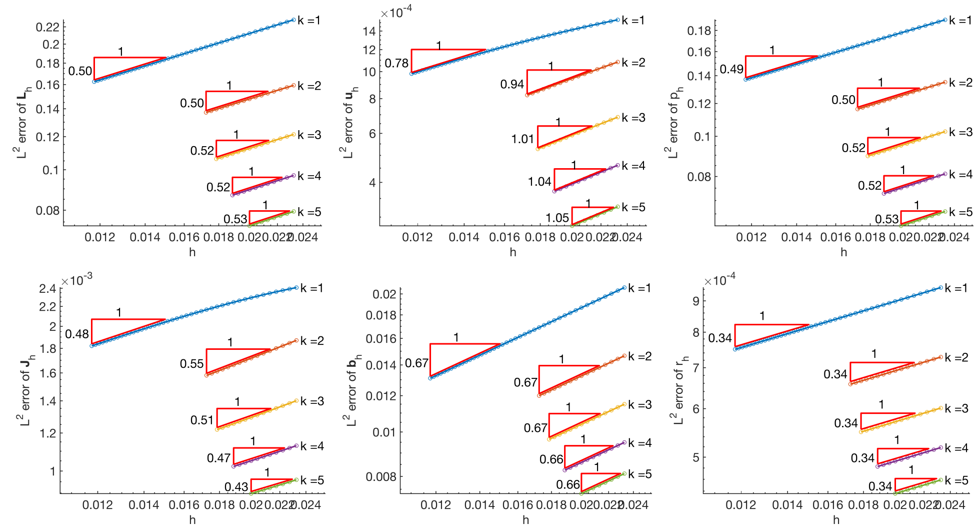

FIG. 3. *Singular solution problem: L^2 convergence.*

this case. We set $\Omega = [0,1]^3$, $\boldsymbol{w} = (1,2,-4)$, $\boldsymbol{d} = (-3,1,5)$ and choose the forcing function such that the exact solution is given by

$$\boldsymbol{u} = \boldsymbol{b} = \begin{pmatrix} \sin(2\pi x_1)\sin(2\pi x_2)\sin(2\pi x_3) \\ \sin(2\pi x_1)\cos(2\pi x_2)\cos(2\pi x_3) \\ \cos(2\pi(x_1 - x_3))\sin(2\pi x_2) \end{pmatrix},$$

$$p = e^{\left(x_1-\frac{1}{2}\right)^2+\left(x_2-\frac{1}{2}\right)^2+\left(x_3-\frac{1}{2}\right)^2} - \pi^{\frac{3}{2}}\mathrm{erf}\left(\frac{1}{2}\right)^3, \qquad r = 0.$$

We use Dirichlet boundary conditions on $\partial\Omega$ for $\boldsymbol{u}$, r, and for the tangential components of $\boldsymbol{b}$.

Figure 4 shows the convergence of all quantities as the mesh is refined. As can be seen, p and r converge optimally with a rate of $(k+1)$, while the convergence rate of $\boldsymbol{u}$ is suboptimal by a half order. For $\boldsymbol{L}$, $\boldsymbol{J}$, and $\boldsymbol{b}$, the convergence rate is, however, suboptimal by almost one order. This example clearly shows that rigorous results for simplicial meshes may not be extended to quadrilateral and hexahedral meshes.

6. Conclusions.

In this paper we proposed an HDG method for a linearization of the incompressible resistive MHD equations and carried out the a priori error analysis. Problems with both smooth and singular solutions are presented to examine the performance of the proposed HDG method. For problems with smooth solutions, the numerical convergence rates agree with the theoretical predictions. For singular solutions, the numerical results show that a high order method is still beneficial in terms of accuracy, though the convergence rate is limited by the regularity of the solution. For hexahedral meshes, for which our analysis is not applicable, the numerical results indicate that the convergence rate can be suboptimal by one order. Ongoing work includes 3D computation on parallel computers for large-scale problems, and extensions of our HDG method to nonlinear time-dependent MHD equations.

Appendix A. Auxiliary results.

In this appendix we collect some technical results that are useful for our analysis.

