NEW ANALYTICAL TOOLS FOR HDG IN ELASTICITY, WITH APPLICATIONS TO ELASTODYNAMICS

SHUKAI DU AND FRANCISCO-JAVIER SAYAS

Abstract. We present some new analytical tools for the error analysis of hybridizable discontinuous Galerkin (HDG) method for linear elasticity. These tools allow us to analyze more variants of HDG method using the projection-based approach, which renders the error analysis simple and concise. The key result is a tailored projection for the Lehrenfeld-Schöberl type HDG (HDG+ for simplicity) methods. By using the projection we recover the error estimates of HDG+ for steady-state and time-harmonic elasticity in a simpler analysis. We also present a semi-discrete (in space) HDG+ method for transient elastic waves and prove it is uniformly-in-time optimal convergent by using the projection-based error analysis. Numerical experiments supporting our analysis are presented at the end.

1. Introduction

The paper is devoted to present some new techniques for the a priori error analysis of a new class of hybridizable discontinuous Galerkin methods. The methods in this class use a special type of stabilization function that was first introduced by Lehrenfeld and Schöberl in [15] and we will call them HDG+ methods for simplicity. Instead of attempting to reach for maximal generality, we will focus on linear elasticity on tetrahedral meshes.

We begin by reviewing some existing works. The first HDG method for linear elasticity was proposed in [22]. The method enforces the symmetry of the stress strongly and uses order \( k \) polynomial spaces for all variables. It was then proved in [12] that the method is optimal for displacement but only suboptimal for stress (order of \( k + \frac{1}{2} \)); they also showed the order is sharp on triangular meshes in the numerical experiments. To recover optimal convergence (based on which the superconvergence by post-processing is possible), there are mainly three approaches. The first approach relaxes the strong symmetry condition to weak symmetry [10]. In general, mixed finite element methods based on weak symmetric stress formulations are relatively easier to implement but use more degrees of freedom and therefore can be more costly to compute. The second and the third approaches are all based on strong symmetric stress formulations, where the conservation of angular momentum is automatically preserved. The second approach [3] uses \( M \)-decomposition [4] to enrich the approximation space for stress by adding some basis functions. The approach recovers optimal convergence and also provides an associated tailor projection as an useful tool for error analysis. However, the added basis can be rational functions instead of polynomials and therefore can lead to some difficulties in implementation.

Date: March 29, 2019.

2010 Mathematics Subject Classification. 65N30, 65M60, 74B05.

Key words and phrases. discontinuous Galerkin method, hybridization, optimal convergence, elastic waves.

This work was partially supported by the NSF grant DMS-1818867.
The focus of this paper is the third approach, which is to use HDG+ method for linear elasticity. The method was originally proposed in [15] for diffusion problems, then applied to steady-state linear elasticity in [18]. It uses only polynomial basis functions, achieves one order higher convergence rate for the displacement without post-processing, and its computational complexity is the same to the standard HDG method (order $k$ for both stress and displacement) for their global systems. However, the existing error analysis of HDG+ methods are all based on using orthogonal projections [14, 17, 18, 19, 20], which make the analysis slightly more complicated (it requires a bootstrapping argument to prove convergence of all variables, as opposed to consecutive energy and duality proofs), and detached from the existing projection-based error analysis of HDG methods [2, 5, 6, 7, 8, 9, 10, 13], where specifically constructed projections are used to make the analysis simple and concise. This motivates us to find a new kind of projection for HDG+ for elasticity. The goal of the projection is double: first of all, it takes care of all the off-diagonal terms in the matrix form of the equations (except for a $2 \times 2$ block which is considered as a single diagonal term), and allows us to do a simple energy estimate for some of the variables; second, it facilitates a duality argument where the adjoint equation is fed with the missing error terms to estimate, by using the adjoint projection (consisting of a simple change of sign in the stabilization parameter $\tau$). This has been done in [11] for diffusion problems. A novelty of this paper is the fact that we complement the projection with an error term that does not affect the error bounds or the simplicity of their proofs. We have attached this error term to the projection to make the arguments simpler.

In summary, we have devised a projection for the HDG+ methods for linear elasticity. The projection enables us to: (1) recycle existing projection-based error analysis techniques for the analysis of HDG+ methods; (2) make the error analysis simple and concise; (3) build connections between $M$-decompositions [3, 4] and HDG+ methods. To be more specific, we present a semi-discrete HDG+ method for transient elastic waves that has a uniform-in-time optimal convergence. We show that the proof for optimal convergence can be easily obtained by using the new projection and some existing techniques in traditional HDG methods for evolutionary equations [5]. Moreover, we recover the error estimates for steady-state elasticity [18] and frequency domain elastodynamics [14] by using the projection-based analysis, and we show that the analysis can be simplified in both cases. Since the construction of the HDG+ projection involves first constructing a projection associated to an $M$-decomposition, it also shed some light upon the connections between these two kinds of methods.

To provide a more intuitive view, we put the main procedures of constructing the projection in the flow chart Figure 1. The associated boundary remainder term related to the projection behaves like interpolation error and it depends only on the local projection. As we will see later in the applications, the boundary remainder together with the stabilization parameter play a key role in the optimal convergence of the HDG+ methods, allowing us more flexibility, since now we only need to find a projection such that its associated boundary remainder is small enough to guarantee optimal convergence, instead of enforcing it to vanish, which is the case of the standard projection.

The rest of the paper is organized as follows. In Section 2 we present the main theorem about the projection and its properties. In Section 3 we show the procedures for constructing the projection on the reference element. In Section 4 we develop a systematic approach of changes of variables to obtain the projection on the physical element. In
Since the main tool and one of the principal novelties of this article is the new HDG projection for elasticity, we first introduce its main properties in Theorem 2.1. The reader just interested in the applications can skip the sections devoted to its construction and analysis (Section 3 and 4) and jump directly to how it is used (Section 5, 6 and 7).

To speed up the introduction of the projection we give a quick notation list to be used throughout the article.

- For a domain \( \mathcal{O} \subset \mathbb{R}^3 \), the respective inner products of \( L^2(\mathcal{O}; \mathbb{R}^3) \) and \( L^2(\mathcal{O}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \) will be denoted
  \[
  (u, v)_\mathcal{O} := \int_\mathcal{O} u \cdot v, \quad (\sigma, \rho)_\mathcal{O} := \int_\mathcal{O} \sigma : \rho,
  \]
  where in the latter the colon denotes the Frobenius product of matrices and \( \mathbb{R}^{3 \times 3}_{\text{sym}} \) is the space of symmetric \( 3 \times 3 \) matrices. The norm of both spaces will be denoted \( \| \cdot \|_\mathcal{O} \).

- For a Lipschitz domain \( \mathcal{O} \), the inner product in \( L^2(\partial \mathcal{O}; \mathbb{R}^3) \) will be denoted
  \[
  \langle \mu, \eta \rangle_{\partial \mathcal{O}} := \int_{\partial \mathcal{O}} \mu \cdot \eta
  \]
  and the associated norm will be denoted \( \| \cdot \|_{\partial \mathcal{O}} \).

- The Sobolev seminorm in \( H^m(\mathcal{O}; \mathbb{R}^3) \) and \( H^m(\mathcal{O}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \) will be denoted \( | \cdot |_{m, \mathcal{O}} \).

- The symmetric gradient operator (linearized strain) is given by
  \[
  \varepsilon(u) := \frac{1}{2}(Du + (Du)^\top),
  \]
and the divergence operator \( \text{div} \) will be applied to symmetric-matrix-valued functions by acting on their rows, outputting a column vector-valued function.

- When \( \mathcal{O} \) is a Lipschitz domain, we will consider the space
  \[
  H(\mathcal{O}, \text{div}; \mathbb{R}^{3 \times 3}_{\text{sym}}) := \{ \sigma \in L^2(\mathcal{O}; \mathbb{R}^{3 \times 3}_{\text{sym}}) : \text{div} \sigma \in L^2(\mathcal{O}; \mathbb{R}^3) \},
  \]
  and the normal traction operator \( \gamma_n : H(\mathcal{O}, \text{div}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \to H^{-1/2}(\partial \mathcal{O}; \mathbb{R}^3) \), defined by Betti’s formula
  \[
  \langle \gamma_n \sigma, \nu \rangle_{\partial \mathcal{O}} := (\sigma, \varepsilon(\nu))_\mathcal{O} + (\text{div} \sigma, \nu)_\mathcal{O} \quad \forall \nu \in H^1(\mathcal{O}; \mathbb{R}^3).
  \]
  Here \( \gamma \) is the trace operator, \( H^{-1/2}(\partial \mathcal{O}; \mathbb{R}^3) \) is the dual space of \( H^{1/2}(\partial \mathcal{O}; \mathbb{R}^3) \), and the angled bracket denotes their duality product that extends the \( L^2(\partial \mathcal{O}; \mathbb{R}^3) \) inner product.

For discretization we will consider a sequence of tetrahedral meshes \( \mathcal{T}_h \) and the following notation:

- \( K \) is a tetrahedron, of diameter \( h_K \) and inradius at most \( ch_K \) for a fixed shape-regularity constant \( c > 0 \).
- \( \mathcal{F}(K) \) is the set of faces of \( K \).
- \( \mathcal{P}_k(K; X) \) is the space of \( X \)-valued polynomial functions of degree up to \( k \geq 0 \), where \( X \in \{ \mathbb{R}^3, \mathbb{R}^{3 \times 3}_{\text{sym}} \} \).
- \( \mathcal{R}_k(\partial K; X) := \prod_{F \in \mathcal{F}(K)} \mathcal{P}_k(F; X) \) is the space of piecewise polynomial functions on the boundary of \( K \), with \( X \) as above.
- \( \mathcal{P}_k : L^2(\mathcal{K}; X) \to \mathcal{P}_k(K; X) \) is the orthogonal projection onto the image space, with \( X \) as above.
- \( \mathcal{P}_M : L^2(\partial K; \mathbb{R}^3) \to \mathcal{R}_k(\partial K; \mathbb{R}^3) \) is the orthogonal projection onto the image space.
  This operator will often be applied on the trace of a function in \( H^1(K; \mathbb{R}^3) \) on \( \partial K \).
- \( \tau \in \mathcal{R}_0(\partial K; \mathbb{R}^{3 \times 3}_{\text{sym}}) \) is a so-called stabilization function satisfying
  \[
  C_1 h_K^{-1} \| \mu \|_{\partial K}^2 \leq \langle \tau \mu, \mu \rangle_{\partial K} \leq C_2 h_K^{-1} \| \mu \|_{\partial K}^2 \quad \forall \mu \in L^2(\partial K; \mathbb{R}^3),
  \]
  for fixed positive constants \( C_1 \) and \( C_2 \), independent of \( h \) (i.e., of the particular mesh).

- The wiggled inequality sign \( a \lesssim b \) hides a constant \( a \leq C b \) that is independent of \( h \), while \( a \approx b \) means \( a \leq b \leq a \).

**Theorem 2.1.** For \( k \geq 1 \), there exists a family of projections and associated boundary remainder operators

\[
\Pi : H^1(K; \mathbb{R}^{3 \times 3}_{\text{sym}}) \times H^1(K; \mathbb{R}^3) \to \mathcal{P}_k(K; \mathbb{R}^{3 \times 3}_{\text{sym}}) \times \mathcal{P}_{k+1}(K; \mathbb{R}^3),
\]

\[
R : H^1(K; \mathbb{R}^{3 \times 3}_{\text{sym}}) \times H^1(K; \mathbb{R}^3) \to \mathcal{R}_k(\partial K; \mathbb{R}^3),
\]

depending on \( \tau \in \mathcal{R}_0(\partial K; \mathbb{R}^{3 \times 3}_{\text{sym}}) \), where \( \tau \) is a constant positive definite matrix on each face of \( K \), and if \( (\sigma_K, u_K) = \Pi(\sigma, u; \tau) \) and \( \delta_K = R(\sigma, u; \tau) \), the following conditions hold:

\[
(2.2a) \quad \langle u_K - u, v \rangle_K = 0 \quad \forall v \in \mathcal{P}_{k-1}(K; \mathbb{R}^3),
\]

\[
(2.2b) \quad -\langle \text{div} (\sigma_K - \sigma), w \rangle_K + \langle \tau \mathcal{P}_M (u_K - u), w \rangle_{\partial K} = \langle \delta_K, w \rangle_{\partial K} \quad \forall w \in \mathcal{P}_{k+1}(K; \mathbb{R}^3),
\]

\[
(2.2c) \quad -\langle (\sigma_K - \sigma) n - \tau (u_K - u), \mu \rangle_{\partial K} = \langle \delta_K, \mu \rangle_{\partial K} \quad \forall \mu \in \mathcal{R}_k(\partial K; \mathbb{R}^3).
\]
Moreover, if \( \boldsymbol{\tau} \) satisfies (2.1) and \((\sigma, u) \in H^m(K; \mathbb{R}^{3 \times 3}) \times H^{m+1}(K; \mathbb{R}) \) with \( 1 \leq m \leq k+1 \), then we have the estimates:

\[
\|\sigma_K - \sigma\|_K + h_K^{1/2} \|u_K - u\|_K + h_K^{1/2} \|\delta_K\|_K \leq C h_K^n (\|\sigma\|_{m,K} + \|u\|_{m+1,K}).
\]

The constant \( C \) depends only on the polynomial degree \( k \), the constants \( C_1 \) and \( C_2 \) in (2.1) and the shape-regularity constant \( c \). Hence, the ‘adjoint’ projection can be defined as

\[
\Pi(\sigma, u; -\boldsymbol{\tau}) := (\sigma_K, u_K), \quad \text{where} \quad (\sigma_K, -u_K) := \Pi(\sigma, -u; \boldsymbol{\tau}).
\]

This projection satisfies the properties (2.2) with the same \( \delta_K \), if we substitute \( \boldsymbol{\tau} \) by \(-\boldsymbol{\tau}\).

Note that conditions (2.2) are not enough to define the projection \( \Pi \) but are exactly the ones that will be needed for the applications. As a final note, notice that combining (2.2b) and (2.2c), we have

\[
(\sigma_K - \sigma, \varepsilon(u))_K = 0 \quad \forall u \in \mathcal{P}_k(K; \mathbb{R}^3).
\]

Since \( k \geq 1 \) and \( \mathcal{P}_0(K; \mathbb{R}^{3 \times 3}) \subset \varepsilon(\mathcal{P}_1(K; \mathbb{R}^3)) \), we have

\[
(\sigma_K - \sigma, \boldsymbol{\tau})_K = 0 \quad \forall \boldsymbol{\tau} \in \mathcal{P}_0(K; \mathbb{R}^{3 \times 3}).
\]

3. The Projection in the Reference Element

3.1. Preparatory work on the reference element. In this section we will work on the reference tetrahedron \( \hat{K} := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3, 1 - x_1 - x_2 - x_3 > 0\} \). The trace for vector-valued functions on the reference element will be denoted \( \hat{\gamma} \), and the normal traction operator on the reference element will be denoted \( \hat{\gamma}_n \). We will use shortened notation for the following spaces:

\[
\begin{align*}
\hat{\mathcal{V}} &:= \mathcal{P}_k(\hat{K}; \mathbb{R}^3), & \hat{\mathcal{V}}_- &:= \mathcal{P}_{k-1}(\hat{K}; \mathbb{R}^3), & \hat{\mathcal{V}}_+ &:= \mathcal{P}_{k+1}(\hat{K}; \mathbb{R}^3), \\
\hat{\mathcal{S}} &:= \mathcal{P}_k(\hat{K}; \mathbb{R}^{3 \times 3}), & \hat{\mathcal{S}}_- &:= \{\sigma \in \hat{\mathcal{S}} : \text{div} \, \sigma = 0, \quad \hat{\gamma}_n \sigma = 0\} + \varepsilon(\hat{\mathcal{V}}).
\end{align*}
\]

A new space \( \hat{\mathcal{S}}_+ \) will be defined once we have introduced some tools for it. Note that these constructions can be done directly on any tetrahedron \( K \).

The first of these constructions is a lifting of the traction operator. It will act on the space

\[
L^2_\mathcal{M}(\partial \hat{K}; \mathbb{R}^3) := \{\mu \in L^2(\partial \hat{K}; \mathbb{R}^3) : \langle \mu, \hat{\gamma}_n m \rangle_{\partial \hat{K}} = 0 \quad \forall m \in \mathcal{M}\},
\]

where \( \mathcal{M} \) is the six-dimensional space of infinitesimal rigid motions

\[
\mathcal{M} := \{m(x) := b + Ax : A^T = -A, \quad b \in \mathbb{R}^3\} = \{m \in H^1(\hat{K}; \mathbb{R}^3) : \varepsilon(m) = 0\}.
\]

Note that the trace space of \( \mathcal{M} \) is also six-dimensional, i.e., if \( m \in \mathcal{M} \) vanishes on the boundary of \( \hat{K} \), then \( m = 0 \). We thus define the operator \( \hat{\gamma}_n^+ : L^2_\mathcal{M}(\partial \hat{K}; \mathbb{R}^3) \rightarrow H(\hat{K}, \text{div}; \mathbb{R}^{3 \times 3}) \) by

\[
\begin{align*}
(\hat{\gamma}_n^+ m, \varepsilon(u))_{\hat{K}} &= \langle m, \hat{\gamma}_n v \rangle_{\partial \hat{K}} \quad \forall v \in H^1(\hat{K}; \mathbb{R}^3), \\
(\varepsilon(u), \varepsilon(v))_{\hat{K}} &= \langle \mu, \hat{\gamma}_n v \rangle_{\partial \hat{K}} \quad \forall v \in H^1(\hat{K}; \mathbb{R}^3),
\end{align*}
\]

where \( \mu := \varepsilon(u) \).
or equivalently

\begin{equation}
\hat{\gamma}_n^+ \mu := \sigma, \quad \text{where} \quad \sigma \in \mathcal{E}(H^1(\hat{K}; \mathbb{R}^3)),
\end{equation}

\begin{equation}
(\sigma, \varepsilon(\nu))_{\hat{K}} = \langle \mu, \hat{\gamma} \nu \rangle_{\partial \hat{K}} \quad \forall \nu \in H^1(\hat{K}; \mathbb{R}^3).
\end{equation}

The definition (3.2) is correct since it involves the solution of a coercive variational problem on a quotient space, due to Korn’s Second Inequality. From (3.3) it is clear that

\begin{equation}
\text{div} \hat{\gamma}_n^+ \mu = 0, \quad \hat{\gamma}_n^+ \mu = \mu \quad \forall \mu \in L^2(\partial K; \mathbb{R}^3).
\end{equation}

We next consider the spaces

\[ \hat{M} := \mathcal{R}_k(\partial \hat{K}; \mathbb{R}^3), \quad \hat{\Sigma}_S := \{ \sigma \in \hat{\Sigma} : \text{div} \sigma = 0 \}, \]

\[ \hat{\Theta} := \{ \mu \in \hat{M} : \langle \hat{\gamma}_n \sigma + \hat{\gamma} \mu \rangle_{\partial \hat{K}} = 0 \quad \forall (\sigma, \mu) \in \hat{\Sigma}_S \times \hat{M} \}, \]

and

\[ \hat{\Sigma}_{\text{fill}} := \hat{\gamma}_n^+ \hat{\Theta}. \]

**Theorem 3.1.** The following properties hold:

(a) \( \text{div} \sigma = 0 \) for all \( \sigma \in \hat{\Sigma}_{\text{fill}} \),

(b) \( \hat{\gamma}_n^+ \) is an isomorphism between \( \hat{\Theta} \) and \( \hat{\Sigma}_{\text{fill}} \), and its inverse is \( \hat{\gamma}_n \),

(c) \( \hat{\Sigma}_{\text{fill}} \cap \hat{\Sigma} = \emptyset \),

(d) \( \hat{M} = \hat{\Theta} \oplus \hat{\gamma}_n \hat{\Sigma}_S \oplus \gamma \hat{M} \) with orthogonal sum.