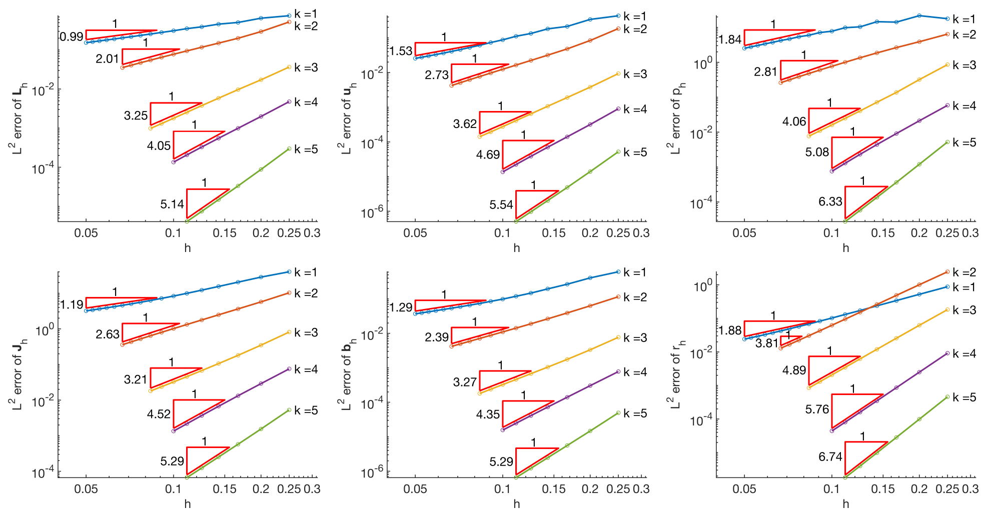

FIG. 4. *3D problem with structured hexahedral mesh: L^2 convergence.*

LEMMA A.1 (inverse inequality [25, Lemma 1.44]). *For $v \in \mathcal{P}_k(K)$ with $K \in \Omega_h$, there exists $C > 0$ independent of h such that $\|\nabla v\|_{0,K} \le C h_K^{-1} \|v\|_{0,K}$.*

LEMMA A.2 (trace inequality [25, Lemma 1.49]). *For $v \in H^1(\Omega_h)$ and for $K \in \Omega_h$ with $e \subset \partial K$, there exists $C > 0$ independent of h such that*

$$\|v\|_{0,e}^2 \le C \left(\|\nabla v\|_{0,K} + h_K^{-1} \|v\|_{0,K} \right) \|v\|_{0,K} .$$

Applying the arithmetic-geometric mean inequality to the right-hand side, we can derive

$$\text{(A.1)} \qquad \|v\|_{0,e} \lesssim \left(h_k^{\frac{1}{2}} \|\nabla v\|_{0,K} + h_K^{-\frac{1}{2}} \|v\|_{0,K} \right).$$

If $v \in H^1(\Omega_h)$ is in piecewise polynomial spaces, we can derive the following inequality from Lemma A.2 and the inverse inequality (Lemma A.1):

$$\text{(A.2)} \qquad \|v\|_{0,e} \lesssim h_K^{-\frac{1}{2}} \|v\|_{0,K}.$$

LEMMA A.3. *Suppose that $\Pi : H^1(K) \to \mathcal{P}_k(K)$ is a bounded interpolation which is a projection on $\mathcal{P}_k(K)$. Then $\|\nabla \Pi \boldsymbol{v}\|_{0,K} \lesssim \|\boldsymbol{v}\|_{1,K}$.*

Proof. See [22] for the proof. ∎

Appendix B. Definition of projections and their properties. In this section we use $\mathcal{P}_k$, $\boldsymbol{\mathcal{P}}_k$, $\tilde{\boldsymbol{\mathcal{P}}}_k$ for spaces of scalar, d-dimensional vector, $d \times d$ matrix-valued polynomials. By $\mathcal{P}_k^{\perp}$, $\boldsymbol{\mathcal{P}}_k^{\perp}$, $\tilde{\boldsymbol{\mathcal{P}}}_k^{\perp}$ we denote the spaces of polynomials of order at most k orthogonal to all polynomials of order at most $(k-1)$. $\boldsymbol{\mathcal{P}}_k^t(e)$ contains the tangential component of all polynomials in $\boldsymbol{\mathcal{P}}_k(e)$. We desire to have error equations conform to the original equations to facilitate the error analysis. To begin, we define a collective interpolation operator $\boldsymbol{\Pi}(\boldsymbol{L}, \boldsymbol{u}, p, \boldsymbol{J}, \boldsymbol{b}, r, \hat{\boldsymbol{u}}, \hat{\boldsymbol{b}}^t, \hat{r})$ implicitly through the interpolation errors $\varepsilon_{\boldsymbol{u}}^I = \boldsymbol{u} - \Pi\boldsymbol{u}$, $\varepsilon_{\boldsymbol{b}}^I = \boldsymbol{b} - \Pi\boldsymbol{b}$, etc., where $\Pi\boldsymbol{u}$, $\Pi\boldsymbol{b}$, etc., are components of the collective interpolator $\boldsymbol{\Pi}$ on $\boldsymbol{u}, \boldsymbol{b}$, etc. Specifically, the collective interpolation operator is defined by the following:

- L^2-projections on $e \in \mathcal{E}_h$ or on $K \in \Omega_h$ are defined as

(B.1a) $$\left\langle \varepsilon^I_{\hat{r}}, \gamma \right\rangle_e = 0, \quad \gamma \in \mathcal{P}_k\left(e\right),$$

(B.1b) $$\left\langle \varepsilon^I_{\hat{\boldsymbol{u}}}, \boldsymbol{\mu} \right\rangle_e = 0, \quad \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}_k\left(e\right),$$

(B.1c) $$\left\langle \varepsilon^I_{\hat{\boldsymbol{b}}^t}, \boldsymbol{\lambda}^t \right\rangle_e = 0, \quad \boldsymbol{\lambda}^t \in \boldsymbol{\mathcal{P}}^t_k\left(e\right),$$

(B.1d) $$\left(\varepsilon^I_{\boldsymbol{J}}, \boldsymbol{H}\right)_K = 0, \quad \boldsymbol{H} \in \left[\mathcal{P}_k\left(K\right)\right]^{\tilde{d}},$$

(B.1e) $$\left(\varepsilon^I_p, q\right)_K = 0, \quad q \in \mathcal{P}_k\left(K\right).$$

- On each $K \in \Omega_h$ and $e \in \mathcal{E}_h$, $e \subset \partial K$, $\Pi\boldsymbol{b}$ and Πr are defined as

(B.1f) $$(\varepsilon^I_{\boldsymbol{b}}, \boldsymbol{c})_K = 0, \quad \boldsymbol{c} \in \boldsymbol{\mathcal{P}}_{k-1}(K),$$

(B.1g) $$(\varepsilon^I_r, s)_K = 0, \quad s \in \mathcal{P}_{k-1}(K),$$

(B.1h) $$\left\langle \varepsilon^I_{\boldsymbol{b}} \cdot \boldsymbol{n} + \alpha_3 \varepsilon^I_r, \gamma \right\rangle_e = 0, \quad \gamma \in \mathcal{P}_k(e).$$

- On each $K \in \Omega_h$ and $e \in \mathcal{E}_h$, $e \subset \partial K$, $\Pi\boldsymbol{L}$ and $\Pi\boldsymbol{u}$ are defined as

(B.1i) $$-\left(\varepsilon^I_{\boldsymbol{L}}, \boldsymbol{G}\right)_K + \left(\varepsilon^I_{\boldsymbol{u}} \otimes \boldsymbol{w}, \boldsymbol{G}\right)_K = 0, \quad \boldsymbol{G} \in \tilde{\boldsymbol{\mathcal{P}}}_{k-1}\left(K\right),$$

(B.1j) $$\left(\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{v}\right)_K = 0, \quad \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}\left(K\right),$$