**Proof.** Properties (a) and (b) are easy consequences of (3.4). By (a), \( \hat{\Sigma}_{\text{fill}} \cap \hat{\Theta} = \hat{\Sigma}_{\text{fill}} \cap \hat{\Sigma}_S \) and, therefore, if \( \sigma \in \hat{\Sigma}_{\text{fill}} \cap \hat{\Sigma}_S \), then \( \hat{\gamma}_n \sigma \in \hat{\Theta} \cap \hat{\gamma}_n \hat{\Sigma}_S = \{0\} \) (the latter two sets are orthogonal to each other by definition of \( \hat{\Theta} \)), but then (b) proves that \( \sigma = 0 \), which proves (c). Finally, if \( \sigma \in \hat{\Sigma}_S \) and \( \mu \in \hat{M} \), then

\[ \langle \hat{\gamma}_n \sigma, \gamma \mu \rangle_{\partial \hat{K}} = (\text{div} \sigma, \mu)_{\hat{K}} + (\sigma, \varepsilon(\mu))_{\hat{K}} = 0, \]

which shows that the sum \( \hat{\gamma}_n \hat{\Sigma}_S + \gamma \hat{M} \) is orthogonal. Since \( \hat{\Theta} \) is the orthogonal complement of the latter set, (d) is proved.

The Cockburn-Fu discrete pair for elasticity \[ \hat{\Sigma} \] is given by the spaces

\[ \hat{\Sigma}_+ := \hat{\Sigma} \oplus \hat{\Sigma}_{\text{fill}} \]

(the sum is direct because of Theorem 3.1) and \( \hat{V} \). In their context of \( M \)-decompositions for elasticity, the following result is a key one, that we will need to work with our extended pair \( \hat{\Sigma}_+ \times \hat{\Sigma}_- \).

**Theorem 3.2.** The following properties hold:

(a) \( \varepsilon(\hat{V}) \subset \hat{\Sigma}_- \),

(b) \( \text{div} \hat{\Sigma}_+ \subset \hat{\Sigma}_- \),

(c) \( \hat{\gamma}_n \hat{\Sigma}_+ + \hat{\gamma} \hat{V} \subset \hat{M} \),

(d) \( \dim \hat{M} = \dim \hat{\Sigma}_- + \dim \hat{V}_- \), where the orthogonal complement \( \hat{\Sigma}_- \) is taken in \( \hat{\Sigma}_- \), while \( \hat{V}_- \) is taken in \( \hat{V} \),

(e) \( \hat{M} = \hat{\gamma}_n \hat{\Sigma}_- \oplus \gamma \hat{V}_- \), with orthogonal sum,

(f) The map \( \hat{\Sigma}_- \times \hat{\Sigma}_- \ni (\sigma, \mu) \mapsto \hat{\gamma}_n \sigma + \hat{\gamma} \mu \in \hat{M} \) is an isomorphism.
Proof. Property (a) follows by definition and (b) is a simple consequence of Theorem 3.1(a). To show (c), note simply that \( \hat{\gamma}_n \hat{\Sigma} = \hat{\Theta} \subset \hat{M} \), while all other elements are polynomials of degree less than or equal to \( k \).

To prove (d) to (f), it will be convenient to identify the set

\[
(3.5) \quad \hat{\Sigma}_S^0 := \{ \sigma \in \hat{\Sigma} : \text{div} \sigma = 0, \ \hat{\gamma}_n \sigma = 0 \} = \ker \hat{\gamma}_n \hat{\Sigma}_S^0,
\]

which appeared in the definition of \( \hat{\Sigma}_- \) (3.1).

By Theorem 3.1(d) and (3.5), we have

\[
(3.6) \quad \dim \hat{M} = \dim \hat{\Theta} + \dim \hat{\gamma}_n \hat{\Sigma}_S + \dim \gamma \mathcal{M} = \dim \hat{\Sigma}^\dagger + \dim \hat{\Sigma}_S - \dim \hat{\Sigma}_S^0 + \dim \mathcal{M}.
\]

Using the definitions of \( \hat{\Sigma}_\pm \) we have

\[
(3.7) \quad \dim \hat{\Sigma}_\pm^\dagger = \dim \hat{\Sigma} - \dim \hat{\Sigma}_\pm = \dim \hat{\Sigma} + \dim \hat{\Sigma}^\dagger - (\dim \hat{\Sigma}_S^0 + \dim \varepsilon(\hat{V}))
\]

\[
= \dim \hat{V}_+ + \dim \hat{\Sigma}_S + \dim \hat{\Sigma}^\dagger - (\dim \hat{\Sigma}_S^0 + \dim \hat{V} - \dim \mathcal{M}),
\]

where in the last equality we have applied that \( \text{div} : \hat{\Sigma} \to \hat{V}_- \) is onto (this is easy to prove), its kernel is \( \hat{\Sigma}_S \), and the kernel of \( \varepsilon : \hat{V} \to \varepsilon(\hat{V}) \) is \( \mathcal{M} \). The equalities (3.6) and (3.7) prove (d).

Due to parts (a) and (b) of this theorem, we have

\[
\langle \hat{\gamma}_n \sigma, \hat{\gamma} u \rangle_{\partial \hat{\mathcal{K}}} = 0 \quad \forall \sigma \in \hat{\Sigma}_\pm^\dagger, \ u \in \hat{V}_-, 
\]

which proves that the sum of \( \hat{\gamma}_n \hat{\Sigma}_\dagger \) and \( \hat{\gamma} \hat{V}_- \) is orthogonal. Since \( \hat{\gamma} : \hat{V}_- \to \hat{M} \) is injective (this result is known; see, for instance, [21] and [11]), the property (e) will follow from having proved that \( \hat{\gamma}_n : \hat{\Sigma}_\dagger \to \hat{M} \) is injective.

We first prove the following technical result: if \( \sigma \in \hat{\Sigma}_+ \) satisfies \( \text{div} \sigma = 0 \) and \( \hat{\gamma}_n \sigma = 0 \), then \( \sigma \in \hat{\Theta} \), i.e., the component in \( \hat{\Sigma}^\dagger \) vanishes. To do that, take \( \sigma = \sigma_1 + \sigma_2 \in \hat{\Theta} + \hat{\Sigma}^\dagger \) and note that Theorem 3.1(a) shows that if \( \text{div} \sigma = 0 \), then \( \text{div} \sigma_1 = 0 \). Since

\[
0 = \hat{\gamma}_n \sigma = \hat{\gamma}_n \sigma_1 + \hat{\gamma}_n \sigma_2 \in \hat{\gamma}_n \hat{\Sigma}_S \oplus \hat{\Theta},
\]

by Theorem 3.1(d), it follows that \( \hat{\gamma}_n \sigma_1 = \hat{\gamma}_n \sigma_2 = 0 \), which proves that \( \sigma_2 = 0 \) by Theorem 3.1(b).

The proof of injectivity of \( \hat{\gamma}_n | \hat{\Sigma}_\pm \) is then simple. If \( \sigma \in \hat{\Sigma}_- \) satisfies \( \hat{\gamma}_n \sigma = 0 \), then by part (a)

\[
(\text{div} \sigma, v)_{\hat{\mathcal{K}}} = -(\sigma, \varepsilon(v))_{\hat{\mathcal{K}}} = 0 \quad \forall v \in \hat{V}
\]

and, taking \( v = \text{div} \sigma \) (by part (b)), we prove that \( \text{div} \sigma = 0 \). Therefore, \( \sigma \in \hat{\Sigma}_S \subset \hat{\Sigma}_- \) and hence \( \sigma = 0 \). This finishes the proof of (e) and (f).

For the rest of this section we fix the stabilization parameter \( \tau \in \mathcal{R}_0(\partial \hat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \) such that

\[
(3.8) \quad \tau |_{\mathcal{F}} \text{ is positive definite} \quad \forall F \in \mathcal{F}(\hat{\mathcal{K}}).
\]
3.2. A projection on an extended space. Given \((\sigma, u) \in H^1(\hat{K}; \mathbb{R}^{3 \times 3}) \times H^1(\hat{K}, \mathbb{R}^3)\), we look for

\[
\hat{\Pi}_0(\sigma, u; \tau) := (\sigma_{\hat{K}}, u_{\hat{K}}) \in \hat{\Sigma}_+ \times \hat{V}_+
\]
satisfying

\[
(3.9a)
(\sigma_{\hat{K}} - \sigma, \theta)_{\hat{K}} = 0 \quad \forall \theta \in \hat{\Sigma}_-,
\]

\[
(3.9b)
(\sigma_{\hat{K}} - \sigma, \theta)_{\hat{K}} = 0 \quad \forall \theta \in \hat{\Sigma}_-,
\]

\[
(3.9c)
\langle \gamma_y (\sigma_{\hat{K}} - \sigma) - \tau_{\hat{K}} \gamma_y u_{\hat{K}} - u, \mu \rangle_{\hat{K}} = 0 \quad \forall \mu \in \hat{M},
\]

\[
(3.9d)
- (\text{div} (\sigma_{\hat{K}} - \sigma), w)_{\hat{K}} + \langle \tau_{\hat{K}} \gamma_y (u_{\hat{K}} - u), \gamma_y w \rangle_{\hat{K}} = 0 \quad \forall w \in \hat{V}_+,
\]

where \(\hat{V}^\perp\) is the orthogonal complement of \(\hat{V}\) in \(\hat{V}_+\). Note that (3.9c) is equivalent to

\[
(\sigma_{\hat{K}} - \sigma, \varepsilon(w))_{\hat{K}} - \langle \gamma_y (\sigma_{\hat{K}} - \sigma) - \tau_{\hat{K}} \gamma_y u_{\hat{K}} - u, \gamma_y w \rangle_{\hat{K}} = 0 \quad \forall w \in \hat{V}_+,
\]

and, since \(\varepsilon(\hat{V}) \subset \hat{\Sigma}_-\) and \(\gamma \hat{V} \subset \hat{M}\), equations (3.9c) and (3.9d) imply that we can substitute (3.9c) by the condition

\[
(3.10)
- (\text{div} (\sigma_{\hat{K}} - \sigma), w)_{\hat{K}} + \langle \tau_{\hat{K}} \gamma_y (u_{\hat{K}} - u), \gamma_y w \rangle_{\hat{K}} = 0 \quad \forall w \in \hat{V}_+,
\]

which contains redundant restrictions already imposed in the other equations. Note also that the projection \(\tau_{\hat{K}} \gamma_y\) can be eliminated in (3.9d) but not in (3.9c) or in (3.10), and that the bilinear form

\[
\langle \tau_{\hat{K}} \gamma_y \mu, \eta \rangle_{\hat{K}}
\]
is symmetric, bounded, and positive definite in \(L^2(\partial \hat{K}; \mathbb{R}^3)\). We next prove that \(\hat{\Pi}_0\) is actually a well-defined projection onto \(\hat{\Sigma}_+ \times \hat{V}_+\).

**Proposition 3.3.** The process of defining the projection \(\hat{\Pi}_0\) in (3.9) is equivalent to a square invertible linear system.

**Proof.** Note first that by Theorem 3.2(d)

\[
\dim \hat{V}_- + \dim \hat{\Sigma}_- + \dim \hat{M} + \dim \hat{V}^\perp = \dim \hat{V}_- + \dim \hat{\Sigma}_- + \dim \hat{V}^\perp + \dim \hat{\Sigma}_+ - \dim \hat{V}
\]

\[= \dim \hat{\Sigma}_+ + \dim \hat{V}_+,
\]

which proves that (3.9) is equivalent to a linear system with as many equations as unknowns. We thus only need to prove uniqueness of solution.

A homogeneous solution of (3.9) is a pair \((\sigma_{\hat{K}}, u_{\hat{K}}) \in \hat{\Sigma}_+ \times \hat{V}_+\) satisfying (recall how (3.10) is a consequence of (3.9))

\[
(3.11a)
(u_{\hat{K}}, v)_{\hat{K}} = 0 \quad \forall v \in \hat{V}_-,
\]

\[
(3.11b)
(\sigma_{\hat{K}}, \theta)_{\hat{K}} = 0 \quad \forall \theta \in \hat{\Sigma}_-,
\]

\[
(3.11c)
\langle \gamma_y (\sigma_{\hat{K}} - \sigma) - \tau_{\hat{K}} \gamma_y u_{\hat{K}} - u, \mu \rangle_{\hat{K}} = 0 \quad \forall \mu \in \hat{M},
\]

\[
(3.11d)
- (\text{div} \sigma_{\hat{K}}, w)_{\hat{K}} + \langle \tau_{\hat{K}} \gamma_y u_{\hat{K}}, \gamma_y w \rangle_{\hat{K}} = 0 \quad \forall w \in \hat{V}_+.
\]
We now take $w = u_\widehat{\mathcal{K}} \in \hat{\mathbf{V}}^\perp_-$ (with the complement in $\hat{\mathbf{V}}_+$; also see (3.11a)) in (3.11d), recalling that $\text{div} \sigma_\widehat{\mathcal{K}} \in \hat{\mathbf{V}}_-$ (cf. Theorem 3.2(b)) and obtain

$$\langle \tau \widehat{\mathbf{M}} \gamma u_\widehat{\mathcal{K}}; \gamma u_\widehat{\mathcal{K}} \rangle_{\partial \widehat{\mathcal{K}}} = \langle \tau \widehat{\mathbf{M}} \gamma u_\widehat{\mathcal{K}}; \widehat{\mathbf{M}} \gamma u_\widehat{\mathcal{K}} \rangle_{\partial \widehat{\mathcal{K}}} = 0.$$

This argument uses that $\tau$ is piecewise constant, so that multiplication by $\tau$ is an endomorphism in $\widehat{\mathbf{M}}$. Since $\tau$ is positive definite on each face, this proves that $\widehat{\mathbf{M}} \gamma u_\widehat{\mathcal{K}} = 0$.

Using the above conclusion and taking $\mu = \widehat{\gamma}_n \sigma_\widehat{\mathcal{K}} \in \widehat{\mathcal{M}}$ in (3.11c) (cf. Theorem 3.2(c)), it follows that $\widehat{\gamma}_n \sigma_\widehat{\mathcal{K}} = 0$. Given the fact that $\sigma_\widehat{\mathcal{K}} \in \hat{\Sigma}^\perp$ (by (3.11b)) and Theorem 3.2(f), this proves that $\sigma_\widehat{\mathcal{K}} = 0$.

Note finally that $\text{div} \varepsilon(u_\widehat{\mathcal{K}}) \in \hat{\mathbf{V}}_-$ (Theorem 3.2(b)) and $\widehat{\gamma}_n \varepsilon(u_\widehat{\mathcal{K}}) \in \widehat{\mathcal{M}}$ (Theorem 3.2(c)), so that

$$\langle \varepsilon(u_\widehat{\mathcal{K}}), \varepsilon(u_\widehat{\mathcal{K}}) \rangle_{\widehat{\mathcal{K}}} = - \langle \text{div} \varepsilon(u_\widehat{\mathcal{K}}), u_\widehat{\mathcal{K}} \rangle_{\widehat{\mathcal{K}}} + \langle \widehat{\gamma}_n \varepsilon(u_\widehat{\mathcal{K}}), \gamma u_\widehat{\mathcal{K}} \rangle_{\partial \widehat{\mathcal{K}}} = 0,$$

and therefore $u_\widehat{\mathcal{K}} \in \mathcal{M}$. This implies that $\gamma u_\widehat{\mathcal{K}} = \widehat{\mathbf{M}} \gamma u_\widehat{\mathcal{K}} = 0$ and, therefore, $u_\widehat{\mathcal{K}} = 0$, which finishes the proof. \hfill $\Box$

Looking at the proof of Proposition 3.3 it is clear that we could have also defined the projection for any $\tau \in \mathcal{R}_0(\partial \widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}})$ such that $\langle \tau \rangle_F$ is negative definite on each face.

**Proposition 3.4 (Stability).** For any $\tau_{\max} \geq \tau_{\min} > 0$ and $k \geq 1$, there exists $C = C(\tau_{\max}, \tau_{\min}, k)$ such that if $(\sigma_\widehat{\mathcal{K}}, u_\widehat{\mathcal{K}}) \in \widehat{\Pi}_0(\sigma, u, \tau)$, then

$$\|\sigma_\widehat{\mathcal{K}}\|_{\widehat{\mathcal{K}}} + \|u_\widehat{\mathcal{K}}\|_{\widehat{\mathcal{K}}} \leq C(\|\sigma\|_{1, \widehat{\mathcal{K}}} + \|u\|_{1, \widehat{\mathcal{K}}}) \quad \forall (\sigma, u) \in H^1(\widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}} \times \mathbb{R}^3),$$

whenever $\tau \in \mathcal{R}_0(\partial \widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}})$ satisfies

$$\tau_{\min} \leq \mu \cdot \langle \tau \rangle_F \mu \leq \tau_{\max} \quad \forall \mu \in \mathbb{R}^3, |\mu| = 1, \quad \forall F \in \mathcal{F}(\widehat{\mathcal{K}}).$$

**Proof.** Let $l$ be the number of faces of $\widehat{\mathcal{K}}$. Numbering the faces of $\widehat{\mathcal{K}}$ and the entries of a symmetric matrix with indices from one to six, we can identify $\mathcal{R}_0(\partial \widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \equiv (\mathbb{R}^{3 \times 3}_{\text{sym}})^l \equiv \mathbb{R}^{6l}$. We can thus make the identification

$$\{\tau \in \mathcal{R}_0(\partial \widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}}) : \tau \text{ satisfies (3.8)}\} \equiv \mathcal{O} \subset \mathbb{R}^{6l},$$

where $\mathcal{O}$ is an open set. We can also identify

$$\{\tau \in \mathcal{R}_0(\partial \widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}}) : \tau \text{ satisfies (3.13)}\} \equiv \mathcal{K} \subset \mathbb{R}^{6l},$$

where $\mathcal{K} \subset \mathcal{O}$ is compact. Consider now the continuous trial and test spaces

$$\mathcal{U} := H^1(\widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \times H^1(\widehat{\mathcal{K}}; \mathbb{R}^3),$$

$$\mathbf{V} := L^2(\widehat{\mathcal{K}}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \times L^2(\widehat{\mathcal{K}}; \mathbb{R}^3) \times L^2(\partial \widehat{\mathcal{K}}; \mathbb{R}^3) \times H^1(\widehat{\mathcal{K}}; \mathbb{R}^3),$$

and their discrete counterparts

$$\mathbf{U}_k := \hat{\Sigma}_+ \times \hat{\mathbf{V}}_+,$$

$$\mathbf{V}_k := \hat{\mathbf{V}}_- \times \hat{\Sigma}_- \times \widehat{\mathbf{M}} \times \hat{\mathbf{V}}^\perp_+.$$

Equations (3.3) can be rephrased in the following form: given $U \in \mathbf{U}$, find $\widehat{\Pi}_0 U = U_k$ such that

$$a_0(\widehat{\Pi}_0 U - U, V) + \sum_{j=1}^{6l} x_j a_j(\widehat{\Pi}_0 U - U, V) = 0 \quad \forall V \in \mathbf{V}_k.$$
where
\[ a_j : \mathcal{U} \times \mathcal{V} \rightarrow \mathbb{R} \quad j = 0, \ldots, 6l \]
are bounded \( \tau \)-independent (but dependent on \( k \) through the operator \( P_{M,K} \)) bilinear forms and the identification \( \tau \equiv x = (x_1, \ldots, x_{6l}) \) is done as above. Equations (3.14) are uniquely solvable for every \( x \in \mathcal{O} \) (this is a restatement of Proposition 3.3) and define a bounded linear operator \( T_k(x) : \mathcal{U} \rightarrow \mathcal{U}_k \). The function \( T_k \), from \( \mathcal{O} \) to the space of bounded linear operators from \( \mathcal{U} \) to \( \mathcal{U}_k \), is rational and therefore bounded on the compact set \( \mathcal{K} \).

We can thus bound
\[ \| T_k(x)U \| \leq C(k, \mathcal{K}, \tau) \| U \| \quad \forall U \in \mathcal{U}, \]
where \( \| \cdot \| \) is any norm we choose in \( \mathcal{U}_k \). If we select the norm
\[ \|(\sigma, u)\| := \|\sigma\|_K + \|u\|_K, \]
in \( \mathcal{U}_k \), then (3.13) becomes (3.12).

A simple argument shows that algebraic condition (3.13) is equivalent to asking that the spectrum of \( \tau|_F \) is contained in \([\tau_{\min}, \tau_{\max}]\) for all \( F \in \mathcal{F}(\hat{K}) \) and also to the inequality
\[ \tau_{\min} \| \mu \|^2_{c\hat{K}} \leq \langle \tau \mu, \mu \rangle_{c\hat{K}} \leq \tau_{\max} \| \mu \|^2_{c\hat{K}} \quad \forall \mu \in L^2(\partial \hat{K}; \mathbb{R}^3). \]

3.3. The projection and the remainder. Let now \( P_{\Sigma} : L^2(K; \mathbb{R}^{3 \times 3}_{\text{sym}}) \rightarrow \hat{\Sigma} \) be the orthogonal projection onto \( \hat{\Sigma} \). For \( (\sigma, u) \in H^1(\hat{K}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \times H^1(\hat{K}; \mathbb{R}^3) \) we define
\[ \hat{\Pi}(\sigma, u; \tau) := (\sigma_{cK}^*, u_{cK}^*) \in \hat{\Sigma} \times \hat{V}_+, \quad \hat{R}(\sigma, u; \tau) := \hat{\gamma}_n(\sigma_{cK} - \sigma_{cK}^*) \in \hat{M}, \]
where
\[ (\sigma_{cK}^*, u_{cK}^*) := P_{\Sigma}^* \sigma_{cK}, \quad (\sigma_{cK}^*, u_{cK}^*) = \hat{\Pi}_0(\sigma, u; \tau). \]

Since \( \hat{\Sigma} \times \hat{V}_+ \subset \hat{\Sigma}_+ \times \hat{V}_+ \), it follows that \( \hat{\Pi}_0(\theta, v; \tau) = (\theta, v) \) for all \( (\theta, v) \in \hat{\Sigma} \times \hat{V}_+ \) and therefore
\[ \hat{\Pi}(\theta, v; \tau) = (\theta, v), \quad \hat{R}(\theta, v; \tau) = 0 \quad \forall (\theta, v) \in \hat{\Sigma} \times \hat{V}_+. \]

In particular, \( \hat{\Pi} \) is a projection onto \( \hat{\Sigma} \times \hat{V}_+ \). Note that we do not give a set of equations to define \( \hat{\Pi} \), which is given as the application of \( \hat{\Pi}_0 \) followed by an orthogonal projection applied to the first component of the output. However, the following equations relate the projection \( \hat{\Pi} \) and the associated remainder \( \hat{R} \).

**Proposition 3.5.** Let \( \tau \in \mathcal{R}_0(\partial \hat{K}; \mathbb{R}^{3 \times 3}_{\text{sym}}) \) satisfy (3.8), \( (\sigma_{cK}^*, u_{cK}^*) := \hat{\Pi}(\sigma, u; \tau) \) and \( \delta_{cK} := \hat{R}(\sigma, u; \tau) \) satisfy
\[ \begin{align*}
(3.19a) \quad (u_{cK}^* - u, \text{div } \theta)_{cK} &= 0 \quad \forall \theta \in \hat{\Sigma}, \\
(3.19b) \quad -(\text{div } (\sigma_{cK}^* - \sigma), w)_{cK} + \langle \tau P_{M} \hat{\gamma}(u_{cK} - u), \hat{\gamma} w \rangle_{cK} &= \langle \delta_{cK} - \hat{\gamma} w, \delta_{cK} \rangle_{cK} \quad \forall w \in \hat{V}_+, \\
(3.19c) \quad -\langle \hat{\gamma}_n(\sigma_{cK}^* - \sigma) - \tau P_{M} \hat{\gamma}(u_{cK} - u), \mu \rangle_{cK} &= \langle \delta_{cK} - \hat{\gamma} \mu, \delta_{cK} \rangle_{cK} \quad \forall \mu \in \hat{M}. 
\end{align*} \]

**Proof.** Following the definition (3.17), we introduce the intermediate projection \( (\sigma_{cK}^*, u_{cK}^*) = \hat{\Pi}_0(\sigma, u; \tau) \) so that \( \sigma_{cK}^* = P_{\Sigma}^* \sigma_{cK} \) and \( \delta_{cK} = \hat{\gamma}_n(\sigma_{cK} - \sigma_{cK}^*). \)
Since $\nabla \hat{\Sigma} \subset \hat{V}_-$ (this is a direct consequence of the definitions), (3.19a) is a consequence of (3.9b). Note also that $\varepsilon(\hat{V}_+) \subset \hat{\Sigma}$ (again by definition) and therefore
\[
(\sigma^c_{\hat{K}} - \sigma_{\hat{K}}, \varepsilon(w))_{\hat{K}} = 0 \quad \forall w \in \hat{V}_+.
\]
The identity (3.19c) is a direct consequence of (3.9d). Finally, by (3.10) and (3.20)
\[
-(\nabla (\sigma^c_{\hat{K}} - \sigma), w)_{\hat{K}} + \langle \tau P_{\hat{M}} \hat{\gamma}(u_{\hat{K}} - u), \hat{\gamma}w \rangle_{\hat{K}} = -(\nabla (\sigma^c_{\hat{K}} - \sigma), w)_{\hat{K}}
\]
\[
= -\langle \gamma_n(\sigma^c_{\hat{K}} - \sigma_{\hat{K}}), \hat{\gamma}w \rangle_{\hat{K}},
\]
which proves (3.19b).

\[\square\]

**Proposition 3.6** (Estimate in the reference element). For any $\tau_{\max} \geq \tau_{\min} > 0$, and integers $k \geq 1$, $1 \leq m \leq k + 1$, there exists $C = C(\tau_{\max}, \tau_{\min}, k, m)$ such that if
\[
(\sigma^c_{\hat{K}}, u_{\hat{K}}) = \hat{\Pi}_0(\sigma, u; \eta \tau), \quad (\sigma, u) \in H^m(\hat{K}; \mathbb{R}^{3 \times 3}) \times H^{m+1}(\hat{K}; \mathbb{R}^3) \quad 0 \neq \eta \in \mathbb{R},
\]
\[
\delta_{\hat{K}} = \hat{R}(\sigma, u; \eta \tau),
\]
with $\tau \in \mathcal{R}_0(\hat{K}; \mathbb{R}^{3 \times 3})$ satisfying (3.13), then
\[
\|\sigma - \sigma^c_{\hat{K}}\|_{\hat{K}} + |\eta| \|u - u_{\hat{K}}\|_{\hat{K}} + \|\delta_{\hat{K}}\|_{\hat{K}} \leq C(\|\sigma\|_{1,\hat{K}} + |\eta| \|u\|_{1,\hat{K}}).
\]

**Proof.** Let $(\sigma^c_{\hat{K}}, u_{\hat{K}}) = \hat{\Pi}_0(\sigma, u; \eta \tau)$. Note that $(\sigma^c_{\hat{K}}, \eta u_{\hat{K}}) = \hat{\Pi}_0(\sigma, \eta u; \tau)$ (this is obvious from the equations that define $\hat{\Pi}_0$, namely (3.3)) and therefore
\[
(\sigma^c_{\hat{K}}, \eta u_{\hat{K}}) = \hat{\Pi}_0(\sigma, \eta u; \tau), \quad \delta_{\hat{K}} = \hat{R}(\sigma, \eta u; \tau).
\]
By Proposition 3.3 we have
\[
\|\sigma^c_{\hat{K}}\|_{\hat{K}} + |\eta| \|u_{\hat{K}}\|_{\hat{K}} \leq C(\|\sigma\|_{1,\hat{K}} + |\eta| \|u\|_{1,\hat{K}}),
\]
for some constant $C$ depending only on $\tau_{\min}$, $\tau_{\max}$ and $k$. Since $\hat{\gamma}_n : \hat{\Sigma}_+ \rightarrow \hat{M}$ is bounded, there exists a constant $D = D(k)$ such that
\[
\|\delta_{\hat{K}}\|_{\hat{K}} \leq D \|\sigma - \sigma^c_{\hat{K}}\|_{\hat{K}} \leq 2D \|\sigma - \sigma^c_{\hat{K}}\|_{\hat{K}} + D \|\sigma - P_{\hat{\Sigma}} \sigma\|_{\hat{K}}.
\]
Taking now $\theta \in \hat{\Sigma}$ and $v \in \hat{V}_+$, and applying (3.18) to the pair $(\theta, \eta v)$, we have
\[
\|\sigma - \sigma^c_{\hat{K}}\|_{\hat{K}} + |\eta| \|u - u_{\hat{K}}\|_{\hat{K}} \leq (1 + C)(\|\sigma - \theta\|_{1,\hat{K}} + |\eta| \|u - v\|_{1,\hat{K}}),
\]
and
\[
\|\delta_{\hat{K}}\|_{\hat{K}} \leq 2D(1 + C)(\|\sigma - \theta\|_{1,\hat{K}} + |\eta| \|u - v\|_{1,\hat{K}}) + D \|\sigma - P_{\hat{\Sigma}} \sigma\|_{\hat{K}}.
\]
Note that $P_{\hat{\Sigma}}$ is the $L_2$ projection onto $\hat{\Sigma}$. Therefore, there exists a constant $C' = C'(\tau_{\min}, \tau_{\max}, k)$ such that
\[
\|\sigma - \sigma^c_{\hat{K}}\|_{\hat{K}} + |\eta| \|u - u_{\hat{K}}\|_{\hat{K}} + \|\delta_{\hat{K}}\|_{\hat{K}} \leq C'(\inf_{\theta \in \hat{\Sigma}} \|\sigma - \theta\|_{1,\hat{K}} + |\eta| \inf_{v \in \hat{V}_+} \|u - v\|_{1,\hat{K}}).
\]
Finally, notice that $\sigma - \sigma^c_{\hat{K}} = \sigma - P_{\hat{\Sigma}} \sigma + P_{\hat{\Sigma}} (\sigma - \sigma^c_{\hat{K}})$. The result follows now by a compactness argument (Bramble-Hilbert lemma).
\[\square\]
4. The projection in the physical elements

4.1. Pull-backs and push-forwards. In this section we derive a systematic approach to changes of variables for vector- and matrix-valued functions from a general shape-regular tetrahedron to the reference element. The language mimics that of [21] (or [11]). Let \( K \) be a tetrahedron and \( F : \tilde{K} \to K \) be an invertible affine map from the reference element to \( K \). We will denote \( B := DF \) and \( J := \det B \). We also consider the piecewise constant function \( a : \partial \tilde{K} \to (0, \infty) \) such that for all integrable \( \phi \),

\[
\int_{\partial \tilde{K}} \phi = \int_{\partial K} |a| \phi \circ F.
\]

The trace and normal trace operators on \( K \) will be denoted \( \gamma \) and \( \gamma_n \). Given \( u, u^* : K \to \mathbb{R}^3 \), \( \sigma, \sigma^* : K \to \mathbb{R}^{3 \times 3}_\text{sym} \), \( \mu, \mu^* : \partial K \to \mathbb{R}^3 \), we define

\[
\hat{u} := B^T u \circ F : \tilde{K} \to \mathbb{R}^3, \quad \hat{\sigma} := |J|B^{-1}(\sigma \circ F)B^{-T} : \tilde{K} \to \mathbb{R}^{3 \times 3}_\text{sym}, \quad \hat{\mu} := B^T \mu \circ F : \partial \tilde{K} \to \mathbb{R}^3.
\]

where \( B^{-T} = (B^T)^{-1} \). The following properties are easy to prove.

**Proposition 4.1** (Change of variables in integrals). We have the following identities

\[
\begin{align*}
(4.1a) \quad & (u, u^*)_K = (\hat{u}, \hat{u}^*)_\tilde{K} \quad \forall u, u^* \in L^2(K; \mathbb{R}^3), \\
(4.1b) \quad & (\sigma, \sigma^*)_K = (\hat{\sigma}, \hat{\sigma}^*)_\tilde{K} \quad \forall \sigma, \sigma^* \in L^2(K; \mathbb{R}^{3 \times 3}_\text{sym}), \\
(4.1c) \quad & \langle \mu, \mu^* \rangle_{\partial K} = \langle \hat{\mu}, \hat{\mu}^* \rangle_{\partial \tilde{K}} \quad \forall \mu, \mu^* \in L^2(\partial K; \mathbb{R}^3).
\end{align*}
\]

The next group of results about changes of variables involve the interaction of integrals with differential operators or trace operators.

**Proposition 4.2** (Change of variables in bilinear forms). We have the following identities:

\[
\begin{align*}
(4.2a) \quad & (\sigma, \varepsilon(u))_K = (\hat{\sigma}, \hat{\varepsilon}(\hat{u}))_{\tilde{K}} \quad \forall \sigma \in L^2(K; \mathbb{R}^{3 \times 3}_\text{sym}), \forall u \in H^1(K; \mathbb{R}^3), \\
(4.2b) \quad & (u, \text{div} \, \sigma)_K = (\hat{u}, \text{div} \, \hat{\sigma})_{\tilde{K}} \quad \forall \sigma \in H^1(K; \mathbb{R}^{3 \times 3}_\text{sym}), \forall u \in L^2(K; \mathbb{R}^3), \\
(4.2c) \quad & \langle \gamma_n \sigma, \gamma u \rangle_{\partial K} = \langle \hat{\gamma}_n \hat{\sigma}, \hat{\gamma} \hat{u} \rangle_{\partial \tilde{K}} \quad \forall \sigma \in H^1(K; \mathbb{R}^{3 \times 3}_\text{sym}), \forall u \in H^1(K; \mathbb{R}^3), \\
(4.2d) \quad & \langle \gamma_n \sigma, \mu \rangle_{\partial K} = \langle \hat{\gamma}_n \hat{\sigma}, \hat{\mu} \rangle_{\partial \tilde{K}} \quad \forall \sigma \in H^1(K; \mathbb{R}^{3 \times 3}_\text{sym}), \forall \mu \in L^2(\partial K; \mathbb{R}^3), \\
(4.2e) \quad & \langle \tau \gamma u, \mu \rangle_{\partial K} = \langle \hat{\tau} \hat{\gamma} \hat{u}, \hat{\mu} \rangle_{\partial \tilde{K}} \quad \forall u \in H^1(K; \mathbb{R}^3), \forall \mu \in L^2(\partial K; \mathbb{R}^3), \\
(4.2f) \quad & \langle \tau \gamma u, \gamma v \rangle_{\partial K} = \langle \hat{\tau} \hat{\gamma} \hat{u}, \hat{\gamma} \hat{v} \rangle_{\partial \tilde{K}} \quad \forall u, v \in H^1(K; \mathbb{R}^3).
\end{align*}
\]

Here \( \tau \in \mathcal{R}_0(\partial K; \mathbb{R}^{3 \times 3}_\text{sym}) \) and

\[
\hat{\tau} := |a|B^{-1}(\tau \circ F)B^{-T} \in \mathcal{R}_0(\partial \tilde{K}; \mathbb{R}^{3 \times 3}_\text{sym}).
\]

**Proof.** Using the definitions, it is easy to prove that

\[
\varepsilon(u) = \hat{\varepsilon}(\hat{u}), \quad \hat{\gamma} \hat{u} = \gamma \hat{u}, \quad \hat{\tau} \hat{\mu} = \tau \hat{\mu}.
\]
Then (4.2a), (4.2c), and (4.2b) are easy consequences of Proposition 4.1. Let now \( B = (b_{ij})_{i,j=1}^3 \) and \( A = (a_{ij})_{i,j=1}^3 = B^{-1} \). Using implicit summation for repeated indices, we have for each \( i \)

\[
(\text{div} \, \hat{\sigma})_i = \partial_{\hat{\xi}} \hat{\sigma}_{ki} = |J|a_{kl}a_{im} \partial_{\hat{\xi}} (\sigma_{im} \circ F) \\
= |J|a_{kl}a_{im}b_{jk} (\partial_{\hat{\xi}} \sigma_{im} \circ F) = |J|\delta_{ij}a_{im} (\partial_{\hat{\xi}} \sigma_{im} \circ F) \\
= |J|a_{im} (\partial_{\hat{\xi}} \sigma_{jm} \circ F) = |J|a_{im} (\text{div} \, \sigma)_m \circ F \\
= |J|(B^{-1} \text{div} \, \sigma)_i \circ F = (\text{div} \, \sigma)_i
\]

and therefore \( \text{div} \, \hat{\sigma} = \text{div} \, \sigma \). This and (4.1a) prove (4.2b). Note next that, using identities we have already proved and Proposition 4.1 we have

\[
\langle \gamma_n \sigma, \hat{\gamma} \hat{u} \rangle_{\partial K} = \langle \gamma_n \sigma, \gamma \hat{u} \rangle_{\partial K} = \langle \gamma_n \sigma, \gamma u \rangle_{\partial K} = (\text{div} \, \sigma, u)_K + (\sigma, \nabla u)_K \\
= (\text{div} \, \hat{\sigma}, \hat{u})_\hat{K} + (\hat{\sigma}, \hat{\nabla} \hat{u})_\hat{K} = \langle \gamma_n \sigma, \gamma \hat{u} \rangle_{\partial \hat{K}},
\]

and (4.2c) is thus proved. Since \( H^{1/2}(\partial K; \mathbb{R}^3) \) is dense in \( L^2(\partial K; \mathbb{R}^3) \), (4.2d) follows from (4.2c).

The final result of this section contains all scaling inequalities. Shape-regularity can be rephrased as the asymptotic equivalences (recall that we are in three dimensions)

\[
\|B\| \approx h_K, \quad \|B^{-1}\| \approx h_K^{-1}, \quad |J| \approx h_K^3, \quad |a| \approx h_K.
\]

Therefore

\[
\|\phi\|_K \approx h_K^{3/2} \|\phi \circ F\|_{\partial \hat{K}}, \quad \|\phi\|_{\partial K} \approx h_K \|\phi \circ F\|_{\partial \hat{K}}.
\]

For the hat transformations, we have the following scaling rules, which can be easily proved using (4.6), (4.7), and the chain rule.