(B.1k) $$\left\langle -\varepsilon^I_{\boldsymbol{L}}\boldsymbol{n} + \left(m + \alpha_1\right)\varepsilon^I_{\boldsymbol{u}}, \boldsymbol{\mu} \right\rangle_e = -\left\langle \varepsilon^I_p \boldsymbol{n} + \frac{1}{2}\kappa \boldsymbol{d} \times \left(\boldsymbol{n} \times \left(\varepsilon^I_{\boldsymbol{b}^t} + \varepsilon^I_{\hat{\boldsymbol{b}}^t}\right)\right), \boldsymbol{\mu} \right\rangle_e, \quad \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}_k\left(e\right).$$

The well-definedness and optimality of the L^2-projections are clear. The coupled projector $\Pi\left(\boldsymbol{b}, r\right) := \left(\Pi\boldsymbol{b}, \Pi r\right)$ has been studied in [8], and in particular we have

(B.2a) $$\|\varepsilon^I_{\boldsymbol{b}}\|_{0,K} \lesssim h^{k+1}\left\|\boldsymbol{b}\right\|_{k+1,K} + \alpha_3 h^{k+1}\left\|r\right\|_{k+1,K},$$

(B.2b) $$\|\varepsilon^I_r\|_{0,K} \lesssim \alpha_3^{-1} h^{k+1}\left\|\nabla \cdot \boldsymbol{b}\right\|_{k,K} + h^{k+1}\left\|r\right\|_{k+1,K},$$

where, again, for simplicity we choose the same solution order k for all the unknowns. Here, we assume that $\boldsymbol{b}$ and r are sufficiently smooth, that is, $\boldsymbol{b} \in \left[H^{k+1}\left(\Omega\right)\right]^d$ and $r \in H^{k+1}\left(\Omega\right)$.

LEMMA B.1 (estimation for $\varepsilon^I_{\boldsymbol{u}}$). *Suppose* $\boldsymbol{u} \in \left[H^{k+1}\left(\Omega\right)\right]^d$, $\boldsymbol{L} \in \left[H^{k+1}\left(\Omega\right)\right]^{d\times d}$, $r \in H^{k+1}\left(\Omega\right)$, $\boldsymbol{b} \in \left[H^{k+1}\left(\Omega\right)\right]^d$, *and* $p \in H^{k+1}\left(\Omega\right)$. *Then*

$$\begin{aligned}\|\varepsilon^I_{\boldsymbol{u}}\|_0 \lesssim C(\alpha_1, \boldsymbol{w})[(\alpha_1 + \left\|\boldsymbol{w}\right\|_{L^\infty} + h\left\|\boldsymbol{w}\right\|_{W^{1,\infty}})h^{k+1}\|\boldsymbol{u}\|_{k+1} \\ + h^{k+1}\|\nabla \cdot \boldsymbol{L} - \nabla p\|_k + \kappa h^{k+1}\|\boldsymbol{d}\|_{L^\infty}\left(\|\boldsymbol{b}\|_{k+1} + \alpha_3\|r\|_{k+1}\right)]\end{aligned}$$

holds with $C(\alpha_1, \boldsymbol{w}) = 1/(\alpha_1 - \frac{1}{2}\left\|\boldsymbol{w}\right\|_{L^\infty})$.

Proof. See [22] for the proof. ∎

LEMMA B.2 (estimation for $\varepsilon^I_{\boldsymbol{L}}$). *Suppose* $\boldsymbol{u} \in \left[H^{k+1}\left(\Omega\right)\right]^d$, $\boldsymbol{L} \in \left[H^{k+1}\left(\Omega\right)\right]^{d\times d}$, $r \in H^{k+1}\left(\Omega\right), \boldsymbol{b} \in \left[H^{k+1}\left(\Omega\right)\right]^d$, *and* $p \in H^{k+1}\left(\Omega\right)$. *Furthermore, suppose the trace of the tensor* $\boldsymbol{L}$ *vanishes, i.e.,* $\operatorname{tr}\boldsymbol{L} = 0$. *There holds*