**Proposition 4.3 (Scaling equivalences).** With hidden constants depending on shape-regularity and on \( m \geq 0 \), we have

\[
\begin{align*}
(4.8a) & \quad |\hat{u}|_{m,\hat{K}} \approx h_K^{m-1/2} |u|_{m,K} \quad \forall u \in H^m(K; \mathbb{R}^3), \\
(4.8b) & \quad |\hat{\sigma}|_{m,\hat{K}} \approx h_K^{m-1/2} |\sigma|_{m,K} \quad \forall \sigma \in H^m(K; \mathbb{R}^{3 \times 3}), \\
(4.8c) & \quad \|\hat{\mu}\|_{\partial \hat{K}} \approx \|\mu\|_{\partial K} \approx \|\mu\|_{\partial \hat{K}} \quad \forall \mu \in L^2(\partial K; \mathbb{R}^3).
\end{align*}
\]

4.2. **The projection and the remainder on \( K \).** Consider the spaces

\[
\Sigma(K) := \{ \sigma \in L^2(K; \mathbb{R}^{3 \times 3}) : \hat{\sigma} \in \hat{\Sigma} \} = \mathcal{P}_k(K; \mathbb{R}^{3 \times 3}), \\
V_+(K) := \{ u \in L^2(K; \mathbb{R}^3) : \hat{u} \in \hat{V}_+ \} = \mathcal{P}_{k+1}(K; \mathbb{R}^3), \\
M(\partial K) := \{ \mu \in L^2(\partial K; \mathbb{R}^3) : \hat{\mu} \in \hat{M} \}
\]

\[
= \{ \mu \in L^2(\partial K; \mathbb{R}^3) : \mu \in M(\partial K) \} = \mathcal{R}_k(\partial K; \mathbb{R}^3).
\]

The projection and the remainder are defined by a pull-back process: given \((\sigma, u) \in H^1(K; \mathbb{R}^{3 \times 3}) \times H^1(K; \mathbb{R}^3)\), we define

\[
(4.9a) \quad \Pi(\sigma, u; \tau) := (\sigma_K, u_K) \in \Sigma(K) \times V_+(K), \quad R(\sigma, u; \tau) := \delta_K \in M(\partial K),
\]
by the relations

\[(4.9b) \quad (\hat{\sigma}_K, \hat{u}_K) = \hat{\Pi}(\sigma, \hat{u}; \hat{\tau}), \quad \hat{\tau} := |a|B^{-1}(\tau \circ F)B^{-\top},\]

\[(4.9c) \quad \tilde{\delta}_K = \hat{\Pi}(\sigma, \hat{u}; \hat{\tau}).\]

We now prove Theorem 2.1. We start by proving a technical lemma, continue showing

\[
\mathcal{R}_0(\partial K; \mathbb{R}^{3\times 3}_{\text{sym}}), \quad \text{we have}
\]

\[
\langle \tau P_M \gamma u, \mu \rangle_{\partial K} = \langle \tau P_M \hat{\gamma} \hat{u}, \hat{\mu} \rangle_{\partial \hat{K}} \quad \forall \mu \in L^2(\partial K; \mathbb{R}^3),
\]

\[
\langle \tau P_M \gamma u, \gamma v \rangle_{\partial K} = \langle \tau P_M \hat{\gamma} \hat{u}, \hat{\gamma} \hat{v} \rangle_{\partial \hat{K}} \quad \forall \mu, \nu \in H^1(K; \mathbb{R}^3),
\]

with \(\hat{\tau}\) defined in (4.9b).

**Proof.** It follows from the definitions that \(\hat{P}_M \mu = \hat{M}\) and that \(\mu^* \in M(K)\) if and only if \(\hat{\mu}^* \in \hat{M}\). Therefore,

\[
\langle \hat{P}_M \mu, \mu^* \rangle_{\partial \hat{K}} = \langle P_M \mu, \mu^* \rangle_{\partial K} = \langle \mu, \mu^* \rangle_{\partial K} = \langle \hat{\mu}, \hat{\mu}^* \rangle_{\partial \hat{K}} = \langle P_M \hat{\mu}, P_M \hat{\mu}^* \rangle_{\partial \hat{K}} \quad \forall \mu^* \in \hat{M},
\]

and (4.10) is proved. The next two identities in the statement follow from (4.10) and Proposition 4.1 and Lemma 4.4.

**Proposition 4.5.** Assume that \(\tau \in \mathcal{R}_0(\partial K; \mathbb{R}^{3\times 3}_{\text{sym}})\) satisfies

\[
\langle \tau \mu, \mu \rangle_{\partial K} > 0 \quad \forall \mu \in L^2(\partial K; \mathbb{R}^3), \quad \mu \neq 0.
\]

If \((\sigma_K, u_K) = \Pi(\sigma, u; \tau)\) and \(\delta_K = \Delta(\sigma, u; \tau)\), then the following equations hold:

\[
(4.11a) \quad (u_K - u, \text{div } \theta)_K = 0 \quad \forall \theta \in \Sigma(K),
\]

\[
(4.11b) \quad -\langle \text{div } (\sigma_K - \sigma), w \rangle_K + \langle \tau P_M \gamma (u_K - u), \gamma w \rangle_{\partial K} = \langle \delta_K, \gamma w \rangle_{\partial K} \quad \forall w \in V_+(K),
\]

\[
(4.11c) \quad -\langle \gamma u(\sigma_K - \sigma) - \tau P_M \gamma (u_K - u), \mu \rangle_{\partial K} = \langle \delta_K, \mu \rangle_{\partial K} \quad \forall \mu \in M(\partial K).
\]

**Proof.** The equality (4.11a) follows from changing variables to the reference element, applying (3.19a) and (4.2b). Similarly, (4.11b) follows from (3.19b) using (4.2b), Lemma 4.4, and (4.4). Finally, (4.11c) follows from (3.19c) applying (4.2d), Lemma 4.4, and (4.4).

Assume now that \(\tau\) is of order \(h_K^{-1}\), as expressed in (2.1). By (4.6) and Proposition 4.3, we can write (take \(\eta = \hat{\mu}\) for \(\mu \in L^2(\partial K; \mathbb{R}^3)\))

\[
(4.12) \quad \alpha_1 C h_K^{-1} \|\eta\|^2_{\partial \hat{K}} \leq \|\hat{\tau} \eta\|^2_{\partial \hat{K}} \leq \alpha_2 C h_K^{-1} \|\eta\|^2_{\partial \hat{K}} \quad \forall \eta \in L^2(\partial \hat{K}; \mathbb{R}^3),
\]

where \(\alpha_1\) and \(\alpha_2\) are constants related to shape-regularity of \(K\). We are now ready to apply Proposition 3.6 with \(\tau_{\min} = \alpha_1 C_1\), \(\tau_{\max} = \alpha_2 C_2\) and \(\eta = h_K^{-1}\) (compare (4.12).
with\ (3.16)\). Using the definition of the projection\ (1.9), Proposition\ 4.3 for the scaling properties, and Proposition\ 3.6 for the estimates, we have
\[
\| \sigma - \sigma_K \|_K + h_K^{-1}\| u - u_K \|_K + h_K^{1/2}\| \delta_K \|_K \approx h_K^{1/2}(\| \sigma - \sigma_K \|_K + h_K^{-1}\| u - u_K \|_K + \| \delta_K \|_{\partial K})
\]
\[
\lesssim h_K^{1/2}(\| \sigma \|_{m, K} + h_K^{-1}\| u \|_{m+1, K})
\]
and\ (2.3)\ is thus proved.

5. Steady-state elasticity

5.1. Method and convergence estimates. From now on, we shift our attention from the construction of the projection to its applications. We begin by introducing more notation for the rest of the paper. Let $\Omega$ be a Lipschitz polyhedral domain in $\mathbb{R}^3$. We denote the density function by $\rho \in W^{1, \infty}(\Omega; \mathbb{R})$ and the compliance tensor by $A \in L^{\infty}(\Omega; \mathcal{B}(\mathbb{R}^{3 \times 3}_{\text{sym}}))$, where $\mathcal{B}(\mathbb{R}^{3 \times 3}_{\text{sym}})$ is the space of linear maps from $\mathbb{R}^{3 \times 3}_{\text{sym}}$ to itself. We assume $\rho \geq \rho_0$ for some positive constant $\rho_0$, and $A$ is uniformly symmetric and positive almost everywhere on $\Omega$, i.e., there exists $C_0 > 0$ such that for almost all $x \in \Omega$,

\[
(\mathcal{A}(x)\xi) : \chi = \xi : (\mathcal{A}(x)\chi)
\]
\[
(\mathcal{A}(x)\xi) : \xi \geq C_0 \xi : \xi
\]

Let $T_h$ be a family of conforming tetrahedral partitions of $\Omega$, which we assume to be shape-regular. Namely, there exists a fixed constant $c_0 > 0$ such that $h_K / \rho_K \leq c_0$ for all $K \in T_h$, where $\rho_K$ denotes the inradius of $K$. We denote $h := \max_{K \in T_h} h_K$ as the mesh size and set the following discrete spaces:

\[
V_h := \prod_{K \in T_h} \mathcal{P}_k(K; \mathbb{R}^{3 \times 3}_{\text{sym}}), \quad W_h := \prod_{K \in T_h} \mathcal{P}_{k+1}(K; \mathbb{R}^3), \quad M_h := \prod_{K \in T_h} \mathcal{P}_k(\partial K; \mathbb{R}^3).
\]

The related discrete inner products are denoted as

\[
(\sigma, \tau)_T := \sum_{K \in T_h} (\sigma, \tau)_K, \quad (u, v)_T := \sum_{K \in T_h} (u, v)_K, \quad \langle \mu, \lambda \rangle_{\partial T} := \sum_{K \in T_h} \langle \mu, \lambda \rangle_{\partial K}.
\]

Finally, we use the following notation for the discrete and the weighted discrete norms:

\[
\| \cdot \|_T := (\cdot, \cdot)_T^{1/2}, \quad \| \cdot \|_{\partial T} := (\cdot, \cdot)^{1/2}_{\partial T},
\]
\[
\| \cdot \|_{*1} := (\star_1 \cdot, \cdot)_T^{1/2}, \quad \| \cdot \|_{*2} := (\star_2 \cdot, \cdot)^{1/2}_{\partial T},
\]

where $\star_1 = \mathcal{A}$ or $\rho$, and $\star_2 = \tau$ or $\tau^{-1}$. Here $\tau \in \prod_{K \in T_h} \mathcal{R}_0(\partial K; \mathbb{R}^{3 \times 3}_{\text{sym}})$ is the stabilization function which we assume to satisfy\ (2.1)\ on every element $K$.

In this section, we give a projection-based error analysis to the HDG+ method introduced in\ [13]. We will show that the analysis can be simplified by using Theorem\ 2.1. To begin with, we review the steady-state linear elasticity equations:

\[
A\sigma - \varepsilon(u) = 0 \quad \text{in } \Omega,
\]
\[
-\text{div}\sigma = f \quad \text{in } \Omega,
\]
\[
\gamma u = g \quad \text{on } \Gamma := \partial \Omega,
\]

(5.1a)
where \( f \in L^2(\Omega;\mathbb{R}^3) \) and \( g \in H^{1/2}(\Gamma;\mathbb{R}^3) \). The HDG+ method for (5.1) is: find \((\sigma_h, u_h, \hat{u}_h) \in V_h \times W_h \times M_h\) such that

\[
\begin{align}
(A \sigma_h, \theta)_{T_h} + (u_h, \text{div} \theta)_{T_h} - \langle \hat{u}_h, \theta \rangle_{\partial T_h} &= 0 \quad \forall \theta \in V_h, \\
-\langle \text{div} \sigma_h, u \rangle_{T_h} + \langle \tau P_M (u_h - \hat{u}_h), u \rangle_{\partial T_h} &= \langle f, u \rangle_{T_h} \quad \forall u \in W_h, \\
\langle \sigma_h n - \tau P_M (u_h - \hat{u}_h), \mu \rangle_{\partial T_h \setminus \Gamma} &= 0 \quad \forall \mu \in M_h, \\
\langle \hat{u}_h, \mu \rangle_{\Gamma} &= \langle g, \mu \rangle_{\Gamma} \quad \forall \mu \in M_h.
\end{align}
\]

Since we will use a duality argument to estimate the convergence of \( u_h \), here we write down the adjoint equations for (5.1):

\[
\begin{align}
(A^* \hat{\sigma}, \chi)_{T_h} &= 0 \quad \forall \chi \in V_h, \\
D \hat{\Psi} &= \hat{\Theta} \quad \forall \hat{\Theta} \in W_h, \\
\gamma \hat{\Phi} &= 0 \quad \forall \hat{\Phi} \in M_h,
\end{align}
\]

with input data \( \Theta \in L^2(\Omega;\mathbb{R}^3) \), and we assume the additional elliptic regularity estimate

\[
\| \Psi \|_{1,\Omega} + \| \Phi \|_{2,\Omega} \leq C_{\text{reg}} \| \Theta \|_{\Omega},
\]

where \( C_{\text{reg}} \) is a constant depending only on \( A \) and \( \Omega \). The following convergence theorem will be proved in Subsection 5.2.

**Theorem 5.1.** For \( k \geq 1 \) and \( 1 \leq m \leq k + 1 \), we have

\[
\| \sigma - \sigma_h \|_A \leq C_1 h^m (\| \sigma \|_{m,\Omega} + \| u \|_{m+1,\Omega}).
\]

If (5.4) holds, then we also have

\[
\| u - u_h \|_{\Omega} \leq C_2 h^{m+1} (\| \sigma \|_{m,\Omega} + \| \sigma \|_{m+1,\Omega}).
\]

Here, \( C_1 \) depends only on the polynomial degree \( k \) and the shape-regularity of \( T_h \), while \( C_2 \) depends also on \( A \) and \( C_{\text{reg}} \).

Note that in [18, Theorem 2.1] the meshes \( T_h \) are assumed to be quasi-uniform, whereas here we only require \( T_h \) to be shape-regular.

5.2. **Proof of Theorem 5.1.** To begin with, we use the element-by-element projection defined in (4.9) (we will only use Theorem 2.1 not the definition) on the solution \((\sigma, u)\):

\[
(\Pi \sigma, \Pi u) := \prod_{K \in T_h} \Pi(\sigma|_K, u|_K; \tau|_{\partial K}), \quad \delta := \prod_{K \in T_h} \tau(\sigma|_K, u|_K; \tau|_{\partial K}),
\]

and define the error terms and the approximation terms:

\[
\begin{align}
\varepsilon_h^u &:= \Pi u - u_h, \\
\varepsilon_h^{\sigma} &:= \Pi \sigma - \sigma_h, \\
\hat{\varepsilon}_h &:= \hat{\tau} P_M u - \hat{u}_h, \\
e_{\sigma} &:= \Pi \sigma - \sigma, \\
e_u &:= \Pi u - u.
\end{align}
\]

We aim to control the terms \( \varepsilon_h^* \) by the terms \( e_* \) and \( \delta \).

**Proposition 5.2** (Energy estimate). The following energy identity holds

\[
(A \varepsilon_h^{\sigma}, \varepsilon_h^{\sigma})_{T_h} + \langle \tau P_M (\varepsilon_h^u - \hat{\varepsilon}_h), (\varepsilon_h^{\sigma} - \hat{\varepsilon}_h) \rangle_{\partial T_h} = (A e_{\sigma}, e_{\sigma})_{T_h} + \langle \delta, (\varepsilon_h^u - \hat{\varepsilon}_h) \rangle_{\partial T_h}.
\]

Consequently,

\[
\| \varepsilon_h^{\sigma} \|_A^2 + \| P_M (\varepsilon_h^u - \hat{\varepsilon}_h) \|_{\tau}^2 \leq \| e_{\sigma} \|_A^2 + \| \delta \|_{\tau}^2.
\]
Proof. The proof here will be similar to the proof of [6, Lemma 3.1 & 3.2]. By Theorem 2.1, we first obtain a set of projection equations satisfied by $\Pi \sigma$, $\Pi u$ and $\delta$. We then subtract (5.2) from the projection equations and obtain the following error equations:

\begin{align}
(5.7a) \quad (A_\epsilon e^\theta, \theta)_{T_h} + (\epsilon_h^n, \text{div} \theta)_{T_h} - \langle \epsilon_h^n, \theta n \rangle_{\partial T_h} &= (A e, \theta)_{T_h}, \\
(5.7b) \quad -(\text{div} \epsilon_h^n, w)_{T_h} + \langle \tau P_M (\epsilon_h^n - \epsilon_h^n), w \rangle_{\partial T_h} &= \langle \delta, w \rangle_{\partial T_h}, \\
(5.7c) \quad \langle \epsilon_h^n, \tau P_M (\epsilon_h^n - \epsilon_h^n), \mu \rangle_{\partial T_h | \Gamma} &= -\langle \delta, \mu \rangle_{\partial T_h | \Gamma}, \\
(5.7d) \quad \langle \epsilon_h^n, \mu \rangle_{\Gamma} &= 0.
\end{align}

By (5.7c) and (5.7d) we have

\begin{align}
(5.8) \quad \langle \epsilon_h^n, \tau P_M (\epsilon_h^n - \epsilon_h^n), \mu \rangle_{\partial T_h | \Gamma} = -\langle \delta, \epsilon_h^n \rangle_{\partial T_h | \Gamma}.
\end{align}

Now taking $\theta = \epsilon_h^n$ in (5.7a), $w = \epsilon_h^n$ in (5.7b), and then adding (5.7a), (5.7b) and (5.8), we obtain the energy identity (5.5), from which the estimate (5.6) follows easily.

**Proposition 5.3** (Estimate by duality). If $k \geq 1$ and (5.4) holds, then

\[ \| \epsilon_h^n \|_{T_h} \leq h \mathcal{C} \left( \| e_\sigma \|_{T_h} + \| \delta \|_{\tau^{-1}} \right). \]

Here, $\mathcal{C}$ is independent of $h$, but depends on $A$ and $C_{\text{reg}}$.

Proof. The proof here will be similar to the proof of [6, Lemma 4.1 & Theorem 4.1]. Consider the adjoint equations (5.3). We take $\Theta = \epsilon_h^n$ as the input data and apply the projection on $(P, \Phi)$ with $-\tau$ as the stabilization function. Namely,

\[ (\Pi \Psi, \Pi \Phi) := \prod_{K \in \mathcal{T}_h} \Pi (\Psi|_{K}, \Phi|_{K}; -\tau|_{\partial K}), \quad \Delta := \prod_{K \in \mathcal{T}_h} \Pi (\Psi|_{K}, \Phi|_{K}; -\tau|_{\partial K}). \]

By Theorem 2.1, $\Pi \Psi$, $\Pi \Phi$ and $\Delta$ satisfy

\begin{align}
(A \Pi \Psi, \theta)_{T_h} - (\Pi \Phi, \text{div} \theta)_{T_h} + (P_M \Phi, \theta n)_{\partial T_h} &= (A (\Pi \Psi - \Psi), \theta)_{T_h}, \\
(\text{div} \Pi \Psi, \Phi)_{T_h} + \langle \tau P_M (\Pi \Phi - \Phi), \Phi \rangle_{\partial T_h} &= \langle \delta, \Phi \rangle_{\partial T_h}, \\
-\langle \Pi \Psi \nabla, \tau P_M (\Pi \Phi - \Phi), \mu \rangle_{\partial T_h | \Gamma} &= \langle \Delta, \mu \rangle_{\partial T_h | \Gamma}, \\
\langle P_M \Phi, \mu \rangle_{\Gamma} &= 0.
\end{align}

We now test the above equations with $\theta = \epsilon_h^n$, $w = \epsilon_h^n$ and $\mu = \epsilon_h^n$, then test the error equations (5.7) with $\theta = \Pi \Psi$, $w = \Pi \Phi$ and $\mu = P_M \Phi$. By comparing the two sets of equations, we obtain

\[ \langle A (\Pi \Psi - \Psi), \epsilon_h^n \rangle_{T_h} + \| \epsilon_h^n \|^2_{T_h} - \langle \Delta, P_M \epsilon_h^n - \epsilon_h^n \rangle_{\partial T_h} = (A e_\sigma, \Pi \Psi)_{T_h} + \langle \delta, P_M (\Pi \Phi - \Phi) \rangle_{\partial T_h}. \]

After rearranging terms, we have

\[ \| \epsilon_h^n \|^2_{T_h} = -(A (\Pi \Psi - \Psi), \sigma - \sigma_h)_{T_h} - (\Pi \sigma - \sigma, \epsilon(\Phi))_{T_h} \\
+ \langle \Delta, P_M \epsilon_h^n - \epsilon_h^n \rangle_{\partial T_h} + \langle \delta, P_M (\Pi \Phi - \Phi) \rangle_{\partial T_h}. \]

By Theorem 2.1 (with $m = 1$) we have

\[ \| \Pi \Psi - \Psi \|_{T_h} + \| \Delta \|_{\tau^{-1}} \leq h (\| \Psi \|_{1, \Omega} + \| \Phi \|_{2, \Omega}). \]

By (2.5) we have

\[ (\Pi \sigma - \sigma, \epsilon(\Phi))_{T_h} = (\Pi \sigma - \sigma, \epsilon(\Phi) - P_0 \epsilon(\Phi))_{T_h} \leq h \| e_\sigma \|_{T_h} \| \Phi \|_{2, \Omega}. \]
Finally, (2.1) implies
\[ \|P_M(\Pi \Phi - \Phi)\|_r \approx \left( \sum_{K \in T_h} h_K^{-1/2} \|P_M(\Pi \Phi - \Phi)\|_{\partial K}^2 \right)^{1/2} \lesssim h(\|\Psi\|_{1,\Omega} + \|\Phi\|_{2,\Omega}). \]

We next use (5.4) to control the term \( \|\Psi\|_{1,\Omega} + \|\Phi\|_{2,\Omega} \) and use Proposition 5.2 to control the terms \( \|A(\sigma - \sigma_h)\|_{T_h} \) and \( \|P_M \varepsilon_h^{\sigma} - \varepsilon_h^{\sigma}\|_r \). The proof is thus finished. \( \square \)

Combining Proposition 5.2, Proposition 5.3 and Theorem 2.1, Theorem 5.1 follows readily.

6. Frequency-domain elastodynamics

6.1. Method and convergence estimates. In this section we give new proofs of error estimates to the second HDG+ method studied in [14] by using projection based analysis for the following time-harmonic linear elasticity problem (second order in frequency, first order in space)

\begin{align}
(6.1a) & \quad A\sigma - \varepsilon(u) = 0 \quad \text{in } \Omega, \\
(6.1b) & \quad -\text{div } \sigma - \kappa^2 \rho u = f \quad \text{in } \Omega, \\
(6.1c) & \quad \gamma u = g \quad \text{on } \Gamma := \partial \Omega,
\end{align}

where \( f \in L^2(\Omega; \mathbb{C}^3) \) and \( g \in H^{1/2}(\Gamma; \mathbb{C}^3) \). We assume that the wave number \( \kappa > 0 \) and \( \kappa^2 \) is not a Dirichlet eigenvalue so that (6.1) is well-posed. Note that because of the different proof techniques, the dependence on the wave numbers of the error estimates will be different and not easy to compare. The emphasis here will be the simplified error analysis thanks to the introduction of the projection.

The HDG+ method for (6.1) is: find \( (\sigma_h, u_h, \hat{u}_h) \in V_h \times W_h \times M_h \) such that

\begin{align}
(6.2a) & \quad (A\sigma_h, \theta)_{T_h} + (u_h, \text{div } \theta)_{T_h} - \langle \hat{u}_h, \theta n \rangle_{\partial T_h} = 0, \\
(6.2b) & \quad -\langle \text{div } \sigma_h, w \rangle_{T_h} + \langle \tau P_M(u_h - \hat{u}_h), w \rangle_{\partial T_h} - (\kappa^2 \rho u_h, w)_{T_h} = (f, w)_{T_h}, \\
(6.2c) & \quad \langle \sigma_h n - \tau P_M(u_h - \hat{u}_h), \mu \rangle_{\partial T_h \backslash \Gamma} = 0, \\
(6.2d) & \quad \langle \hat{u}_h, \mu \rangle_{\Gamma} = \langle g, \mu \rangle_{\Gamma},
\end{align}

for all \( \theta \in V_h, w \in W_h \) and \( \mu \in M_h \). Here, the spaces \( V_h, W_h \) and \( M_h \) are the same as those defined in Section 5.1 except now the functions contained in these spaces take complex values.

For \( \Theta \in L^2(\Omega; \mathbb{R}^3) \), we assume that the solution to the adjoint equations for (6.1)

\begin{align}
(6.3a) & \quad A\Psi + \varepsilon(\Phi) = 0 \quad \text{in } \Omega, \\
(6.3b) & \quad \text{div } \Psi - \kappa^2 \rho \Phi = \Theta \quad \text{in } \Omega, \\
(6.3c) & \quad \gamma \Phi = 0 \quad \text{on } \Gamma,
\end{align}

satisfy

\[ \|\Phi\|_{2,\Omega} + \|\Psi\|_{1,\Omega} \leq C_\kappa \|\Theta\|_{\Omega}, \]

where we make the dependence on \( \kappa \) explicit for the constant \( C_\kappa \). For the rest of this section, the wiggles sign \( \lesssim \) will hide constants independent of \( h \) and \( \kappa \). We aim to prove the following theorem.
Proof. The proof of Proposition 5.3, we first apply the projection on $\Box$ on the identity, we have

$$\|u - u_h\|_\Omega \lesssim (1 + C_n) h^{m+1} + C_n(\kappa^2 + \kappa) h^{m+2} \left( \|\sigma\|_{m,\Omega} + \|u\|_{m+1,\Omega} \right),$$

$$\|\sigma - \tau_h\|_\Omega \lesssim (h^m + (1 + C_n)\kappa h^{m+1} + C_n(\kappa^3 + \kappa^2) h^{m+2} \left( \|\sigma\|_{m,\Omega} + \|u\|_{m+1,\Omega} \right),$$

with $1 \leq m \leq k + 1$.

6.2. Proof of Theorem 6.1. Since the solution $(\sigma, u)$ in (6.1) can take complex values, applying directly Theorem 2.1 is not feasible. However, this can be easily fixed by defining a new complex projection based on the original one: for $(\sigma, u) \in H^1(K; \mathbb{C}^3 \times H^1(K; \mathbb{C}^3)$, we define

$$\Pi(\sigma, u) := \Pi(\text{Re} \sigma, \text{Re} u) + i\Pi(\text{Im} \sigma, \text{Im} u).$$

It is easy to show that this complex projection also satisfies (2.2) and (2.3) (the only difference is that now the test functions $\theta, w, \mu$ in (2.2) can take complex values).

Similar to the beginning of Section 5.2, we first define the projections $(\Pi \sigma, \Pi u)$, the remainder term $\delta$, the error $\varepsilon_h^\sigma, \varepsilon_h^\mu, \varepsilon^\mu$, and approximation terms $e_{\sigma, e_{\mu}}$. 

Proposition 6.2 (Garding-type identity). The following energy identity holds

$$\|\varepsilon_h^\sigma\|_{A}^2 + \|P_M(\varepsilon_h^\sigma - \varepsilon_h^\mu)\|_\tau^2 - \kappa^2\|\varepsilon_h^\mu\|_\tau^2 = (Ae_{\sigma}, e_{\sigma})_{T_h} + (\delta, e_{\sigma}^\mu - \varepsilon_h^\mu)_{\partial T_h} - \kappa^2(\rho e_{\sigma}, e_{\sigma}^\mu)_{T_h}.$$

Consequently,

$$\|\varepsilon_h^\sigma\|_{T_h} + \|P_M(\varepsilon_h^\sigma - \varepsilon_h^\mu)\|_\tau \lesssim \kappa\|\varepsilon_h^\sigma\|_{T_h} + \kappa\|e_{\mu}\|_{T_h} + \|e_{\sigma}\|_{T_h} + \|\delta\|_{\tau^{-1}}. \tag{6.6}$$

Proof. By similar ideas used in the proof of Proposition 5.2 we obtain the following error equations:

$$(Ae_{\sigma}, \varepsilon_h^\sigma)_{T_h} + (\varepsilon_h^\mu, \text{div} \varepsilon_h^\mu)_{T_h} - (\varepsilon_h^\mu, \theta n)_{\partial T_h} = (Ae_{\sigma}, \varepsilon_h^\mu)_{T_h}, \tag{6.7a}$$

$$(\text{div} \varepsilon_h^\mu, w)_{T_h} + (\tau P_M(\varepsilon_h^\mu - \varepsilon_h^\mu), w)_{\partial T_h} = (\delta, w)_{\partial T_h} - (\kappa^2\varepsilon_h^\mu, w)_{T_h}, \tag{6.7b}$$

$$(\varepsilon_h^\mu n - \tau P_M(\varepsilon_h^\mu - \varepsilon_h^\mu), \mu)_{\partial T_h \setminus \Gamma} = - (\delta, \mu)_{\partial T_h \setminus \Gamma}, \tag{6.7c}$$

$$(\varepsilon_h^\mu, \mu)_{\Gamma} = 0, \tag{6.7d}$$

for all $\theta \in V_h, w \in W_h$ and $\mu \in M_h$. Taking $\theta = \varepsilon_h^\mu, w = \varepsilon_h^\mu$ and $\mu = \varepsilon_h^\mu$, then adding the equations, we have the energy identity.

Denoting $A^2 := \|\varepsilon_h^\sigma\|_{T_h}^2 + \|P_M(\varepsilon_h^\sigma - \varepsilon_h^\mu)\|_\tau^2$, and applying the Cauchy-Schwarz inequality on the identity, we have

$$A^2 \lesssim \kappa^2\|\varepsilon_h^\sigma\|_{T_h}^2 + \kappa^2\|\varepsilon_h^\mu\|_{T_h}^2 + A(\|e_{\sigma}\|_{T_h} + \|\delta\|_{\tau^{-1}}).$$

The estimate (6.6) now follows by using Young’s inequality.

Proposition 6.3 (Estimate by bootstrapping). Suppose $k \geq 1$ and (6.4) holds. If $(h^2 \kappa^2 + h\kappa)C_k$ is small enough, then

$$\|\varepsilon_h^\mu\|_{T_h} \lesssim h(\kappa^2 + \kappa)C_k \|e_{\mu}\|_{T_h} + h C_n(\|\delta\|_{\tau^{-1}} + \|e_{\sigma}\|_{T_h}),$$

$$\|\varepsilon_h^\mu\|_{T_h} + \|P_M(\varepsilon_h^\mu - \varepsilon_h^\mu)\|_\tau \lesssim (1 + h\kappa C_n)(\|e_{\sigma}\|_{T_h} + \|\delta\|_{\tau^{-1}} + (\kappa + h C_n(\kappa^3 + \kappa^2))\|e_{\mu}\|_{T_h}).$$

Proof. Consider the adjoint equations (6.3) and take $\Theta = \varepsilon_h^\mu$. Following similar ideas in the proof of Proposition 5.3, we first apply the projection on $(\Psi, \Phi)$ with stabilization function $-\tau$ to obtain the projection equations (about $(\Pi \Psi, \Pi \Phi, \Delta)$). We next test the
projection equations with $e^P_h$, $e^u_h$ and $e^\phi_h$, test the error equations (6.7) with $\Pi \Psi$, $\Pi \bar{\Psi}$ and $P_M \Phi$, and compare the two sets of equations. Then we obtain (define $e_\psi := \Pi \Psi - \Psi$ and $e_\phi := \Pi \Phi - \Phi$ for convenience)

$$(6.8) \quad \langle A \epsilon_\sigma, \Pi \Psi \rangle_{\mathcal{T}_h} + \langle \delta, \Pi \Phi - P_M \Phi \rangle_{\mathcal{T}_h} + \kappa^2(\rho(u - u_h), \Pi \Psi)_{\mathcal{T}_h} + \kappa^2(\rho(\bar{\Phi}, u - \Pi u)_{\mathcal{T}_h} + \langle A(\sigma_h - \sigma_h), \Pi \Psi \rangle_{\mathcal{T}_h} - (e_\sigma, \Pi \bar{\Psi})_{\mathcal{T}_h}$$

After rearranging terms, we have

$$\|e^u_h\|_{\mathcal{T}_h}^2 = \kappa^2(\epsilon_\sigma, \rho(u - u_h))_{\mathcal{T}_h} + \kappa^2(\rho(\bar{\Phi}, u - \Pi u)_{\mathcal{T}_h} + \langle A(\sigma_h - \sigma), \Pi \Psi \rangle_{\mathcal{T}_h} - (e_\sigma, \Pi \bar{\Psi})_{\mathcal{T}_h}$$

By Theorem 2.1 (with $m = 1$), we have

$$h^{-1}\|e_\phi\|_{\mathcal{T}_h} + \|e_\psi\|_{\mathcal{T}_h} + \|\Pi \Phi - P_M \Phi\|_{\mathcal{T}_h} + \|\delta\|_{\mathcal{T}_h} \lesssim h(\|\Psi\|_{1,\Omega} + \|\Phi\|_{2,\Omega})$$

and also

$$\kappa^2(\rho(\bar{\Phi}, u - \Pi u)_{\mathcal{T}_h} \lesssim h\kappa^2(\rho(\bar{\Phi}, u - \Pi u)_{\mathcal{T}_h} \lesssim h\kappa^2(\epsilon_\sigma, \Pi \Psi - P_0(\bar{\Phi}))_{\mathcal{T}_h} \lesssim h\|e_\sigma\|_{\mathcal{T}_h} + \|\Phi\|_{2,\Omega}.$$ 

By (2.3) we have $(e_\sigma, \Pi \bar{\Psi})_{\mathcal{T}_h} = (e_\sigma, \Pi \bar{\Psi} - P_0(\bar{\Phi}))_{\mathcal{T}_h} \lesssim h\|e_\sigma\|_{\mathcal{T}_h} + \|\Phi\|_{2,\Omega}$. Now we use (6.6) in (6.8) to obtain

$$\|e^u_h\|_{\mathcal{T}_h} \lesssim h^2\kappa^2C_\kappa\|u - u_h\|_{\mathcal{T}_h} + h\kappa^2C_\kappa\|\Pi u - u_h\|_{\mathcal{T}_h} + h\kappa^2C_\kappa\|\sigma - \sigma_h\|_{\mathcal{T}_h} + hC_\kappa\|e_\sigma\|_{\mathcal{T}_h} + hC_\kappa\|e_\phi\|_{\mathcal{T}_h}$$

Next we use (6.6) and bound

$$\|e^u_h\|_{\mathcal{T}_h} \lesssim (h^2\kappa^2 + h\kappa)C_\kappa\|e^u_h\|_{\mathcal{T}_h} + \|\epsilon_\sigma\|_{\mathcal{T}_h} + hC_\kappa(\|e_\sigma\|_{\mathcal{T}_h} + \|\delta\|_{\mathcal{T}_h})$$

Therefore, when $(h^2\kappa^2 + h\kappa)C_\kappa$ is small enough, we have

$$\|e^u_h\|_{\mathcal{T}_h} \lesssim (h\kappa + h\kappa^2)C_\kappa\|e_\sigma\|_{\mathcal{T}_h} + h\kappa(\|e_\sigma\|_{\mathcal{T}_h} + \|\delta\|_{\mathcal{T}_h}).$$

The first inequality is thus proved. The second inequality can be proved by combining (6.6) and the above inequality.

Theorem 6.1 now follows easily from Proposition 6.3 and Theorem 2.1.

7. Transient elastodynamics

7.1. The semi-discrete HDG+ method. In this section, we present a semi-discrete (in space) HDG+ method for transient elastic waves and prove it is uniformly-in-time optimal convergent. The equations we consider are

$$(7.1a) \quad A \sigma(t) - \epsilon(u(t)) = 0 \quad \text{in } \Omega \times [0, T],$$

$$(7.1b) \quad \rho \ddot{u}(t) - \text{div } \sigma(t) = f(t) \quad \text{in } \Omega \times [0, T],$$

$$(7.1c) \quad \gamma u(t) = g(t) \quad \text{on } \Gamma \times [0, T],$$

$$(7.1d) \quad u(0) = u_0 \quad \text{on } \Omega,$$

$$(7.1e) \quad \dot{u}(0) = v_0 \quad \text{on } \Omega,$$
where $f \in C([0, \infty); L^2(\Omega; \mathbb{R}^3))$ and $g \in C([0, \infty); H^{1/2}(\Gamma; \mathbb{R}^3))$. For the initial conditions, we assume $u_0, v_0 \in L^2(\Omega; \mathbb{R}^3)$.