$$\begin{aligned}\left\|\varepsilon^I_{\boldsymbol{L}}\right\|_0 \lesssim h^{k+1}\left\|p\right\|_{k+1} + h^{k+1}\left\|\boldsymbol{L}\right\|_{k+1} + \kappa\left\|\boldsymbol{d}\right\|_{L^\infty}\left(h^{k+1}\left\|\boldsymbol{b}\right\|_{k+1} + \alpha_3 h^{k+1}\left\|r\right\|_{k+1}\right) \\ + \left(\alpha_1 + \left\|\boldsymbol{w}\right\|_{L^\infty} + h\left\|\boldsymbol{w}\right\|_{W^{1,\infty}}\right)\left\|\varepsilon^I_{\boldsymbol{u}}\right\|_0 + \left(\alpha_1 + \left\|\boldsymbol{w}\right\|_{L^\infty}\right)h^{k+1}\left\|\boldsymbol{u}\right\|_{k+1}.\end{aligned}$$

Proof. See [22] for the proof. ∎

We now define the adjoint projection $\Pi^*(\boldsymbol{L}^*, \boldsymbol{u}^*, p^*, \boldsymbol{J}^*, \boldsymbol{b}^*, r^*, \hat{\boldsymbol{u}}^*, (\hat{\boldsymbol{b}}^*)^t, \hat{r}^*)$. As in the splitting of errors with Π, we define $\varepsilon^I_{\sigma^*} = \sigma^* - \Pi^*\sigma^*$ for an adjoint unknown σ^*. We first define $\Pi^*\boldsymbol{J}^*$, Π^*p^*, $\Pi^*\hat{\boldsymbol{u}}^*$, $\Pi^*(\hat{\boldsymbol{b}}^*)^t$, $\Pi^*\hat{r}^*$ as L^2-projections into relevant polynomial spaces, and define $\Pi^*\boldsymbol{b}^*$, Π^*r^* to satisfy

$$(\varepsilon^I_{\boldsymbol{b}^*}, \boldsymbol{c})_K = 0 \quad \forall \boldsymbol{c} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \tag{B.3a}$$

$$(\varepsilon^I_{r^*}, s)_K = 0 \quad \forall s \in \mathcal{P}_{k-1}(K), \tag{B.3b}$$

$$\left\langle -\varepsilon^I_{\boldsymbol{b}^*} \cdot \boldsymbol{n} + \alpha_3 \varepsilon^I_{r^*}, \gamma \right\rangle_e = 0 \quad \forall \gamma \in \mathcal{P}_k(e). \tag{B.3c}$$

We then choose $\Pi^*\boldsymbol{L}^*$, $\Pi^*\boldsymbol{u}^*$ to satisfy

$$(\varepsilon^I_{\boldsymbol{L}^*}, \boldsymbol{G})_K + (\varepsilon^I_{\boldsymbol{u}^*} \otimes \boldsymbol{w}, \boldsymbol{G})_K = 0 \quad \forall \boldsymbol{G} \in \tilde{\boldsymbol{\mathcal{P}}}_{k-1}(K), \tag{B.4a}$$

$$(\varepsilon^I_{\boldsymbol{u}^*}, \boldsymbol{v})_K = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \tag{B.4b}$$

$$\left\langle -\varepsilon^I_{\boldsymbol{L}^*}\boldsymbol{n} + \alpha_1 \varepsilon^I_{\boldsymbol{u}^*}, \boldsymbol{\mu} \right\rangle_e = \langle \boldsymbol{g}, \boldsymbol{\mu} \rangle_e \quad \forall \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}_k(e), \tag{B.4c}$$

where $\boldsymbol{g} = \varepsilon^I_{p^*}\boldsymbol{n} - \frac{1}{2}\kappa \boldsymbol{d} \times (\boldsymbol{n} \times (-(\varepsilon^I_{\boldsymbol{b}^*})^t + \varepsilon^I_{(\hat{\boldsymbol{b}}^*)^t}))$.