The HDG+ method for (7.1) looks for
\[\sigma_h, u_h, \hat{u}_h : [0, \infty) \to V_h \times W_h \times M_h,\]
such that for all $t \geq 0$

(7.2a) \[(A\sigma_h(t), \theta)_{T_h} + (u_h(t), \text{div}\theta)_{T_h} - \langle \hat{u}_h(t), \theta n \rangle_{\partial T_h} = 0,\]
(7.2b) \[(\rho \ddot{u}_h(t), w)_{T_h} - (\text{div}\sigma_h(t), w)_{T_h} + \langle \tau P_M(u_h(t) - \hat{u}_h(t)), w \rangle_{\partial T_h} = (f(t), w)_{T_h},\]
(7.2c) \[\langle \sigma_h(t) n - \tau P_M(u_h(t) - \hat{u}_h(t)), \mu \rangle_{\partial T_h} = 0,\]
(7.2d) \[\langle \hat{u}_h(t), \mu \rangle_{\Gamma} = \langle g(t), \mu \rangle_{\Gamma},\]

for all $\theta \in V_h$, $w \in W_h$ and $\mu \in M_h$. For the initial conditions of the method, we use ideas from (5). The initial velocity $\dot{u}_h(0)$ is defined by using the projection in Theorem 2.1

\[(\dot{u}_h(0)) = \Pi(A^{-1} \varepsilon(v_0), v_0; \tau),\]
and the initial displacement $u_h(0)$ is defined to be the solution of the HDG+ discretization of the steady-state system

(7.4a) \[A\sigma(0) - \varepsilon(u(0)) = 0 \quad \text{in } \Omega,\]
(7.4b) \[-\text{div}\sigma(0) = -\text{div}(A^{-1} \varepsilon(u_0)) \quad \text{in } \Omega,\]
(7.4c) \[u(0) = g(0) \quad \text{on } \Gamma.\]

Namely, we find $(\sigma_h(0), u_h(0), \hat{u}_h(0)) \in V_h \times W_h \times M_h$ such that

(7.5a) \[(A\sigma_h(0), \theta)_{T_h} + (u_h(0), \text{div}\theta)_{T_h} - \langle \hat{u}_h(0), \theta n \rangle_{\partial T_h} = 0,\]
(7.5b) \[-(\text{div}\sigma_h(0), w)_{T_h} + \langle \tau P_M(u_h(0) - \hat{u}_h(0)), w \rangle_{\partial T_h} = (\text{div}(A^{-1} \varepsilon(u_0)), w)_{T_h},\]
(7.5c) \[\langle \sigma_h(0) n - \tau P_M(u_h(0) - \hat{u}_h(0)), \mu \rangle_{\partial T_h} = 0,\]
(7.5d) \[\langle \hat{u}_h(0), \mu \rangle_{\Gamma} = \langle g(0), \mu \rangle_{\Gamma},\]

for all $(\theta, w, \mu) \in V_h \times W_h \times M_h$. For notational convenience, we define the following space-time norms
\[\| \cdot \|_{[0,T]} : = \left( \int_0^T \| \cdot \|_{\ast,p}^p \, dt \right)^{1/p},\]
where $\ast$ can be replaced by $H^m(\Omega)$, $\Omega$, $T_h$, $A$, $\rho$, $\partial T_h$, $\tau$ or $\tau^{-1}$. The parameter $p$ takes values in $\{1, 2, \infty\}$, and when $p = \infty$, we consider the supreme norm in time instead of $L^p$ integration. Now we define the projections and the remainder terms for all $t \geq 0:$

\[(\Pi\sigma(t), \Pi u(t)) := \prod_{K \in T_h} \Pi(\sigma(t)|_{K}, u(t)|_{K}, \tau|_{\partial K}), \quad \delta(t) := \prod_{K \in T_h} R(\sigma(t)|_{K}, u(t)|_{K}, \tau|_{\partial K}).\]

The related error and approximation terms (for all $t \geq 0$) are denoted
\[\varepsilon_h^\sigma(t) := \Pi\sigma(t) - \sigma_h(t), \quad \varepsilon_h^u(t) := \Pi u(t) - u_h(t), \quad \varepsilon_h^\hat{u}(t) := P_M u(t) - \hat{u}_h(t),\]
\[e_\sigma(t) := \Pi\sigma(t) - \sigma(t), \quad e_u(t) := \Pi u(t) - u(t).\]

For the rest of this section, the wigglers sign $\lesssim$ will hide constants independent of $h$ and $T$. The main results in this section are Theorems (7.1) and (7.2).
Theorem 7.1. For $k \geq 1$, the following estimates hold
\[
\|\varepsilon_h^u(T)\|_A + \|P_M(\varepsilon_h^u(T) - \varepsilon_h^u(T))\|_\tau + \|\tilde{\varepsilon}_h^u(T)\|_\rho \\
\leq \|\varepsilon_\sigma(0)\|_A + \|\dot{\varepsilon}_\sigma\|_{A_{\Omega,1}} + \|\tilde{\varepsilon}_u\|_{\rho_{\Omega,1}} + \|\delta\|_{\tau^{-1,\infty}} + \|\tilde{\delta}\|_{\tau^{-1,1}},
\]
\[
\|\varepsilon_h^u(T)\|_A + \|P_M(\varepsilon_h^u(T) - \varepsilon_h^u(T))\|_\tau + \|\tilde{\varepsilon}_h^u(T)\|_\rho \\
\leq \|\dot{\varepsilon}_\sigma(0)\|_A + \|\dot{\varepsilon}_u(0)\|_\rho + \|\tilde{\varepsilon}_\sigma\|_{A_{\Omega,1}} + \|\tilde{\varepsilon}_u\|_{\rho_{\Omega,1}} + \|\delta\|_{\tau^{-1,\infty}} + \|\tilde{\delta}\|_{\tau^{-1,1}}.
\]

For the next theorem, we need to assume that the elliptic regularity condition
\[
\Phi \in H^1(\Omega; \mathbb{R}^3)
\]
holds for any $\Phi \in H^1(\Omega; \mathbb{R}^3)$ such that the right hand side of the above inequality is finite. Note that (7.6) is the same as (5.4). We rephrase it here to have it in the form we will use it.

Theorem 7.2. If $k \geq 1$ and (7.6) holds, then
\[
\|\varepsilon_h^u(T)\|_\Omega \leq h(1 + T)^2 \left( \|\varepsilon_\sigma(0)\|_\Omega + \|\dot{\varepsilon}_\sigma\|_{\Omega_{\tau,\infty}} + \|\tilde{\varepsilon}_\sigma\|_{\Omega_{\tau,\infty}} \\
+ \|\delta(0)\|_{\tau^{-1}} + \|\tilde{\delta}\|_{\tau^{-1,\infty}} + \|\tilde{\delta}\|_{\tau^{-1,\infty}} \right) \\
+ (1 + T)^2 \left( \|\varepsilon_\sigma(0)\|_\Omega + \|\dot{\varepsilon}_\sigma\|_{\Omega_{\tau,\infty}} + \|\dot{\varepsilon}_\sigma\|_{\Omega_{\tau,\infty}} \right).
\]

Therefore, for $1 \leq m \leq k + 1$,
\[
\|\varepsilon_h^u(T)\|_\Omega \leq h^{m+1}(1 + T)^2 \sum_{i=0}^3 \left( \|\sigma^{(i)}\|_{H^m(\Omega),\infty} + \|u^{(i)}\|_{H^{m+1}(\Omega),\infty} \right).
\]

We next give the proofs for the above two theorems in the following two subsections respectively.

7.2. Energy estimates. In this subsection, we give a proof to Theorem 7.1. We first present two lemmas, which give the error equations when $t \geq 0$ and the error equations when $t = 0$, respectively.

Lemma 7.3. For all $t \geq 0$, the following error equations
\[
(7.7a) \quad (A\varepsilon_h^u(t), \theta)_{T_h} + (\varepsilon_h^u(t), \div \theta)_{T_h} - \langle \varepsilon_h^u(t), \theta\nabla \varepsilon_h^u(t) \rangle_{\partial T_h} = (A\varepsilon_\sigma(t), \theta)_{T_h},
\]
\[
(7.7b) \quad (\rho \varepsilon_h^u(t), w)_{T_h} - (\div \varepsilon_h^u(t), w)_{T_h} + \langle \tau P_M(\varepsilon_h^u(t) - \varepsilon_h^u(t)), w \rangle_{\partial T_h} = (\rho \varepsilon_u(t), w)_{T_h} + \langle \delta(t), w \rangle_{\partial T_h},
\]
\[
(7.7c) \quad \langle \varepsilon_h^u(t), n - \tau(\varepsilon_h^u(t) - \varepsilon_h^u(t)), \mu \rangle_{\partial T_h \setminus \Gamma} = -\langle \delta(t), \mu \rangle_{\partial T_h \setminus \Gamma},
\]
\[
(7.7d) \quad \langle \varepsilon_h^u(t), \mu \rangle_{\Gamma} = 0,
\]
\[
(7.7e) \quad \varepsilon_h^u(0) = \Pi u_0 - u_h(0),
\]
\[
(7.7f) \quad \dot{\varepsilon}_h^u(0) = 0,
\]
hold for all $(\theta, w, \mu) \in V_h \times W_h \times M_h$.

Proof. Use (7.1), (7.2), (7.3), (7.5), and Theorem 2.1. \qed
Lemma 7.4. The following error equations

\begin{align}
(7.8a) & \quad (\mathcal{A} \varepsilon_h^0(t), \theta)_{\mathcal{T}_h} + (\varepsilon_h^0(t), \text{div} \theta)_{\mathcal{T}_h} - \langle \varepsilon_h^0(t), \theta n \rangle_{\partial \mathcal{T}_h} = (\mathcal{A} e_\sigma(0), \theta)_{\mathcal{T}_h}, \\
(7.8b) & \quad - (\text{div} \varepsilon_h^0(t), w)_{\mathcal{T}_h} + \langle \tau P_M (e_h^0(0) - \varepsilon_h^0(t)), w \rangle_{\partial \mathcal{T}_h} = \langle \delta(0), w \rangle_{\partial \mathcal{T}_h}, \\
(7.8c) & \quad \langle \varepsilon_h^0(t), n - \tau (\varepsilon_h^u(0) - \varepsilon_h^0(0)), \mu \rangle_{\partial \mathcal{T}_h \setminus \Gamma} = - \langle \delta(0), \mu \rangle_{\partial \mathcal{T}_h \setminus \Gamma}, \\
(7.8d) & \quad \langle \varepsilon_h^0(t), \mu \rangle_{\Gamma} = 0,
\end{align}

hold for all \((\theta, w, \mu) \in V_h \times W_h \times M_h\).

Proof. Use (7.4), (7.5), and Theorem 2.1

The next proposition gives estimates to the error terms when \(t = 0\).

Proposition 7.5. The following estimates hold:

\begin{align}
(7.9) & \quad \| \varepsilon_h^0(0) \|_A^2 + \| P_M (\varepsilon_h^0(t) - \varepsilon_h^0(0)) \|_T^2 \leq \| e_\sigma(0) \|_A^2 + \| \delta(0) \|_{\mathcal{T}_h}^2, \\
(7.10) & \quad \| \dot{\varepsilon}_h^0(0) \|_A^2 + \| P_M (\dot{\varepsilon}_h^0(t) - \dot{\varepsilon}_h^0(0)) \|_T^2 \leq \| \dot{e}_\sigma(0) \|_A^2 + \| \dot{\delta}(0) \|_{\mathcal{T}_h}^2, \\
(7.11) & \quad \| \ddot{\varepsilon}_h^0(0) \|_\rho \leq \| \ddot{e}_u(0) \|_\rho.
\end{align}

Proof. Taking \(\theta = \varepsilon_h^0(0), w = \varepsilon_h^0(0)\) and \(\mu = \varepsilon_h^0(0)\) in the error equations (7.8) and adding the equations, we have

\[ \| \varepsilon_h^0(0) \|_A^2 + \| P_M (\varepsilon_h^0(t) - \varepsilon_h^0(0)) \|_T^2 = (\mathcal{A} e_\sigma(0), \varepsilon_h^0(0))_{\mathcal{T}_h} + \langle \delta(0), \varepsilon_h^0(t) - \varepsilon_h^0(0) \rangle_{\partial \mathcal{T}_h}. \]

The first estimate (7.9) then follows from the latter identity.

Consider the error equations (7.7). We take the first order derivative of (7.7a) - (7.7c) and test the equations with \(\theta = \dot{\varepsilon}_h^0, w = \dot{\varepsilon}_h^0, \mu = \ddot{\varepsilon}_h^0\). We next add the equations, evaluate them at \(t = 0\), and use the fact that \(\dot{\varepsilon}_h^0(0) = 0\). Then

\[ \| \dot{\varepsilon}_h^0(0) \|_A^2 + \| P_M (\dot{\varepsilon}_h^0(t) - \dot{\varepsilon}_h^0(0)) \|_T^2 = (\mathcal{A} \dot{e}_\sigma(0), \dot{e}_h^0(0))_{\mathcal{T}_h} + \langle \dot{\delta}(0), \dot{e}_h^0(t) - \dot{e}_h^0(0) \rangle_{\partial \mathcal{T}_h}, \]

from which the second estimate (7.10) follows.

Finally, taking \(t = 0\) in (7.7b) and subtracting (7.8b), then taking \(w = \ddot{e}_h^0(0)\), we have

\[ (\ddot{\varepsilon}_h^0(0), \ddot{e}_h^0(0))_{\rho} = (\ddot{e}_u(0), \ddot{e}_h^0(0))_{\rho}, \]

which implies (7.11).

Note that the boundary remainder operator \(R\) defined in Theorem 2.1 is linear, thus commuting with the time derivative

\[ \dot{\delta}(t) = R(\sigma(t), \dot{u}(t); \tau). \]

This commutativity holds for the projection II as well for similar reasons.

Proposition 7.6. For \(t \geq 0\), we have

\begin{align}
12 dt & \quad \left( \| \varepsilon_h^0(t) \|_A^2 + \| P_M (\varepsilon_h^u(t) - \varepsilon_h^0(t)) \|_T^2 + \| \dot{\varepsilon}_h^0(t) \|_\rho \right) \\
= & \quad (\rho \ddot{e}_u(t), \ddot{e}_h^0(t)) + (\mathcal{A} \dot{e}_\sigma(t), \ddot{e}_h^0(t)) + \langle \dot{\delta}(t), \ddot{e}_h^0(t) - \dot{e}_h^0(t) \rangle, \\
12 dt & \quad \left( \| \dot{e}_h^0(t) \|_A^2 + \| P_M (\dot{e}_h^u(t) - \dot{e}_h^0(t)) \|_T^2 + \| \ddot{e}_h^0(t) \|_\rho \right) \\
= & \quad (\rho \dddot{e}_u(t), \dddot{e}_h^0(t)) + (\mathcal{A} \ddot{e}_\sigma(t), \dddot{e}_h^0(t)) + \langle \dddot{\delta}(t), \dddot{e}_h^0(t) - \ddot{e}_h^0(t) \rangle.
\end{align}
Proof. Taking the first order derivative of (7.7a) and testing it with \( \theta = \varepsilon^\mu_h(t) \), then choosing \( w = \dot{\varepsilon}^\mu_h(t) \) in (7.7b) and \( \mu = \dot{\varepsilon}^\mu_h(t) \) in (7.7c), and finally taking the first order derivative of (7.7d) and adding the equations, we obtain (7.12).

Taking the second order derivative of (7.7a) and testing it with \( \theta = \varepsilon^\mu_h(t) \), then taking the first order derivative of (7.7b) and testing it with \( w = \ddot{\varepsilon}^\mu_h(t) \), taking the first order derivative of (7.7c) and testing it with \( \mu = \ddot{\varepsilon}^\mu_h(t) \), and finally taking the second order derivative of (7.7d) and adding the equations, we obtain (7.13).

The final ingredient we need is a Grönwall type inequality.

**Lemma 7.7.** Suppose \( \phi, \beta, l \) are continuous and positive functions defined on \([0, \infty)\) and \( r \) is a constant. If

\[
\phi^2(t) \leq r + 2 \int_0^t \phi(s) \beta(s) ds + \phi(t)l(t),
\]

then

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

**Proof.** Consider the interval \([0, t]\) and let \( \phi(t^*) \) maximize \( \phi \) in the interval. Then we have

\[
\phi^2(t^*) \leq r + \phi(t^*) \left( 2 \int_0^t \beta(s) ds + \sup_{s \in [0,t]} l(s) \right)
\]

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

\( \square \)

With Proposition 7.5, Proposition 7.6, and Lemma 7.7, we are ready to prove Theorem 7.1.

**Proof of Theorem 7.1.** Integrating (7.12) from 0 to \( T \), we have

\[
\frac{1}{2} \left( \| \varepsilon^\mu_h(t) \|_A^2 + \| P_M(\varepsilon^\mu_h(t) - \varepsilon^\mu_h(T)) \|_\tau^2 + \| \dot{\varepsilon}^\mu_h(t) \|_\rho^2 \right)_{t=0}^{t=T} = \langle \delta(t), P_M(\varepsilon^\mu_h(t) - \varepsilon^\mu_h(T)) \rangle_{\partial \mathcal{T}_h} \bigg|_{t=0}^{t=T}
\]

\[
+ \int_0^T \left( \langle \rho \dot{\varepsilon}_u(t), \dot{\varepsilon}^\mu_h(t) \rangle_{\mathcal{T}_h} + (A \varepsilon_{\sigma}(t), \varepsilon^\mu_h(t))_{\mathcal{T}_h} - \langle \delta(t), P_M(\varepsilon^\mu_h(t) - \varepsilon^\mu_h(T)) \rangle_{\partial \mathcal{T}_h} \right) dt
\]

\[
\leq \| \delta(T) \|_{\tau^{-1}} \| P_M(\varepsilon^\mu_h(T) - \varepsilon^\mu_h(T)) \|_{\tau} + \| \delta(0) \|_{\tau^{-1}} \left( \| \varepsilon_{\sigma}(0) \|_A + \| \delta(0) \|_{\tau^{-1}} \right)^{1/2}
\]

\[
+ \int_0^T \left( \| \varepsilon^\mu_h(t) \|_A^2 + \| \varepsilon^\mu_h(t) \|_A^2 + \| P_M(\varepsilon^\mu_h(t) - \varepsilon^\mu_h(T)) \|_\tau^2 \right)^{1/2}
\]

\[
\left( \| \varepsilon_u(t) \|_\rho^2 + \| \varepsilon_{\sigma}(t) \|_A^2 + \| \delta(t) \|_{\tau^{-1}} \right)^{1/2} dt,
\]

where we used (7.3) to estimate \( \| P_M(\varepsilon^\mu_h(0) - \varepsilon^\mu_h(T)) \|_{\tau} \) in the last step.
Now we define
\[
\phi^2(t) := \|\hat{\epsilon}_h^u(t)\|_p^2 + \|\varepsilon_h^\sigma(t)\|_A^2 + \|P_M(\hat{\epsilon}_h^u(t) - \hat{\epsilon}_h^u(t))\|_{\tau}^2,
\]
\[
\beta^2(t) := \|\hat{\epsilon}_u(t)\|_p^2 + \|\varepsilon_h^\sigma(t)\|_A^2 + \|\delta(t)\|_{\tau-1}^2,
\]
\[
l(t) := 2\|\delta(t)\|_{\tau-1},
\]
\[
r := 2\|\varepsilon_\sigma(0)\|_A^2 + 3\|\delta(0)\|_{\tau-1}^2.
\]
Note that by (7.9) and (7.11), we have \(\phi^2(0) \leq \|\varepsilon_\sigma(0)\|_A^2 + \|\delta(0)\|_{\tau-1}^2\), and therefore,
\[
\phi^2(T) \leq r + 2 \int_0^T \phi(t)\beta(t)dt + \phi(T)l(T).
\]
By Lemma 7.7 we have
\[
\phi^2(T) \leq 2r + \left(2 \int_0^T \beta(t)dt + \sup_{t \in [0,T]} l(t)\right)^2 \leq r + \sup_{t \in [0,T]} l^2(t) + \left(\int_0^T \beta(t)dt\right)^2
\]
\[
\leq \|\delta(0)\|_{\tau-1}^2 + \|\varepsilon_\sigma(0)\|_A^2 + \left(\|\delta\|_{0,T}^2\right)^2
\]
\[
+ \left(\int_0^T \left(\|\hat{\epsilon}_u(t)\|_p + \|\varepsilon_h^\sigma(t)\|_A + \|\delta(t)\|_{\tau-1}\right) dt\right)^2
\]
\[
\leq \left(\|\varepsilon_\sigma(0)\|_A + \|\delta\|_{0,T} + \|\hat{\varepsilon}_\sigma\|_A + \|\delta\|_{\tau-1}\right)^2.
\]
Similarly, for the second estimate, we integrate (7.13) from 0 to \(T\), then use (7.10) to estimate \(\|P_M(\hat{\epsilon}_h^u(0) - \hat{\epsilon}_h^u(0))\|_{\tau}\), and obtain
\[
\frac{1}{2} \left(\|\hat{\epsilon}_h^u(t)\|_A^2 + \|P_M(\hat{\epsilon}_h^u(t) - \hat{\epsilon}_h^u(t))\|_{\tau}^2 + \|\hat{\epsilon}_h^u(t)\|_p^2\right)_{t=0}^{t=T}
\]
\[
\leq \|\delta(T)\|_{\tau-1} \|P_M(\hat{\epsilon}_h^u(T) - \hat{\epsilon}_h^u(T))\| + \|\delta(0)\|_{\tau-1} \left(\|\hat{\varepsilon}_u(0)\|_A^2 + \|\delta(0)\|_{\tau-1}^2\right)^{1/2}
\]
\[
+ \int_0^T \left(\|\varepsilon_h^u(t)\|_A^2 + \|P_M(\varepsilon_h^u(t) - \varepsilon_h^u(t))\|_{\tau}^2 + \|\varepsilon_h^u(t)\|_p^2\right)^{1/2} \left(\|\hat{\varepsilon}_u(t)\|_p^2 + \|\varepsilon_h^\sigma(t)\|_A^2 + \|\delta(t)\|_{\tau-1}^2\right)^{1/2} dt.
\]
Now we define
\[
\phi^2(t) := \|\hat{\epsilon}_h^u(t)\|_A^2 + \|P_M(\hat{\epsilon}_h^u(t) - \hat{\epsilon}_h^u(t))\|_{\tau}^2 + \|\hat{\epsilon}_h^u(t)\|_p^2,
\]
\[
\beta^2(t) := \|\hat{\epsilon}_u(t)\|_p^2 + \|\varepsilon_h^\sigma(t)\|_A^2 + \|\delta(t)\|_{\tau-1}^2,
\]
\[
l(t) := 2\|\delta(t)\|_{\tau-1},
\]
\[
r := 2\|\hat{\varepsilon}_u(0)\|_A^2 + 3\|\delta(0)\|_{\tau-1}^2 + \|\hat{\varepsilon}_u(0)\|_p^2.
\]
Since by (7.10) and (7.11) we have \(\phi^2(0) \leq \|\varepsilon_\sigma(0)\|_A^2 + \|\delta(0)\|_{\tau-1}^2 + \|\hat{\varepsilon}_u(0)\|_p^2\), and it follows that
\[
\phi^2(T) \leq r + 2 \int_0^T \phi(t)\beta(t)dt + \phi(T)l(T).
\]
Using Lemma 7.7 again we obtain the second estimate. □
7.3. **Duality argument.** In this subsection, we give a proof to Theorem 7.2 by using the duality argument. To begin with, we consider the adjoint equations of (7.1):

\[ A\Psi + \varepsilon(\Phi) = 0 \quad \text{in } \Omega \times [0, T], \]

\[ \rho \Phi + \text{div} \Psi = 0 \quad \text{in } \Omega \times [0, T], \]

\[ \gamma \Phi = 0 \quad \text{on } \Gamma \times [0, T], \]

\[ \Phi(T) = 0 \quad \text{on } \Omega, \]

\[ \Phi(T) = \varepsilon_h^w(T) \quad \text{on } \Omega. \]

For a time-dependent function \( f : [0, \infty) \to X \), we write

\[ \int_t^T f(s)ds. \]

The following proposition allows us to control the solution of (7.14) in certain energy norms. Similar results can be found in [9].