Assuming that $(\boldsymbol{L}^*, \boldsymbol{u}^*, p^*, \boldsymbol{J}^*, \boldsymbol{b}^*, r^*, \hat{\boldsymbol{u}}^*, (\hat{\boldsymbol{b}}^*)^t, \hat{r}^*)$ are sufficiently regular, we can show that the interpolation $\boldsymbol{\Pi}^*$ is well-defined and provides optimal approximations. Due to the similarity between $(\Pi\boldsymbol{b}, \Pi r)$ and $(\Pi^*\boldsymbol{b}^*, \Pi^* r^*)$, we can conclude that

$$\left\| \varepsilon^I_{\boldsymbol{b}^*} \right\|_{0,K} \lesssim h^{k+1} \left\| \boldsymbol{b}^* \right\|_{k+1,K} + \alpha_3 h^{k+1} \left\| r^* \right\|_{k+1,K}, \tag{B.5a}$$

$$\left\| \varepsilon^I_{r^*} \right\|_{0,K} \lesssim \alpha_3^{-1} h^{k+1} \left\| \nabla \cdot \boldsymbol{b}^* \right\|_{k,K} + h^{k+1} \left\| r^* \right\|_{k+1,K}. \tag{B.5b}$$

It can also be shown that

$$\begin{aligned} \|\varepsilon^I_{\boldsymbol{u}^*}\|_0 \lesssim \left(\alpha_1 - \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty}\right)^{-1} &\left[(\alpha_1 + h\|\boldsymbol{w}\|_{W^{1,\infty}}) h^{k+1} \|\boldsymbol{u}^*\|_{k+1} \right. \\ &\left. + h^{k+1} \|\nabla \cdot \boldsymbol{L}^* + \nabla p^*\|_k + \kappa h^{k+1} \|\boldsymbol{d}\|_{L^\infty} \left(\|\boldsymbol{b}^*\|_{k+1} + \alpha_3 \|r^*\|_{k+1}\right)\right], \\ \left\|\varepsilon^I_{\boldsymbol{L}^*}\right\|_0 \lesssim h^{k+1} \|p^*\|_{k+1} + h^{k+1} &\|\boldsymbol{L}^*\|_{k+1} + \kappa h^{k+1} \|\boldsymbol{d}\|_{L^\infty} \left(\|\boldsymbol{b}^*\|_{k+1} + \alpha_3 \|r^*\|_{k+1}\right) \\ &+ (\alpha_1 + h\|\boldsymbol{w}\|_{W^{1,\infty}}) \left\|\varepsilon^I_{\boldsymbol{u}^*}\right\|_0 + \alpha_1 h^{k+1} \|\boldsymbol{u}^*\|_{k+1}, \end{aligned}$$

assuming $\operatorname{tr} \boldsymbol{L}^* = 0$. The proofs are analogous to those for the $\boldsymbol{\Pi}$ projections. As a consequence, from the elliptic regularity assumption (4.13), we have

$$\max\left\{ \left\|\varepsilon^I_{\boldsymbol{L}^*}\right\|_0, \left\|\varepsilon^I_{\boldsymbol{u}^*}\right\|_0, \left\|\varepsilon^I_{p^*}\right\|_0, \left\|\varepsilon^I_{\boldsymbol{b}^*}\right\|_0, \left\|\varepsilon^I_{r^*}\right\|_0 \right\} \lesssim h \|\boldsymbol{\sigma}, \boldsymbol{\theta}\|_0, \tag{B.6}$$

and the implicit constant is independent of h.