**Proposition 7.8.** The following inequality holds

\[ \|\Phi\|_{H^1(\Omega),\infty}^{[0,T]} + \|\Phi\|_{H^2(\Omega),\infty}^{[0,T]} \leq C\|\varepsilon_h^w(T)\|_\Omega. \]

If (7.6) holds, then

\[ \|\Psi\|_{H^1(\Omega),\infty}^{[0,T]} + \|\Phi\|_{H^2(\Omega),\infty}^{[0,T]} \leq C\|\varepsilon_h^w(T)\|_\Omega. \]

**Proof.** By conservation of energy we have

\[ \|\Phi(t)\|_{\rho}^2 + \|\varepsilon(t)\|_{A^{-1}}^2 = \|\varepsilon_h^w(T)\|_{\rho}^2, \]

for all \( t \in [0, T] \). Now (7.15) follows by using Korn’s second inequality.

Integrating (7.14a) and (7.14b) from \( t \) to \( T \), we have

\[ A\Psi(t) + \varepsilon(\Phi(t)) = 0, \quad \text{div} \Psi(t) = \rho \dot{\Phi}(t) - \rho \dot{\Phi}(T), \]

for all \( t \in [0, T] \). Combining the latter equations with (7.6) and (7.15), we obtain (7.16).

Now we define the dual projections and boundary remainder terms for the adjoint problem (7.14) (for all \( t \in [0, T] \))

\[ (\Pi\Psi(t), \Pi\Phi(t)) := \prod_{K \in T_h}(\Psi(t)|_K, \Phi(t)|_K; -\tau|_{\partial K}), \]

\[ \Delta(t) := \prod_{K \in T_h}(\Psi(t)|_K, \Phi(t)|_K; -\tau|_{\partial K}). \]

We also define the corresponding approximation terms

\[ e_\Phi(t) := \Pi\Phi(t) - \Phi(t), \quad e_\Psi(t) := \Pi\Psi(t) - \Psi(t). \]

**Proposition 7.9.** The following identity holds

\[ \|\varepsilon_h^w(T)\|_{\rho}^2 = \sum_{i=1}^7 T_i, \]
where

\[
T_1 := (\rho \dot{\Phi}(0), \Pi u_0 - u_h(0))_{\mathcal{T}_h}, \quad T_2 := -\int_0^T \langle \delta(t), e_{\Phi}(t) \rangle_{\partial\mathcal{T}_h} dt,
\]

\[
T_3 := -\int_0^T \langle \Delta(t), \varepsilon_h^n(t) - \varepsilon_h^u(t) \rangle_{\partial\mathcal{T}_h} dt, \quad T_4 := \int_0^T \langle \mathcal{A}\varepsilon_{\Psi}(t), \sigma(t) - \sigma_h(t) \rangle_{\mathcal{T}_h} dt,
\]

\[
T_5 := -\int_0^T \langle \mathcal{A}e_{\sigma}(t), \Psi(t) \rangle_{\mathcal{T}_h} dt, \quad T_6 := -\int_0^T \langle \rho \Phi(t), \ddot{e}_u(t) \rangle_{\mathcal{T}_h} dt,
\]

\[
T_7 := \int_0^T (\rho(\ddot{u}(t) - \ddot{u}_h(t)), e_{\Phi}(t))_{\mathcal{T}_h} dt.
\]

**Proof.** From the adjoint equations (7.14) tested with \((\varepsilon_h^n(t), \varepsilon_h^u(t), \varepsilon_h^\Phi(t))\), and from the properties of the adjoint projection (see Theorem 2.1), we obtain

\[
(\mathcal{A} \Pi \Psi(t), \varepsilon_h^n(t))_{\mathcal{T}_h} - (\Pi \Phi(t), \nabla \varepsilon_h^n(t))_{\mathcal{T}_h} + \langle P_M \Phi(t), \varepsilon_h^n(t) \nabla \rangle_{\partial\mathcal{T}_h} = (\mathcal{A} \varepsilon_{\Psi}(t), \varepsilon_h^n(t))_{\mathcal{T}_h},
\]

\[
(\nabla \Pi \Psi(t), \varepsilon_h^n(t))_{\mathcal{T}_h} + (\tau P_M (\Pi \Phi(t) - \Phi(t)), \varepsilon_h^n(t))_{\mathcal{T}_h} = -\langle \Delta(t), \varepsilon_h^n(t) \rangle_{\partial\mathcal{T}_h} - \langle \rho \ddot{\Phi}(t), \varepsilon_h^n(t) \rangle_{\mathcal{T}_h},
\]

\[
-\langle \Pi \Psi(t) \nabla + \tau (\Pi \Phi(t) - \Phi(t)), \varepsilon_h^u(t) \rangle_{\partial\mathcal{T}_h \setminus \Gamma} = \langle \Delta, \varepsilon_h^u(t) \rangle_{\partial\mathcal{T}_h \setminus \Gamma},
\]

\[
\langle P_M \Phi(t), \varepsilon_h^n(t) - \tau P_M \varepsilon_h^n(t), \varepsilon_h^n(t) \rangle_{\mathcal{T}_h} = 0.
\]

(7.17) Taking now \(\theta = \Pi \Psi(t)\), \(w = \Pi \Phi(t)\), \(\mu = P_M \Phi(t)\) and \(\mu = \Pi \Psi(t) \nabla + \tau P_M (\Pi \Phi(t) - \Phi(t))\) in the error equations (7.7a to 7.7d), and comparing the two sets of equations, we have

\[
-\langle \Delta(t), \varepsilon_h^n(t) - \varepsilon_h^u(t) \rangle_{\partial\mathcal{T}_h} + (\mathcal{A} \varepsilon_{\Psi}(t), \varepsilon_h^n(t) - \rho \ddot{\Phi}(t), \varepsilon_h^n(t))_{\mathcal{T}_h}
\]

\[
= (\mathcal{A} e_{\sigma}(t), \Pi \Psi(t))_{\mathcal{T}_h} + \langle \delta(t), \Pi \Phi(t) - \Phi(t) \rangle_{\partial\mathcal{T}_h} + (\rho (\ddot{u}_h(t) - \ddot{u}(t)), \Pi \Phi(t))_{\mathcal{T}_h}.
\]

After rearranging terms, we have

\[
(\rho \ddot{\Phi}(t), \varepsilon_h^n(t))_{\mathcal{T}_h} = -\langle \Delta(t), \varepsilon_h^n(t) - \varepsilon_h^u(t) \rangle_{\partial\mathcal{T}_h} - \langle \delta(t), \Pi \Phi(t) - \Phi(t) \rangle_{\partial\mathcal{T}_h}
\]

\[
- (\rho (\ddot{u}_h(t) - \ddot{u}(t)), \Pi \Phi(t))_{\mathcal{T}_h} + (\mathcal{A} \varepsilon_{\Psi}(t), \sigma(t) - \sigma_h(t))_{\mathcal{T}_h}
\]

\[
- (\mathcal{A} e_{\sigma}(t), \Psi(t))_{\mathcal{T}_h}.
\]

Defining now

\[
\eta(t) := (\rho \ddot{\Phi}(t), \varepsilon_h^n(t))_{\mathcal{T}_h} - (\rho \Phi(t), \ddot{e}_h^n(t))_{\mathcal{T}_h},
\]

which satisfies

\[
\eta(T) = (\rho \varepsilon_h^n(T), \varepsilon_h^n(T))_{\mathcal{T}_h}, \quad \eta(0) = (\rho \ddot{\Phi}(0), \Pi u_0 - u_h(0))_{\mathcal{T}_h},
\]

due to \(\Phi(T) = 0\), \(\dot{\Phi}(T) = \varepsilon_h^n(T), \varepsilon_h^n(0) = \Pi u_0 - u_h(0)\) and \(\dot{e}_h^n(0) = 0\) (see (7.7) and (7.14)). By (7.17), we have

\[
\dot{\eta}(t) = (\rho \ddot{\Phi}(t), \varepsilon_h^n(t))_{\mathcal{T}_h} - (\rho \Phi(t), \ddot{e}_h^n(t))_{\mathcal{T}_h}
\]

\[
= -\langle \Delta(t), \varepsilon_h^n(t) - \varepsilon_h^u(t) \rangle_{\partial\mathcal{T}_h} - \langle \delta(t), e_{\Phi}(t) \rangle_{\partial\mathcal{T}_h}
\]

\[
+ (\mathcal{A} \varepsilon_{\Psi}(t), \sigma(t) - \sigma_h(t))_{\mathcal{T}_h} - (\mathcal{A} e_{\sigma}(t), \Psi(t))_{\mathcal{T}_h}
\]

\[
+ (\rho (\ddot{u}(t) - \ddot{u}_h(t)), e_{\Phi}(t))_{\mathcal{T}_h} - (\rho \Phi(t), \ddot{e}_u(t))_{\mathcal{T}_h}.
\]
Therefore
\[ \| \rho^{1/2} \varepsilon_h^n(T) \|_{T_h}^2 = \int_0^T \dot{\varepsilon}(t) dt + (\rho \Phi(0), \Pi u_0 - u_h(0))_{T_h}, \]
and the proof is finished.

The next proposition gives an estimate for the term \( T_7 \) in Proposition 7.9.

**Proposition 7.10.** If (7.6) holds, then
\[ |T_7| \leq h \| \varepsilon_h^n(T) \|_{T_h} \left( \| \varepsilon_h^n(0) \|_{T_h} + T \| \varepsilon_h^n \|_{T_h, \infty} \right) \]
\[ + \| P_M \varepsilon_h^n(0) - \varepsilon_h^n(0) \|_T + T \| \varepsilon_h^n - \varepsilon_h^n \|_{T_h, \infty} \]
\[ + h \| \dot{\varepsilon}(0) \|_{T_h} + hT \| \varepsilon(0) \|_{T_h, \infty} + \| \delta(0) \|_T + T \| \dot{\delta} \|_{T_h, \infty} \right) \]

**Proof.** In the coming arguments, for the sake of shortening some estimates, we will prove bounds in terms of the quantity
\[ \Theta(T) := \sup_{t \in [0, T]} |\Phi(t)| \Omega + \sup_{t \in [0, T]} |\Psi(t)| \Omega, \]
which we have shown in Proposition 7.8 that, assuming (7.6), we have the estimate
\[ \Theta(T) \leq \| \varepsilon_h^n(T) \|_\Omega. \]

Note that
\[ T_7 = \int_0^T (\rho \dot{\varepsilon}_h^n(t), e_\Phi(t))_{T_h} dt - \int_0^T (\rho \ddot{e}_u(t), e_\Phi(t))_{T_h} dt. \]

For the second term of (7.20), we have
\[ \left| \int_0^T (\rho \dot{e}_u(t), e_\Phi(t))_{T_h} dt \right| = \left| (\rho \dot{e}_u(0), e_\Phi(0))_{T_h} + \int_0^T (\rho \dot{e}_u(t), e_\Phi(t))_{T_h} dt \right| \]
\[ \leq h^2 \left( \| \dot{e}_u(0) \|_{T_h} + T \| \dot{e}_u \|_{T_h, \infty} \right) \Theta(T). \]

We next estimate the remaining term in (7.20). Since \( \ddot{e}_h^n(t) \in P_{k+1}(K; \mathbb{R}^3) \) for all \( K \), we have
\[ \int_0^T (\rho \ddot{e}_h^n(t), e_\Phi(t))_{T_h} dt = \int_0^T (\rho \ddot{e}_h^n(t), P_{k+1} e_\Phi(t))_{T_h} dt, \]
where, \( P_{k+1} \) is the \( \rho \)-weighted \( L^2 \) projection onto \( P_{k+1}(K; \mathbb{R}^3) \). Testing the second error equation (7.7b) with \( w = P_{k+1} e_\Phi(t) \) we have
\[ (\rho \ddot{e}_h^n(t), P_{k+1} e_\Phi(t))_{T_h} = (\text{div} \varepsilon_h^n(t), P_{k+1} e_\Phi(t))_{T_h} - (\tau P_M (\varepsilon_h^n(t) - \varepsilon_h^n(t)), P_{k+1} e_\Phi(t))_{T_h} \]
\[ + (\rho \dot{e}_u(t), P_{k+1} e_\Phi(t))_{T_h} + (\delta(t), P_{k+1} e_\Phi(t))_{T_h} \]
\[ =: Q_1(t) + Q_2(t) + Q_3(t) + Q_4(t). \]
Now we use integration by part and obtain
\[
\int_0^T Q_1(t) \, dt = \langle \Theta, T \rangle_{\mathcal{T}_h} + \int_0^T \langle \Theta, T \rangle_{\mathcal{T}_h} \, dt,
\]
\[
\int_0^T Q_2(t) \, dt = -\langle \Theta, T \rangle_{\mathcal{T}_h} + \int_0^T \langle \Theta, T \rangle_{\mathcal{T}_h} \, dt,
\]
\[
\int_0^T Q_3(t) \, dt = \langle \Theta, T \rangle_{\mathcal{T}_h} + \int_0^T \langle \Theta, T \rangle_{\mathcal{T}_h} \, dt,
\]
\[
\int_0^T Q_4(t) \, dt = \langle \Theta, T \rangle_{\mathcal{T}_h} + \int_0^T \langle \Theta, T \rangle_{\mathcal{T}_h} \, dt.
\]
Note that
\[
P_{k+1} \mathbf{e}_q(t) = P_{k+1} (\Pi \Phi(t) - \Phi(t)) = \Pi \Phi(t) - P_{k+1} \Phi(t) + \Phi(t).
\]
Combining the above with the convergence properties about \( \Pi \Psi(t) \) and \( \Pi \Phi(t) \) (see Theorem 2.1) we have
\[
\|P_{k+1} \mathbf{e}_q(t)\|_K \lesssim h^2_K (\|\Phi(t)\|_{2,K} + \|\Psi(t)\|_{1,K}),
\]
\[
\|T^{1/2} P_{k+1} \mathbf{e}_q(t)\|_\mathcal{K} \lesssim h_K (\|\Phi(t)\|_{2,K} + \|\Psi(t)\|_{1,K}).
\]
Now back to the estimate of \( \int_0^T Q_i(t) \, dt \), we have
\[
\int_0^T Q_1(t) \, dt \lesssim h(\|\mathbf{e}_q(0)\|_{\mathcal{T}_h} + T\|\dot{\mathbf{e}}_q\|_{[0,T]} \|\Theta(T)\|_{\mathcal{T}_h} + T\|\dot{\Theta}(T)\|_{\mathcal{T}_h}),
\]
\[
\int_0^T Q_2(t) \, dt \lesssim h(\|P_M \mathbf{e}_q(0) - \dot{\mathbf{e}}_q(0)\|_{\mathcal{T}_h} + T\|P_M \dot{\mathbf{e}}_q - \ddot{\mathbf{e}}_q\|_{[0,T]} \|\Theta(T)\|_{\mathcal{T}_h} + T\|\dot{\Theta}(T)\|_{\mathcal{T}_h}),
\]
\[
\int_0^T Q_3(t) \, dt \lesssim h^2 (\|\dot{\mathbf{e}}_q(0)\|_{\mathcal{T}_h} + T\|\ddot{\mathbf{e}}_q\|_{[0,T]} \|\Theta(T)\|_{\mathcal{T}_h} + T\|\dot{\Theta}(T)\|_{\mathcal{T}_h}),
\]
\[
\int_0^T Q_4(t) \, dt \lesssim h(\|\delta(0)\|_{\tau^{-1}} + T\|\dot{\delta}(0)\|_{[0,T]} \|\Theta(T)\|_{\mathcal{T}_h} + T\|\dot{\Theta}(T)\|_{\mathcal{T}_h}),
\]
where we used the fact that \( \|\mathbf{e}_q(0)\|_{\mathcal{T}_h} \lesssim h^{-1}_K \|\mathbf{e}_q\|_K \) for any \( \mathbf{e}_q \in \mathcal{P}_k(K; \mathbb{R}^{3 \times 3}) \). Finally, we use (7.19) to bound \( \Theta(T) \) and the proof is finished. 

Now we are ready to prove Theorem 7.2.

Proof of Theorem 7.2. Consider Proposition 7.9. We will give estimates for the terms \( T_i \) for \( i = 1 \rightarrow 7 \).

For \( T_1, \) by (7.13) we have
\[
|T_1| = \left| \langle \rho \dot{\Phi}(0), \Pi \mathbf{u}_0 - \mathbf{u}_h(0) \rangle_{\mathcal{T}_h} \right| \lesssim \|\mathbf{e}_q(T)\|_{\mathcal{T}_h} (\|\Pi \mathbf{u}_0 - \mathbf{u}_0\|_{\mathcal{T}_h} + \|\mathbf{u}_0 - \mathbf{u}_h(0)\|_{\mathcal{T}_h}).
\]
Note that \( \mathbf{u}_h(0) \) is the solution of the HDG+ scheme (7.5). By Theorem 5.1 in Section 5.1 we have
\[
\|\mathbf{u}_h(0) - \mathbf{u}_0\|_{\mathcal{T}_h} \lesssim h(\|\mathbf{e}_q(0)\|_{\mathcal{T}_h} + \|\delta(0)\|_{\tau^{-1}}).
\]
Therefore
\[ \left| (\rho \Phi(0), \Pi u_0 - u_h(0))_{\mathcal{V}_h} \right| \leq \| \varepsilon_h^{\mu}(T) \|_{\mathcal{V}_h} (\| e_\sigma(0) \|_{\mathcal{V}_h} + h \| e_\sigma(0) \|_{\mathcal{V}_h} + h \| \delta(0) \|_{\mathcal{T}_{-1}}). \]

For $T_2$, we have
\[
|T_2| = \left| \int_0^T \langle \delta(t), \Pi \Phi(t) - \Phi(t) \rangle_{\Omega_{\mathcal{V}_h}} dt \right|
\leq \| \delta(0) \|_{\mathcal{T}_{-1}} \| \Pi \Phi(0) - \Phi(0) \|_{\mathcal{V}_h} + \int_0^T \| \dot{\delta}(t) \|_{\mathcal{V}_h} \| \Pi \Phi(t) - \Phi(t) \|_{\mathcal{V}_h} dt
\leq h \| \varepsilon_h^{\mu}(T) \|_{\mathcal{V}_h} \left( \| \delta(0) \|_{\mathcal{T}_{-1}} + T \| \delta \|_{[0,T]} \right),
\]
where we used the convergence properties about $\Pi \Phi$ (by Theorem 2.1) and then equation (7.16) to bound $\Theta(T)$ (see (7.18) for the definition).

Using similar ideas, for $T_3$ we have
\[
|T_3| = \left| \int_0^T \langle \Delta(t), \varepsilon_h^{\mu}(t) - \hat{\varepsilon}^u_h(t) \rangle_{\Omega_{\mathcal{V}_h}} dt \right|
\leq \| \Delta(0) \|_{\mathcal{T}_{-1}} \| P_M(\varepsilon_h^{\mu}(0) - \hat{\varepsilon}^u_h(0)) \|_{\mathcal{V}_h} + \int_0^T \| \Delta(t) \|_{\mathcal{T}_{-1}} \| P_M(\hat{\varepsilon}^u_h(t) - \hat{\varepsilon}^u_h(t)) \|_{\mathcal{V}_h} dt
\leq h \| \varepsilon_h^{\mu}(T) \|_{\mathcal{V}_h} \left( \| P_M(\varepsilon_h^{\mu}(0) - \hat{\varepsilon}^u_h(0)) \|_{\mathcal{V}_h} + T \| P_M(\hat{\varepsilon}^u_h - \hat{\varepsilon}^u_h) \|_{[0,T]} \right),
\]
and for $T_4$, we have
\[
|T_4| = \left| \int_0^T \langle A e_{\psi}(t), \sigma(t) - \sigma_h(t) \rangle_{\mathcal{V}_h} dt \right|
\leq |(A(\Pi \Psi(0) - \Psi(0)), \sigma(0) - \sigma_h(0))_{\mathcal{V}_h}|
+ \left| \int_0^T (A(\Pi \Psi(t) - \Psi(t)), \dot{\sigma}(t) - \dot{\sigma}_h(t))_{\mathcal{V}_h} dt \right|
\leq h \| \varepsilon_h^{\mu}(T) \|_{\mathcal{V}_h} \left( \| \sigma(0) - \sigma_h(0) \|_{\mathcal{V}_h} + T \| \dot{\sigma} - \dot{\sigma}_h \|_{[0,T]} \right).
\]

For $T_5$, by (2.5) and (7.16), we have
\[
|T_5| = \left| \int_0^T (e_{\sigma}(t), \varepsilon(\Phi)(t))_{\mathcal{V}_h} dt \right|
\leq |(\Pi \sigma(0) - \sigma(0), \varepsilon(\Phi)(0) - P_0 \varepsilon(\Phi)(0))_{\mathcal{V}_h}|
+ \left| \int_0^T (\Pi \dot{\sigma}(s) - \dot{\sigma}(t), \varepsilon(\Phi)(t) - P_0 \varepsilon(\Phi)(t))_{\mathcal{V}_h} dt \right|
\leq h \| \varepsilon_h^{\mu}(T) \|_{\mathcal{V}_h} \left( \| \Pi \sigma(0) - \sigma(0) \|_{\mathcal{V}_h} + T \| \Pi \dot{\sigma} - \dot{\sigma} \|_{[0,T]} \right).
For $T_6$, we simply use (7.15) and obtain
\[
|T_6| = \left| \int_0^T (\rho \Phi, \ddot{u})_{\tau_h} \right| \leq \|\Phi\|_{\Omega_{1, \infty}} \|\ddot{u}\|_{\Omega_{1, 1}} \leq T \|\varepsilon_h^u(T)\|_{\tau_h} \|\ddot{e}_u\|_{\Omega_{1, \infty}}.
\]

Now we use Proposition 7.10 to estimate $T_7$, and finally obtain
\[
\|\varepsilon_h^u(T)\|_{\tau_h} \lesssim (\|\varepsilon_h^u(0)\|_{\tau_h} + h\|\varepsilon_h^\sigma(0)\|_{\tau_h} + h\|\delta(0)\|_{\tau^{-1}})
+ h \left( \|\delta(0)\|_{\tau^{-1}} + T\|\delta\|_{\tau^{-1: \infty}} \right)
+ h \left( \|\varepsilon_h^\sigma(0)\|_{\tau^{-1}} + T\|\varepsilon_h^\sigma\|_{\tau^{-1: \infty}} \right)
+ h \left( \|\varepsilon_h^\sigma(0)\|_{\tau^{-1}} + T\|\varepsilon_h^\sigma\|_{\tau^{-1: \infty}} \right)
+ h \left( \|\varepsilon_h^\sigma(0)\|_{\tau^{-1}} + T\|\varepsilon_h^\sigma\|_{\tau^{-1: \infty}} \right).
\]

Combining the above estimates with Proposition 7.5 and Theorem 7.1, the proof is finished. \qed

8. Numerical experiments

In this section, we present some numerical experiments to support our error estimates in Section 7. Note that there are experiments for the steady-state and time-harmonic cases in [18] and [14] respectively.

![Figure 2. Visualization of meshes used for error tests.](image)

The experiments are carried out on a cubic domain $\Omega = (0,1)^3$ and a non-convex polyhedral domain (we will refer to it as the chimney; see Figure 2), the time interval is $[0, T]$ with $T = \frac{3}{2}$, and we aim to estimate the relative $L^2$ errors
\[
E_{\varepsilon} := \frac{\|\varepsilon_h(T) - \varepsilon(T)\|_\Omega}{\|\varepsilon(T)\|_\Omega}, \quad E_u := \frac{\|u_h(T) - u(T)\|_\Omega}{\|u(T)\|_\Omega}.
\]

We consider a non-homogeneous isotropic material
\[
\mathcal{A} \sigma = \frac{1}{2\mu} \sigma - \frac{\lambda}{2\mu(2\mu + 3\lambda)} \text{tr} \sigma I,
\]
with Lamé parameters
\[
\lambda = \frac{2 + x^2 + y^2 + z^2}{1 + x^2 + y^2 + z^2}, \quad \mu = 3 + \cos(xyz),
\]
and the mass density is constant \( \rho = 1 \). As exact solution we take \( u(x, t) := U(x)H(t) \),
where
\[
U = (\cos(\pi x) \sin(\pi y) \cos(\pi z), 5x^2yz + 4xy^2z + 3xyz^2 + 17, \cos(2x) \cos(3y) \cos(z)),
\]
and the temporal part is \( H(t) = t^3(1 - t)^2 \). The input data \( f \) and \( g \) are chosen so that \( (7.1) \) is satisfied. Note that we have chosen the exact solution \( u \) that has vanishing initial conditions. This simplifies the calculations of \( u_h(0) \) and \( u_h(0) \) (they are automatically 0 by \( (7.3) \) and \( (7.5) \)). When \( u \) does not have vanishing initial conditions, the calculation of \( u_h(0) \) involves projecting \( u(0) \) to a space enriched by the \( M \)-decomposition spaces (the Cockburn-Fu filling), which we have not found an easy way to implement. Finding easier ways of calculating \( u_h(0) \) for non-vanishing initial conditions will constitute our future works.

For the numerical schemes, we use the HDG+ method \( (7.2) \) for space discretization and Trapezoidal Rule Convolution Quadrature (TRCQ) (see \[1, 16\]) for time integration. This is equivalent to using Trapezoidal Rule time-stepping in the semidiscrete system. The time interval \([0, T]\) is equally divided and each timestep is of length \( \kappa \). Since the error from the TRCQ is \( \mathcal{O}(\kappa^2) \), we choose \( \kappa \approx h^{(k+2)/2} \) so that the error from the time discretization does not pollute the order of convergence of the space discretization.

From Figure 3 or Table 1, we observe that the orders of convergence for \( \sigma(T) \) and \( u(T) \) are \( \mathcal{O}(h^{k+1}) \) and \( \mathcal{O}(h^{k+2}) \) respectively, agreeing the estimates in Theorem 7.1 and Theorem 7.2.

| \( k \) | \( h \) | Error | Order | Error | Order | Error | Order | Error | Order |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.6329 | 2.82E-1 | - | 4.32E-2 | - | 1.27E-1 | - | 3.93E-2 | - |
| 0.8164 | 7.06E-2 | 2.00 | 6.16E-3 | 2.81 | 3.12E-2 | 2.02 | 5.36E-3 | 2.87 |
| 0.4082 | 1.93E-2 | 1.87 | 5.44E-4 | 3.50 | 8.16E-3 | 1.93 | 4.92E-4 | 3.44 |
| 0.2041 | 4.79E-3 | 2.01 | 5.65E-5 | 3.27 | 1.99E-3 | 2.03 | 5.09E-5 | 3.28 |
| 2 | 1.6329 | 1.30E-1 | - | 1.84E-2 | - | 6.44E-2 | - | 1.68E-2 | - |
| 0.8164 | 1.72E-2 | 2.92 | 1.52E-3 | 3.59 | 6.83E-3 | 3.24 | 1.23E-3 | 3.78 |
| 0.4082 | 2.31E-3 | 2.89 | 5.92E-5 | 4.68 | 8.96E-4 | 2.93 | 4.74E-5 | 4.70 |
| 0.2041 | 2.86E-4 | 3.02 | 2.52E-6 | 4.55 | - | - | - | - |
| 3 | 1.6329 | 4.36E-2 | - | 9.50E-3 | - | 2.18E-2 | - | 7.74E-3 | - |
| 0.8164 | 3.90E-3 | 3.48 | 3.35E-4 | 4.83 | 1.47E-3 | 3.89 | 2.59E-4 | 4.90 |
| 0.4082 | 2.56E-4 | 3.93 | 6.34E-6 | 5.72 | 9.54E-5 | 3.94 | 4.88E-6 | 5.73 |

Table 1. History of convergence for \( \sigma(T) \) and \( u(T) \) with sequence of uniform refinements in space and over-refinements in time.

9. Conclusion and remarks

In this paper, we propose some new tools for the error analysis of HDG methods. Let us just mention here the two most important ones: (1) the extended projection constructed
by enforcing weak commutativity on a higher order polynomial space (see (3.9)); (2) the boundary remainder reflecting the discrepancy between the normal trace of the $M$-decomposition projection and the normal trace of a composite projection (see (3.17)). These tools allow us to flexibly devise projections for more variants of HDG methods. We have demonstrated this by constructing the projection for the Lehrenfeld-Schöberl type HDG (HDG+) method for elasticity and showing that we can recover the existing error estimates in a simpler analysis and prove the optimal convergence of a semi-discrete HDG+ scheme for the transient elastic waves.

Figure 3. History of convergence for $\sigma(T)$ and $u(T)$ with sequence of uniform refinements in space and over-refinements in time. Top two figures for cubic meshes and bottom two for chimney.
We remark that there are more ways of using these tools to construct projections. First of all, there is an alternative way of constructing the HDG+ projection – instead of first constructing the projection on the reference element then using a push-forward argument, we can construct the projection and the boundary remainder directly on the physical element. This is feasible since the $M$-decomposition can be constructed on general polyhedral elements [4, 3]. This alternative approach has the potential of extending our results to the more generalized setting of general polyhedral element. (This is the aim of future work.) Secondly, the HDG+ projection can be constructed for elliptic diffusion as well. We have explored this in [11]. Finally, the boundary remainder technique can be used before constructing the extended projection. To be more specific, if we construct the boundary remainder based on the space $P_{sym}^k \Sigma_{fill} \hat{P}_k$ instead of on the extended space $(P_{sym}^k \Sigma_{fill}) \times P_{k+1}$ (see Figure 1), we can construct a projection (together with the remainder) for the standard HDG for elasticity proposed in [12]. Note that in this case, the boundary remainder contributes $O(h^{k+1/2})$ to the convergence rate and cannot be compensated by the stabilization function (where $\tau = O(1)$), therefore the numerical scheme is sub-optimal. To see this more clearly, we put here the energy identity from Proposition 5.2:

\[
(\mathcal{A}e^u_h, e^\sigma_h)_{\mathcal{T}_h} + \langle \tau P_M (e^u_h - e^\tilde{u}_h), (\varepsilon^\sigma_h - e^\tilde{u}^\sigma_h) \rangle_{\partial \mathcal{T}_h} = (\mathcal{A}e^\sigma_h, e^\sigma_h)_{\mathcal{T}_h} + \langle \delta, e^u_h - e^\tilde{u}_h \rangle_{\partial \mathcal{T}_h}.
\]

With the above energy identity in mind, we next put a table for the relevant quantities of the three versions of HDG methods for linear elasticity - the $P_{sym}^k \times P_k$ standard HDG [12], the $(P_{sym}^k \Sigma_{fill}) \times P_k$ M-decomposition HDG [3], and the $P_{sym}^k \times P_{k+1}$ HDG+ [18]. A simple calculation shows that, with the help of boundary remainder, we can potentially recover the error estimates of the three versions in a unified way. No matter whether these schemes are optimal convergent or not, the boundary remainder will capture the best convergence rate these schemes can offer.

|         | $\tau$ | $e^\sigma_h$ | $\delta$ | $e^\sigma_h$ |
|---------|--------|--------------|----------|--------------|
| Standard HDG | $O(1)$ | $O(h^{k+1})$ | $O(h^{k+1/2})$ | $O(h^{k+1/2})$ |
| M-decomposition | $-$ | $O(h^{k+1})$ | $0$ | $O(h^{k+1})$ |
| HDG+     | $O(h^{k+1})$ | $O(h^{k+1})$ | $O(h^{k+1/2})$ | $O(h^{k+1/2})$ |

Table 2. Relevant quantities for three versions of HDG methods for steady state elasticity.

References

[1] Lehel Banjai. Multistep and multistage convolution quadrature for the wave equation: algorithms and experiments. *SIAM J. Sci. Comput.*, 32(5):2964–2994, 2010.
[2] Brandon Chabaud and Bernardo Cockburn. Uniform-in-time superconvergence of HDG methods for the heat equation. *Math. Comp.*, 81(277):107–129, 2012.
[3] Bernardo Cockburn and Guosheng Fu. Devising superconvergent HDG methods with symmetric approximate stresses for linear elasticity by $M$-decompositions. *IMA J. Numer. Anal.*, 38(2):566–604, 2018.
[4] Bernardo Cockburn, Guosheng Fu, and Francisco Javier Sayas. Superconvergence by $M$-decompositions. Part I: General theory for HDG methods for diffusion. *Math. Comp.*, 86(306):1609–1641, 2017.
[5] Bernardo Cockburn, Zhixing Fu, Allan Hungria, Liangyue Ji, Manuel A. Sánchez, and Francisco-Javier Sayas. Stormer-Numerov HDG methods for acoustic waves. *J. Sci. Comput.*, 75(2):597–624, 2018.

[6] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Francisco-Javier Sayas. A projection-based error analysis of HDG methods. *Math. Comp.*, 79(271):1351–1367, 2010.

[7] Bernardo Cockburn and Kassem Mustapha. A hybridizable discontinuous Galerkin method for fractional diffusion problems. *Numer. Math.*, 130(2):293–314, 2015.

[8] Bernardo Cockburn, Weifeng Qiu, and Ke Shi. Conditions for superconvergence of HDG methods for second-order elliptic problems. *Math. Comp.*, 81(279):1327–1353, 2012.

[9] Bernardo Cockburn and Vincent Quenneville-Bélair. Uniform-in-time superconvergence of the HDG methods for the acoustic wave equation. *Math. Comp.*, 83(285):65–85, 2014.

[10] Bernardo Cockburn and Ke Shi. Superconvergent HDG methods for linear elasticity with weakly symmetric stresses. *IMA J. Numer. Anal.*, 33(3):747–770, 2013.

[11] Shukai Du and Francisco-Javier Sayas. *An invitation to the theory of the Hybridizable Discontinuous Galerkin Method*. Submitted to SpringerBriefs in Mathematics.

[12] Guosheng Fu, Bernardo Cockburn, and Henryk K. Stolarski. Analysis of an HDG method for linear elasticity. *Internat. J. Numer. Methods Engrg.*, 102(3-4):551–575, 2015.

[13] Roland Griesmaier and Peter Monk. Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation. *J. Sci. Comput.*, 49(3):291–310, 2011.

[14] Allan Hungria, Daniele Prada, and Francisco-Javier Sayas. HDG methods for elastodynamics. *Comput. Math. Appl.*, 74(11):2671–2690, 2017.

[15] Christoph Lehrenfeld. Hybrid discontinuous Galerkin methods for solving incompressible flow problems. *Rheinisch-Westfälischen Technischen Hochschule Aachen*, 2010.

[16] Christian Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.*, 52(2):129–145, 1988.

[17] Isssei Oikawa. A hybridized discontinuous Galerkin method with reduced stabilization. *J. Sci. Comput.*, 65(1):327–340, 2015.

[18] Weifeng Qiu, Jiguang Shen, and Ke Shi. An HDG method for linear elasticity with strong symmetric stresses. *Math. Comp.*, 87(309):69–93, 2018.

[19] Weifeng Qiu and Ke Shi. An HDG method for convection diffusion equation. *J. Sci. Comput.*, 66(1):346–357, 2016.

[20] Weifeng Qiu and Ke Shi. A superconvergent HDG method for the incompressible Navier-Stokes equations on general polyhedral meshes. *IMA J. Numer. Anal.*, 36(4):1943–1967, 2016.

[21] Francisco-Javier Sayas. From Raviart-Thomas to HDG. *arXiv:1307.2491*, 2013.

[22] See-Chew Soon, Bernardo Cockburn, and Henryk K. Stolarski. A hybridizable discontinuous Galerkin method for linear elasticity. *Internat. J. Numer. Methods Engrg.*, 80(8):1058–1092, 2009.

**University of Delaware**

*E-mail address: shukaidu@udel.edu*

**University of Delaware**

*E-mail address: fjsayas@udel.edu*