REFERENCES

[1] D. N. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 7–32.
[2] P. B. Bochev and A. C. Robinson, *Matching algorithms with physics: Exact sequences of finite element spaces*, in Collected Lectures on the Preservation of Stability under Discretization, D. Estep and S. Tavener, eds., Proc. 109, SIAM, 2002, pp. 145–165.
[3] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer, 2008.
[4] T. Bui-Thanh, *From Godunov to a unified hybridized discontinuous Galerkin framework for partial differential equations*, J. Comput. Phys., 295 (2015), pp. 114–146.
[5] A. Cesmelioglu, B. Cockburn, N. C. Nguyen, and J. Peraire, *Analysis of HDG methods for Oseen equations*, J. Sci. Comput., 55 (2013), pp. 392–431.
[6] L. Chacón, D. A. Knoll, and J. M. Finn, *An implicit nonlinear reduced resistive MHD solver*, J. Comput. Phys., 178 (2002), pp. 15–36.
[7] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365, https://doi.org/10.1137/070706616.
[8] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comput., 79 (2010), pp. 1351–1367.
[9] B. Cockburn, N. C. Nguyen, and J. Peraire, *A comparison of HDG methods for Stokes flow*, J. Sci. Comput., 45 (2010), pp. 215–237.
[10] B. Cockburn and F.-J. Sayas, *Divergence-conforming HDG methods for Stokes flows*, Math. Comp., 83 (2014), pp. 1571–1598.
[11] R. Codina and N. Hernández-Silva, *Approximation of the thermally coupled MHD problem using a stabilized finite element method*, J. Comp. Phys., 230 (2011), pp. 1281–1303.
[12] M. Costabel and M. Dauge, *Singularities of electromagnetic fields in polyhedral domains*, Arch. Ration. Mech. Anal., 151 (2000), pp. 221–276.
[13] M. Costabel and M. Dauge, *Weighted regularization of Maxwell equations in polyhedral domains*, Numer. Math., 93 (2002), pp. 239–277.
[14] L. C. Evans, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, 1998.
[15] J.-F. Gerbeau, *A stabilized finite element method for the incompressible magnetohydrodynamic equations*, Numer. Math., 87 (2000), pp. 83–111.
[16] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms,* Springer Ser. Comput. Math. 5, Springer-Verlag, 1986.
[17] H. Goedbloed and S. Poedts, *Principles of Magnetohydrodynamics with Applications to Laboratory and Astrophysical Plasmas*, Cambridge University Press, 2004.
[18] M. Gunzburger, A. Meir, and J. Peterson, *On the existence, uniqueness, and finite element approximation of solutions of the equations of stationary, incompressible magnetohydrodynamics*, Math. Comp., 56 (1991), pp. 523–563.
[19] P. Houston, D. Schötzau, and X. Wei, *A mixed DG method for linearized incompressible magnetohydrodynamics*, J. Sci. Comput., 40 (2009), pp. 281–314.
[20] K. Hu, Y. Ma, and J. Xu, *Stable finite element methods preserving $\nabla \cdot B = 0$ exactly for MHD models*, Numer. Math., 135 (2017), pp. 371–396.
[21] J. Hyman and M. Shashkov, *Adjoint operators for the natural discretizations of the divergence, gradient and curl on logically rectangular grids*, Appl. Numer. Math., 25 (1997), pp. 413–442.
[22] J. J. Lee, S. Shannon, T. Bui-Thanh, and J. N. Shadid, *Analysis of an HDG Method for Linearized Incompressible Resistive MHD Equations*, preprint, https://arxiv.org/abs/1702.05124, 2018.
[23] J. Nedelec, *Mixed finite elements in* $\mathbf{R}^3$, Numer. Math., 35 (1980), pp. 315–341.
[24] N. C. Nguyen, J. Peraire, and B. Cockburn, *A hybridizable discontinuous Galerkin method for Stokes flow*, Comput Methods Appl. Mech. Engrg., 199 (2010), pp. 582–597.
[25] D. A. D. Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer, 2012.
[26] N. B. Salah, A. Soulaimani, W. G. Habashi, and M. Fortin, *A conservative stabilized finite element method for the magento-hydrodyanamics equations*, Int. J. Numer. Methods Fluids, 29 (1999), pp. 535–554.
[27] D. Schötzau, *Mixed finite element methods for stationary incompressible magnetohydrodynamics*, Numer. Math., 96 (2004), pp. 771–800.

[28] J. N. Shadid, R. P. Pawlowski, J. W. Banks, L. Chacón, P. T. Lin, and R. S. Tuminaro, *Towards a scalable fully-implicit fully-coupled resistive MHD formulation with stabilized FE methods*, J. Comput. Phys., 229 (2010), pp. 7649–7671.

[29] J. N. Shadid, R. P. Pawlowski, E. C. Cyr, R. S. Tuminaro, L. Chacón, and P. D. Weber, *Scalable implicit incompressible resistive MHD with stabilized FE and fully-coupled Newton-Krylov-AMG*, Comput. Methods Appl. Mech. Engrg., 304 (2016), pp. 1–25.